


ORIGINAL ARTICLE

Using data-driven phenotyping to investigate the impact of sex on 3D human facial surface morphology

Harold S. Matthews^{1,2} | Soha Mahdi² | Anthony J. Penington^{3,4,5} | Mary L. Marazita^{6,7} | John R. Shaffer^{6,7} | Susan Walsh⁸ | Mark D. Shriver⁹ | Peter Claes^{1,2,3,10} | Seth M. Weinberg^{6,7} 

¹Department of Human Genetics, KU Leuven, 3000, Leuven, Belgium

²Medical Imaging Research Center, UZ Leuven, Herestraat 49, 3000, Leuven, Belgium

³Facial Sciences Research Group, Murdoch Children's Research Institute, Parkville, 3052, Australia

⁴Department of Plastic and Maxillofacial Surgery, Royal Children's Hospital, Melbourne, 3052, Australia

⁵Department of Pediatrics, University of Melbourne, Melbourne, 3052, Australia

⁶Center for Craniofacial and Dental Genetics, Department of Oral and Craniofacial Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15219, USA

⁷Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

⁸Department of Biology, Indiana University Purdue University Indianapolis, Indianapolis, Indiana 46202, USA

⁹Department of Anthropology, Pennsylvania State University, State College, Pennsylvania 16802, USA

¹⁰Department of Electrical Engineering, ESAT/PSI, KU Leuven, 3000, Leuven, Belgium

Correspondence

Harold S. Matthews, Department of Human Genetics, KU Leuven, 3000 Leuven, Belgium.

Email: harry.matthews@kuleuven.be

Seth M. Weinberg, Center for Craniofacial and Dental Genetics, Department of Oral and Craniofacial Sciences, University of Pittsburgh, Pittsburgh, PA 15219, USA.

Email: smwst46@pitt.edu

Funding information

National Institute of Dental and Craniofacial Research, Grant/Award Number: R01-DE016148, R01-DE027023 and U01-DE020078; Research Fund KU Leuven, Grant/Award Number: BOF-C1 and C14/20/081; Royal Children's Hospital Foundation, Grant/Award Number: 2013-127; Science Foundation of Ireland Walton Fellowship, Grant/Award Number: 04.W4/B643; US National Institute of Justice, Grant/Award Number: 2008-DN-BX-K125, 2014-DN-BX-K031, 2018-DU-BX-0219 and 2018-DU-BX-0219

Abstract

The effects of sex on human facial morphology have been widely documented. Because sexual dimorphism is relevant to a variety of scientific and applied disciplines, it is imperative to have a complete and accurate account of how and where male and female faces differ. We apply a comprehensive facial phenotyping strategy to a large set of existing 3D facial surface images. We investigate facial sexual dimorphism in terms of size, shape, and shape variance. We also assess the ability to correctly assign sex based on shape, both for the whole face and for subregions. We applied a predefined data-driven segmentation to partition the 3D facial surfaces of 2446 adults into 63 hierarchically linked regions, ranging from global (whole face) to highly localized subparts. Each facial region was then analyzed with spatially dense geometric morphometrics. To describe the major modes of shape variation, principal components analysis was applied to the Procrustes aligned 3D points comprising each of the 63 facial regions. Both nonparametric and permutation-based statistics were then used to quantify the facial size and shape differences and visualizations were generated. Males were significantly larger than females for all 63 facial regions. Statistically significant sex differences in shape were also seen in all regions and the effects tended to be more pronounced for the upper lip and forehead, with more subtle changes emerging as the facial regions became more granular. Males also showed greater

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Journal of Anatomy* published by John Wiley & Sons Ltd on behalf of Anatomical Society.

levels of shape variance, with the largest effect observed for the central forehead. Classification accuracy was highest for the full face (97%), while most facial regions showed an accuracy of 75% or greater. In summary, sex differences in both size and shape were present across every part of the face. By breaking the face into subparts, some shape differences emerged that were not apparent when analyzing the face as a whole. The increase in facial shape variance suggests possible evolutionary origins and may offer insights for understanding congenital facial malformations. Our classification results indicate that a high degree of accuracy is possible with only parts of the face, which may have implications for biometrics applications.

KEYWORDS

3D surface imaging, facial shape, geometric morphometrics, sexual dimorphism

1 | INTRODUCTION

Sex differences in human facial soft-tissue morphology have been reported in multiple studies (Borman et al., 1999; Evison et al., 2010; Farkas et al., 2004; Ferrario et al., 1993, 1994, 1999; Hennessy et al., 2002; Kau et al., 2006; Kesterke et al., 2016; Koudelová et al., 2015; Sforza, Grandi, Binelli, et al., 2010; Sforza, Grandi, De Menezes, et al., 2010; Toma et al., 2008; Velemínská et al., 2012) and are estimated to account for approximately 13% of among-individual variation in facial shape (Claes et al., 2014). Although most pronounced after puberty, the effects of biological sex on facial surface features are already present in early postnatal life (Matthews et al., 2016). Documenting facial sex differences across the lifespan is relevant to clinical genetics and dysmorphology (Hammond & Suttie, 2012; Matthews et al., 2021), ecological and evolutionary studies (Kleisner et al., 2021), psychology and cognitive science (Summersby et al., 2022), forensics (Franklin et al., 2013; Shrimpton et al., 2014), orthodontics (Dean et al., 2000; Ursi et al., 1993), and craniofacial reconstructive surgery (Bannister et al., 2022), among others. Thus, it is important to have a complete and accurate account of the effects of sex on facial appearance.

From prior studies analyzing 3D facial surfaces, the consensus is that mature male faces are larger on average compared with female faces and tend to exhibit a general pattern of increased midfacial, supraorbital, and chin prominence coupled with relative retrusion of the orbital region, cheek/malar region, and forehead. These studies employ both traditional landmark-based approaches (Kesterke et al., 2016; Liu et al., 2014; Tanikawa et al., 2016) and newer spatially dense geometric morphometric approaches (Claes, Walters, Shriver, et al., 2012; Ekrami et al., 2021; Hennessy et al., 2005; Matthews, Penington, Hardiman, et al., 2018; Mydlová et al., 2015; Shrimpton et al., 2014), where the surface is represented by densely sampled “quasi-landmarks.” Methods using sparse landmarks are well established but provide limited information about facial form due to an inability to cover broad, feature-deficient regions like the cheeks and forehead. Quasi- and to a lesser extent semi-landmark methods can overcome this limitation by incorporating many more

data points from across the entire facial surface. However, with these approaches, localized shape effects can sometimes be difficult to detect especially when they explain only a small portion of the overall variation. Further, geometric morphometric methods employing Procrustes superimposition are limited, as the superimposition of landmark configurations is a globally optimal fitting of one configuration onto another or onto a consensus configuration. As such, observed shape differences in a local region (e.g., the nasal tip) are not independent of the rest of the face, making statistical inferences about localized shape differences difficult. As a result, questions remain about which parts of the face are most and least influenced by sex.

