OXFORD

Cerebral Cortex, 2022, **32**, 3206–3223
https://doi.org/10.1093/cercor/bhab410
Advance access publication date: 24 December 2021
Original Article

Influence of gonadal steroids on cortical surface area in infancy

Ann Mary Alex[1], Tom Ruvio[1], Kai Xia[2], Shaili C. Jha[3], Jessica B. Girault[2,4], Li Wang[5], Gang Li[5], Dinggang Shen[6,7], Emil Cornea[2], Martin A. Styner[2,8], John H. Gilmore[2], Rebecca C. Knickmeyer[1,2,9,10]

[1]Neuroengineering Division, Institute for Quantitative Health Sciences and Engineering, Michigan State University, East Lansing, MI 48824, USA,
[2]Department of Psychiatry, University of North Carolina Chapel Hill, Chapel Hill, NC 27599, USA,
[3]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA,
[4]Carolina Institute for Developmental Disabilities, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA,
[5]Department of Radiology and Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA,
[6]School of Biomedical Engineering, ShanghaiTech University, Shanghai 201210, China,
[7]Department of Artificial Intelligence, Korea University, Seoul 02841, Republic of Korea,
[8]Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA,
[9]Department of Pediatrics and Human Development, Michigan State University, East Lansing, MI 48824, USA,
[10]Center for Research in Autism, Intellectual, and Other Neurodevelopmental Disabilities, Michigan State University, East Lansing, MI 48824, USA

*Address correspondence to Rebecca C. Knickmeyer, Institute for Quantitative Health Science and Engineering, 775 Woodlot Dr, East Lansing, MI 48824, USA. Email: knickmey@msu.edu.

Sex differences in the human brain emerge as early as mid-gestation and have been linked to sex hormones, particularly testosterone. Here, we analyzed the influence of markers of early sex hormone exposure (polygenic risk score (PRS) for testosterone, salivary testosterone, number of CAG repeats, digit ratios, and PRS for estradiol) on the growth pattern of cortical surface area in a longitudinal cohort of 722 infants. We found PRS for testosterone and right-hand digit ratio to be significantly associated with surface area, but only in females. PRS for testosterone at the most stringent *P* value threshold was positively associated with surface area development over time. Higher right-hand digit ratio, which is indicative of low prenatal testosterone levels, was negatively related to surface area in females. The current work suggests that variation in testosterone levels during both the prenatal and postnatal period may contribute to cortical surface area development in female infants.

Key words: cortical surface area; digit ratio; neuroimaging; PRS; testosterone.

Introduction

Although men and women are similar in many ways, certain health risks and behaviors are more common in one sex as compared with the other. For example, certain psychiatric illnesses including eating disorders and depression are more common in women, while certain neurodevelopmental disorders including autism and ADHD are more common in males (Green et al. 2018). The reasons for these differences are undoubtedly complex and multifactorial. However, they may reflect, in part, sex differences in human brain structure. The adult human brain is 9–12% larger in males when compared with females (Ruigrok et al. 2014; Ritchie et al. 2018; Kaczkurkin et al. 2019). Differences in overall brain size are accompanied by substantial sex differences in cortical surface area with males having greater surface area than females. Cortical thickness, however, is higher for females (Ritchie et al. 2018). Regional sex differences in cortical volume, surface area, and thickness, and the volumes of subcortical brain structures such as the hippocampus, nucleus accumbens, amygdala, caudate nucleus, putamen, and thalamus are primarily a reflection of these global differences, although there are some regions that demonstrate sex differences even when adjusting for global measures (Ritchie et al. 2018; Wierenga, Sexton, et al. 2018b). How these differences emerge across human development, as well as the mechanisms responsible for these differences, remains an active area of investigation and is the primary focus of the current manuscript.

Some sex differences emerge during adolescence and may be linked to the surge in sex hormone production that accompanies activation of the hypothalamic–pituitary–gonadal (HPG) axis in puberty (Vigil et al. 2016). Many different research teams have reported associations between circulating sex steroids and brain morphometry, or brain morphometry change during adolescence (Neufang et al. 2009; Peper et al. 2009; Bramen et al. 2011, 2012; Nguyen et al. 2013; Herting et al. 2014, 2015; Koolschijn et al. 2014; Wierenga, Bos, et al. 2018a), though the pattern of results differs somewhat between studies in terms of the specific regions affected and in the direction of effect. Other sex differences emerge much earlier

Received: April 9, 2021. **Revised:** September 13, 2021. **Accepted:** September 14, 2021

in life. For example, sex differences in cortical volume can be observed in midgestation (Studholme et al. 2020) and 2 weeks postnatal (Dean et al. 2018; Jha et al. 2019). Male brains are about 6% larger than female brains at birth and continue to grow faster postnatally (Holland et al. 2014; Wierenga et al. 2014; Gilmore et al. 2018). As in adults, greater brain size in males appears to be more closely related to cortical surface area than cortical thickness. Approximately 4% of variation in surface area at 1 month of age can be accounted by sex (Jha et al. 2019). In contrast, sex is not a significant predictor of cortical thickness in early infancy (Jha et al. 2019), and male and female children exhibit similar trajectories of cortical thickness development when examining individual lobes or the entire brain (Raznahan et al. 2011; Koolschijn and Crone 2013).

Sex differences in the early development of cortical surface area may be the result of gonadal steroids acting in the prenatal or early postnatal period. In human males, placental choriogonadotropin induces testicular mesenchymal cells to differentiate into Leydig cells and stimulates testosterone production beginning in the ninth gestational week. The male HPG axis is active by midgestation, producing a peak in gonadotropin levels and a surge in testosterone production, with the highest levels reached between 11 and 17 weeks (Reyes et al. 1974; Shigeo et al. 1977; Tapanainen et al. 1981). Human female fetuses also experience this surge in gonadotropin production and produce higher levels of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) than human male fetuses (Clements et al. 1976; Kaplan and Grumbach 1976; Debieve et al. 2000). However, estrogen exposure is similar between the sexes, since the ovaries produce relatively little in comparison to the placenta (Lanciotti et al. 2018). Increases in placental estrogen eventually suppress the activity of the fetal HPG axis in both males and females. Immediately postbirth, males experience a brief surge in LH, followed by a transient rise in testosterone, which lasts about 12 h. About 1 week later, male testosterone levels begin to increase again, peaking at 1–3 months (Forest et al. 1973; Andersson et al. 1998; Kuiri-Hänninen et al. 2011). In females, estradiol increases postbirth and remains high until at least the sixth month of life (Schmidt et al. 2002; Chellakooty et al. 2003; Kuiri-Hänninen et al. 2013). This transient activation of the HPG axis in the first 6 months of infancy is referred to as the minipuberty in both males and females (Lanciotti et al. 2018; Vasung et al. 2019; Becker and Hesse 2020), and its potential effect on brain development is poorly understood.

A rich body of preclinical research, primarily conducted in rodents, shows that prenatal and neonatal sex steroids influence numerous neurodevelopmental processes and are ideal candidates for exerting epigenetic effects on the developing brain (McCarthy and Crews 2008). However, complementary studies in human infants are lacking. This is, in large part, due to the challenges of measuring hormonal exposure in the prenatal and early postnatal period, as well as the challenges of imaging infants and young children. There are several small studies that attempted to link measurements of the prenatal hormonal milieu to neuroimaging phenotypes assessed in late childhood or adulthood. Kallai et al. (2005) reported that the 2D:4D digit ratio, a putative indicator of the intrauterine ratio of testosterone to estrogen, was associated with the volume of certain hippocampal subregions in 40 adult women but was not associated with volumes of the amygdala or cerebral cortex (Kallai et al. 2005). Darnai et al. (2016) reported significant positive correlations between digit ratio and total cerebral cortex, total cerebellar white matter, and total cerebellar cortex in 32 adult males but did not observe significant correlations in females for any of the variables they examined, which included total intracranial volume, total cerebral white matter, total cerebral cortex, total cerebellar white matter, total cerebellar cortex, and total ventricular volume (Darnai et al. 2016). Lombardo et al. (2012) reported that amniotic testosterone levels predict increased local gray matter (GM) volume in the right temporoparietal junction/posterior superior temporal sulcus (RTPJ/pSTS) and bilateral somatosensory, motor, and premotor cortex, and decreased local GM volume along both sides of the Sylvian fissure in 28 boys between 8 and 11 years of age (Lombardo et al. 2012). Our own group has tested whether 2D:4D digit ratio or efficiency of the androgen receptor, indexed by the number of CAG repeats in androgen receptor (AR) gene, predicts global and local brain volumes around 2 weeks of age and found minimal, sex-specific effects on local gray matter volume (Knickmeyer, Wang, Zhu, Geng, Woolson, Hamer, Konneker, Styner, et al. 2014b). We have also reported that genetic variation in the estrogen receptor is associated with neonatal brain volume (Knickmeyer, Wang, Zhu, Geng, Woolson, Hamer, Konneker, Lin, et al. 2014a).

The current study significantly expands upon our earlier work by evaluating five different markers of early hormonal exposure in relation to cortical surface area (SA) in a large, longitudinal cohort of infants. These markers include 1) a polygenic risk score (PRS) for serum testosterone, 2) salivary testosterone levels measured at 3 months of age, 3) number of CAG repeats in the androgen receptor (AR) gene, 4) a PRS for serum estradiol, and 5) 2D:4D digit ratios. We focused on cortical surface area because it shows robust sex differences in this cohort and, similar to adults, appears to be the primary driver of sex differences in overall brain volumes in infancy (Jha et al. 2019). We hypothesized that these markers would be associated with cortical surface area measured at 1 month postbirth, 1 year, and 2 years of age.

Materials and Methods

Participants

A total of 722 infants are included in this study (367 singletons and 355 twins). All infants were participants

in the UNC Early Brain Development Study (EBDS), a prospective, longitudinal cohort study that has collected neuroimaging data around 1, 12, and 24 months of age (Knickmeyer et al. 2008, 2017; Gilmore et al. 2010). Some EBDS participants also provided buccal swabs for the analysis of DNA, saliva samples for hormone analysis, and/or black and white photocopies of the left- and right hand for the calculation of the 2D:4D digit ratio, an anthropometric proxy for prenatal gonadal steroid exposure. Infants with these ancillary samples and high-quality measurements of cortical surface area from at least one study visit are the focus of the current report. Mothers were recruited during the second trimester of pregnancy from the outpatient obstetrics and gynecology clinics at UNC hospitals. Exclusion criteria at enrollment were the presence of abnormalities on fetal ultrasound or major medical or psychotic illness in the mother. Written consent of each subject's mother or father was obtained prior to experimental procedures. The study was conducted with the approval of the Institutional Review Board of the University of North Carolina (UNC) School of Medicine and the Institutional Review Board of Michigan State University and complies with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

MRI Data Acquisition and Processing

Images were acquired on a Siemens Allegra head-only 3T scanner [$N = 456$ (1 month); $N = 331$ (12 months); $N = 245$ (24 months)] or a Siemens TIM Trio 3T scanner [$N = 84$ (1 month); $N = 69$ (12 months); $N = 62$ (24 months)] (Siemens Medical Systems, Erlangen, Germany). Children were scanned unsedated while asleep, fitted with ear protection, and with their heads secured in a vacuum-fixation device. T1-weighted images, which are used for cortical reconstruction at 12 and 24 months of age, were obtained with 3D magnetization prepared rapid gradient echo sequences (MP-RAGE). T1 acquisition parameters for the Allegra were TR = 1820 ms, inversion time = 1100 ms, echo time = 4.38 ms, flip angle = 7°, resolution = 1 × 1 × 1 mm ($N = 576$). T1 acquisition parameters for the Trio were TR = 1860–1900 ms, TE = 3.74 ms, flip angle = 7°, spatial resolution = 1 mm × 1 mm × 1 mm ($N = 131$). Proton density and T2-weighted images, which are used for cortical reconstruction at 1 month, were obtained on the Allegra with turbo-spin echo sequences. Our standard TSE parameters were TR = 6200 ms, TE1 = 20 ms, TE2 = 119 ms, flip angle = 150°, resolution = 1.25 × 1.25 × 1.95 mm ($N = 225$). A "fast" turbo-spin echo sequence was collected on the Allegra for neonates who had trouble sleeping through the scan session, using a decreased TR, a smaller image matrix, and fewer slices (TR range = 5270–5690 ms, TE1 range = 20–21 ms, TE2 range = 119–124 ms, flip angle = 150°, spatial resolution = 1.25 mm × 1.25 mm × 1.95 mm, $N = 226$). For acquisition in TIM Trio, some children were scanned using a TSE protocol (TR = 6200 ms, TE1 = 17, TE2 = 116 ms, flip angle = 150°, spatial resolution = 1.25 mm × 1.25 mm × 1.95 mm, $N = 8$), while the rest were scanned using a 3-D T2 SPACE protocol (TR = 3200 ms, TE = 406, flip angle = 120°, spatial resolution = 1 mm × 1 mm × 1 mm, $N = 79$). All images in this study were visually checked and rated for motion artifacts using a 4-point visual scale based on (Blumenthal et al. 2002), where none = 1, mild = 2, moderate = 3, and severe = 4.

