PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: Whitehouse AJO *et al*. 2015 Prenatal testosterone exposure is related to sexually dimorphic facial morphology in adulthood. *Proc. R. Soc. B* **282**: 20151351.
http://dx.doi.org/10.1098/rspb.2015.1351

Received: 4 June 2015
Accepted: 1 September 2015

Subject Areas:
physiology, behaviour, developmental biology

Keywords:
face shape, testosterone, hormones, prenatal, sexual dimorphism

Author for correspondence:
Andrew J. O. Whitehouse
e-mail: andrew.whitehouse@telethonkids.org.au

†These authors contributed equally to this work.

THE ROYAL SOCIETY PUBLISHING

Prenatal testosterone exposure is related to sexually dimorphic facial morphology in adulthood

Andrew J. O. Whitehouse[1,†], Syed Zulqarnain Gilani[2,†], Faisal Shafait[2,7], Ajmal Mian[2], Diana Weiting Tan[1,3], Murray T. Maybery[3], Jeffrey A. Keelan[4], Roger Hart[4], David J. Handelsman[8], Mithran Goonawardene[5] and Peter Eastwood[6]

[1]Telethon Kids Institute, University of Western Australia, 100 Roberts Road, Subiaco, Perth, Western Australia 6008, Australia
[2]School of Computer Science and Software Engineering, [3]Neurocognitive Development Unit, School of Psychology, [4]School of Women's and Infants' Health, [5]School of Dentistry/Oral Health Centre of Western Australia, and [6]Centre for Sleep Science, School of Anatomy, Physiology and Human Biology, University of Western Australia, 35 Stirling Highway, Crawley, Perth, Western Australia 6009, Australia
[7]School of Electrical Engineering and Computer Science, National University of Science and Technology, Islamabad, Pakistan
[8]ANZAC Research Institute, University of Sydney, Concord Hospital, New South Wales 2139, Australia

Prenatal testosterone may have a powerful masculinizing effect on postnatal physical characteristics. However, no study has directly tested this hypothesis. Here, we report a 20-year follow-up study that measured testosterone concentrations from the umbilical cord blood of 97 male and 86 female newborns, and procured three-dimensional facial images on these participants in adulthood (range: 21–24 years). Twenty-three Euclidean and geodesic distances were measured from the facial images and an algorithm identified a set of six distances that most effectively distinguished adult males from females. From these distances, a 'gender score' was calculated for each face, indicating the degree of masculinity or femininity. Higher cord testosterone levels were associated with masculinized facial features when males and females were analysed together ($n = 183$; $r = -0.59$), as well as when males ($n = 86$; $r = -0.55$) and females ($n = 97$; $r = -0.48$) were examined separately (p-values < 0.001). The relationships remained significant and substantial after adjusting for potentially confounding variables. Adult circulating testosterone concentrations were available for males but showed no statistically significant relationship with gendered facial morphology ($n = 85$, $r = 0.01$, $p = 0.93$). This study provides the first direct evidence of a link between prenatal testosterone exposure and human facial structure.

1. Background

The human face is sexually dimorphic, with the average male face differing from the average female face in the size and shape of, and distance between, the jaws, lips, eyes, nose and cheekbones [1,2]. Even within sex, there are considerable variations in these dimensions, leading to individuals appearing more or less feminine or masculine than the prototypical gendered face. While the origin of this variability remains unclear, there has been significant interest in the influence of the most abundant androgen, testosterone, in the development of face structure.

Genetic sex is determined at conception, but gonadal hormones play a vital role in the differentiation of male and female phenotypes throughout human development [3]. To date, studies investigating the influence of testosterone exposure on facial morphology have focused predominantly on hormone exposure during adolescence. Testosterone levels surge during puberty, and concentrations are 20–30-fold higher in males than females [4], which has been postulated to explain

the contemporaneous increased sexual dimorphism in facial appearance. For example, among 12–18-year-olds, a positive correlation is present between the concentration of testosterone in saliva and the dimensions of several male-typical facial characteristics, such as a broader forehead, chin, jaw and nose [5]. Furthermore, the administration of testosterone to adolescents with delayed puberty accelerates craniofacial growth in these same features [6]. Experimental studies of adults have also revealed a link between contemporaneous testosterone levels and subjective ratings of facial masculinity [7], as well as more objective measurements, such as increased facial width-to-height ratio (fWHR) [8].

However, there is accumulating evidence that the influence of androgens on the development of secondary sex characteristics begins much earlier in life. Biosynthesis of testosterone commences at approximately nine weeks gestation. While testosterone is made in the adrenal cortex and ovary of females, it is produced in far greater amounts by the Leydig cells of the testes in males, and a sex difference in prenatal testosterone level exposure emerges during the second trimester [9] and persists until birth [10]. Studies of non-human mammals have found that exposure to supraphysiological levels of testosterone has a masculinizing effect on physical development, in the form of increased anogenital distance, reduced nipple number, a smaller vaginal opening and increased body growth [11–14]. Sex differences in facial morphology are apparent in six-month-old infants [15], and increase steadily across childhood [16]. While testosterone surges and associated growth spurts during puberty exaggerate existing sex differences in face structure, it has been postulated that the pre-existing sexual dimorphism in face structure is related to differences in prenatal testosterone exposure [17].

Investigating the proposed link between testosterone and face development in humans has been hampered by two methodological challenges. The first has been the valid and reliable measurement of prenatal testosterone exposure. As it is unethical to manipulate hormone concentrations in the human fetus, the majority of clinical research has used surrogate measures of the prenatal hormone environment, such as the ratio of the second digit to the fourth digit (2D : 4D) [18] and the observation of fetuses exposed to aberrant hormone environments *in utero* (e.g. females with congenital adrenal hyperplasia) [19,20]. However, limitations of these approaches have been widely documented, with reports of poor correlations between 2D : 4D and more direct measures of the prenatal hormone environment [21,22], as well as concerns about the extent to which data from clinical populations can be extrapolated to the general population.

An alternative approach for sampling the prenatal androgen environment is investigating the hormonal milieu of umbilical cord blood. Cord blood can be obtained at delivery in normal pregnancy, and therefore randomly selected participant samples are more likely to be representative of the general population. Several studies have reported higher testosterone levels in cord blood samples from male versus female fetuses [10,23], and thus these samples are thought to reflect fetal circulation during late gestation. A limitation of this approach is that testosterone levels in cord blood may not reflect concentrations during the first and second trimester, in particular gestational weeks 8–24, which has traditionally been regarded as a 'sensitive period' for the maximal effects of sex steroids on human development [24]. However, there is increasing recognition that there may be multiple sensitive periods, and animal studies have found that postnatal development may be affected by hormones at different times throughout prenatal development [25]. Human studies have also started to link sex steroids from umbilical cord blood to a range of childhood behaviours, including language development [26,27], internalizing and externalizing behaviours [28], and spatial abilities [29]. To date, no studies have examined the relationship between cord blood testosterone and human physical development.