It has been reported that human faces are both more variable and less intercorrelated than other parts of the body (Sheehan & Nachman, 2014), which suggests that treating the face as a singular structure in morphometric studies may not be ideal. In 2018, we introduced an innovative extension of spatially dense geometric morphometrics involving the segmentation of 3D facial surfaces into 63 regions or modules based on patterns of multivariate correlation among the ~8000 aligned quasi-landmarks (Claes et al., 2018). The segmentation is data driven (not defined a priori) and works by exploiting the natural covariance present in morphologically integrated structures like the face. Further, the segmentation is hierarchically linked in a global-to-local pattern. This allows statistical analysis and inference to be conducted simultaneously at the level of global facial variation as well as on increasing localized regions. By effectively breaking the facial surface into discrete yet interlinked regions and then conducting spatially dense geometric morphometric analysis on each region separately, patterns of localized shape change become easier to detect statistically. In the context of facial dimorphism, this approach may reveal subtle sex effects missed by earlier studies.

In the present report, we will apply spatially dense modular facial phenotyping to test several hypotheses relevant to the impact of biological sex on human faces. We investigate whether male faces differ from female faces in terms of both size and shape. We will also investigate the effect of sex on shape variance explicitly, based on

conflicting reports that male faces show greater levels of variation than female faces (Kleisner et al., 2013, 2021; Milella et al., 2021). Finally, we assess the practical ability to correctly assign sex based on 3D facial shapes. Importantly, for each of these analyses, we look not just at the face as a monolithic structure, but across the entire range of modular facial phenotypes.

2 | MATERIALS AND METHODS

2.1 | Sample description

The total analysis sample included 2446 adults. Individuals were selected from several large data sets with available 3D facial images: the 3D facial norms repository ($n = 674$), the IUPUI facial data set ($n = 353$), the Penn State ADAPT study ($n = 1415$), and the Melbourne AHEAD study ($n = 4$). These data sets have been extensively described and used in prior studies of facial genetics (White et al., 2021) or morphology (Matthews, Penington, Clement, et al., 2018; Matthews, Penington, Hardiman, et al., 2018). Individuals self-identified as male or female, and in 99.8% of cases, the self-designation was confirmed by genetic analysis (the four individuals from the AHEAD study did not have genetic follow-up). Individuals were selected from these data sets to ensure equal numbers of males and females with a balanced age distribution; both sexes had 1223 individuals with a mean age of 27.8 years. Figure 1 shows the distribution of males and females. All individuals are self-reported as having recent European ancestry. Because these cohorts were originally collected with the goal of studying normal-range human facial variation, individuals were screened for a history of craniofacial birth defects, trauma, or surgery. All participants provided written informed consent prior to participation, and all research was approved by each institution's ethics oversight committee.

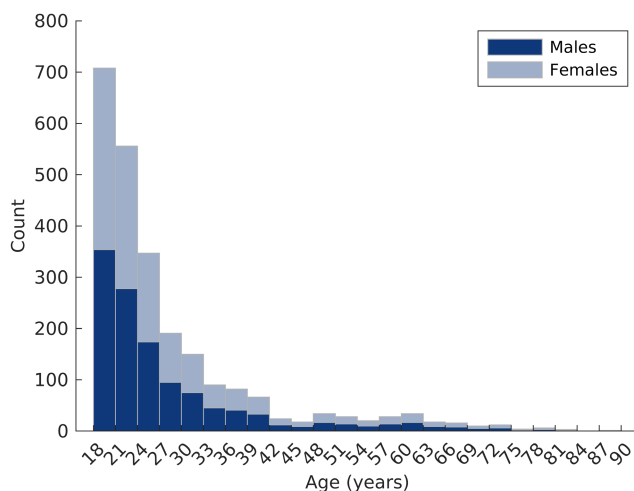


FIGURE 1 The age distributions of males and females used in this study.

2.2 | Spatially dense landmarking and 3D facial segmentation

3D facial surface images for all participants were acquired through digital stereophotogrammetry using standardized protocols described previously (Heike et al., 2010). After the initial cleaning, the 3D surface images were placed into dense correspondence, such that each 3D face was aligned across of set of 7160 densely spaced quasi-landmarks. This was accomplished automatically by nonrigid mapping a specially constructed template face or anthropometric mask (Claes, Walters, & Clement, 2012) onto each image using the MeshMonk image-processing pipeline (White et al., 2019). To symmetrize the dense landmark configurations, a copy of each configuration was reflected about the x axis and each bilateral quasi-landmark was relabeled as the corresponding quasi-landmarks on the opposite side of the midline. The copy was then superimposed onto the original via a least-squares Procrustes superimposition and the average of each corresponding landmark was computed. These averages constitute the symmetrical version of the initial configuration.

Each full-face dense landmark configuration was then segmented according to a previously defined hierarchical segmentation (White et al., 2021). As shown in Figure 2, starting with the complete 3D surface, the face is split into two large segments (or modules) representing the nose/upper lip and the rest of the face. This partitioning is then repeated for four more rounds, resulting in 63 total facial segments linked in a hierarchical manner. The facial segmentation is data driven and based on patterns of intercorrelation among the 3D quasi-landmarks; points that exhibit strong covariance will clump together. As is evident in Figure 2, this approach naturally breaks the face into four main quadrants: nose, upper lip, upper face (forehead and periorbital region), and lower face (mandible and cheeks).

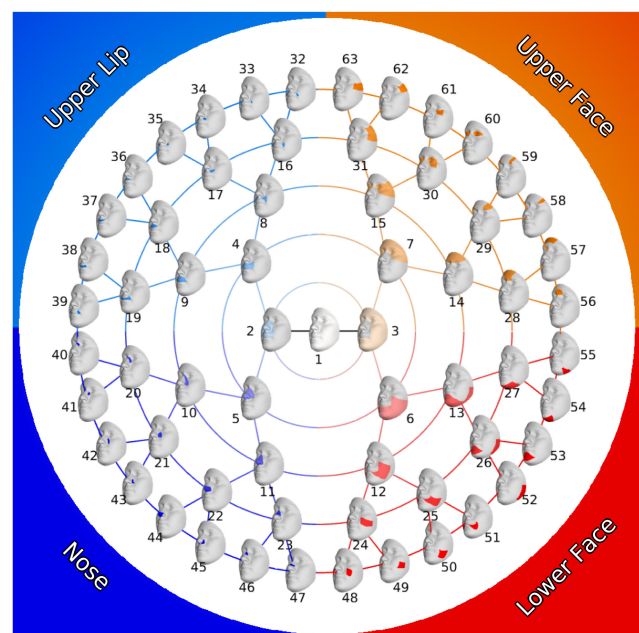


FIGURE 2 Modular facial segmentation scheme.