For deriving SA measures, a previously described pipeline was used (Li et al. 2016; Jha et al. 2019). Images were resampled into $1 \times 1 \times 1$ mm^3 resolution, skull-stripped, and then had the cerebellum and brain stem removed (Shi et al. 2012). A standalone infant-specific patch–driven coupled level sets method for segmentation of gray matter, white matter, and cerebrospinal fluid (CSF) during infancy was jointly applied (Wang et al. 2014). Inner and outer cortical surfaces were first constructed using the process described in Li et al. with the inner cortical surface defined by the interface between white matter and gray matter, and the outer cortical surface defined by the interface between gray matter and CSF (Li et al. 2012, 2014). SA was measured after the surface reconstruction. There were 615 infants with high-quality measurements of global SA at 1 month, 410 infants with high-quality measurements of global SA at 12 months, and 314 infants with high-quality measurements of global SA at 24 months.

Salivary Testosterone Measurement

Saliva was collected approximately 3 months postbirth ($N = 291$). Collection visits were scheduled for 9:00 in the morning. About 1 mL of passive drool was collected from each participant using a suction catheter (Centurion Healthcare Products, Howell, MO). All samples were frozen within 4 h and stored in an −80 °C freezer. Enzyme immunoassay using a commercially available kit (Salimetrics, State College, PA) was used for the measurement of salivary testosterone. The intra-assay precision for samples with low testosterone levels (mean 18.12 pg/mL) is 6.7%; for high testosterone levels (mean 188.83 pg/mL), it is 2.5%. Interassay precision for samples with low testosterone levels (mean 19.6 pg/mL) is 14.05%; for high testosterone levels (mean 199.08 pg/mL), it is 5.6%. Percent recovery for this assay varies from 92% to 111.4%. The minimal concentration of testosterone that can be distinguished from 0 is <1.0 pg/mL. The saliva–serum correlation is stronger for males, $r = 0.91$, than for females, $r = 0.61$ (Nahoul et al. 1986). All samples were evaluated for blood contamination using the Salimetrics Salivary Blood Contamination Enzyme Immunoassay kit, which quantitatively measures transferrin. Transferrin is a large protein that is present in abundance in blood, but that is normally present in only trace amounts in saliva. Intra-assay precision for samples with high (3.88 mg/dL) transferrin levels is 10.2%, and for samples with low (0.42 mg/dL) transferrin levels, it is 4.9%. Interassay precision is 7.1% for low (1.02 mg/dL) and 7.2% for high (4.93 mg/dL) transferrin levels. Percent recovery

varies from 91.9% to 101.5%. The minimal concentration of transferrin that can be distinguished from zero is 0.08 mg/dL. Mean transferrin levels in the samples were higher than expected, which suggests a possibility for blood contamination. Furthermore, the untransformed transferrin variable showed a high level of skewness. To overcome these issues, we excluded three individuals who were outliers for transferrin level (above 6.68) and included log(10) transferrin as a covariate in all analyses that included salivary testosterone.

Digit Ratio Measurements

Digit ratio measurements were taken at three visits at 1 month postbirth ($N = 436$), at 12 months of age ($N = 354$), and at 24 months ($N = 352$). Black and white photocopies of both the left- and right hand of participating children were collected. Measurements of 2nd and 4th digit length were taken from photocopies using Vernier calipers with an accuracy of ±0.02 mm and repeatability of 0.01 mm. The 2D:4D ratio was calculated from digit length measured from the basal crease of the digit proximal to the palm to the tip of the digit. Each digit was measured twice with a minimum of 1 day between measurements. Two raters performed the 2D:4D measurements. The raters were blind to subjects' sex. While the raters were not given explicit information as regards subject age or ethnicity, the photocopies themselves provide some limited information relevant to these parameters (e.g., size and shape of hand and skin tone). Right versus left hand could be determined from the photocopies themselves. The measurement had high inter-rater reliability for both right- and left hand as evidenced from ICC at 1 month (0.89, 0.87), 12 months (0.91, 0.87), and 24 months (0.91, 0.9). ICC for comparison of right- and left-hand digit ratios across years showed moderate correlation at 1 month (0.39), 12 months (0.39), and 24 months (0.36). ICC across ages for the right-hand digit ratio was 0.41 and the left-hand digit ratio 0.37, which suggests a medium correlation over the time points.

Genotyping

CAG Repeats

CAG repeats were genotyped using buccal DNA samples. DNA was extracted using standard methods as described in the Puregene DNA Purification Kit (Gentra Systems) using supplies from Qiagen ($N = 401$). After extraction, samples were aliquoted into 2 (23ul) tubes and stored at −80 °C. Prior to freezing, DNA quantity and quality (indexed by the 260/280 nm ratio) were assessed by spectrophotometer (Beckman DU640, Beckman-Coulter, Brea, CA). CAG repeat polymorphism was genotyped using the protocol adapted from Allen et al. (1992). First, a PCR spanning the CAG repeat polymorphism was performed. The forward primer was 1, 5′-FAM-GCTGTGAAGGTTGCTGTTCCTCAT-3′; the reverse primer was 2, 5′-TCCAGAATCTGTTCCAGAGCGTGC-3′ The PCR was carried out on an MJ Research PTC-2000 Thermocycler. The PCR product was purified using a Promega Wizard SV 96 PCR clean-up system using the standard protocol. The sizes of the PCR fragments for each sample were measured using a 3730 DNA Analyzer (Applied Biosystems, Foster City, CA) using the 600LIZ size standard for comparison. Only PCR products with a FAM labeled end are detected by the analyzer. Data were viewed using GeneMapper software (Applied Biosystems) to determine peak fragment length for each sample. Genotypes were called by visual observation of the peak fragment length within GeneMapper software. Peak fragment lengths were then converted to CAG repeat numbers.

Genome-Wide Genotyping

Genome-wide genotyping was performed using Affymetrix Axiom Genome-Wide LAT and Exome arrays on DNA extracted from buccal cells. During the early phase of the study, DNA was extracted as described in the prior section (CAG Repeats). During the later phase of the study, DNA was extracted using phenol/chloroform, quantitated via picogreen, aliquoted into multiple Sarstedt 1.5 mL tubes, and stored at −80 °C. Samples were randomized across 96-well plates. A common control sample was included in each plate. Principal component analysis (PCA) was used for assessing population stratification (Price et al. 2006). Imputation was performed with MACH-Admix (Liu et al. 2013) using the 1000 Genomes Project (1000G) reference panel (phase1_release_v3.20101123) (Altshuler et al. 2010, 2012). Mean R^2 for varying minor allele frequency (MAF) categories and R^2 cutoffs was calculated for evaluating the quality of imputed SNPs (Wright et al. 2014). SNPs with average $R^2 > 0.8$ were retained for further analysis. Samples with low DishQC (<0.82 for LAT array and <0.79 for Exome array), low call rates (<95%), outliers for homozygosity, sex or zygosity from genotypes inconsistent with reported phenotypes, ancestry outliers, excessive relatedness, and unexpected relatedness were removed. Individual SNPs that deviated from Hardy–Weinberg equilibrium (PHWE $< 1 \times 10^{-8}$) had low call rate (<95%), high Mendelian error rate (>0.1, based on five parent–child trios), and high deviation of allele frequency compared with European American and African American subsets from the 1000 Genomes Project (1000G, either $P < 1 \times 10^{-5}$ and frequency difference > 0.07 or frequency difference > 0.15) and that did not match 1000G EUR/AFR founders were also removed (Xia et al. 2017).

Polygenic Risk Score

Polygenic risk scores for serum testosterone and estradiol were generated based on a genome-wide association analysis (GWAS) of UK Biobank samples of unrelated individuals of European ancestry ($N = 361\,194$ with 194 174 females and 167 020 males in the age range 40–69 years), controlling for age, age^2, and principal coordinates 1–20 (Walters et al. 2019). Higher risk scores are indicative of higher levels of serum testosterone and serum estradiol in an individual. Because the

genetic component underlying individual variation in testosterone and estradiol levels differs between males and females with limited overlap of genome-wide significant signals and opposing effects at many loci (Ruth et al. 2020), we generated sex-specific PRS for individuals based on genetic determinants. Because the predictive power of PRS is reduced in individuals whose genetic ancestry differs substantially from the ancestry of individuals in the original GWAS (Martin et al. 2019), we restricted our analyses to individual of Caucasian ancestry ($n = 430$). PRSice version 2 software was used for generation of PRSs (Choi and O'Reilly 2019). The generation followed a series of processes. SNPs present in only the base UK Biobank GWAS or target cohort were removed. Info score filtering threshold was kept at 0.8. SNPs with minor allele frequency less than 0.01 were excluded. Ambiguous (A/T or C/G) SNPs were also removed. SNPs in linkage disequilibrium ($r^2 > 0.1$) were removed using a process called clumping, leaving a single SNP in each 250 kb LD window with the smallest P value from the GWAS. The remaining SNPs were used to calculate PRSs for 6 P value thresholds ($5*10^{-8}$, 0.001, 0.01, 0.05, 0.5, 1). We used multiple thresholds because this approach addresses noise arising from null SNPs that may be incorporated, especially at large P value thresholds, while capturing true, subtle signals that increase power for association. PRS for testosterone and estradiol was calculated using additive model for all individuals of Caucasian ancestry with genotype data.

Statistical Analysis

Statistical analysis was performed using R version 3.6.3. In the demographics, for categorical variables, we calculated frequency distributions, and for continuous variables, means and standard deviations were calculated. Fisher's exact test and the Kruskal–Wallis test were used for analyzing differences between males and females in baseline categorical and continuous variables, respectively. For the dependent and predictor variables, we used a linear mixed model for analyzing sex effects controlling for maternal ethnicity. For sex effects on testosterone, we also controlled for log transferrin values.

Analysis of Surface Area

Total surface area followed a quadratic growth curve in infancy. Consequently, we used an orthogonal polynomial with order 2 for the regression analysis. Orthogonal regression allows for the evaluation of independent estimates of the importance of the linear component, which corresponds to constant rate of change, and the nonlinear component, which corresponds to varying rate of change (Dutka and Ewens 1971). We used mixed models to test for associations between our predictor variables (all of which relate to gonadal steroid exposure or sensitivity) and surface area. We chose mixed models for this longitudinal study because it allows one to use all data points, even when each subject has a different number of measurements. This models the relationship between the predictor variables and surface area measures while accounting for within- and between-individual variances. Because our study includes a substantial number of twins, we included a nested random-effects term that modeled within-twin and within-individual dependence of observations. Birthweight, gestational age, maternal ethnicity, scanner type, and the average motion score for T1 and T2 for each scan were included as control variables in all our regression models. For testing the relationship between PRSs for testosterone and estradiol with total surface area, we also included the first ten principal components as covariates to control for possible population stratification. We also examined the interaction of predictors with the linear and quadratic age terms. Analyses were performed separately in males and females. The basic statistical model used for analysis is given below.

$$\text{Total Surface Area} = \beta\text{intercept} + \beta\text{predictor} + \beta\text{poly (age at MRI, 2)} + \beta\,(\text{predictor} * \text{poly (age at MRI, 2)}) + \beta\text{birthweight} + \beta\text{gestational age} + \beta\text{maternal ethnicity} + \beta\text{T1motion correction} + \beta\text{T2 motion correction} + \beta\text{scanner} + \left(1|\frac{\text{twinid}}{\text{id}}\right) + \varepsilon$$

predictor is the variable under investigation for its effect on total surface area (PRS for testosterone at various significance thresholds, salivary testosterone level, mean CAG repeat number, right-hand digit ratio, left-hand digit ratio, right–left 2D:4D (Dr-l), and PRS for estradiol at various significance thresholds. Standardized beta values are reported for all predictor variables. Since some of the independent variables in the analysis are correlated, we used the method proposed by Li and Ji (2005) to calculate the effective number of independent tests performed (Meff_pred) using (https://gump.qimr.edu.au/general/daleN/matSpD/). The level of P value significance was determined as 0.05 divided by 2* Meff_pred. Meff_pred was multiplied by two to account for separate analysis of males and females. Analysis of the predictor variables resulted in 14 independent tests, giving a multiple-comparison corrected P value threshold of $P < 0.002$.

Results

Demographics

Table 1 shows both demographic and obstetric history features of the sample. Males and females were highly similar for all these features, although there is a nominally significant difference for maternal ethnicity. Details for the subsample used for PRS analysis are provided in Supplementary Table 1.