The second methodological challenge has been the accurate measurement of facial shape. Previous studies have focused on either subjective masculinity ratings of two-dimensional (2D) facial stimuli [30], or measurements between standardized landmarks on 2D photographs [17,31–34]. Distances based on 2D photographs are unreliable as they change with the camera distance, angle and optics. Some researchers have used three-dimensional (3D) models to calculate Euclidean (linear) distances between anatomical landmarks. However, the use of these measurement techniques may miss important information since Euclidean distances 'cut through' the facial features and thus do not take the underlying morphology into account. For example, calculation of some characteristics of facial dimorphism using these distances does not correlate well with human perceptual measurements [35]. In contrast, geodesic (3D surface) distances are known to model the 3D facial structure in a better way as compared with the linear distances. For example, one study has reported that objective scores for masculinity calculated from 3D facial images correlate more highly with subjective (perceptual) masculinity scores if the objective scores are based on Euclidean and geodesic distances, rather than on Euclidean distances alone [35].

The current study represents the first to investigate the potential link between prenatal testosterone exposure and objective measurements of 3D facial masculinity in adults. The current study spanned a 25-year period, in which umbilical cord blood was collected at birth (1989–1991) and facial morphology was examined on the same participants in early adulthood (2012–2014). A total of 97 males and 86 females (all Caucasian) had both cord blood and adult facial morphology data available, and thus provided the data for the current study. Blood collected in early adulthood was available for the male participants only, and testosterone concentrations were measured in these samples. The 2D : 4D measurements from both hands of the adult participants were also procured and included in the analyses.

2. Material and methods

(a) Participants

Participants were part of the Western Australian Pregnancy Cohort (Raine) Study [36], which is a longitudinal investigation of women between 16 and 20 weeks pregnancy from the public antenatal clinic at King Edward Memorial Hospital (Perth, Australia) or surrounding private clinics, between May 1989 and November 1991. Approximately 100 unselected antenatal patients per month were enrolled during this period, with a final sample of 2900 women. The inclusion criteria were a gestational age between 16 and 20 weeks, English language skills sufficient to understand the study demands, an expectation to deliver at King Edward Memorial Hospital and an intention to remain in Western Australia to enable future follow-up of their child. By the end of the recruitment period, 2868 live births (96%) were available for follow-up.

Umbilical cord blood was obtained at the birth of 861 singleton deliveries (selected randomly). The participants in this study were the 97 Caucasian males and 86 Caucasian females who volunteered to have their faces photographed in early adulthood (age range: 21–24 years). There was no statistically significant difference in age at the time of facial photography between males ($M = 22.81$ years, s.d. = 0.61) and females ($M = 22.91$, s.d. = 0.60; $t_{181} = 1.04$, $p = 0.31$). Compared with the remainder of the Raine cohort, the participants who volunteered for the current study were more likely to have come from families who had a family income above the Australian government-designated 'poverty line' (annual household income of 24 000 Australian dollars) at the time of their birth (current study participants: $n = 118/180$ [65.6%] with available data; remainder of cohort: $n = 1394/2500$ [55.8%] with available data; $\chi^2 = 6.553$, d.f. = 1, $p = 0.01$), and whose mothers had completed secondary school (current study participants: $n = 85/183$ [46.4%] with available data; remainder of cohort: $n = 1024/2649$ [38.5%] with available data; $\chi^2 = 4.533$, d.f. = 1, $p = 0.03$). All participants in the current study had a normal obstetric history and no history of neurological disorder or other conditions known to affect facial morphology.

(b) Umbilical cord blood procurement, storage and analysis

At birth, mixed arterial and venous umbilical cord blood was obtained. Blood samples were immediately centrifuged, plasma isolated and then stored at −80°C without thawing. Detailed sequence analysis of DNA obtained from 10 randomly selected maternal–child pairs confirmed that the cord blood samples were not contaminated by maternal blood. In January 2010, these serum samples were thawed and analysed for androgen content. Total testosterone was measured by liquid chromatography-tandem mass spectrometry (LCMS/MS) after solvent extraction as described in detail by Keelan *et al.* [10]. The limit of quantitation was 0.025 ng ml^{-1} (0.08 nmol l^{-1}). Inter-batch variation was low at 6–11% ($n = 24$); recovery from cord serum was 93–98%. Sex hormone-binding globulin (SHBG) was measured by ELISA using a commercial kit (IBL International, Hamburg, Germany) according to the manufacturer's instructions. All samples were measured in duplicate by a single operator using assay kits from the same batch. The inter-assay imprecision was less than 4.5% ($n = 25$). Intra-assay variation was 5.2% ($n = 861$). Samples with an initial replicate coefficient of variation of more than 10% were reanalysed. BioT (nmol l^{-1}), representing the fraction of total testosterone either free (unsequestered by SHBG) or bound to serum albumin, was calculated using the formula: BioT = (free testosterone) + (albumin-bound testosterone) [37]. Free testosterone was calculated using the empirical method and formula described by Sartorius *et al.* [38]. Albumin levels were adjusted using published reference values to take into account the decrease in serum albumin concentrations with gestational age [39].

(c) Adult face, digit ratio and testosterone measurements and analysis

(i) Adult facial photography

The 3dMD face system (3dMD LLC, Atlanta, GA, USA) is a non-invasive imaging technology that produces 3D images (180° ear-to-ear frontal view) using random infra-red light projection on the face. The technology measures the 3D shape of the face to millimetre precision and maps the colour texture over it. Images captured by a 3dMD face system have been shown to be highly precise and replicable [40]. Participants sat in front of the 3dMD face scanner with the distance between the chair and scanner adjusted so their faces appeared in the middle of the computer screen. During the imaging process, participants fixated their gaze on a sticker pasted on the wall behind the scanner, maintained a neutral facial expression and kept their mouths closed. No accessories were worn for the imaging process and loose hair was pinned back from the face.

(ii) Adult digit ratio

Digit ratio was measured on the participants at the same testing session during which facial photography was conducted. The lengths of the second and fourth fingers on the ventral surface of the left and right hand were measured from the basal crease of the digit to its tip using Vernier callipers, working from purpose-collected hand photocopies (Lanier LD 122 photocopier) under standardized and supervised conditions [41]. All measurements were made by one observer and repeated by a second observer blinded to the findings of the first. Digit ratio was calculated by dividing the length of the second digit by the length of the fourth digit. Separate ratios were calculated for the left and right hand.