The quasi-landmarks comprising each module were aligned to the sample mean and scaled to unit size with generalized Procrustes analysis and the centroid size for each participant for each module was recorded. To describe the major modes of shape variation, principal components analysis was applied to the Procrustes aligned quasi-landmarks comprising each of the 63 facial modules. For each segment, the components required to explain 98% of the shape variation were retained. The number of components retained for each module is shown in Supplementary Figure 1.

2.3 | Impact of sex on facial size

The median centroid sizes of males and females were compared using the Wilcoxon rank-sum test, which was performed separately in each facial module. To account for the multiple testing burden the nominal significance threshold of 0.05 was divided by the effective number of comparisons ($n = 12$) estimated from the eigenvalues of the 63×63 matrix of correlations of centroid sizes in each facial module following Li and Ji (2005). This resulted in a p value threshold of 0.0042 (0.05/12). The effect size was measured by the A statistic (Ruscio & Mullen, 2012). This is a nonparametric estimator of the probability that a randomly selected instance of class A (males) is larger than a randomly selected instance of class B (females). Intuitively, this reflects the degree of overlap between the two distributions and is identical to the area under the receiver operator characteristic curve. We report $A-0.5$. The statistic then ranges from -0.5 (all females have larger centroid sizes than all males) to $+0.5$ (all males have larger centroid sizes than all females).

2.4 | Impact of sex on facial shape

Sexual dimorphism of shape was assessed by a partial least squares regression of each facial module's PC scores onto sex (coded as male = 1; female = 2). Following Shrimpton et al. (2014), the effect size was partial R^2 , and statistical significance was estimated via a permutation test on this statistic with 100,000 permutations. Centroid size of the full configuration of facial quasi-landmarks and BMI were included as covariates in the regression model. Including age as a covariate showed no appreciable effect on the results, as expected from the strict age matching between the samples of each sex (data not shown). Regression coefficients for each facial module were projected onto the surface normal of the average shape to be visualized as color-coded heat maps showing the change from males to females in the locally inward-outward direction. 3D morphs of hypermasculine and hyperfeminine shapes for each module were generated by transforming the average shape equivalently by both adding (feminine) and subtracting (masculine) the regression coefficients for sex multiplied by a scalar to the vertex coordinates of the average shape. The scalar was determined for each module to make the root mean square difference between the vertices of the two morphs equal 0.15. The value 0.15 was chosen by trial and error

to make morphs, where the shape differences are visible. The effective number of comparisons (29) was estimated using a multivariate extension of the method above (Claes et al., 2018) from the eigenvalues of the 63×63 matrix of RV coefficients between PC scores in each pair of facial modules after statistical adjustment for covariates. This adjustment was done by taking the residuals of a PLS regression model that regresses PC scores onto the covariates. Therefore, the p -value threshold was set at 0.0017 (0.05/29).

2.5 | Quantifying and comparing facial shape variance

To test for differences in multivariate shape variance between males and females for each facial module, we used the log of the ratio of the trace of the covariance matrix of males to the trace of the covariance matrix of females as a test statistic:

$$\log\left(\frac{\text{tr}(\text{cov}(X_M))}{\text{tr}(\text{cov}(X_F))}\right),$$

where X_M and X_F are the n (observations) \times k (principal components) matrix of the principal component scores of males and females, respectively. Prior to this, the PC scores were adjusted for centroid size of the full facial landmarks and BMI by taking the residuals of a PLS regression of the scores of both groups together onto the covariates. Additional tests that included age as a covariate showed little appreciable effect on the results, except that the p values associated with two small modules of the nasal tip and alar regions (Figure 2, modules 44 and 45) were reduced slightly pushing them over the significance threshold. Statistical significance was calculated via a permutation test on the test statistic. The significance threshold of 0.05 was adjusted to 0.05/63 ($p \leq 0.0008$). This is a stringent adjustment. Although dependency between the tests will exist and the effective number of comparisons will be lower than 63, we are not aware of a method to estimate the effective number of comparisons in this context.

2.6 | Sex classification based on facial shape

We assessed the ability of PC scores, describing morphology in each facial module, to classify males and females. We used linear discriminant analysis as the classifier with empirical prior probabilities calculated as the proportion of each group in the training set (as implemented in MATLAB's `fitcdiscr` function). The linear discriminant analysis model was tested using 10-fold stratified cross-validation and we report the percentage correct classification. We further investigated what properties may contribute to the classification accuracy. To do this, we correlated the percentage classified correctly in each module with the surface area (the area of all complete triangles in the module on the template mesh) and shape complexity, measured as the number of principal components required to explain 98% of the shape variation.

3 | RESULTS

Males showed significantly higher centroid size than females in all modules (Figure 3). We also ran an alternate version of the facial size analysis with BMI included as a covariate, but the results were unchanged (data not shown), suggesting that the effects we observed were not driven by overall body size.

Males also showed greater levels of shape variance in all except three modules (modules 27, 32, and 54, but these were not statistically significant). Although each of the main facial quadrants contained modules where males showed significant increases in variation, the effect size was largest in the central forehead region (modules 28 and 57) and smallest in the upper lip region. In the lower face, the difference in variance was most evident in the buccal region (the branch beginning at module 12), as opposed to the chin. In the nose, the difference was most pronounced along the branch beginning at module 10. In the upper lip area, the difference was most evident around the vermillion portion and adjacent superior cutaneous upper lip (modules 38 and 37).

Statistically significant sex differences in shape were seen in all 63 modules (Figure 3). These differences tended to be more pronounced in the full face, upper lip, and forehead modules. Looking at the full face (module 1), the pattern is clear and striking—the entire central portion of the face stretching from the forehead to the chin is more inwardly projecting in females, while the lateral regions including the zygoma and cheeks are more outwardly projecting (Figure 4). The overall impression would be of a more anteriorly projecting male face and a flatter female face.

Examining additional modules in Figure 4 offers more nuanced information. For example, the forehead in the full face module is uniformly outwardly projecting in males; however, subtleties emerge when considered in isolation (e.g., modules 7, 14, 28, and 56), with the brow ridge projecting outwardly and the more superior portion of the frontal bone sloping posteriorly. The nose also shows a

complex series of shape changes as the modules focus on more localized regions. In the full face, the nose, along with the entire midface shows outward projection in males. In more localized modules (e.g., 5, 11), the dorsum of the nasal bridge is less prominent relative to the nasal tip. This trend continues through modules 11, 22, and 44, where the inferior-most part of the bridge now reverses to become slightly retrusive. We can interpret this as showing overall midface advancement in males, but within the nose itself, there are more subtle shape changes taking place along the dorsum of the nasal bridge and nasal tip. The result is a slight narrowing and down-turning of the nasal tip in males. The upper lip and philtrum also show some subtle shape changes. In the context of the full face, the entire upper lip region is projecting outward in males. However, in more localized lip modules (e.g., 9, 18, 37), the central part of the philtrum is positioned more inwardly in males resulting in a deeper and more pronounced philtrum, which is not visible in the analysis of the full face. Interestingly, one of the strongest sex differences is an increased concavity of the nasolabial fold in females, seen on the branch beginning at module 8. Colormaps and 3D morphs for all facial modules can be found in [Supplementary File 1](#).