Table 2 includes the descriptive statistics for predictor and outcome variables between males and females in the study sample. Salivary testosterone levels were higher in males than females, but the difference was not statistically significant (P value = 0.062). The correlation of salivary–serum testosterone levels is known to be higher for males than females (Shirtcliff et al. 2002) such that

Table 1. Demographic and obstetric history data

		Total n (%/SD)	Males n (%/SD)	Females n (%/SD)	P value*
Sample size		722	386 (53%)	336 (47%)	
Mean gestational age (days)		260.77 (19.62)	260.54 (19.5)	261.04 (19.78)	0.614
Mean birthweight in grams		2831.43 (710.96)	2877.92 (716.47)	2778.01 (701.86)	0.074
Maternal age		29.83 (5.85)	29.73 (5.76)	29.95 (5.97)	0.512
Paternal age		32.41 (6.75)	32.17 (6.7)	32.68 (6.81)	0.369
Singletons/twins		367/355	196/190	171/165	0.694
Total family income	Low	277 (38%)	148 (38%)	129 (38%)	0.375
	Medium	339 (47%)	183 (47%)	156 (46%)	
	High	63 (9%)	28 (7%)	35 (10%)	
Maternal education	Secondary	227 (31%)	119 (31%)	108 (32%)	0.748
	Tertiary	494 (68%)	266 (69%)	228 (68%)	
Paternal education	Primary	1		1	0.271
	Secondary	257 (36%)	130 (34%)	127 (38%)	
	Tertiary	439 (61%)	240 (62%)	199 (59%)	
Maternal ethnicity	Asian	13 (2%)	11 (3%)	2 (1%)	0.028
	Black	170 (24%)	89 (23%)	81 (24%)	
	American Indian or Alaskan Native	4 (1%)	4 (1%)	0	
	White	535 (74%)	282 (73%)	253(75%)	
Paternal ethnicity	Asian	20 (3%)	12 (2%)	8 (2%)	0.108
	Black	187 (26%)	101 (26%)	86 (26%)	
	American Indian or Alaskan Native	5 (1%)		5 (1%)	
	White	499 (69%)	265 (69%)	234 (70%)	

*P values are based on Fisher's exact test for categorical variables and Kruskal–Wallis test for continuous variables.

the predicted average serum testosterone value for males in our sample is 2.33 as opposed to 0.36 in our females. The mean number of CAG repeats in both males and females was similar. In general, the digit ratios were lower in males compared with females. Sex was a significant predictor for digit ratios in right- (P value = 0.04) and left (P value = 0.001) hand after controlling for age and maternal ethnicity. Age was not a significant factor in predicting digit ratio in this study (Supplementary Figs 1 and 2). Correlation matrix of the predictor variables is given as Supplementary Figure 3. Total surface area was larger in males when compared with females (Fig. 1). Age, sex, and their interaction were significant contributors in determining total surface area after controlling for birthweight, gestational age, and maternal ethnicity (P value <0.001) (Fig. 2) (Supplementary Table 2).

PRS for Testosterone

In females, we observed a significant positive interaction between PRS at the stringent threshold (5×10^{-8}) and the linear age term. This association remained significant after multiple corrections (Table 4). PRS for serum testosterone significantly moderated age-related differences in surface area growth in females such that females with high PRS scores showed more surface area expansion from birth to age 2 than females with low PRS scores (Fig. 3). We also observed nominally significant inverse associations between surface area and PRS for testosterone in females at less stringent thresholds (0.01, 0.05, 0.5, and 1) and nominally significant interactions of PRS with age (linear component) at less stringent thresholds (0.05, 0.5, and 1) (Table 3). However, these associations did not survive correction for multiple comparisons. The PRS at various P value thresholds was not significantly associated with total surface area in males. This also remained the case for the interaction term of PRS for testosterone with age in males (Supplementary Table 3). We also tested the interaction of PRS at the stringent threshold for testosterone with sex and age to confirm sex-specific effects. The results were statistically significant for the interaction of PRS with sex and the linear age term (Supplementary Table 4).

Salivary Testosterone

The effect of testosterone in minipuberty was assessed using salivary testosterone levels at 3 months. We did not observe any association between the testosterone levels or their interaction with age (Table 4) and total surface area in infancy in either males or females. We tested if the change in surface area from 1 to 12 months was associated with testosterone levels during the minipuberty. We found a nominally significant association between the testosterone levels and change in surface area in females. However, this association does not withstand correction for multiple comparisons.

CAG Repeats

To assess the effect of sensitivity of androgen receptors (AR) on surface area development, we investigated the role of a polymorphic nucleotide repeat (CAG). There was

Table 2. Descriptive statistics of salivary testosterone, mean CAG repeat length, digit ratios, and brain surface area by sex

Variable	Total mean (SD)	Males	Females	*P* values*
Testosterone	40.98(15.23)	42.02(14.99)	39.72(15.46)	0.062
CAG repeats	19.51(2.7)	19.55(3.04)	19.47(2.28)	0.919
R2R4 Y0	0.93(0.04)	0.93(0.04)	0.94(0.04)	0.007
L2L4 Y0	0.93(0.04)	0.92(0.04)	0.93(0.04)	0.0126
R2R4 Y1	0.92(0.05)	0.92(0.05)	0.93(0.06)	0.22
L2L4 Y1	0.92(0.04)	0.91(0.04)	0.92(0.05)	0.156
R2R4 Y2	0.93(0.05)	0.93(0.04)	0.94(0.05)	0.024
L2L4 Y2	0.93(0.04)	0.93(0.04)	0.93(0.04)	0.191
TSA Y0	79055.61 (9385.34)	80785.79 (9055.76)	77012.51 (9372.43)	<0.001
TSA Y1	142 188 (13567.31)	147731.7 (11845.78)	136193.9 (12760.15)	<0.001
TSA Y2	168873.3 (15229.91)	173826.4 (13663.08)	162637.4 (14846.96)	<0.001

*P values are based on linear mixed models controlled for maternal ethnicity. Age and transferrin levels were also controlled for in the analysis of salivary testosterone. Visit age, birthweight, and gestational age were controlled for in the analysis of total surface area.

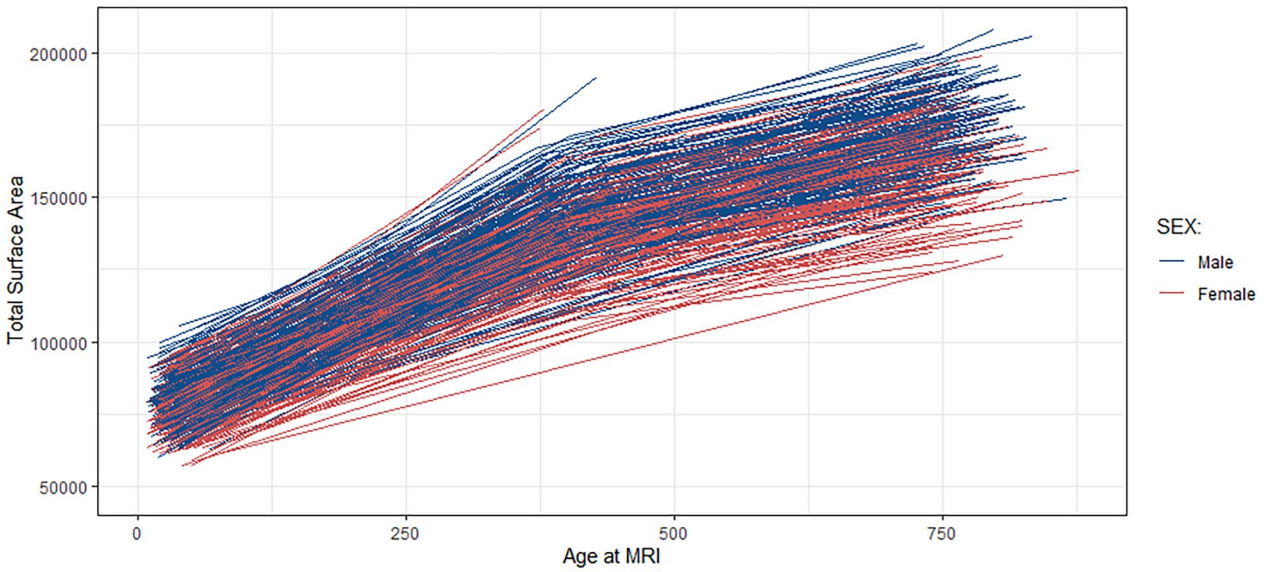

Fig. 1. Spaghetti plots demonstrating within subject changes over time for surface area.

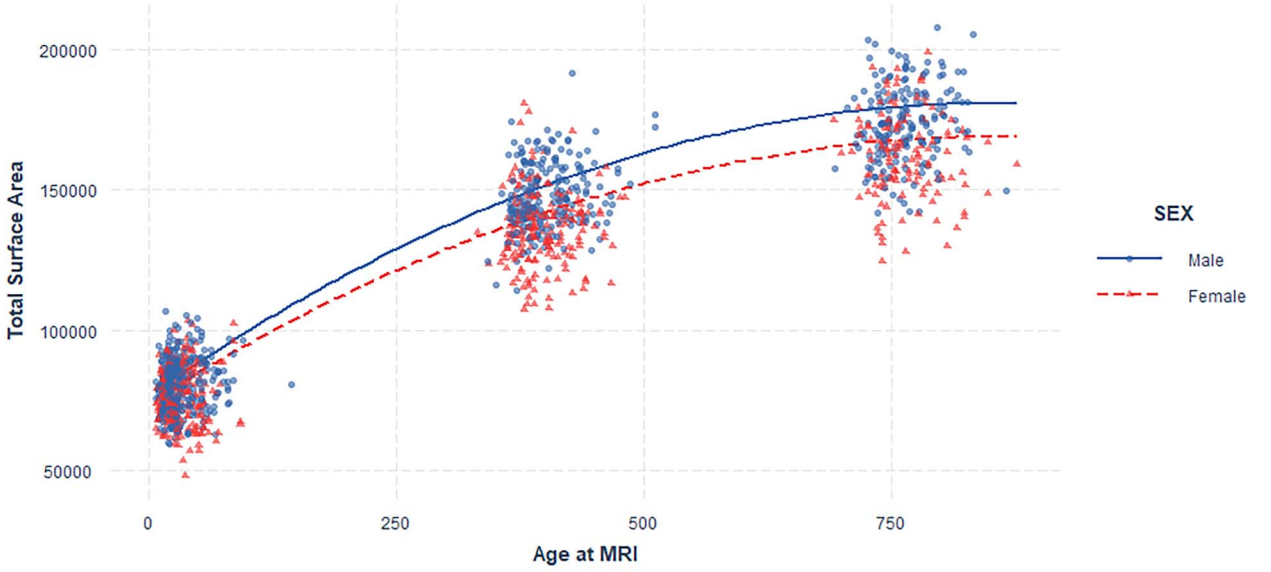

Fig. 2. Age–sex interaction and total surface area. Sex significantly moderates the age-related difference in total surface area**.** Plot shows quadratic mixed effect regression line for the effect of age at MRI*Sex.

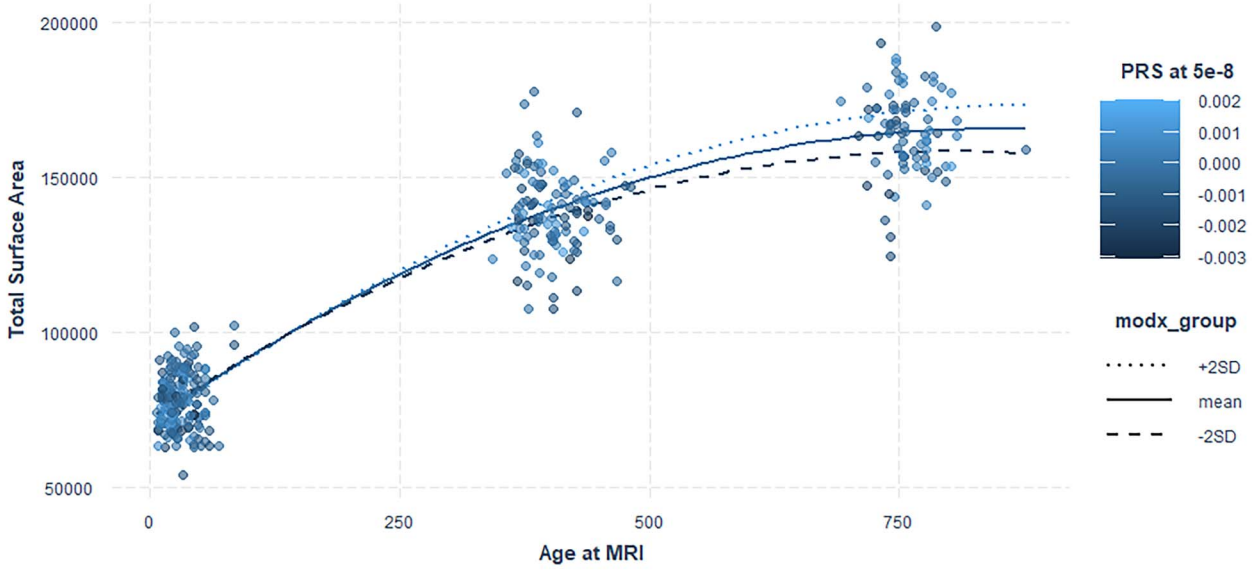

Fig. 3. PRS testosterone–age interaction and total surface area. PRS for serum testosterone significantly moderated age-related differences in surface area growth in females. Plot shows quadratic mixed effect regression line for the effect of PRS at threshold 5×10^{-8} with age. modx_group shows the values for PRS for which the lines are plotted. Actual data points are plotted as scatterplot on top of the interaction lines and the color of the dots are based on the PRS scores of the subjects.