(iii) Adult testosterone concentrations

At a previous testing session approximately 3 years earlier ($M = 20.00$ years, s.d. = 0.53), 85 of the male participants had blood procured via venepuncture as part of the protocol for an unrelated study. To allow an investigation of the effect of adult testosterone levels on facial morphology, testosterone concentrations were measured in serum taken from these samples by LCMS/MS as previously described [42], with limits of quantitation of 0.025 ng ml^{-1}. While hormone concentrations can vary depending upon the time of day and individual circumstance (e.g. level of activity, competition, stress), there is evidence that a single blood sample provides an accurate estimate of circulating testosterone levels in males over longer periods [43]. To minimize variation in measurement error between individuals, all blood samples were collected in the morning (i.e. prior to midday). Blood samples from female participants were not available for analysis.

(d) Statistical analyses

Analyses first focused on the facial images. Twenty-one facial landmarks defined by Farkas [1] were annotated on each image (figure 1). Next, 21 Euclidean (linear) and 20 geodesic 3D distances were measured between these landmarks. We defined geodesic distance as the length of the 3D curve that is generated by orthogonal projection of the Euclidean line between the points on the 3D facial surface. The gradient-based efficient feature selection (GEFS) algorithm [44] was used to select the features that maximally distinguish between male and female faces. The algorithm re-orders the features in each iteration to maximize the classification accuracy. A mathematical model is then used to train a linear discriminant analysis (LDA) classifier [35] with the selected features using a 10-fold validation technique. A 'gender score' is then created by comparing the sexually dimorphic features of each face with the mean measurements of both males and females in the LDA space. Figure 2 provides a full description of this procedure. Gender scores vary from 0 (highly masculine facial features) to 1 (highly feminine facial feature). To facilitate comparisons with previous studies, we also calculated fWHR according to published guidelines [8].

We then examined the relationships between the facial gender scores and a direct measure of prenatal testosterone exposure (BioT from umbilical cord blood), a proxy measure of prenatal testosterone exposure (adult digit ratio) and a direct measure of adult testosterone exposure (nM concentration from adult blood). Correlations with BioT were also calculated using free testosterone and total testosterone levels. These analyses were conducted for males and females separately, and also with data from both

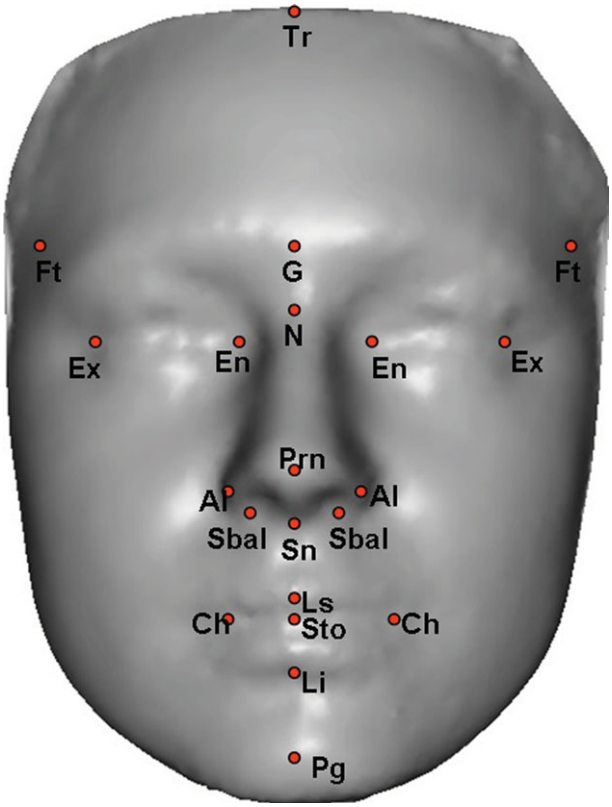

Figure 1. Three-dimensional image annotated with 21 facial landmarks. (Online version in colour.)

sexes combined. An alpha value of $p < 0.05$ was used to indicate statistical significance.

3. Results

(a) Facial gender score

GEFS selected five linear (forehead width, nasal bridge length, nasal tip protrusion, upper lip height and nose width) and seven geodesic (intercanthal width, forehead width, outer canthal width, nasal bridge length, nasal tip protrusion, upper lip height and nose width) distances as the most discriminating features between the two sexes. These distances are outlined in figure 3. A mathematical model was generated based on these distances (described in figure 2) and was found to correctly classify male and female faces with 99.47% accuracy.

(b) Cord blood testosterone and adult facial morphology

Cord blood samples were analysed for bioactive concentrations of testosterone (BioT). A significantly higher level of BioT was present in the cord blood of males ($M = 0.16$, s.d. $= 0.12$) compared with females ($M = 0.07$, s.d. $= 0.05$; $t_{181} = 6.0$, $p < 0.001$). Table 1 presents correlations between testosterone, hand and facial measurements. There was a significant negative association between BioT levels and gender scores when males and females were analysed together ($r = -0.59$), as well significant correlations when males ($r = -0.55$) and females ($r = -0.48$) were examined separately (all p-values < 0.001). Figure 4 presents the correlations between BioT and gender scores for males and females, and shows two outlying data points for each sex. Pearson correlation coefficients calculated after excluding these data points remained statistically significant for both the male ($r = -0.34$, $p = 0.001$) and female data ($r = -0.40$, $p < 0.001$), and also when males and female data were combined ($r = -0.60$, $p < 0.001$). There was no statistically significant correlations between BioT concentrations and fWHR for males ($r = 0.17$, $p = 0.09$), females ($r = 0.16$, $p = 0.15$), nor when the data from both sexes were combined ($r = 0.10$, $p = 0.17$).

Correlations of similar magnitude and statistical significance ($p < 0.001$) were observed between free testosterone levels from cord blood and facial gender scores for females ($r = -0.43$), males ($r = -0.52$) and the combined sample ($r = -0.51$). Correlations between facial gender scores and total testosterones for males ($r = -0.40$), females ($r = -0.25$) and combined samples ($r = -0.52$) were reduced in magnitude but statistically significant at the level of $p < 0.05$.

To further examine the significant associations between BioT and facial gender scores, hierarchical multiple regression analyses were performed. In the first block, we included the potentially confounding variables of gestational age at birth, the presence of labour at birth [10], maternal smoking and alcohol consumption during pregnancy, maternal age at birth, family income and the total facial area measured by 3D photography (in millimetres). In the second block, we included BioT. For females, BioT concentrations were significantly associated with facial gender scores ($B = -1.07$, 95% confidence interval $= -1.51$, -0.63, $p < 0.001$). The inclusion of BioT concentrations significantly improved the regression model ($F_{1,75} = 23.28$, $p < 0.001$) for the change in R^2, accounting for an additional 20.6% of the variation in gender score. The same analyses were conducted for males, except that we added adult testosterone concentrations into the first block of the regression model. There was a significant relationship between BioT and gender scores in males ($B = -0.46$, 95% CI $= -0.69$, -0.24, $p < 0.001$), and the inclusion of BioT concentrations significantly improved the regression model ($F_{1,72} = 16.93$, $p < 0.001$) for the change in R^2, accounting for an additional 13.2% of the variation in gender score. When the two outlying data points were excluded for each sex (figure 4), improvements in the regression models with the addition of BioT concentrations remained significant (females: $F_{1,73} = 10.90$, $p < 0.001$; males: $F_{1,70} = 8.14$, $p < 0.01$).