Figure 5 shows classification accuracy per facial module and the nonparametric correlations between (1) accuracy and module surface area, (2) accuracy and shape complexity, and (3) surface area and shape complexity. The greatest accuracy was achieved using the full face. However, the majority (84%) of facial modules showed an accuracy of 75% or greater. Upper and lower facial modules tended to perform better than the nose and upper lip modules. In general, classification accuracy decreased as facial modules decreased in size (as measured in terms of the proportional surface area captured by each module) and shape complexity (as determined by the number of principal components needed to account for 98% of the variation). As expected, module surface area and module shape complexity were strongly correlated, since larger modules tend to require more principal components to capture the shape variation present.

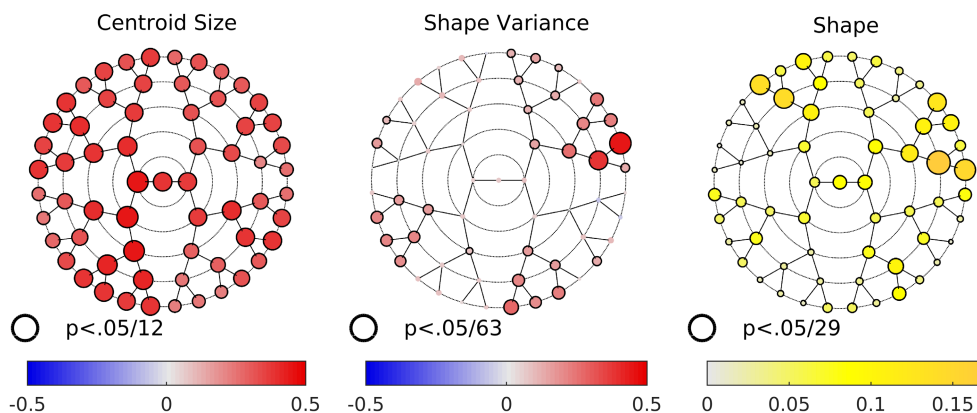


FIGURE 3 Sex differences in centroid size, variance, and shape. For centroid size, the size and color of the nodes are indexed to A-0.5, where A is the nonparametric estimate of the probability of superiority. 0.5 (red) indicates all men have larger centroid sizes than all women -0.5 indicates all women have centroid sizes larger than all men. For shape variance color and size are indexed to the log of the ratio of male variance to female variance. Shape, color, and size are indexed to the variance explained by sex in each module. For all three, black borders indicate a statistically significant difference after adjustment for multiple comparisons (see text).

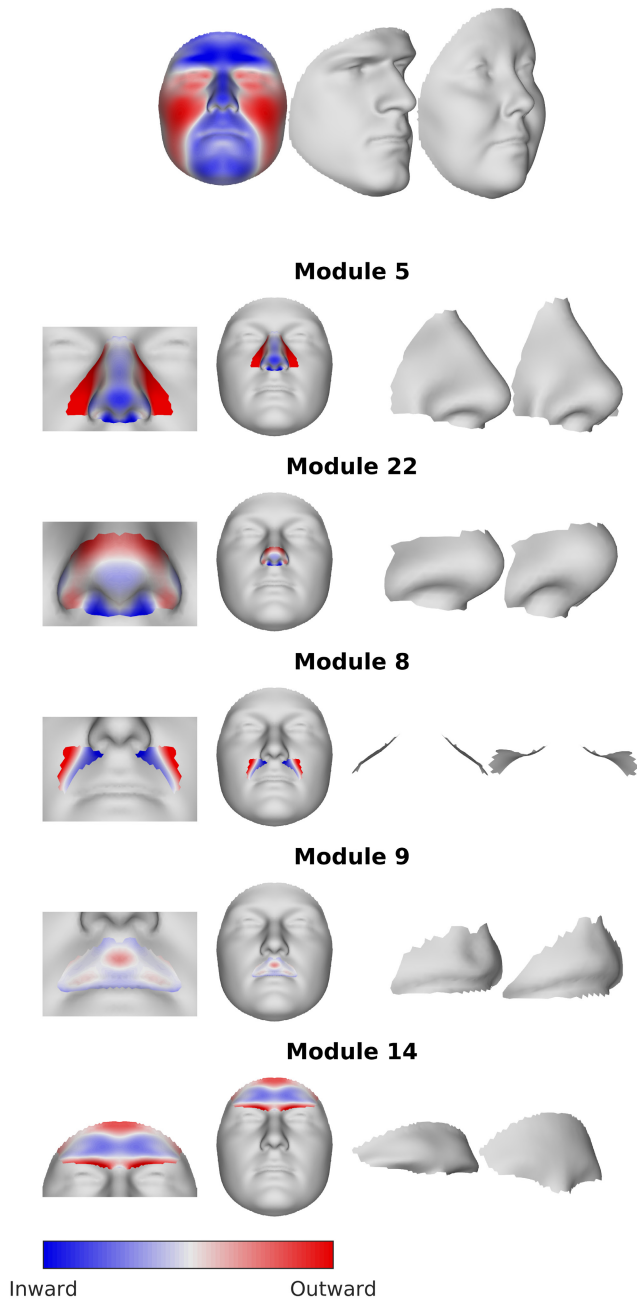


FIGURE 4 Visualization of the sex differences in shape. Color maps show the regression coefficients associated with the effect of sex at each vertex, projected onto the surface normal. Red indicates the vertex is more locally outward in females than males, blue indicates the point is more locally inward. The static morphs illustrate the difference between the sexes, by showing an exaggerated hypermasculine and hyperfeminine face (see text).

4 | DISCUSSION

Sexual dimorphism has received considerable attention from multiple fields of study. Studying patterns of sex differences can yield insights into the action of natural (including sexual) selection on morphology and inform forensic investigations and precision medicine. The majority of past studies investigated facial sexual dimorphism with linear

distances or in terms of sparse landmark configurations (Kesterke et al., 2016; Liu et al., 2014; Tanikawa et al., 2016), which provides a limited representation of the face, or else employed spatially dense image analysis of the entire face or face and head together (Claes, Walters, Shriver, et al., 2012; Ekrami et al., 2021; Hennessy et al., 2005; Matthews, Penington, Hardiman, et al., 2018; Mydlová et al., 2015; Shrimpton et al., 2014). While the latter analysis can comprehensively indicate how different parts are spatially interrelated, it can also obscure more subtle and localized shape differences, which may nevertheless have important implications. In this study, we used a large, well-balanced sample of adults of European descent and employed a unique data-driven global-to-local facial segmentation to study sex differences in size, shape, and shape variance.