Table 3. Regression results of testing the effect of PRS for serum testosterone and total surface area in females

Predictors	Estimates	SE	Std. Beta	Standardized SE	Standardized CI	Statistic	*P* *	Number of SNPs
Main effect								
$5*10^{-8}$	1127320.85	892343.65	0.02	0.02	−0.01—0.06	1.26	0.206	84
0.001	−354207.1	8332570.19	0	0.02	−0.04—0.03	−0.04	0.966	1917
0.01	−50 883 254	22447030.5	−0.04	0.02	−0.08—−0.01	−2.27	**0.023***	9808
0.05	−105 906 398	44735865.3	−0.04	0.02	−0.08—−0.01	−2.37	**0.018***	32 467
0.5	−367 182 317	141 164 366	−0.05	0.02	−0.08—−0.01	−2.6	**0.009***	164 249
1	−477 743 843	191 461 063	−0.05	0.02	−0.08 – −0.01	−2.5	**0.013***	230 215
Interaction with age (linear term)								
$5*10^{-8}$	35106253.42	9758461.83	0.7	0.19	0.32–1.08	3.6	<**0.001****	84
0.001	46762013.06	104 969 673	0.1	0.22	−0.33– 0.52	0.45	0.656	1917
0.01	−465 262 904	277 494 893	−0.36	0.22	−0.79– 0.06	−1.68	0.094	9808
0.05	−1 141 413 442	507 680 927	−0.44	0.2	−0.83 – −0.06	−2.25	**0.025***	32 467
0.5	−4 795 299 290	1 615 227 570	−0.59	0.2	−0.98 – −0.20	−2.97	**0.003***	164 249
1	−6 190 370 444	2 179 198 099	−0.57	0.2	−0.96 – −0.17	−2.84	**0.005***	230 215
Interaction with age (quadratic term)								
$5*10^{-8}$	296415.22	9193904.88	0.01	0.19	−0.36– 0.37	0.03	0.974	84
0.001	13222079.3	92336579.3	0.03	0.19	−0.35– 0.40	0.14	0.886	1917
0.01	−97 506 765	235 388 392	−0.08	0.19	−0.44– 0.29	−0.41	0.679	9808
0.05	402739260.5	477 674 540	0.16	0.19	−0.21– 0.53	0.84	0.399	32 467
0.5	1 584 337 524	1 497 253 898	0.2	0.19	−0.17– 0.56	1.06	0.29	164 249
1	2 113 042 144	2 026 616 501	0.2	0.19	−0.17– 0.56	1.04	0.297	230 215
N	188_{id}							
	153_{twinid}							
Observations	343							

P value <0.05 "" and <0.002 "**" (*P* value for significance after correction for multiple comparison).

no significant relationship between the length of the CAG repeats and total surface area development in males and females. The interaction of CAG repeat with the linear term of age variable was nominally significant in males (Table 5). However, this association does not withstand correction for multiple comparisons.

Digit Ratios

To explore the effect of digit ratio on total surface area, we analyzed both right- and left-hand digit ratios separately. There was a nominally significant inverse association between the right-hand digit ratio and surface area in female infants, suggesting that a high ratio of

Table 4. Regression results of testing the effect of salivary testosterone on total surface area

Predictors	Estimates	SE	Std. Beta	Standardized SE	Standardized CI	Statistic	*P* *
Males *(N = 160 id; 122 twinid; Observations = 306)*							
Testosterone	12.41	53.01	0.00	0.02	−0.03– 0.04	0.23	0.815
Testosterone: age (linear)	121.86	520.48	0.05	0.19	−0.32– 0.42	0.23	0.815
Testosterone: age (quadratic)	248.65	493.50	0.09	0.18	−0.27– 0.45	0.50	0.614
Females *(N = 131 id; 103 twinid; Observations = 255)*							
Testosterone	−85.97	60.40	−0.03	0.02	−0.08– 0.01	−1.42	0.155
Testosterone: age (linear)	−476.88	542.68	−0.17	0.20	−0.571– 0.23	−0.88	0.38
Testosterone: age (quadratic)	469.84	490.05	0.18	0.19	−0.19– 0.55	0.96	0.338
Effect of salivary testosterone on change of surface area from birth to year 1							
Males *(N = 68 twinid; Observations = 89)*							
Testosterone	4.66	78.64	0.01	0.11	−0.2 – 0.21	0.06	0.953
Females *(N = 59 twinid; Observations = 72)*							
Testosterone	−142.91	69.09	−0.19	0.09	−0.38 – −0.01	−2.07	**0.039***

P value <0.05 "" and <0.002 "**" (P value for significance after correction for multiple comparison).

Table 5. Regression results of testing the effect of number of CAG repeats and total surface area

Predictors	Estimates	SE	Std. Beta	Standardized SE	Standardized CI	Statistic	*P* *
Males *(N = 210; 169 twinid; Observations = 396)*							
CAGrepeat	32.32	246.86	0.00	0.02	−0.03– 0.04	0.13	0.896
CAGrepeat: age (linear)	5656.71	2614.84	0.39	0.18	0.04–0.75	2.16	**0.031***
CAGrepeat: age (quadratic)	−1133.32	2473.77	−0.08	0.18	−0.43– 0.27	−0.46	0.647
Females *(N = 191; 150 twinid; Observations = 363)*							
CAGrepeat	−100.65	339.58	−0.01	0.02	−0.04– 0.03	−0.30	0.767
CAGrepeat: age (linear)	4809.25	3440.29	0.27	0.20	−0.12– 0.66	1.40	0.162
CAGrepeat: age (quadratic)	−4413.30	3412.23	−0.26	0.20	−0.64– 0.13	−1.29	0.196

P value <0.05 "" and <0.002 "**" (P value for significance after correction for multiple comparison).

testosterone to estrogen in the prenatal period correlates with greater cortical surface area at birth and throughout infancy (Fig. 4). However, this association did not survive correction for multiple comparisons. The interaction of digit ratio with age was also not significant for either males or females (Table 6). We also tested for the relationship between right–left digit ratio (Dr-l) and cortical surface area. The results were not significant in either males or females (Supplementary Table 5). We also tested the influence of interaction of digit ratio with sex and age in surface area growth to show the sex-specific effects (Supplementary Table 6).

PRS for Estradiol

In our study, we did not observe any significant relationship between PRS for estradiol, or their interaction with age and total surface area in males or females (Supplementary Tables 7 and 8).

Discussion

Although there is substantial overlap in the distributions of males and females, sex differences have been reported for multiple aspects of brain structure. Many of these differences appear to be linked to male–female differences in total brain size (Ritchie et al. 2018). Sex differences in brain size are detectable by midgestation (Studholme et al. 2020), increase throughout infancy, and are accompanied by more rapid expansion of cortical surface area in human males (Gilmore et al. 2018). The mechanisms responsible for these patterns are currently unknown, but many researchers have hypothesized that prenatal and/or early postnatal exposure to gonadal sex steroids might be involved. The current study addressed this question by examining associations between different markers of prenatal and postnatal hormonal exposure and brain surface area in a large, prospective neuroimaging study of human infants. We found that brain surface area in infancy had a positive relationship with the interaction of linear age term and polygenic risk scores for testosterone in females. This relationship remained significant after correction for multiple comparisons. The results suggest testosterone promotes surface area expansion postnatally in females. But this relationship is only observed when we use the PRS derived from SNPs that are most strongly associated with testosterone in adult females. We also observed inverse relationships between PRS at less stringent thresholds and total surface in females. However, these relationshipsneither survived correction for multiple comparisons nor were they evident when examining other indices of early testosterone effects, namely salivary testosterone measured during the minipuberty and the number of CAG repeats in androgen receptor gene. However, we did observe a nominally significant negative association between salivary testosterone levels and change in surface area from birth to year one.

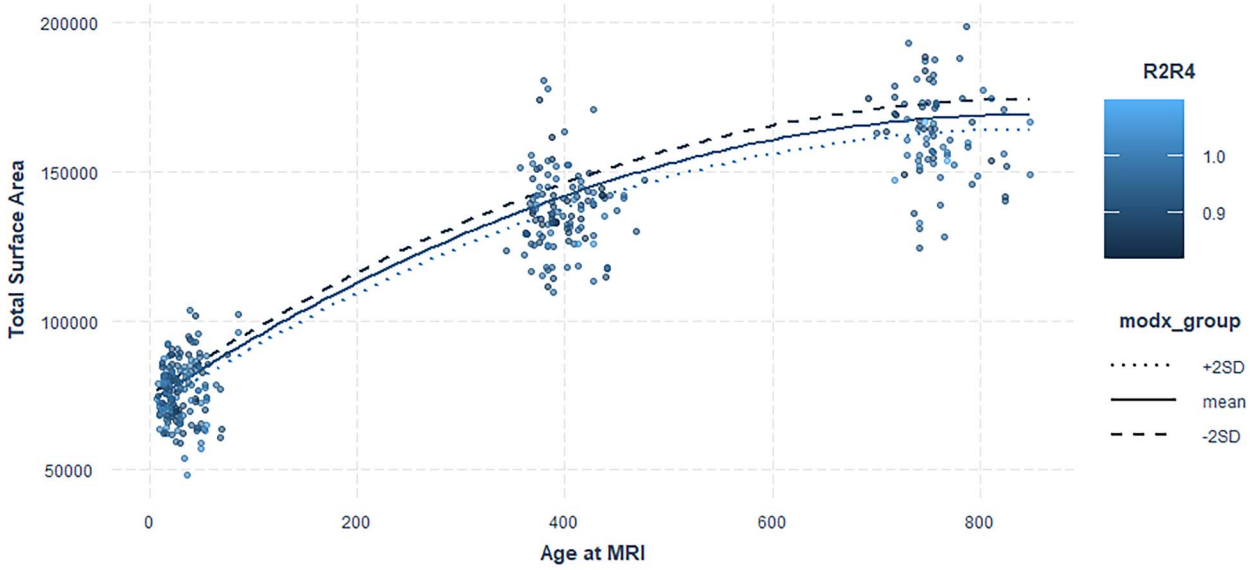

Fig. 4. Right-hand digit ratio age interaction and total surface area. Right-hand digit ratio moderates age-related differences in surface area growth in females. Plot shows quadratic mixed effect regression line for the effect of right-hand digit ratio with age. modx_group shows the values for digit ratio for which the lines are plotted. Actual data points are plotted as scatterplot on top of the interaction lines and the color of the dots are based on their right-hand digit ratio of the subjects.

Table 6. Regression results of testing the effect of digit ratios and total surface area

Predictors	Estimates	SE	Std. Beta	Standardized SE	Standardized CI	Statistic	P *
Males *(N = 278; 222 twinid; Observations = 387)*							
R2/R4	−2072.86	13302.33	0.00	0.01	−0.03– 0.02	−0.16	0.876
L2/L4	10723.36	13225.16	0.01	0.01	−0.01– 0.04	0.81	0.417
Interaction with age (linear term)							
R2/R4:age (linear)	24797.96	246551.86	0.01	0.24	−0.45– 0.48	0.10	0.92
L2/L4:age (linear)	323563.71	257849.07	0.30	0.23	−0.16– 0.76	1.25	0.21
Interaction with age (quadratic term)							
R2/R4:age (quadratic)	−265400.82	244393.14	−0.26	0.24	−0.73– 0.21	−1.09	0.277
L2/L4:age (quadratic)	93638.27	256261.68	0.09	0.24	−0.38– 0.55	0.37	0.715
Females *(N = 237;187 twinid; Observations = 339)*							
R2/R4	−42040.74	13926.49	−0.05	0.02	−0.08 – −0.02	−3.02	**0.003***
L2/L4	2264.39	13365.29	0.00	0.01	−0.03– 0.03	0.17	0.865
Interaction with age (linear term)							
R2/R4:age (linear)	−233487.13	247570.22	−0.25	0.26	−0.76– 0.27	−0.94	0.346
L2/L4:age (linear)	178959.41	256496.01	0.18	0.28	−0.36– 0.72	0.70	0.485
Interaction with age (quadratic term)							
R2/R4:age (quadratic)	47652.64	242929.51	0.05	0.26	−0.47– 0.57	0.20	0.844
L2/L4:age (quadratic)	−272510.82	241978.83	−0.30	0.27	−0.82– 0.22	−1.13	0.26

P value <0.05 "" and <0.002 "**" (P value for significance after correction for multiple comparison).