(c) Adult digit ratio (2D : 4D) and adult facial morphology

Consistent with expectations, small but statistically significant sex differences were also observed for 2D : 4D ratios. Males had significantly smaller 2D : 4D ratios than females for both the left hand (males: $n = 82$, $M = 0.96$, s.d. $= 0.03$; females: $n = 0.97$, $M = 0.97$, s.d. $= 0.03$; $t_{156} = 2.54$, $p = 0.01$) and the right hand (males: $n = 0.83$, $M = 0.97$, s.d. $= 0.03$; females: $n = 76$, $M = 0.98$, s.d. $= 0.04$; $t_{181} = 2.97$, $p = 0.003$). When data from males and females were combined, there were statistically significant correlations between facial gender scores and the 2D : 4D ratio on the left hand ($n = 158$, $r = 0.20$, $p = 0.01$), and a trend with the 2D : 4D ratio on the right hand ($n = 159$, $r = 0.14$, $p = 0.07$). However, table 1 shows that there were no statistically significant correlations between digit ratios and gender scores when males (left hand: $r = -0.02$, $p = 0.85$; right hand: $r = -0.08$, $p = 0.46$) and females (left hand: $r =$

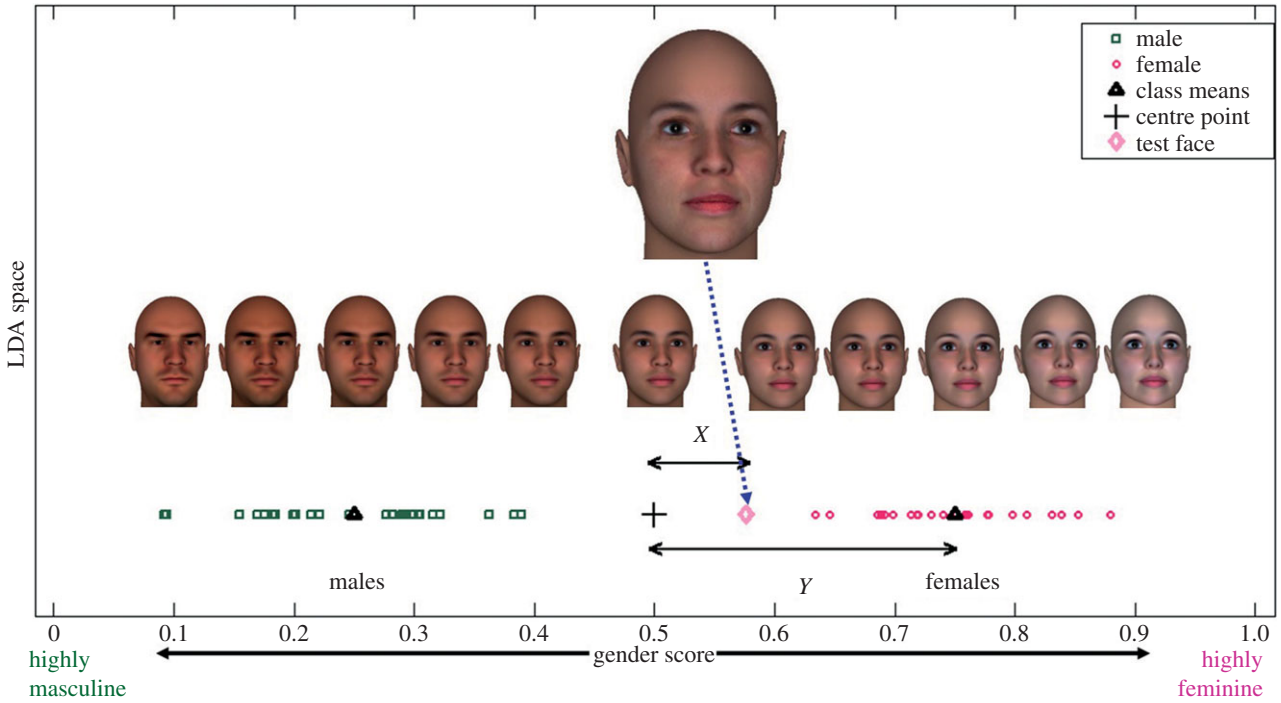

Figure 2. Creation of the 'gender score' for each face. The selected features of each 3D face in the training data are projected in the LDA space, which separates the two classes of males and females (figure 3). We find the mean of both classes and the centre point between these means in the LDA space. These are shown in figure 3 as black triangles and a black cross. The selected features of each test face (pink diamond) are then projected in the LDA space. The algorithm finds the distance between the test face and the centre of the mean of the two classes (the black cross) and ascribes a 'gender score' as $G = 1 - X/2Y$. The further this test face is from the centre, the higher the masculinity or femininity. In the particular example, the test face (pink diamond) lies between the centre point and the mean for females, and so will have a 'gender score' that represents low femininity. The synthetic faces shown in the figure depict the varying masculinity of the same identity. (Online version in colour.)

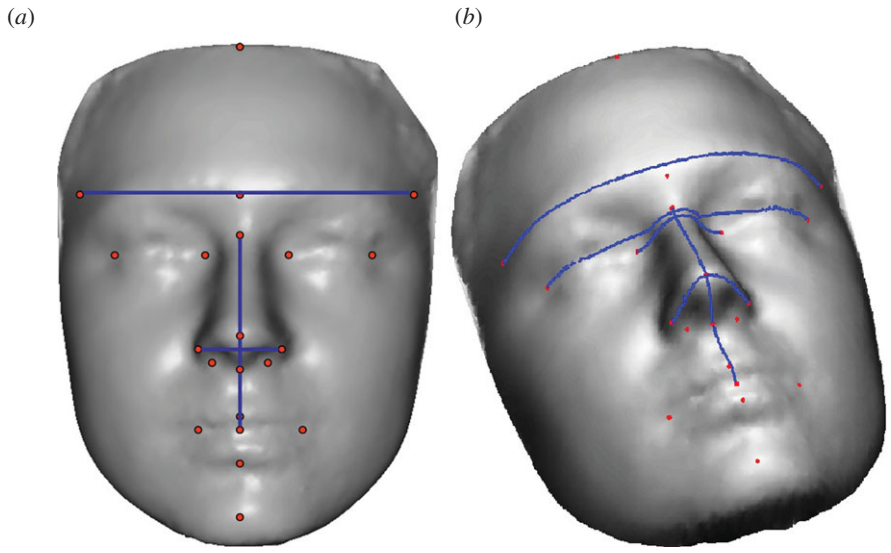

Figure 3. (*a*) Five Euclidean and (*b*) seven geodesic distances that maximally separate males and females in the LDA space. (Online version in colour.)