Not surprisingly, our study confirmed that male faces are larger than female faces. This was true even when accounting for overall body size (BMI). Notably, we demonstrate the size effect was present across modules spanning all levels of the global-to-local hierarchy. In terms of shape, at the level of the full face, we observed a more anteriorly projecting male face and a relatively flatter female face, broadly consistent with the findings of prior studies of 3D facial morphology (Bannister et al., 2022; Hennessy et al., 2005; Kau et al., 2006; Kesterke et al., 2016; Koudelová et al., 2015; Toma et al., 2008). By virtue of the segmentation approach, we were able to reveal that sex-related differences are greatest in the upper face (particularly the forehead) and portions of the upper lip (particularly the nasolabial groove and philtrum); this is evidenced by the partial R^2 values in Figure 3. Analyzing shape differences at successively finer scales also revealed shape differences that were not evident in the analysis of the full face. For example, modules involving the forehead showed evidence of retrusion (sloping) superior to the more projecting brow ridges that are characteristic in males, shape changes in nose-specific modules revealed the upturned nasal tip characteristic of females, and males exhibited a more well-defined (deeper) philtrum than females.

Revealing these more subtle changes helps complete the picture of how sex impacts facial structure and provides objective evidence to corroborate well-known sex effects that were previously difficult to capture fully. While analysis of the full face can provide a comprehensive picture of how various facial parts interrelate, some localized effects may be lost—masked by more global aspects of shape variation. Partitioning the face and analyzing each part separately can reveal those effects. A similar phenomenon was recently shown by Bannister et al. (2022) in their analysis of facial sexual dimorphism that divided the facial surface into seven subregions. Unlike the present study, however, these subregions were predefined based on their relevance for surgical planning. Overall, the findings of these two studies agree. Here, however, we find some additional evidence for shape differences in the upper lip region, particularly involving the philtrum. This may be because the 3D lip segment in Bannister et al. (2022) was too large to capture some of these very subtle shape effects.

Milella et al. (2021) found that males showed increased shape and size variance across several different regions of the skull, although

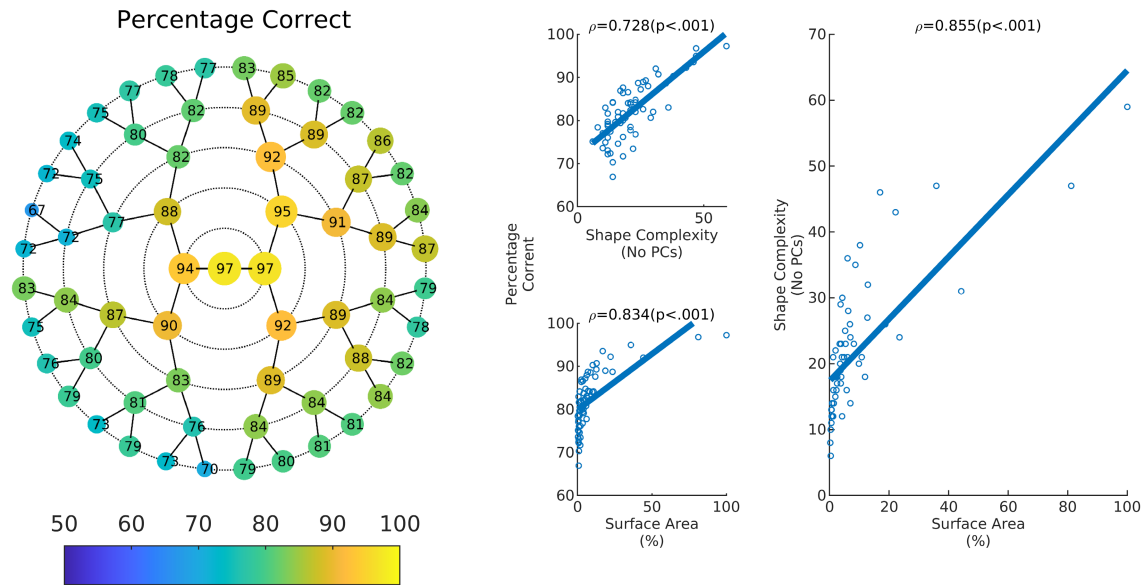


FIGURE 5 Classification accuracy. The left shows the percentage of correct classification in each module. Right explores some predictors of classification accuracy. This shows proportion correct as a function of shape complexity and surface area of the module. The legend shows spearman's rho and the associated p values. Shape complexity as a function of the surface area is also plotted.

the difference was only significant for total cranial size in their study. Interestingly, the one region where females showed slightly higher (but nonsignificant) levels of shape variance was the face. The lack of statistical significance in their study may be related to the small sample sizes that were available. In contrast to Milella et al. (2021), we found that males showed statistically greater levels of shape variance in multiple facial regions including the forehead, nose, and lower cheek regions. There were no facial modules where females showed significantly higher levels of shape variance. There are several factors that might account for these different study outcomes. One obvious difference in our study is that we focus exclusively on soft-tissue surface morphology. In addition, our cohorts were collected within the past decade, whereas Milella and colleagues used skeletal remains spanning the 19th and 20th centuries. Levels of craniofacial variability may have shifted over time due to improved nutrition and reduced infant mortality. Two studies by Kleisner et al. (2013, 2021) also reported that males showed excess variance in facial shape compared with females from multiple populations, corroborating our general findings. However, because these results were based on the morphometric analysis of discrete landmarks derived from 2D frontal photographs, it is not possible to determine to what degree sex differences in shape variance impacted different parts of the face. These study differences make direct comparisons a challenge.

The increased facial shape variation observed in males fits within a basic pattern documented for morphological traits across numerous animal taxa (Zajitschek et al., 2020). Sexual selection pressure (Wainer, 2007) and heterogamy (Reinhold & Engqvist, 2013) have both been put forth as possible explanations for the propensity of males to show excess variation in physical traits. Zajitschek et al. (2020) found that traits with the greatest levels of sexual

dimorphism in mice also showed the greatest sex disparities in variance—a prediction that would follow from the sexual selection hypothesis. Milella et al. (2021) found partial evidence to support this claim in their human skull data. The facial surface data in the current study also provide some support for this. There is a weak but positive and statistically significant (Spearman's $\rho(61) = 0.313$, $p = 0.013$ (two tails)) correlation between the variance ratio statistic and partial R^2 associated with sex in the 63 facial modules (Supplementary Figure 2).