We also observed an inverse relationship between right 2D:4D digit ratio and surface area in females, suggesting that the ratio of testosterone to estrogen might influence cortical development with higher prenatal testosterone level predicting surface area in females. Although this relationship did not survive corrections for multiple comparisons, this is in line with the observation of PRS for testosterone in females, where we saw testosterone promotes surface area expansion. The interaction of digit ratio with age was not significant. The lack of age effects suggests the association of digit ratio reflects prenatal effects, while the positive association of PRS for testosterone interaction with age may reflect the action of postnatal testosterone. PRS for estradiol was not associated with surface area in either males or females. Overall, our results suggest that testosterone exposure may contribute to sex differences in surface area during infancy but plays a greater role in explaining interindividual variation in females than in males. Two questions naturally follow: 1) what cellular and molecular mechanisms might link testosterone exposure to cortical surface area development, and 2) Why did we observe potential relationships in females only?

With regard to question 1, growth of cortical surface area is influenced by genetic (Chen et al. 2013; Jha et al. 2018) and environmental factors (Jha et al. 2019) and shows distinct changes across stages of development (Wierenga et al. 2014; Lyall et al. 2015; Amlien et al. 2016). Prenatally, cortical surface expansion is mainly driven by neurogenesis within columnar units as postulated

in Rakic's radial unit hypothesis (Rakic 1988). Postnatal development is likely shaped by synaptogenesis, gliogenesis, dendritic arborization, and intracortical myelination (Mrzljak et al. 1991; Rakic et al. 1994; Petanjek et al. 2011). Evidence from both mouse and human neural stem cell studies shows that testosterone leads to increased proliferation and neuronal differentiation (Ransome and Boon 2015; Quartier et al. 2018). This is further supported by research in animal models where testosterone was found to influence adult neurogenesis in rodent (Spritzer and Galea 2007) and avian models (Rasika et al. 1994; Barker et al. 2014). Several observational studies in humans also suggest that testosterone may promote cortical expansion. In human males, fetal testosterone levels positively correlated with regional gray matter volume in school-age children (Lombardo et al. 2012), and individuals with congenital adrenal hyperplasia (CAH), who are exposed to high prenatal androgen levels (including testosterone), show increased regional surface area (van't Westeinde et al. 2020).

With regard to question 2 (why did we observe potential relationships in females only), one possibility is that the proxy measures that showed associations with SA in infant females (PRS for testosterone and right 2D:4D digit ratio) are simply more reliable proxies in females than males. With regard to the PRS scores, it is important to remember that our PRS scores were based on GWAS conducted in adults, as no large-scale GWAS of infant hormone levels are currently available. It is possible that the genetic architecture underlying variation in adult females is similar to that underlying variation in infant females, while the genetic architecture underlying variation in adult males differs from that underlying variation in infant males. We note that the heritability of testosterone levels varies across the lifespan and between males and females (Koenis et al. 2013; Grotzinger, Briley, et al. 2018a). For example, a large twin study in adolescents reported variation in testosterone levels in females is influenced by genetic factors in contrast to males where it is primarily influenced by environmental factors (Van Hulle et al. 2015). However, in most studies, heritability estimates for testosterone are much lower for females compared with males (Meikle et al. 1986; Harris et al. 1998; Pritchard et al. 1998; Hong et al. 2001; Ring et al. 2005; Kuijper et al. 2007; Bogaert et al. 2008; Franks et al. 2008; Harden et al. 2014; Grotzinger, Mann, et al. 2018b). With regard to digit ratios, several studies have reported that amniotic testosterone level is significantly associated with newborn female's digit ratio only and not with males (Ventura et al. 2013; Richards et al. 2019). In addition, heritability estimates for digit ratio from mother–daughter pairs suggest the trait to be X-linked although this observation is not consistent (Richards et al. 2017). If this is the case, our digit ratio analyses might be capturing an effect of one of more X-chromosome genes on surface area development in females, rather than gonadal steroid effects.

We note that assessing individual variation in gonadal steroid exposure during the prenatal and early postnatal period presents a number of challenges. While testosterone and estradiol can be directly measured in amniotic fluid or in fetal blood, these techniques are medically invasive and usually performed only in high-risk pregnancies. Blood sampling during the minipuberty is certainly possible but may be a recruitment barrier for concerned parents. Consequently, the current study used saliva sampling and a variety of proxy measures to assess exposure including a PRS for serum testosterone, a PRS for serum estradiol, a specific genetic polymorphism in the androgen receptor (AR) gene, and 2D:4D digit ratios. Each of these methods has specific benefits and specific limitations that we will discuss in the following paragraphs.

Salivary measures offer a relatively convenient and noninvasive method for obtaining sex hormone data. These measures represent free, biologically available hormone levels and provide information relevant to neurological and behavioral studies. However, some limitations are known. Although there are reports of strong correlations between serum and salivary testosterone in males (Nahoul et al. 1986; Rilling et al. 1996; Lood et al. 2018), such correlations are significantly lower in females (Granger et al. 1999; Miller et al. 2004). This could potentially underestimate the testosterone-associated phenotypes in females. In our study, we observe a significant relation of surface area with PRS for serum testosterone but only a marginal association for salivary testosterone level in females. The lack of strong correlation between the serum–saliva measures could be a reason for variation in results. Other factors that can significantly influence the salivary testosterone measures include collection methods, blood contamination in saliva, and storage conditions (Granger et al. 2004; Al-DujailI and Sharp 2012). We controlled for all these factors by uniformly collecting samples at fixed time in the morning and samples where frozen in a −80 °C freezer within 4 h of collection. Samples were evaluated for blood contamination by measuring transferrin levels and including them in the analysis. There are more limitations in assessing estradiol levels in children. The levels in saliva are often too low to be reliably measured using immunoassays and also require larger quantities for measurement. The serum–saliva correlations for estradiol are also lower than for other salivary hormones (Shirtcliff et al. 2000; Frederiksen et al. 2020). For these reasons, we did not measure estradiol in our salivary samples.

Polygenic risk scores offer an opportunity to capture interindividual variation in highly dynamic traits like sex hormone levels that are otherwise difficult to capture with single measurements. However, there are some limitations to this approach as it was applied in the context of the current study. First, the base data for PRS generation was summary statistics of GWAS results from UK Biobank samples of unrelated *adult* individuals of European ancestry ($N = 361\,194$ with 194 174 females and

167 020 males in the age range 40–69 years). The utility of our PRS scores as proxies for prenatal and minipubertal hormone exposure is based on the assumption that there is substantial overlap in the genetic variants involved in sex hormone production during prenatal, minipubertal, and adult stages, as well as their direction of effect. This has not yet been demonstrated empirically as no large-scale GWAS have been conducted for prenatal or minipubertal hormone levels. However, genetic correlations for testosterone between pre- and postmenopausal women in the UK Biobank sample are close to 1.0, suggesting a high degree of stability across the lifespan in females (Flynn et al. 2021). Furthermore, the fact that we observed a significant association between surface area expansion and PRS scores in female infants suggests at least some developmental stability. Nevertheless, we cannot rule out the possibility that our results would have looked different using a PRS score based on infant data, which do not currently exist. Second, although classical twin and family studies show moderate to high heritability for testosterone in adolescence and adulthood, the estimated SNP heritability estimate for testosterone in the UK Biobank sample ($n = 312\,102$ unrelated individuals of European ancestry) was 0.0771, which suggests relatively little variation in testosterone levels can be attributed to common genetic variation (Walters 2020). If variation in prenatal and minipubertal testosterone levels is primarily the result of environmental factors or rare genetic variation, rather than common genetic variation, this would limit to utility of the PRS as a proxy measure for early testosterone exposure. Twin studies conducted during infancy suggest that individual variability in salivary testosterone levels is primarily explained by environmental factors (Caramaschi et al. 2012; Xia et al. 2014), and no genetic effects on testosterone were observed in male and female twins using umbilical cord blood (Sakai et al. 1991). The fact that we observed a significant association between surface area expansion and PRS scores in female infants suggests that common genetic variation does influence testosterone levels in early life, at least in females, though we cannot rule out the possibility that these genetic variants influenced surface area expansion via some other mechanisms (e.g., they may have pleiotropic effects). We also note that the heritability of prenatal testosterone levels has not been tested. Consequently, the association of the testosterone PRS with surface area expansion in females could represent an effect of prenatal testosterone on postnatal development. Finally, heritability of estradiol is much lower than other sex hormones in both men and women and the interindividual variation is mainly attributed to environmental effects (Ring et al. 2005; Stone et al. 2009; Travison et al. 2014). Consequently, the estradiol PRS scores may not have functioned well as proxies for early life estrogen exposure. We cannot rule out the possibility that measuring infant estradiol levels in blood using highly sensitive mass spectrometry–based assays would have revealed relations with surface area development.

The AR receptor polymorphism we examined is a microsatellite trinucleotide (CAG) repeat sequence located in the transactivation domain within the first exon, which encodes a string of 8 to 35 glutamines, depending on the individual. The length of the repeat is inversely associated with the transactivational activity of the AR (Choong and Wilson 1998; Davey and Grossmann 2016) and a more "masculine" pattern of cortical maturation during adolescence (Raznahan et al. 2010). However, CAG repeats alone are not the only factors that influence the AR sensitivity. There are other genetic factors including a polyglycine tract encoded by a polymorphic GGN repeat in exon 1 of the AR (Bogaert et al. 2009; Westberg et al. 2009) and co-activators and co-repressors with tissue-specific effects on AR transactivation (Choong and Wilson 1998; Zitzmann and Nieschlag 2001). Also, there might be compensatory mechanisms increasing circulating testosterone levels in cases where there is a less effective AR (Crabbe et al. 2007; Stanworth et al. 2008). Furthermore, although androgens predominantly exert their biological effects by binding to the androgen receptor (AR) (Heinlein and Chang 2002), they can also exert effects via nongenomic signaling pathways like membrane receptors or by being aromatized to estrogen and acting on estrogen receptors. If these mechanisms are important for testosterone effects on surface area, the CAG repeat would be irrelevant. Finally, the location of the AR on the X-chromosome has important implications when examining effects in females. When calculating the average CAG repeat length in females, we assume random inactivation such that half of female cells are expressing one allele and half are expressing the other allele. However, in reality, X-chromosome inactivation varies between individuals and between tissues and, in some cases, is highly skewed with one X-chromosome being preferentially inactivated. All of these factors could explain why we did not observe an effect of the CAG polymorphism on surface area.

The ratio between lengths of the second and the fourth digits (index and ring fingers) (2D:4D) is a putative biomarker of prenatal testosterone and estrogen exposure (Manning et al. 1998; Altshuler et al. 2010). Digit ratios offer a noninvasive method for estimating prenatal hormonal exposure, may reflect combined effects throughout gestation, and are less likely to be affected by daily variations than a single direct measurement (Baxter et al. 2018; Lautenbacher and Neyse 2020). A low 2D:4D ratio, common in males, is positively related to prenatal testosterone, whereas a high 2D:4D ratio, common in females, is positively associated with prenatal estrogen (Fink et al. 2003; Manning 2011; Manning and Fink 2018). 2D:4D ratio in the right hand was found to be a better indicator of prenatal testosterone exposure than that in left hand (Hönekopp and Watson 2010). The role of prenatal sex hormones in generating sex differences in 2D:4D is supported by experimental studies in animal models (Zheng and Cohn 2011), correlational studies in humans

(Lutchmaya et al. 2004; Ventura et al. 2013), opposite-sex twin studies (Voracek and Dressler 2007), and studies on individuals with specific genetic conditions like congenital adrenal hyperplasia (CAH) (Brown et al. 2002; Rivas et al. 2014), Klinefelter's syndrome (Manning et al. 2013; Chang et al. 2015), and androgen insensitivity syndrome (AIS) (Berenbaum et al. 2009; Van Hemmen et al. 2017). However, the strength of the association between prenatal hormone exposure and 2D:4D may be relatively small, and individual variation in 2D:4D also reflects other factors including postnatal hormonal levels (Knickmeyer et al. 2011; McIntyre and Alexander 2011), genetic variation (Paul et al. 2006; Voracek and Dressler 2007; Hiraishi et al. 2012; Kalichman et al. 2019), and impaired fetal growth (Ronalds et al. 2002). Nevertheless, considering the difficulties in obtaining direct measurements of prenatal hormonal exposure, digit ratios remain the most accessible measure for prenatal hormonal exposure.