0.16, $p = 0.16$; right hand: $r = -0.07$, $p = 0.51$) were analysed separately.

(d) Adult testosterone and adult facial morphology

Blood samples in adulthood were available for 85 males. Pearson correlations presented in table 1 revealed no statistically significant relationship between testosterone correlations in these samples and facial gender scores ($r = 0.01$, $p = 0.93$) or fWHR ($r = -0.01$, $n = 0.92$). Additional analyses revealed a weak, but statistically significant association between umbilical cord BioT levels and adult testosterone concentrations ($r = 0.23$, $p = 0.04$). However, there was no statistically relationship between adult testosterone concentrations and adult 2D : 4D ratio on the left ($r = 0.06$, $p = 0.63$) and right hand ($r = 0.07$, $p = 0.57$).

4. Discussion

The data collected in this study provide the first direct evidence for a long-hypothesized link between prenatal levels of testosterone and adult facial morphology [17]. After creating an objective algorithm that quantified the

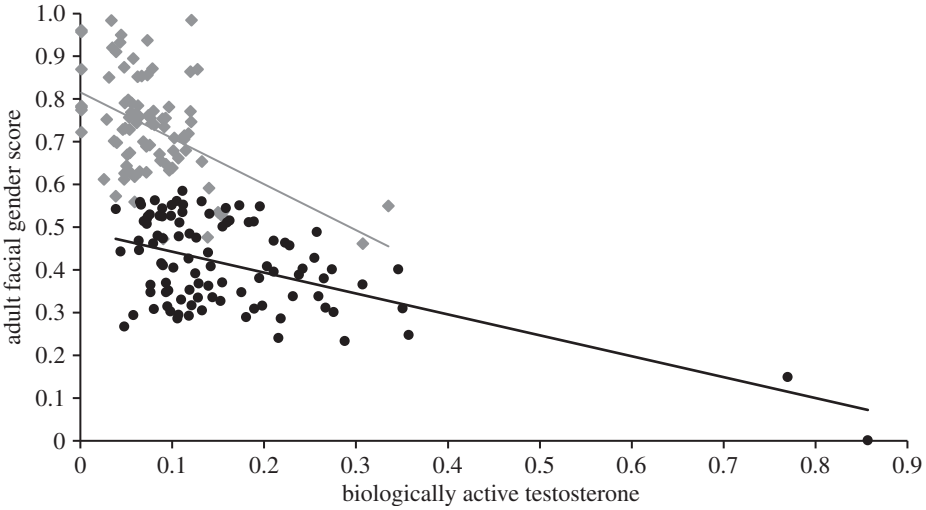

Figure 4. Scatter plot and trend lines showing the relationship between testosterone levels from umbilical cord blood and facial 'gender scores' for females (grey diamonds) and males (black circles). Facial gender scores range from very masculine (score of 0) to very feminine (score of 1).

Table 1. Pearson correlations between testosterone levels in umbilical cord blood (BioT) and adult blood (Adult T), 2D : 4D in the left and right hand, the facial gender score derived by 3D facial photography and the facial width-to-height ratio (fWHR). Correlations for females are below the diagonal, and correlations for males are above the diagonal. The number of data points in each correlation is included in parentheses. No data on adult testosterone concentrations were available for females.

	1	2	3	4	5	6
(1) BioT	1	0.23* (85)	0.10 (82)	−0.02 (83)	−0.55** (97)	0.17 (97)
(2) Adult T	—	1	0.06 (78)	0.07 (79)	0.01 (85)	−0.01 (83)
(3) 2D : 4D left hand	−0.29* (76)	—	1	0.71** (82)	−0.02 (82)	−0.01 (82)
(4) 2D : 4D right hand	0.01 (76)	—	0.62** (76)	1	−0.08 (83)	0.11 (83)
(5) facial gender score	−0.48** (86)	—	0.16 (76)	−0.08 (76)	1	−0.01 (97)
(6) fWHR	0.16 (86)	—	−0.22 (76)	−0.12 (76)	0.08 (86)	1

*$p < 0.05$; **$p < 0.01$.

'genderness' of faces based on Euclidean and geodesic distances between facial landmarks, we identified significant associations between biologically active testosterone measured from umbilical cord blood and the sexual dimorphism of a face. Higher levels of testosterone were related to increased facial masculinity both within and between sex, with correlations indicating a medium-to-strong effect size. By contrast, where data were available (males only), there was no significant associations between adult testosterone concentrations and facial gender score. The associations between cord blood BioT and facial phenotypes remained in both sexes even after controlling for potentially confounding antenatal, sociodemographic and physical variables.

Animal studies have provided consistent evidence that prenatal and perinatal testosterone exposure can exert organizational effects on the architecture of the body and brain [11–14,25]. Given the methodological challenges of measuring the hormone environment *in utero*, human studies have typically relied upon 2D : 4D as an index of prenatal testosterone exposure. While several studies have reported a link between reduced 2D : 4D ratio (purportedly reflecting exposure to increased levels of prenatal testosterone) and facial masculinity [31–33], there have been a number of notable failures to replicate these findings [45,46]. We have previously shown no association between testosterone levels in umbilical cord blood and adult 2D : 4D [21,22], and the current study found no within-sex link between 2D : 4D and the facial gender scores for either males or females. Taken together, these data suggest caution when using 2D : 4D as a surrogate measure of prenatal testosterone exposure, and that more direct measurement techniques, such as the biochemical analysis of umbilical cord blood, may reveal greater insight into links between the prenatal hormone environment and postnatal development.

While the current findings indicate that adult facial morphology is more closely related to prenatal testosterone exposure than adult concentrations, the influence of adolescent testosterone levels cannot be ruled out. The surge of serum testosterone concentrations during puberty is 20–30-fold higher in males compared with females [4], and salivary testosterone levels of 12–18-year-olds are known to be related to sexually dimorphic facial features [5]. The mean age of participants in the current study at the time of 'adult' blood collection was approximately 20 years, and all were post-puberty. Previous studies have identified no relationship between prenatal and adolescent testosterone concentrations in females [47], but to the best of our knowledge, there has been no investigations of these variables in males. Future studies can build on the research presented here by investigating the correlation between testosterone concentrations measured in umbilical

cord and adolescent blood, as well as comparing their relationship with adult facial morphology.