Another possible ecological explanation for the origin of sexually dimorphic traits is niche specialization (Slatkin, 1984), which can occur when each sex is adapted to exploit distinct ecological niches and resources. For example, in environments where males predominantly took on the responsibility for dangerous and high-risk activities like hunting and herding large animals, the risk of broken facial bones could result in blindness or broken teeth and jaws. Larger and more robust facial bones achieved through longer growth periods might be expected because these traits would offer greater resistance to injury. This would be consistent with the longer facial growth patterns seen in males compared with females (Ursi et al., 1993). Since male faces continue growing into their mid-twenties, we might also expect to see greater variability in male faces compared with female faces.

Proximate and ultimate causes aside, the greater shape variance in male faces could have some important implications. Increased shape variation, particularly during facial morphogenesis, has been linked to a greater propensity for malformations like orofacial clefts (Hallgrímsson et al., 2004; Parsons et al., 2008). Essentially, an increase in shape variance will result in more individuals with extreme phenotypes. As a result, more individuals will approach and exceed the limits of viable and functional anatomy

thereby increasing the likelihood of disrupting the precise spatio-temporal pattern required for normal facial formation. Like many structural birth defects, orofacial clefts involving both the lip and palate (the most common type of cleft) occur more frequently in males (Mossey et al., 2009). The sex bias in clefting has never been adequately explained, but higher levels of shape variance in the developing faces of male versus female embryos is an intriguing hypothesis. Currently, it is an open question whether male embryos exhibit increased craniofacial variation, as adults do. There are also potential statistical implications for morphometric studies documenting and comparing faces. For example, more males may be needed than females to capture the representative range of facial variation present in a given population, with implications for studies with small sample sizes.

When we investigated classification based on facial shape, the accuracy ranged from 97% (full face) to 67% (module 38 on the lip), and the vast majority of modules showed an accuracy of at least 75%. We found that modules involving the upper and lower face tended to outperform modules in the lip and nose. For example, several upper face modules performed nearly as well as the full face. Similarly, Bigoni et al. (2010) found that sex classification was most accurate for the upper face region in their Central European sample, although this study was based on skull material instead of soft tissue. We also found that larger modules tend to perform better than smaller modules. These observations could have implications for real-world systems that often rely on partial facial images (e.g., surveillance-based systems) to classify individuals (Tome et al., 2013). Our results suggest that even with relatively small portions of the face, reasonably good sex classification is possible. Furthermore, images that include portions of the upper face should result in improved performance, even if other portions of the face are obscured.

Our results demonstrate that adult faces show strong sexual dimorphism for both size and shape and that sex differences are present across all parts of the face. Further, we found that shape variance is higher in male faces and that this effect was most apparent in the upper face. Finally, we found that faces can be classified on the basis of sex with high accuracy, even when considering only parts of the total face. All of these results must be considered within the context of this study's limitations. Our datasets were limited to 3D facial surface morphology. Care must be taken not to assume that our outcomes apply equally to the underlying facial skeleton or to other parts of the craniofacial complex. Our samples were also limited to adults of recent European ancestry. Although the causal factors are unclear (see Kleisner et al., 2021 for a comprehensive discussion), several studies have shown that patterns of sexual dimorphism in facial morphology vary across human populations and that these population differences can be in both magnitude and direction (Garvin & Ruff, 2012; Humphrey et al., 1999; O'Higgins et al., 1990). Therefore, applying this same approach to samples with different ancestral backgrounds or different ages may produce different outcomes. Expanding the scope of this research by improving sample diversity is essential for understanding the impact of sex on facial traits in humans.

AUTHOR CONTRIBUTIONS

Harold S. Matthews and Seth M. Weinberg conceptualized the study and drafted the manuscript; Peter Claes and Soha Mahdi provided input on the analytical design; Harold S. Matthews conducted all the analyses; Peter Claes and Seth M. Weinberg provided oversight and feedback on analyses and their interpretation; Seth M. Weinberg, Peter Claes, Susan Walsh, Mark D. Shriver, John R. Shaffer, Mary L. Marazita, and Anthony J. Penington acquired funding; Seth M. Weinberg, Susan Walsh, Mark D. Shriver, Mary L. Marazita, and Anthony J. Penington oversaw the recruitment efforts for their respective cohorts; Peter Claes, Susan Walsh, Mark D. Shriver, John R. Shaffer, Mary L. Marazita, Harold S. Matthews, Soha Mahdi, and Anthony J. Penington provided edits and comments on the manuscript draft; all authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The authors have nothing to report.

FUNDING INFORMATION

This study was supported by grants from the National Institute of Dental and Craniofacial Research (U01-DE020078, R01-DE016148, and R01-DE027023); the Research Fund KU Leuven (BOF-C1, C14/20/081) and the Research Program of the Research Foundation - Flanders (Belgium) (FWO, G078518N); the Center for Human Evolution and Development at Penn State, the Science Foundation of Ireland Walton Fellowship (04.W4/B643), the US National Institute of Justice (2008-DN-BX-K125; 2018-DU-BX-0219; 2014-DN-BX-K031; 2018-DU-BX-0219), and the University of Illinois Interdisciplinary Innovation Initiative Research Grant; the Royal Children's Hospital Foundation (2013-127). The funders played no part in the conceptualization, execution, writing, or dissemination of this work.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

Data sufficient to conduct the analyses as presented in this report are provided in supplemental data files. These can be found at <http://d-scholarship.pitt.edu/id/eprint/44263>. For the 3D Facial Norms data set, the raw source data (3D facial surface models in .obj format) are available through the FaceBase Consortium (www.facebase.org). Access to these 3D facial surface models requires proper institutional ethics approval and approval from the FaceBase Data Access Committee. The ADAPT, IUPUI, and AHEAD 3D facial datasets were not collected with broad data-sharing consent. Given the highly identifiable nature of facial scans and unresolved issues regarding risks to participants of inherent reidentification, participants were not consented for inclusion in public repositories or the posting of individual data.

KU Leuven provides the MeshMonk spatially dense facial mapping software, free to use for academic purposes (<https://github.com/TheWebMonks/meshmonk>). Matlab implementations of the hierarchical spectral clustering to obtain facial segmentations are

available at <https://figshare.com/articles/dataset/FACIALRECFROMDNA/7649024>.

ETHICS STATEMENT

All individuals whose 3D images and data were used in these analyses provided their written informed consent prior to participation. This informed consent included explicit language concerning the use of 3D facial image data for research purposes. Institutional ethics oversight was at each of the primary data collection and analyses sites: 3D Facial Norms Dataset (University of Pittsburgh Human Research Protection Office: #PRO09060553 and #RB0405013; Seattle Children's Hospital: IRB #12107; UT Health Committee for the Protection of Human Subjects: #HSC-DB-09-0508); University of Iowa Human Subjects Office (IRB: #200912764 and #200710721); ADAPT data set (PSU IRB #13103, PSU IRB #45727, UC IRB #2015-3073, PSU IRB #2503, PSU IRB #44929, PSU IRB #4320, PSU IRB #44929, PSU IRB #1278); IUPUI data set (IUPUI IRB #1409306349); AHEAD data set (Royal Children's Hospital IRB #290081). In addition, KU Leuven had an intuitional oversite for secondary data analysis (Approval: S60568).