We also acknowledge that sexual differentiation in human brain could be influenced by many factors other than gonadal steroid exposure, as evidenced from rodent and human studies. In particular, the activity of genes on the sex chromosomes could have a significant role in determining sex-specific growth patterns of the brain. This is evident from studies in individuals with complete androgen insensitivity syndrome (CAIS) where external body phenotype is female, but the karyotype is XY. Regional brain volumes in XY individuals with CAIS were found to be similar to non-CAIS (XX) females in some areas, whereas it is similar to non-CAIS (XY) males in other areas (Savic et al. 2017). Of particular relevance to the current study, among youth of different karyotypes (between 5 and 25 years of age), increases in the number of X chromosome are associated with a decrease in total brain volume, cortical volume, and cortical surface area, while increases in the number of Y chromosomes are associated with increased total brain volume, cortical volume, and cortical surface area (Raznahan et al. 2016). Specific genes involved in these patterns have not yet been identified but potential candidates include *Sry* (a Y-linked gene that is involved in male sexual development and is also expressed in the developing male brain) (Dewing et al. 2006), *Xist* (a noncoding RNA on the X chromosome that plays a major role in the X-inactivation process (Arnold 2017), and a variety of X-linked escapee genes located outside the pseudoautosomal region (PAR) (Disteche 2012). Sex differences can also arise due to maternal versus paternal imprinting of the X chromosome (Bonthuis et al. 2015), the action of Mullerian inhibiting substance (Wang et al. 2009), and environmental effects that are difficult to separate from biological effects. Ultimately, all these factors act in parallel or in interaction with each other to bring about sexual differentiation.

In conclusion, this is the first study to explore potential associations between gonadal steroid exposure during early life and infant brain surface area. Strengths of the current study include the use of a large prospective, longitudinal cohort, inclusion of state-of-the-art neuroimaging measures, and rigorous quality control of both neuroimaging and genetic data. Limitations are primarily related to the challenges of measuring gonadal steroid exposure during the prenatal and early postnatal period and have already been discussed at length. In summary, our current work suggests that variation in testosterone during both the prenatal and postnatal period may contribute to cortical surface area development and predicts interindividual variation in females, but not in males.

Supplementary Material

Supplementary material can be found at *Cerebral Cortex* online.

Notes

We thank our participating families and staff in the UNC Biomedical Research Imaging Center (BRIC), the UNC Neuro Image Research and Analysis Laboratories (NIRAL), the UNC Early Brain Development Study (EBDS), the UNC Cytokine & Biomarker Analysis Facility, the UNC BioSpecimen Processing Facility (BSP), and the UNC Biobehavioral Laboratory (BBL). *Conflict of Interest*: The authors have no conflicts of interest to declare.

Funding

National Institutes of Health (MH064065, MH070890,, HD053000 to J.H.G., MH092335, MH083045 to R.C.K., HD003110, EB005149 to M.S., MH116225, MH123202 to G.L., MH117943 to L.W. and G.L., and NS007431 to S.J.); the Brain & Behavior Research Foundation, and the John and Polly Sparks Foundation, Bank of America, Trustee (NARSAD Young Investigator Grant to K.X.); and the National Alliance for Autism Research (NAAR), which has since merged with Autism Speaks (to J.H.G. and R.C.K.).

References

Al-DujailI EAS, Sharp MA. 2012. Female salivary testosterone: measurement, challenges and applications. In: Ostojic S, editor. *Steroids-from physiology to clinical medicine*. Rijeka, Croatia: InTech, pp. 129–167.

Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. 1992. Methylation of Hpaii and Hhai sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X-chromosome inactivation. *Am J Hum Genet*. 51:1229–1239.

Altshuler DL, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Collins FS, De La Vega FM, Donnelly P, Egholm M et al. 2010. A map of human genome variation from population-scale sequencing. *Nature*. 467:1061–1073.

Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB et al. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 491:56–65.

Amlien IK, Fjell AM, Tamnes CK, Grydeland H, Krogsrud SK, Chaplin TA, Rosa MGP, Walhovd KB. 2016. Organizing principles of human

cortical development—thickness and area from 4 to 30 years: insights from comparative primate neuroanatomy. *Cereb Cortex*. 26:257–267.

Andersson A-M, Toppari J, Haavisto A-M, Petersen JH, Simell T, Simell O, Skakkebæk NE. 1998. Longitudinal reproductive hormone profiles in infants: peak of inhibin B levels in infant boys exceeds levels in adult men. *J Clin Endocrinol Metab*. 83:675–681.

Arnold AP. 2017. A general theory of sexual differentiation. *J Neurosci Res*. 95:291–300.

Barker JM, Ball GF, Balthazart J. 2014. Anatomically discrete sex differences and enhancement by testosterone of cell proliferation in the telencephalic ventricle zone of the adult canary brain. *J Chem Neuroanat*. 55:1–8.

Baxter A, Wood EK, Jarman P, Cameron AN, Capitanio JP, Higley JD. 2018. Sex differences in rhesus monkeys' digit ratio (2D:4D ratio) and its association with maternal social dominance rank. *Front Behav Neurosci*. 12:213.

Becker M, Hesse V. 2020. Minipuberty: why does it happen? *Horm Res Paediatr*. 93:76–84.

Berenbaum SA, Bryk KK, Nowak N, Quigley CA, Moffat S. 2009. Fingers as a marker of prenatal androgen exposure. *Endocrinology*. 150:5119–5124.

Blumenthal JD, Zijdenbos A, Molloy E, Giedd JN. 2002. Motion artifact in magnetic resonance imaging: implications for automated analysis. *Neuroimage*. 16:89–92.

Bogaert V, Taes Y, Konings P, Van Steen K, De Bacquer D, Goemaere S, Zmierczak H, Crabbe P, Kaufman J-M. 2008. Heritability of blood concentrations of sex-steroids in relation to body composition in young adult male siblings. *Clin Endocrinol (Oxf)*. 69:129–135.

Bogaert V, Vanbillemont G, Taes Y, De Bacquer D, Deschepper E, Van Steen K, Kaufman J-M. 2009. Small effect of the androgen receptor gene GGN repeat polymorphism on serum testosterone levels in healthy men. *Eur J Endocrinol*. 161:171–177.

Bonthuis PJ, Huang WC, Stacher Hörndli CN, Ferris E, Cheng T, Gregg C. 2015. Noncanonical genomic imprinting effects in offspring. *Cell Rep*. 12:979–991.

Bramen JE, Hranilovich JA, Dahl RE, Chen J, Rosso C, Forbes EE, Dinov ID, Worthman CM, Sowell ER. 2012. Sex matters during adolescence: testosterone-related cortical thickness maturation differs between boys and girls. *PLoS One*. 7:e33850.

Bramen JE, Hranilovich JA, Dahl RE, Forbes EE, Chen J, Toga AW, Dinov ID, Worthman CM, Sowell ER. 2011. Puberty influences medial temporal lobe and cortical gray matter maturation differently in boys than girls matched for sexual maturity. *Cereb Cortex*. 21: 636–646.

Brown WM, Hines M, Fane BA, Breedlove SM. 2002. Masculinized finger length patterns in human males and females with congenital adrenal hyperplasia. *Horm Behav*. 42:380–386.

Caramaschi D, Booij L, Petitclerc A, Boivin M, Tremblay RE. 2012. Genetic and environmental contributions to saliva testosterone levels in male and female infant twins. *Psychoneuroendocrinology*. 37:1954–1959.

Chang S, Skakkebæk A, Trolle C, Bojesen A, Hertz JM, Cohen A, Hougaard DM, Wallentin M, Pedersen AD, Østergaard JR et al. 2015. Anthropometry in Klinefelter syndrome—multifactorial influences due to CAG length, testosterone treatment and possibly intrauterine hypogonadism. *J Clin Endocrinol Metab*. 100:E508–E517.

Chellakooty M, Schmidt IM, Haavisto AM, Boisen KA, Damgaard IN, Mau C, Petersen JH, Juul A, Skakkebæk NE, Main KM. 2003. Inhibin a, inhibin B, follicle-stimulating hormone, luteinizing hormone, Estradiol, and sex hormone-binding globulin levels in 473 healthy infant girls. *J Clin Endocrinol Metab*. 88:3515–3520.

Chen CH, Fiecas M, Gutiérrez ED, Panizzon MS, Eyler LT, Vuoksimaa E, Thompson WK, Fennema-Notestine C, Hagler DJ, Jernigan TL et al. 2013. Genetic topography of brain morphology. *Proc Natl Acad Sci USA*. 110:17089–17094.

Choi SW, O'Reilly PF. 2019. PRSice-2: polygenic risk score software for biobank-scale data. *Gigascience*. 8:1–6.

Choong CS, Wilson EM. 1998. Trinucleotide repeats in the human androgen receptor: a molecular basis for disease. *J Mol Endocrinol*. 21:235–257.

Clements JA, Reyes FI, Winter JSD, Faiman C. 1976. Studies on human sexual development. III. Fetal pituitary and serum, and amniotic fluid concentrations of LH, CG, and FSH. *J Clin Endocrinol Metab*. 42:9–19.

Crabbe P, Bogaert V, De Bacquer D, Goemaere S, Zmierczak H, Kaufman JM. 2007. Part of the interindividual variation in serum testosterone levels in healthy men reflects differences in androgen sensitivity and feedback set point: contribution of the androgen receptor polyglutamine tract polymorphism. *J Clin Endocrinol Metab*. 92:3604–3610.

Darnai G, Plózer E, Perlaki G, Orsi G, Nagy SA, Horváth R, Schwarcz A, Kovács N, Altbäcker A, Janszky J et al. 2016. 2D:4D finger ratio positively correlates with total cerebral cortex in males. *Neurosci Lett*. 615:33–36.

Davey RA, Grossmann M. 2016. Androgen receptor structure, function and biology: from bench to bedside. *Clin Biochem Rev*. 37:3–15.

Dean DC, Planalp EM, Wooten W, Schmidt CK, Kecskemeti SR, Frye C, Schmidt NL, Goldsmith HH, Alexander AL, Davidson RJ. 2018. Investigation of brain structure in the 1-month infant. *Brain Struct Funct*. 223:1953–1970.

Debieve F, Beerlandt S, Hubinont C, Thomas K. 2000. Gonadotropins, prolactin, inhibin a, inhibin B, and activin a in human Fetal serum from midpregnancy and term pregnancy. *J Clin Endocrinol Metab*. 85:270–274.

Dewing P, Chiang CWK, Sinchak K, Sim H, Fernagut PO, Kelly S, Chesselet MF, Micevych PE, Albrecht KH, Harley VR et al. 2006. Direct regulation of adult brain function by the male-specific factor SRY. *Curr Biol*. 16:415–420.

Disteche CM. 2012. Dosage compensation of the sex chromosomes. *Annu Rev Genet*. 46:537–560.

Dutka AF, Ewens FJ. 1971. Method for improving the accuracy of polynomial regression analysis. *J Qual Technol*. 3:149–155.

Fink B, Neave N, Manning JT. 2003. Second to fourth digit ratio, body mass index, waist-to-hip ratio, and waist-to-chest ratio: their relationships in heterosexual men and women. *Ann Hum Biol*. 30: 728–738.

Flynn E, Tanigawa Y, Rodriguez F, Altman RB, Sinnott-Armstrong N, Rivas MA. 2021. Sex-specific genetic effects across biomarkers. *Eur J Hum Genet*. 29:154–163.

Forest MG, Cathiard AM, Bertrand JA. 1973. Evidence of testicular activity in early infancy. *J Clin Endocrinol Metab*. 37: 148–151.

Franks S, Webber LJ, Goh M, Valentine A, White DM, Conway GS, Wiltshire S, McCarthy MI. 2008. Ovarian morphology is a marker of heritable biochemical traits in sisters with polycystic ovaries. *J Clin Endocrinol Metab*. 93:3396–3402.

Frederiksen H, Johannsen TH, Andersen SE, Albrethsen J, Landersoe SK, Petersen JH, Andersen AN, Vestergaard ET, Schorring ME, Linneberg A et al. 2020. Sex-specific estrogen levels and reference intervals from infancy to late adulthood determined by LC-MS/MS. *J Clin Endocrinol Metab*. 105:754–768.

Gilmore JH, Knickmeyer RC, Gao W. 2018. Imaging structural and functional brain development in early childhood. *Nat Rev Neurosci*. 19:123–137.