The current findings present a unique contribution to the large body of literature that has linked behavioural characteristics with facial phenotypes. A more masculinized face structure has been associated with aggression [48], risk-taking behaviour [49,50] and dominance [51]. These relationships are in the same directions as the links between prenatal testosterone exposure and these same behavioural characteristics [24], raising the possibility that testosterone underpins both sets of relationships. Exposure to increased levels of prenatal testosterone has also been linked with several neurodevelopmental disorders, in particular autism spectrum disorder (ASD) [52] and developmental language disorder [26]. While these conditions have not traditionally been thought to be characterized by a specific facial phenotype, the current findings would predict a more masculinized facial structure among affected individuals. No study has conducted 3D facial analysis on children with developmental language disorder, but there is emerging evidence that a more masculinized face structure may be associated with the ASD phenotype [53]. For example, Aldridge *et al.* [54] conducted a 3D modelling study of the faces of 65 boys (aged 8–12 years) with ASD and 41 typically developing comparison males. After mapping landmarks to each face, Euclidean distance matrix analyses revealed two distinct facial morphology subgroups within the broader ASD sample. One of these subgroups had an increase in the length and breadth of the face, and greater distance between the mouth and chin, which is a facial phenotype thought to characterize male-type features.

The participants were recruited from a longitudinal pregnancy cohort study that has been active for 25 years. Only a small proportion (6%) of the originally recruited cohort contributed data to the current study, primarily due to umbilical cord blood being collected on only a minority of the cohort (30%), and sample attrition in the two decades since this time. A non-random pattern of attrition is commonly observed in longitudinal cohort studies [55], and the current study found that socially disadvantaged participants were less likely to volunteer for the facial photography in adulthood. While the loss of these participants may have affected the results of this study, the original cohort over-represented socially disadvantaged women [36], and therefore this pattern of attrition may have increased the extent to which the remaining cohort is representative of the general population, and hence the findings can be generalized. Furthermore, the majority of other studies in this area have drawn participant samples from undergraduate university courses [7,17,31,33,45]. These cohorts are likely to have a similar bias against participants from low socio-economic strata [56], which enables a relatively equivalent comparison between studies.

We also highlight that the index of biologically active concentrations of prenatal testosterone (BioT) was not measured directly but derived from our direct measurement of total testosterone and SHBG concentrations in umbilical cord blood, and published norms of albumin concentrations that take into account hormone changes during pregnancy. While we used standard protocols to calculate BioT [38], it is possible that a degree of measurement inaccuracy was introduced into the analyses by these calculations. However, we suggest that this possibility is small given that the correlations between the facial 'gender score' and total and free testosterone concentrations were of a similar magnitude (and in the same direction), and all achieved statistical significance. Finally, it has been suggested that the use of dense sampling of 3D facial landmarks may provide a better representation of sexual dimorphism than the sparse landmark-based approach used in the current study [57,58]. Future studies that examine whether the significant and substantial correlations identified in the current study are also observed with measurements of the full facial complex [58] will provide the research field with critical information regarding how best to quantify facial structure.

In summary, using umbilical cord blood and 3D facial photography, we report the findings from a 20-year study, revealing a significant relationship between exposure to increasing levels of prenatal testosterone and masculinization of the adult face in both males and females. The current findings provide the first direct evidence for a long-hypothesized link between early testosterone exposure and face structure, and provide further support for the organizational effects of the prenatal hormone environment on postnatal development of the human body.

Ethics. Participant recruitment and all follow-ups of the study families were approved by the Human Ethics Committee at King Edward Memorial Hospital and/or Princess Margaret Hospital for Children in Perth. Informed consent was obtained from all mothers and offspring who participated in this study.

Data accessibility. The datasets supporting this article can be accessed at the following online repository: http://datadryad.org/review?-doi=doi:10.5061/dryad.k5nr3.

Authors' contributions. A.J.O.W., S.Z.G., M.T.M. and A.M. formulated the hypotheses for this study. A.J.O.W., M.T.M. and J.A.K. coordinated the analyses of the umbilical cord blood samples, and R.H. and D.J.H. coordinated the analyses of adult blood samples. P.E., M.G. and D.W.T. coordinated the collected of the adult facial photographs, and the facial analyses were conducted by S.Z.G., F.S. and A.M. All authors were closely involved in manuscript preparation, including data analysis, manuscript writing and the provision of critical feedback.

Competing interests. We have no competing interests.

Funding. The core management of the Raine Study is funded by The University of Western Australia (UWA), Curtin University, The Telethon Kids Institute, Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, Women's and Infant's Research Foundation and Edith Cowan University. The 20-year data collection was funded by the NHMRC (Hart *et al.*, ID 634457). A.J.O.W. (1077966) and P.E. (1042341) are funded by Senior Research Fellowships from the NHMRC. S.Z.G. and D.W.T. are supported by International Postgraduate Research Scholarships in Australia. A.M. (DP 110102399) and F.S. (LP 110201008) are supported by the Australian Research Council.

Acknowledgments. We are grateful to the Raine Study participants and their families, and we thank the Raine Study staff for cohort coordination and data collection.

References

1. Farkas LG. 1981 *Anthropometry of the head and face in medicine*. New York, NY: Elsevier.
2. Enlow DH. 1996 *Essential of facial growth*. Philadelphia, PA: Saunders Company.
3. Hines M. 2011 Gender development and the human brain. *Ann. Rev. Neurosci.* **34**,

69–88. (doi:10.1146/annurev-neuro-061010-113654)