ORCID

Seth M. Weinberg  <https://orcid.org/0000-0001-9467-4556>

REFERENCES

- Bannister, J.J., Juszczak, H., Aponte, J.D., Katz, D.C., Knott, P.D., Weinberg, S.M. et al. (2022) Sex differences in adult facial three-dimensional morphology: application to gender-affirming facial surgery. *Facial Plastic Surgery & Aesthetic Medicine*, 24, S24–S30.
- Bigoni, L., Velemínská, J. & Brůžek, J. (2010) Three-dimensional geometric morphometric analysis of cranio-facial sexual dimorphism in a central European sample of known sex. *Homo*, 61, 16–32.
- Borman, H., Özgür, F. & Gürsu, G. (1999) Evaluation of soft-tissue morphology of the face in 1,050 young adults. *Annals of Plastic Surgery*, 42, 280–288.
- Claes, P., Liberton, D.K., Daniels, K., Rosana, K.M., Quillen, E.E., Pearson, L.N. et al. (2014) Modeling 3D facial shape from DNA. *PLoS Genetics*, 10, e1004224.
- Claes, P., Roosenboom, J., White, J.D., Swigut, T., Sero, D., Li, J. et al. (2018) Genome-wide mapping of global-to-local genetic effects on human facial shape. *Nature Genetics*, 50, 414–423.
- Claes, P., Walters, M. & Clement, J. (2012) Improved facial outcome assessment using a 3D anthropometric mask. *International Journal of Oral and Maxillofacial Surgery*, 41, 324–330.
- Claes, P., Walters, M., Shriver, M.D., Puts, D., Gibson, G., Clement, J. et al. (2012) Sexual dimorphism in multiple aspects of 3D facial symmetry and asymmetry defined by spatially dense geometric morphometrics. *Journal of Anatomy*, 221, 97–114.
- Dean, D., Hans, M.G., Bookstein, F.L. & Subramanyan, K. (2000) Three-dimensional Bolton-brush growth study landmark data: ontogeny and sexual dimorphism of the Bolton standards cohort. *Cleft Palate-Craniofacial Journal*, 37, 145–156.
- Ekrami, O., Claes, P., Van Assche, E., Shriver, M.D., Weinberg, S.M., Marazita, M.L. et al. (2021) Fluctuating asymmetry and sexual dimorphism in human facial morphology: a multi-variate study. *Symmetry*, 13, 304.
- Evison, M., Dryden, I., Fieller, N., Mallett, X., Morecroft, L., Schofield, D. et al. (2010) Key parameters of face shape variation in 3D in a large sample. *Journal of Forensic Science*, 55, 159–162.
- Farkas, L.G., Eiben, O.G., Sivkov, S., Tompson, B., Katic, M.J. & Forrest, C.R. (2004) Anthropometric measurements of the facial framework in adulthood: age-related changes in eight age categories in 600 healthy white north Americans of European ancestry from 16 to 90 years of age. *The Journal of Craniofacial Surgery*, 15, 288–298.
- Ferrario, V.F., Sforza, C., Pizzini, G., Vogel, G. & Miani, A. (1993) Sexual dimorphism in the human face assessed by euclidean distance matrix analysis. *Journal of Anatomy*, 183, 593–600.
- Ferrario, V.F., Sforza, C., Poggio, C.E. & Schmitz, J.H. (1999) Soft-tissue facial morphometry from 6 years to adulthood: a three-dimensional growth study using a new modeling. *Plastic and Reconstructive Surgery*, 103, 768–778.
- Ferrario, V.F., Sforza, C., Poggio, C.E., Serrao, G. & Miani, A. (1994) A three-dimensional study of sexual dimorphism in the human face. *The International Journal of Adult Orthodontics and Orthognathic Surgery*, 9, 303–310.
- Franklin, D., Cardini, A., Flavel, A. & Kuliukas, A. (2013) Estimation of sex from cranial measurements in a Western Australian population. *Forensic Science International*, 229, 158.e1–8.
- Garvin, H.M. & Ruff, C.B. (2012) Sexual dimorphism in skeletal brow-ridge and chin morphologies determined using a new quantitative method. *American Journal of Physical Anthropology*, 147, 661–670.
- Hallgrímsson, B., Dorval, C.J., Zelditch, M.L. & German, R.Z. (2004) Craniofacial variability and morphological integration in mice susceptible to cleft lip and palate. *Journal of Anatomy*, 205, 501–517.
- Hammond, P. & Suttie, M. (2012) Large-scale objective phenotyping of 3D facial morphology. *Human Mutation*, 33, 817–825.
- Heike, C.L., Upson, K., Stuhag, E. & Weinberg, S.M. (2010) 3D digital stereophotogrammetry: a practical guide to facial image acquisition. *Head & Face Medicine*, 6, 18.
- Hennessy, R.J., Kinsella, A. & Waddington, J.L. (2002) 3D laser surface scanning and geometric morphometric analysis of craniofacial shape as an index of cerebro-craniofacial morphogenesis: initial application to sexual dimorphism. *Biological Psychiatry*, 51, 507–514.
- Hennessy, R.J., McLearn, S., Kinsella, A. & Waddington, J.L. (2005) Facial surface analysis by 3D laser scanning and geometric morphometrics in relation to sexual dimorphism in cerebral-craniofacial morphogenesis and cognitive function. *Journal of Anatomy*, 207, 283–295.
- Humphrey, L.T., Dean, M.C. & Stringer, C.B. (1999) Morphological variation in great ape and modern human mandibles. *Journal of Anatomy*, 195, 491–513.
- Kau, C.H., Zhurov, A., Richmond, S., Cronin, A., Savio, C. & Mallorie, C. (2006) Facial templates: a new perspective in three dimensions. *Orthodontics and Craniofacial Research*, 9, 10–17.
- Kesterke, M.J., Raffensperger, Z.D., Heike, C.L., Cunningham, M.L., Hecht, J.T., Kau, C.H. et al. (2016) Using the 3D facial norms database to investigate craniofacial sexual dimorphism in healthy children, adolescents, and adults. *Biology of Sex Differences*, 7, 23.
- Kleisner, K., Priplatova, L., Frost, P. & Flegr, J. (2013) Trustworthy-looking face meets brown eyes. *PLoS ONE*, 8, e53285.
- Kleisner, K., Tureček, P., Roberts, S.C., Havlíček, J., Valentova, J.V., Akoko, R.M. et al. (2021) How and why patterns of sexual dimorphism in human faces vary across the world. *Scientific Reports*, 11, 5978.
- Koudelová, J., Brůžek, J., Cagánová, V., Krajiček, V. & Velemínská, J. (2015) Development of facial sexual dimorphism in children aged between 12 and 15 years: a three-dimensional longitudinal study. *Orthodontics and Craniofacial Research*, 18, 175–184.
- Li, J. & Ji, L. (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, 95, 221–227.
- Liu, Y., Kau, C.H., Talbert, L. & Pan, F. (2014) Three-dimensional analysis of facial morphology. *The Journal of Craniofacial Surgery*, 25, 1890–1894.