Gilmore JH, Schmitt JE, Knickmeyer RC, Smith JK, Lin W, Styner M, Gerig G, Neale MC. 2010. Genetic and environmental contributions to neonatal brain structure: a twin study. *Hum Brain Mapp*. 31:1174–1182.

Granger DA, Schwartz EB, Booth A, Arentz M. 1999. Salivary testosterone determination in studies of child health and development. *Horm Behav*. 35:18–27.

Granger DA, Shirtcliff EA, Booth A, Kivlighan KT, Schwartz EB. 2004. The "trouble" with salivary testosterone. *Psychoneuroendocrinology*. 29:1229–1240.

Green T, Flash S, Reiss AL. 2018. Sex differences in psychiatric disorders: what we can learn from sex chromosome aneuploidies. *Neuropsychopharmacology*. 44:9–21.

Grotzinger AD, Briley DA, Engelhardt LE, Mann FD, Patterson MW, Tackett JL, Tucker-Drob EM, Harden KP. 2018a. Genetic and environmental influences on pubertal hormones in human hair across development. *Psychoneuroendocrinology*. 90:76–84.

Grotzinger AD, Mann FD, Patterson MW, Herzhoff K, Tackett JL, Tucker-Drob EM, Paige HK. 2018b. Twin models of environmental and genetic influences on pubertal development, salivary testosterone, and estradiol in adolescence. *Clin Endocrinol (Oxf)*. 88:243–250.

Harden KP, Kretsch N, Tackett JL, Tucker-Drob EM. 2014. Genetic and environmental influences on testosterone in adolescents: evidence for sex differences. *Dev Psychobiol*. 56:1278–1289.

Harris JA, Vernon PA, Boomsma DI. 1998. The heritability of testosterone: a study of dutch adolescent twins and their parents. *Behav Genet*. 28:165–171.

Heinlein CA, Chang C. 2002. The roles of androgen receptors and androgen-binding proteins in nongenomic androgen actions. *Mol Endocrinol*. 16:2181–2187.

Herting MM, Gautam P, Spielberg JM, Dahl RE, Sowell ER. 2015. A longitudinal study: changes in cortical thickness and surface area during pubertal maturation. *PLoS One*. 10:e0119774.

Herting MM, Gautam P, Spielberg JM, Kan E, Dahl RE, Sowell ER. 2014. The role of testosterone and estradiol in brain volume changes across adolescence: a longitudinal structural MRI study. *Hum Brain Mapp*. 35:5633–5645.

Hiraishi K, Sasaki S, Shikishima C, Ando J. 2012. The second to fourth digit ratio (2D:4D) in a Japanese twin sample: heritability, prenatal hormone transfer, and association with sexual orientation. *Arch Sex Behav*. 41:711–724.

Holland D, Chang L, Ernst TM, Curran M, Buchthal SD, Alicata D, Skranes J, Johansen H, Hernandez A, Yamakawa R et al. 2014. Structural growth trajectories and rates of change in the first 3 months of infant brain development. *JAMA Neurol*. 71:1266–1274.

Hönekopp J, Watson S. 2010. Meta-analysis of digit ratio 2D:4D shows greater sex difference in the right hand. *Am J Hum Biol*. 22:619–630.

Hong Y, Gagnon J, Rice T, Pérusse L, Leon AS, Skinner JS, Wilmore JH, Bouchard C, Rao DC. 2001. Familial resemblance for free androgens and androgen glucuronides in sedentary black and white individuals: the HERITAGE family study. *J Endocrinol*. 170: 485–492.

Jha SC, Xia K, Ahn M, Girault JB, Li G, Wang L, Shen D, Zou F, Zhu H, Styner M et al. 2019. Environmental influences on infant cortical thickness and surface area. *Cereb Cortex*. 29:1139–1149.

Jha SC, Xia K, Schmitt JE, Ahn M, Girault JB, Murphy VA, Li G, Wang L, Shen D, Zou F et al. 2018. Genetic influences on neonatal cortical thickness and surface area. *Hum Brain Mapp*. 39:4998–5013.

Kaczkurkin AN, Raznahan A, Satterthwaite TD. 2019. Sex differences in the developing brain: insights from multimodal neuroimaging. *Neuropsychopharmacology*. 44:71–85.

Kalichman L, Batsevich V, Kobyliansky E. 2019. Heritability estimation of 2D:4D finger ratio in a Chuvashian population-based sample. *Am J Hum Biol*. 31:e23212.

Kallai J, Csathó Á, Kövér F, Makány T, Nemes J, Horváth K, Kovács N, Manning JT, Nadel L, Nagy F. 2005. MRI-assessed volume of left and right hippocampi in females correlates with the relative length of the second and fourth fingers (the 2D:4D ratio). *Psychiatry Res Neuroimaging*. 140:199–210.

Kaplan SL, Grumbach MM. 1976. The ontogenesis of human foetal hormones. II. Luteinizing hormone (LH) and follicle stimulating hormone (FSH). *Acta Endocrinol*. 81: 808–829.

Knickmeyer RC, Gouttard S, Kang C, Evans D, Wilber K, Smith JK, Hamer RM, Lin W, Gerig G, Gilmore JH. 2008. A structural MRI study of human brain development from birth to 2 years. *J Neurosci*. 28:12176–12182.

Knickmeyer RC, Wang J, Zhu H, Geng X, Woolson S, Hamer RM, Konneker T, Lin W, Styner M, Gilmore JH. 2014a. Common variants in psychiatric risk genes predict brain structure at birth. *Cereb Cortex*. 24:1230–1246.

Knickmeyer RC, Wang J, Zhu H, Geng X, Woolson S, Hamer RM, Konneker T, Styner M, Gilmore JH. 2014b. Impact of sex and gonadal steroids on neonatal brain structure. *Cereb Cortex*. 24: 2721–2731.

Knickmeyer RC, Woolson S, Hamer RM, Konneker T, Gilmore JH. 2011. 2D:4D ratios in the first 2 years of life: stability and relation to testosterone exposure and sensitivity. *Horm Behav*. 60:256–263.

Knickmeyer RC, Xia K, Lu Z, Ahn M, Jha SC, Zou F, Zhu H, Styner M, Gilmore JH. 2017. Impact of demographic and obstetric factors on infant brain volumes: a population neuroscience study. *Cereb Cortex*. 27:5616–5625.

Koenis MMG, Brouwer RM, van Baal GCM, van Soelen ILC, Peper JS, van Leeuwen M, Delemarre-van de Waal HA, Boomsma DI, Hulshoff Pol HE. 2013. Longitudinal study of hormonal and physical development in young twins. *J Clin Endocrinol Metab*. 98:E518–E527.

Koolschijn PCMP, Crone EA. 2013. Sex differences and structural brain maturation from childhood to early adulthood. *Dev Cogn Neurosci*. 5:106–118.

Koolschijn PCMP, Peper JS, Crone EA. 2014. The influence of sex steroids on structural brain maturation in adolescence. *PLoS One*. 9:83929.

Kuijper EAM, Lambalk CB, Boomsma DI, van der Sluis S, Blankenstein MA, de Geus EJC, Posthuma D. 2007. Heritability of reproductive hormones in adult male twins. *Hum Reprod*. 22:2153–2159.

Kuiri-Hänninen T, Haanpää M, Turpeinen U, Hämäläinen E, Seuri R, Tyrväinen E, Sankilampi U, Dunkel L. 2013. Postnatal ovarian activation has effects in estrogen target tissues in infant girls. *J Clin Endocrinol Metab*. 98:4709–4716.

Kuiri-Hänninen T, Seuri R, Tyrväinen E, Turpeinen U, Hämäläinen E, Stenman U-H, Dunkel L, Sankilampi U. 2011. Increased activity of the hypothalamic-pituitary-testicular axis in infancy results in increased androgen action in premature boys. *J Clin Endocrinol Metab*. 96:98–105.

Lanciotti L, Cofini M, Leonardi A, Penta L, Esposito S. 2018. Up-to-date review about minipuberty and overview on hypothalamic-pituitary-gonadal axis activation in fetal and neonatal life. *Front Endocrinol (Lausanne)*. 9:410.

Lautenbacher LM, Neyse L. 2020. Depression, neuroticism and 2D:4D ratio: evidence from a large, representative sample. *Sci Rep*. 10:11136.

Li G, Nie J, Wang L, Shi F, Gilmore JH, Lin W, Shen D. 2014. Measuring the dynamic longitudinal cortex development in infants

by reconstruction of temporally consistent cortical surfaces. *Neuroimage*. 90:266–279.

Li G, Nie J, Wu G, Wang Y, Shen D. 2012. Consistent reconstruction of cortical surfaces from longitudinal brain MR images. *Neuroimage*. 59:3805–3820.

Li G, Wang L, Shi F, Lyall AE, Ahn M, Peng Z, Zhu H, Lin W, Gilmore JH, Shen D. 2016. Cortical thickness and surface area in neonates at high risk for schizophrenia. *Brain Struct Funct*. 221:447–461.

Li J, Ji L. 2005. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)*. 95: 221–227.

Liu EY, Li M, Wang W, Li Y. 2013. MaCH-admix: genotype imputation for admixed populations. *Genet Epidemiol*. 37:25–37.

Lombardo MV, Ashwin E, Auyeung B, Chakrabarti B, Taylor K, Hackett G, Bullmore ET, Baron-Cohen S. 2012. Fetal testosterone influences sexually dimorphic gray matter in the human brain. *J Neurosci*. 32:674–680.

Lood Y, Aardal-Eriksson E, Webe C, Ahlner J, Ekman B, Wahlberg J. 2018. Relationship between testosterone in serum, saliva and urine during treatment with intramuscular testosterone undecanoate in gender dysphoria and male hypogonadism. *Andrology*. 6:86–93.

Lutchmaya S, Baron-Cohen S, Raggatt P, Knickmeyer R, Manning JT. 2004. 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Hum Dev*. 77:23–28.

Lyall AE, Shi F, Geng X, Woolson S, Li G, Wang L, Hamer RM, Shen D, Gilmore JH. 2015. Dynamic development of regional cortical thickness and surface area in early childhood. *Cereb Cortex*. 25: 2204–2212.

Manning JT. 2011. Resolving the role of prenatal sex steroids in the development of digit ratio. *Proc Natl Acad Sci USA*. 108: 16143–16144.

Manning JT, Fink B. 2018. Sexual dimorphism in the ontogeny of second (2D) and fourth (4D) digit lengths, and digit ratio (2D:4D). *Am J Hum Biol*. 30:e23138.

Manning JT, Kilduff LP, Trivers R. 2013. Digit ratio (2D:4D) in Klinefelter's syndrome. *Andrology*. 1:94–99.

Manning JT, Scutt D, Wilson J, Lewis-Jones DI. 1998. The ratio of 2nd to 4th digit length: a predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Hum Reprod*. 13:3000–3004.

Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. 2019. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 51:584–591.

McCarthy MM, Crews D. 2008. Epigenetics–new frontiers in neuroendocrinology. *Front Neuroendocrinol*. 29:341–343.

McIntyre MH, Alexander GM. 2011. Sex differences in the fingers of 3 to 5 month old infants do not predict concurrent salivary testosterone levels. *Early Hum Dev*. 87:349–351.

Meikle AW, Bishop DT, Stringham JD, West DW. 1986. Quantitating genetic and nongenetic factors that determine plasma sex steroid variation in normal male twins. *Metabolism*. 35:1090–1095.

Miller KK, Rosner W, Lee H, Hier J, Sesmilo G, Schoenfeld D, Neubauer G, Klibanski A. 2004. Measurement of free testosterone in normal women and women with androgen deficiency: comparison of methods. *J Clin Endocrinol Metab*. 89:525–533.

Mrzljak L, Uylings HBM, Van Eden GG, Judáš M. 1991. Chapter 9 neuronal development in human prefrontal cortex in prenatal and postnatal stages. *Prog Brain Res*. 85:185–222.

Nahoul K, Rao LV, Scholler R. 1986. Saliva testosterone time-course response to hCG in adult normal men. Comparison with plasma levels. *J Steroid Biochem*. 24:1011–1015.

Neufang S, Specht K, Hausmann M, Gunturkun O, Herpertz-Dahlmann B, Fink GR, Konrad K. 2009. Sex differences and the impact of steroid hormones on the developing human brain. *Cereb Cortex Commun*. 19:464–473.

Nguyen TV, McCracken JT, Ducharme S, Cropp BF, Botteron KN, Evans AC, Karama S. 2013. Interactive effects of dehydroepiandrosterone and testosterone on cortical thickness during early brain development. *J Neurosci*. 33:10840–10848.

Paul SN, Kato BS, Cherkas LF, Andrew T, Spector TD. 2006. Heritability of the second to fourth digit ratio (2d:4d): a twin study. *Twin Res Hum Genet*. 9:215–219.