4. Fechner PY. 2003 The biology of puberty: new developments in sex differences. In *Gender differences at puberty* (ed. C Hayward), pp. 17–28. Cambridge, UK: Cambridge University Press.
5. Marečková K *et al.* 2011 Testosterone-mediated sex differences in the face shape during adolescence: subjective impressions and objective features. *Horm. Behav*. **60**, 681–690. (doi:10.1016/j.yhbeh.2011.09.004)
6. Verdonck A, Gaethofs M, Carels C, de Zegher F. 1999 Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. *Eur. J. Orthod*. **21**, 137–143. (doi:10.1093/ejo/21.2.137)
7. Penton-Voak IS, Chen JY. 2004 High salivary testosterone is linked to masculine male facial appearance in humans. *Evol. Hum. Behav*. **25**, 229–241. (doi:10.1016/j.evolhumbehav.2004.04.003)
8. Lefevre CE, Lewis GJ, Perrett DI, Penke L. 2013 Telling facial metrics: facial width is associated with testosterone levels in men. *Evol. Hum. Behav*. **34**, 273–279. (doi:10.1016/j.evolhumbehav.2013.03.005)
9. Grumbach MM, Hughes IA, Conte FA. 2003 Disorders of sex differentiation. In *Williams textbook of endocrinology*, 10th edn (eds S Melmed, KS Polonsky, PR Larsen, HM Kronenberg), pp. 842–1002. Philadelphia, PA: Saunders.
10. Keelan JA, Mattes E, Tan H, Dinan A, Newnham JP, Whitehouse AJO, Jacoby P, Hickey M. 2012 Androgen concentrations in umbilical cord blood and their association with maternal, fetal and obstetric factors. *PLoS ONE* **7**, e42827. (doi:10.1371/journal.pone.0042827)
11. Biancardi MF *et al.* In press. Prenatal exposure to testosterone masculinises the female gerbil and promotes the development of lesions in the prostate (Skene's gland). *Reprod. Fertil. Dev*. (doi:10.1071/RD13387)
12. Wolf CJ, Hotchkiss A, Ostby JS, LeBlanc GA, Gray LE. 2002 Effects of prenatal testosterone propionate on the sexual development of male and female rats: a dose–response study. *Toxicol. Sci*. **65**, 71–86. (doi:10.1093/toxsci/65.1.71)
13. Hotchkiss AK, Lambright CS, Ostby JS, Parks-Saldutti L, Vandenbergh JG, Gray LE. 2007 Prenatal testosterone exposure permanently masculinizes anogenital distance, nipple development, and reproductive tract morphology in female Sprague–Dawley rats. *Toxicol. Sci*. **96**, 335–345. (doi:10.1093/toxsci/kfm002)
14. Slob AK, Goy RW, van der Werf ten Bosch JJ. 1973 Sex differences in growth of guinea-pigs and their modification by neonatal gonadectomy and prenatally administered androgen. *J. Endocrinol*. **58**, 11–19. (doi:10.1677/joe.0.0580011)
15. Bulygina E, Mitteroecker P, Aiello L. 2006 Ontogeny of facial dimorphism and patterns of individual development within one human population. *Am. J. Phys. Anthropol*. **131**, 432–443. (doi:10.1002/ajpa.20317)
16. Snodell SF, Nanda RS, Currier GF. 1993 A longitudinal cephalometric study of transverse and vertical craniofacial growth. *Am. J. Orthod. Dentofacial Orthop*. **104**, 471–483. (doi:10.1016/0889-5406(93)70073-W)
17. Neave N, Laing S, Fink B, Manning JT. 2003 Second to fourth digit ratio, testosterone, and perceived male dominance. *Proc. R. Soc. Lond. B* **270**, 2167–2172. (doi:10.1098/rspb.2003.2502)
18. Manning JT, Bundred PE, Flanagan BF. 2002 The ratio of 2nd to 4th digit length: a proxy for transactivation activity of the androgen receptor gene? *Med. Hypotheses* **59**, 334–336. (doi:10.1016/S0306-9877(02)00181-0)
19. Hines M, Hines C. 1984 Prenatal exposure to diethylstilbestrol (DES) and the development of sexually dimorphic cognitive abilities and cerebral lateralization. *Dev. Psychol*. **20**, 81–94. (doi:10.1037/0012-1649.20.1.81)
20. Kelso W, Kelso MER, Nicholls G, Warne M. 2000 Cerebral lateralization and cognitive functioning in patients with congenital adrenal hyperplasia. *Neuropsychology* **14**, 370–378. (doi:10.1037/0894-4105.14.3.370)
21. Hickey M, Doherty DA, Hart R, Norman RJ, Mattes E, Atkinson HC, Sloboda DM. 2010 Maternal and umbilical cord androgen concentrations do not predict digit ratio (2D:4D) in girls: a prospective cohort study. *Psychoneuroendocrinology* **35**, 1235–1244. (doi:10.1016/j.psyneuen.2010.02.013)
22. Hollier LP, Keelan JA, Jamnadass ESL, Maybery MT, Hickey M, Whitehouse AJO. 2015 Adult digit ratio (2D:4D) is not related to umbilical cord androgen or estrogen concentrations, their ratios or net bioactivity. *Early Hum. Dev*. **91**, 111–117. (doi:10.1016/j.earlhumdev.2014.12.011)
23. Herruzo AJ, Mozas J, Alarcon JL, Lopez JM, Molina R, Molto L, Martos J. 1993 Sex-differences in serum hormone levels in umbilical vein blood. *Int. J. Gynaecol. Obstet*. **41**, 37–41. (doi:10.1016/0020-7292(93)90152-M)
24. Cohen-Bendahan C, van de Beek C, Berenbaum S. 2005 Prenatal sex hormone effects on child and adult sex-typed behavior: methods and findings. *Neurosci. Biobehav. Rev*. **29**, 353–384. (doi:10.1016/j.neubiorev.2004.11.004)
25. Roselli C, Estill C, Stadelman H, Meaker M, Stormshak F. 2011 Separate critical periods exist for testosterone-induced differentiation of the brain and genitals in sheep. *Endocrinology* **152**, 409–415. (doi:10.1210/en.2010-1445)
26. Whitehouse A, Mattes E, Maybery M, Sawyer M, Jacoby P, Keelan J, Hickey M. 2012 Sex-specific associations between umbilical cord blood testosterone levels and language delay in early childhood. *J. Child Psychol. Psychiatr*. **53**, 726–734. (doi:10.1111/j.1469-7610.2011.02523.x)
27. Hollier LP, Mattes E, Maybery MT, Keelan JA, Hickey M, Whitehouse AJO. 2013 The association between perinatal testosterone concentration and early vocabulary development: a prospective cohort study. *Biol. Psychol*. **92**, 212–215. (doi:10.1016/j.biopsycho.2012.10.016)
28. Robinson M, Whitehouse AJO, Jacoby P, Mattes E, Sawyer MG, Keelan JA, Hickey M. 2013 Umbilical cord blood testosterone and childhood internalizing and externalizing behavior: a prospective study. *PLoS ONE* **8**, e59991. (doi:10.1371/journal.pone.0059991)
29. Jacklin CN, Wilcox KT, Maccoby EE. 1988 Neonatal sex-steroid hormones and cognitive-abilities at 6 years. *Dev. Psychobiol*. **21**, 567–574. (doi:10.1002/dev.420210607)
30. Ferdenzi C, Lemaître J-F, Leongómez JD, Roberts SC. 2011 Digit ratio (2D:4D) predicts facial, but not voice or body odour, attractiveness in men. *Proc. R. Soc. B* **278**, 3551–3557. (doi:10.1098/rspb.2011.0544)
31. Burriss R, Little A, Nelson E. 2007 2D:4D and sexually dimorphic facial characteristics. *Arch. Sex Behav*. **36**, 377–384. (doi:10.1007/s10508-006-9136-1)
32. Meindl K, Windhager S, Wallner B, Schaefer K. 2012 Second-to-fourth digit ratio and facial shape in boys: the lower the digit ratio, the more robust the face. *Proc. R. Soc. B* **279**, 2457–2463. (doi:10.1098/rspb.2011.2351)
33. Fink B, Grammer K, Mitteroecker P, Gunz P, Schaefer K, Bookstein FL, Manning JT. 2005 Second to fourth digit ratio and face shape. *Proc. R. Soc. B* **272**, 1995–2001. (doi:10.1098/rspb.2005.3179)
34. Stephen ID, Scott IML, Coetzee V, Pound N, Perrett DI, Penton-Voak IS. 2012 Cross-cultural effects of color, but not morphological masculinity, on perceived attractiveness of men's faces. *Evol. Hum. Behav*. **33**, 260–267. (doi:10.1016/j.evolhumbehav.2011.10.003)
35. Gilani SZ, Rooney K, Shafait F, Walters M, Mian A. 2014 Geometric facial gender scoring: objectivity of perception. *PLoS ONE* **9**, e99483. (doi:10.1371/journal.pone.0099483)
36. McKnight CM, Newnham J, Stanley F, Mountain JA, Landau LI, Beilin LJ. 2012 Birth of a cohort: the first 20 years of the Raine study. *Med. J. Aust*. **197**, 608–610. (doi:10.5694/mja12.10698)
37. Vermeulen A, Verdonck L, Kaufman JM. 1999 A critical evaluation of simple methods for the estimation of free testosterone in serum. *J. Clin. Endocrinol. Metab*. **84**, 3666–3672. (doi:10.1210/jcem.84.10.6079)
38. Sartorius G, Ly L, Sikaris K, McLachlan R, Handelsman D. 2009 Predictive accuracy and sources of variability in calculated free testosterone estimates. *Ann. Clin. Biochem*. **46**, 137–143. (doi:10.1258/acb.2008.008171)
39. Zlotkin S, Casselman C. 1987 Percentile estimates of reference values for total protein and albumin in sera of premature infants (less than 37 weeks of gestation). *Clin. Chem*. **33**, 411–413.
40. Aldridge K, Boyadjiev S, Capone G, DeLeon V, Richtsmeier J. 2005 Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. *Am. J. Med. Genet*. **138**, 247–253. (doi:10.1002/ajmg.a.30959)