- Matthews, H.S., Palmer, R.L., Baynam, G.S., Quarrell, O.W., Klein, O.D., Spritz, R.A. et al. (2021) Large-scale open-source three-dimensional growth curves for clinical facial assessment and objective description of facial dysmorphism. *Scientific Reports*, 11, 12175.
- Matthews, H., Penington, A., Clement, J., Kilpatrick, N., Fan, Y. & Claes, P. (2018) Estimating age and synthesising growth in children and adolescents using 3D facial prototypes. *Forensic Science International*, 286, 61–69.
- Matthews, H.S., Penington, A.J., Hardiman, R., Fan, Y., Clement, J.G., Kilpatrick, N.M. et al. (2018) Modelling 3D craniofacial growth trajectories for population comparison and classification illustrated using sex-differences. *Scientific Reports*, 8, 4771.
- Matthews, H., Penington, T., Saey, I., Halliday, J., Muggli, E. & Claes, P. (2016) Spatially dense morphometrics of craniofacial sexual dimorphism in 1-year-olds. *Journal of Anatomy*, 229, 549–559.
- Milella, M., Franklin, D., Belcastro, M.G. & Cardini, A. (2021) Sexual differences in human cranial morphology: is one sex more variable or one region more dimorphic? *Anatomical Record*, 304, 2789–2810.
- Mossey, P.A., Little, J., Munger, R.G., Dixon, M.J. & Shaw, W.C. (2009) Cleft lip and palate. *Lancet*, 374, 1773–1785.
- Mydlová, M., Dupej, J., Koudelová, J. & Velemínská, J. (2015) Sexual dimorphism of facial appearance in ageing human adults: a cross-sectional study. *Forensic Science International*, 257, 519.e1–519.e9.
- O'Higgins, P., Moore, W.J., Johnson, D.R., McAndrew, T.J. & Flinn, R.M. (1990) Patterns of cranial sexual dimorphism in certain groups of extant hominoids. *Journal of Zoology*, 222, 399–420.
- Parsons, T.E., Kristensen, E., Hornung, L., Diewert, V.M., Boyd, S.K., German, R.Z. et al. (2008) Phenotypic variability and craniofacial dysmorphology: increased shape variance in a mouse model for cleft lip. *Journal of Anatomy*, 212, 135–143.
- Reinhold, K. & Engqvist, L. (2013) The variability is in the sex chromosomes. *Evolution*, 67, 3662–3668.
- Ruscio, J. & Mullen, T. (2012) Confidence intervals for the probability of superiority effect size measure and the area under a receiver operating characteristic curve. *Multivariate Behavioral Research*, 47, 201–223.
- Sforza, C., Grandi, G., Binelli, M., Dolci, C., De Menezes, M. & Ferrario, V.F. (2010) Age- and sex-related changes in three-dimensional lip morphology. *Forensic Science International*, 200, e1–e7.
- Sforza, C., Grandi, G., De Menezes, M., Tartaglia, G.M. & Ferrario, V.F. (2010) Age- and sex-related changes in the normal human external nose. *Forensic Science International*, 204, e1–e9.
- Sheehan, M.J. & Nachman, M.W. (2014) Morphological and population genomic evidence that human faces have evolved to signal individual identity. *Nature Communications*, 5, 4800.
- Shrimpton, S., Daniels, K., de Greef, S., Tilotta, F., Willems, G., Vandermeulen, D. et al. (2014) A spatially-dense regression study of facial form and tissue depth: towards an interactive tool for craniofacial reconstruction. *Forensic Science International*, 234, 103–110.
- Slatkin, M. (1984) Ecological causes of sexual dimorphism. *Evolution*, 38, 622.
- Summersby, S., Harris, B., Denson, T.F. & White, D. (2022) Tracking sexual dimorphism of facial width-to-height ratio across the lifespan: implications for perceived aggressiveness. *Royal Society Open Science*, 9, 211500.
- Tanikawa, C., Zere, E. & Takada, K. (2016) Sexual dimorphism in the facial morphology of adult humans: a three-dimensional analysis. *Homo*, 67, 23–49.
- Toma, A.M., Zhurov, A., Playle, R. & Richmond, S. (2008) A three-dimensional look for facial differences between males and females in a British-Caucasian sample aged 151/2 years old. *Orthodontics and Craniofacial Research*, 11, 180–185.
- Tome, P., Fierrez, J., Vera-Rodriguez, R. & Ramos, D. (2013) Identification using face regions: application and assessment in forensic scenarios. *Forensic Science International*, 233, 75–83.
- Ursi, W.J., Trotman, C.A., McNamara, J.A. & Behrents, R.G. (1993) Sexual dimorphism in normal craniofacial growth. *Angle Orthodontics*, 63, 47–56.
- Velemínská, J., Bigoni, L., Krajčiček, V., Borský, J., Šmahelová, D., Cagánová, V. et al. (2012) Surface facial modelling and allometry in relation to sexual dimorphism. *Homo*, 63, 81–93.
- Wainer, H. (2007) The most dangerous equation. *American Scientist*, 95, 249.
- White, J.D., Indencleef, K., Naqvi, S., Eller, R.J., Hoskens, H., Roosenboom, J. et al. (2021) Insights into the genetic architecture of the human face. *Nature Genetics*, 53, 45–53.
- White, J.D., Ortega-Castrillón, A., Matthews, H., Zaidi, A.A., Ekrami, O., Snyders, J. et al. (2019) MeshMonk: open-source large-scale intensive 3D phenotyping. *Scientific Reports*, 9, 6085.
- Zajitschek, S.R., Zajitschek, F., Bonduriansky, R., Brooks, R.C., Cornwell, W., Falster, D.S. et al. (2020) Sexual dimorphism in trait variability and its eco-evolutionary and statistical implications. *eLife*, 9, e63170.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Matthews, H.S., Mahdi, S., Penington, A.J., Marazita, M.L., Shaffer, J.R., Walsh, S. et al. (2023) Using data-driven phenotyping to investigate the impact of sex on 3D human facial surface morphology. *Journal of Anatomy*, 243, 274–283. Available from: <https://doi.org/10.1111/joa.13866>