Peper JS, Brouwer RM, Schnack HG, van Baal GC, van Leeuwen M, van den Berg SM, Delemarre-Van de Waal HA, Boomsma DI, Kahn RS, Hulshoff Pol HE. 2009. Sex steroids and brain structure in pubertal boys and girls. *Psychoneuroendocrinology*. 34:332–342.

Petanjek Z, Judaš M, Šimić G, Rašin MR, Uylings HBM, Rakic P, Kostović I. 2011. Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proc Natl Acad Sci USA*. 108:13281–13286.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 38: 904–909.

Pritchard J, Després J-P, Gagnon J, Tchernof A, Nadeau A, Tremblay A, Bouchard C. 1998. Plasma adrenal, gonadal, and conjugated steroids before and after long term overfeeding in identical Twins1. *J Clin Endocrinol Metab*. 83:3277–3284.

Quartier A, Chatrousse L, Redin C, Keime C, Haumesser N, Maglott-Roth A, Brino L, Le Gras S, Benchoua A, Mandel JL et al. 2018. Genes and pathways regulated by androgens in human neural cells, potential candidates for the male excess in autism spectrum disorder. *Biol Psychiatry*. 84:239–252.

Rakic P. 1988. Specification of cerebral cortical areas. *Science (80-)*. 241:170–176.

Rakic P, Bourgeois JP, Goldman-Rakic PS. 1994. Synaptic development of the cerebral cortex: implications for learning, memory, and mental illness. *Prog Brain Res*. 102:227–243.

Ransome MI, Boon WC. 2015. Testosterone-induced adult neurosphere growth is mediated by sexually-dimorphic aromatase expression. *Front Cell Neurosci*. 9:253.

Rasika S, Nottebohm F, Alvarez-Buylla A. 1994. Testosterone increases the recruitment and/or survival of new high vocal center neurons in adult female canaries. *Proc Natl Acad Sci USA*. 91:7854–7858.

Raznahan A, Lee NR, Greenstein D, Wallace GL, Blumenthal JD, Clasen LS, Giedd JN. 2016. Globally divergent but locally convergent X- and Y-chromosome influences on cortical development. *Cereb Cortex*. 26:70–79.

Raznahan A, Lee Y, Stidd R, Long R, Greenstein D, Clasen L, Addington A, Gogtay N, Rapoport JL, Giedd JN. 2010. Longitudinally mapping the influence of sex and androgen signaling on the dynamics of human cortical maturation in adolescence. *Proc Natl Acad Sci USA*. 107:16988–16993.

Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, Clasen L, Gogtay N, Giedd JN. 2011. How does your cortex grow? *J Neurosci*. 31:7174–7177.

Reyes FI, Boroditsky RS, Winter JSD, Faiman C. 1974. Studies on human sexual development. II. Fetal and maternal serum gonadotropin and sex steroid concentrations. *J Clin Endocrinol Metab*. 38:612–617.

Richards G, Bellin W, Davies W. 2017. Familial digit ratio (2D:4D) associations in a general population sample from Wales. *Early Hum Dev*. 112:14–19.

Richards G, Gomes M, Ventura T. 2019. Testosterone measured from amniotic fluid and maternal plasma shows no significant association with directional asymmetry in newborn digit ratio (2D:4D). In: *Journal of developmental origins of health and disease*. Cambridge, UK: Cambridge University Press, pp. 362–367.

Rilling JK, Worthman CM, Campbell BC, Stallings JF, Mbizva M. 1996. Ratios of plasma and salivary testosterone throughout puberty: production versus bioavailability. *Steroids*. 61:374–378.

Ring HZ, Lessov CN, Reed T, Marcus R, Holloway L, Swan GE, Carmelli D. 2005. Heritability of plasma sex hormones and hormone binding globulin in adult male twins. *J Clin Endocrinol Metab*. 90: 3653–3658.

Ritchie SJ, Cox SR, Shen X, Lombardo MV, Reus LM, Alloza C, Harris MA, Alderson HL, Hunter S, Neilson E et al. 2018. Sex differences in the adult human brain: evidence from 5216 UK Biobank participants. *Cereb Cortex*. 28:2959–2975.

Rivas MP, Moreira LMA, Santo LDE, Marques ACSS, El-Hani CN, Toralles MBP. 2014. New studies of second and fourth digit ratio as a morphogenetic trait in subjects with congenital adrenal hyperplasia. *Am J Hum Biol*. 26: 559–561.

Ronalds G, Phillips DIW, Godfrey KM, Manning JT. 2002. The ratio of second to fourth digit lengths: a marker of impaired fetal growth? *Early Hum Dev*. 68:21–26.

Ruigrok ANV, 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. *Neurosci Biobehav Rev*. 39:34–50.

Ruth KS, Day FR, Tyrrell J, Thompson DJ, Wood AR, Mahajan A, Beaumont RN, Wittemans L, Martin S, Busch AS et al. 2020. Using human genetics to understand the disease impacts of testosterone in men and women. *Nat Med*. 26: 252–258.

Sakai LM, Baker LA, Jacklin CN, Shulman I. 1991. Sex steroids at birth: genetic and environmental variation and covariation. *Dev Psychobiol*. 24:559–570.

Savic I, Frisen L, Manzouri A, Nordenstrom A, Lindén Hirschberg A. 2017. Role of testosterone and Y chromosome genes for the masculinization of the human brain. *Hum Brain Mapp*. 38:1801–1814.

Schmidt IM, Chellakooty M, Haavisto AM, Boisen KA, Damgaard IN, Steendahl U, Toppari J, Skakkebæk NE, Main KM. 2002. Gender difference in breast tissue size in infancy: correlation with serum estradiol. *Pediatr Res*. 52:682–686.

Shi F, Wang L, Dai Y, Gilmore JH, Lin W, Shen D. 2012. LABEL: pediatric brain extraction using learning-based meta-algorithm. *Neuroimage*. 62:1975–1986.

Shigeo T, Takao Y, Katsuo T, Haruhiko O, Fujii TK, Yoji N, Masao S. 1977. Sex differences in fetal gonadotropins and androgens. *J Steroid Biochem*. 8:609–620.

Shirtcliff EA, Granger DA, Likos A. 2002. Gender differences in the validity of testosterone measured in saliva by immunoassay. *Horm Behav*. 42:62–69.

Shirtcliff EA, Granger DA, Schwartz EB, Curran MJ, Booth A, Overman WH. 2000. Assessing estradiol in biobehavioral studies using saliva and blood spots: simple radioimmunoassay protocols, reliability, and comparative validity. *Horm Behav*. 38:137–147.

Spritzer MD, Galea LAM. 2007. Testosterone and dihydrotestosterone, but not estradiol, enhance survival of new hippocampal neurons in adult male rats. *Dev Neurobiol*. 67:1321–1333.

Stanworth RD, Kapoor D, Channer KS, Jones TH. 2008. Androgen receptor CAG repeat polymorphism is associated with serum testosterone levels, obesity and serum leptin in men with type 2 diabetes. *Eur J Endocrinol*. 159:739–746.

Stone J, Folkerd E, Doody D, Schroen C, Treloar SA, Giles GG, Pike MC, English DR, Southey MC, Hopper JL et al. 2009. Familial correlations in postmenopausal serum concentrations of sex steroid hormones and other mitogens: a twins and sisters study. *J Clin Endocrinol Metab*. 94:4793–4800.

Studholme C, Kroenke CD, Dighe M. 2020. Motion corrected MRI differentiates male and female human brain growth trajectories from mid-gestation. *Nat Commun*. 11:1–16.

Tapanainen J, Kellokumpu-Lehtinen P, Pelliniemi L, Huhtaniemi I. 1981. Age-related changes in endogenous steroids of human Fetal testis during early and midpregnancy. *J Clin Endocrinol Metab*. 52: 98–102.

Travison TG, Zhuang WV, Lunetta KL, Karasik D, Bhasin S, Kiel DP, Coviello AD, Murabito JM. 2014. The heritability of circulating testosterone, oestradiol, oestrone and sex hormone binding globulin concentrations in men: the Framingham heart study. *Clin Endocrinol (Oxf)*. 80:277–282.

van't Westeinde A, Karlsson L, Thomsen Sandberg M, Nordenström A, Padilla N, Lajic S. 2020. Altered Gray matter structure and White matter microstructure in patients with congenital adrenal hyperplasia: relevance for working memory performance. *Cereb Cortex*. 30:2777–2788.

Van Hemmen J, Cohen-Kettenis PT, Steensma TD, Veltman DJ, Bakker J. 2017. Do sex differences in CEOAEs and 2D:4D ratios reflect androgen exposure? A study in women with complete androgen insensitivity syndrome. *Biol Sex Differ*. 8:1–10.

Van Hulle CA, Moore MN, Shirtcliff EA, Lemery-Chalfant K, Goldsmith HH. 2015. Genetic and environmental contributions to covariation between DHEA and testosterone in adolescent twins. *Behav Genet*. 45:324–340.

Vasung L, Abaci Turk E, Ferradal SL, Sutin J, Stout JN, Ahtam B, Lin PY, Grant PE. 2019. Exploring early human brain development with structural and physiological neuroimaging. *Neuroimage*. 187: 226–254.

Ventura T, Gomes MC, Pita A, Neto MT, Taylor A. 2013. Digit ratio (2D:4D) in newborns: influences of prenatal testosterone and maternal environment. *Early Hum Dev*. 89:107–112.

Vigil P, del Río JP, Carrera B, Ara'nguiz FC, Rioseco H, Cortés ME. 2016. Influence of sex steroid hormones on the adolescent brain and behavior: an update. *Linacre Q*. 83:308–329.

Voracek M, Dressler SG. 2007. Digit ratio (2D:4D) in twins: heritability estimates and evidence for a masculinized trait expression in women from opposite-sex pairs. *Psychol Rep*. 100:115–126.

Walters R. 2020. *Heritability of >4,000 traits & disorders in UK Biobank [WWW Document]*. https://nealelab.github.io/UKBB_ldsc/index.html.

Walters R, Nivard M, van der Zee M, Baya N, Carey C, Palmer D, Verweij K, Neale B. 2019. Exploring sex differences in the genetics of UK Biobank phenotypes. *Eur Neuropsychopharmacol*. 29:S56.

Wang L, Shi F, Li G, Gao Y, Lin W, Gilmore JH, Shen D. 2014. Segmentation of neonatal brain MR images using patch-driven level sets. *Neuroimage*. 84:141–158.

Wang PY, Protheroe A, Clarkson AN, Imhoff F, Koishi K, McLennan IS. 2009. Müllerian inhibiting substance contributes to sex-linked biases in the brain and behavior. *Proc Natl Acad Sci USA*. 106: 7203–7208.

Westberg L, Henningsson S, Landén M, Annerbrink K, Melke J, Nilson S, Rosmond R, Holm G, Anckarsäter H, Eriksson E. 2009. Influence of androgen receptor repeat polymorphisms on personality traits in men. *J Psychiatry Neurosci*. 34:205–213.

Wierenga LM, Bos MGN, Schreuders E, vd Kamp F, Peper JS, Tamnes CK, Crone EA. 2018a. Unraveling age, puberty and testosterone

effects on subcortical brain development across adolescence. *Psychoneuroendocrinology*. 91:105–114.

Wierenga LM, Langen M, Oranje B, Durston S. 2014. Unique developmental trajectories of cortical thickness and surface area. *Neuroimage*. 87:120–126.

Wierenga LM, Sexton JA, Laake P, Giedd JN, Tamnes CK. 2018b. A key characteristic of sex differences in the developing brain: greater variability in brain structure of boys than girls. *Cereb Cortex*. 28: 2741–2751.

Wright FA, Sullivan PF, Brooks AI, Zou F, Sun W, Xia K, Madar V, Jansen R, Chung W, Zhou YH et al. 2014. Heritability and genomics of gene expression in peripheral blood. *Nat Genet*. 46: 430–437.

Xia K, Yu Y, Ahn M, Zhu H, Zou F, Gilmore JH, Knickmeyer RC. 2014. Environmental and genetic contributors to salivary testosterone levels in infants. *Front Endocrinol (Lausanne)*. 5:187.

Xia K, Zhang J, Ahn M, Jha S, Crowley JJ, Szatkiewicz J, Li T, Zou F, Zhu H, Hibar D et al. 2017. Genome-wide association analysis identifies common variants influencing infant brain volumes. *Transl Psychiatry*. 7:e1188–e1188.

Zheng U, Cohn MJ. 2011. Developmental basis of sexually dimorphic digit ratios. *Proc Natl Acad Sci USA*. 108:16289–16294.

Zitzmann M, Nieschlag E. 2001. Testosterone levels in healthy men and the relation to behavioural and physical characteristics: facts and constructs. *Eur J Endocrinol*. 144:183–197.