41. Caswell N, Manning JT. 2009 A comparison of finger 2D:4D by self-report direct measurement and experimenter measurement from photocopy: methodological issues. *Arch. Sex Behav*. **38**, 143–148. (doi:10.1007/s10508-007-9202-3)
42. Harwood DT, Handelsman DJ. 2009 Development and validation of a sensitive liquid chromatography-tandem mass spectrometry assay to simultaneously measure androgens and estrogens in serum without derivatization. *Clin. Chim. Acta*, **409** 78–84. (doi:10.1016/j.cca.2009.09.003)
43. Vermeulen A, Verdonck G. 1992 Representativeness of a single point plasma testosterone level for the long term hormonal milieu in men. *J. Clin. Endocrinol. Metab*. **74**, 939–942. (doi:10.1210/jc.74.4.939)
44. Gilani SZ, Shafait F, Mian A. 2014 Gradient based efficient feature selection. In *Applications of Computer Vision: IEEE Winter Conference* (*WACV 2014*), *Steamboat Springs, Colorado, 24–26 March 2014*, pp. 191–197. Red Hook, NY: Curran Associates.
45. Koehler N, Simmons LW, Rhodes G. 2004 How well does second-to-fourth-digit ratio in hands correlate with other indications of masculinity in males? *Proc. R. Soc. Lond. B* **271**, S296–S299. (doi:10.1098/rsbl.2004.0163)
46. Valla K, Halazonetis DJ. 2014 Correlation of 2D:4D digit ratio and craniofacial shape in prepubertal children. *Am. J. Hum. Biol*. **26**, 337–346. (doi:10.1002/ajhb.22512)
47. Hickey M, Sloboda DM, Atkinson HC, Doherty DA, Franks S, Norman RJ, Newnham JP, Hart R. 2009 The relationship between maternal and umbilical cord androgen levels and polycystic ovary syndrome in adolescence: a prospective cohort study. *J. Clin. Endocrinol. Metab*. **94**, 3714–3720. (doi:10.1210/jc.2009-0544)
48. Carre JM, McCormick CM. 2008 In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proc. R. Soc. B* **275**, 2651–2656. (doi:10.1098/rspb.2008.0873)
49. Geniole SN, Keyes AE, Carré JM, McCormick CM. 2014 Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Person. Indiv. Diff*. **57**, 59–64. (doi:10.1016/j.paid.2013.09.023)
50. Haselhuhn MP, Wong EM. 2011 Bad to the bone: facial structure predicts unethical behaviour. *Proc. R. Soc. B* **279**, 571–576. (doi:10.1098/rspb.2011.1193)
51. Alrajih S, Ward J. 2014 Increased facial width-to-height ratio and perceived dominance in the faces of the UK's leading business leaders. *Brit. J. Psychol*. **105**, 153–161. (doi:10.1111/bjop.12035)
52. Baron-Cohen S *et al.* 2015 Elevated fetal steroidogenic activity in autism. *Mol. Psychiatry* **20**, 369–376. (doi:10.1038/mp.2014.48)
53. Gilani S *et al.* 2015 Sexually dimorphic facial features vary according to level of autistic-like traits in the general population. *J. Neurodev. Disord*. **7**, 14. (doi:10.1186/s11689-015-9109-6)
54. Aldridge K, George I, Cole K, Austin J, Takahashi TN, Duan Y, Miles JH. 2011 Facial phenotypes in subgroups of prepubertal boys with autism spectrum disorders are correlated with clinical phenotypes. *Mol. Autism* **2**, 15. (doi:10.1186/2040-2392-2-15)
55. Wolke D, Waylen A, Samara M, Steer C, Goodman R, Ford T, Lamberts K. 2009 Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *Br. J. Psychiatr*. **195**, 249–256. (doi:10.1192/bjp.bp.108.053751)
56. Henrich J, Heine SJ, Norenzayan A. 2010 The weirdest people in the world? *Behav. Brain Sci*. **33**, 61–83. (doi:10.1017/S0140525X0999152X)
57. Claes P *et al.* 2012 Sexual dimorphism in multiple aspects of 3D facial symmetry and asymmetry defined by spatially dense geometric morphometrics. *J. Anat*. **221**, 97–114. (doi:10.1111/j.1469-7580.2012.01528.x)
58. Claes P *et al.* 2014 Modeling 3D facial shape from DNA. *PLoS Genet*. **10**, e1004224. (doi:10.1371/journal.pgen.1004224)