Exploring the midline soft tissue surface changes from 12 to 15 years of age in three distinct country population cohorts.

Richmond S, Zhurov AI, Ali ASM, Pirttiniemi P, Heikkinen T, Harila V, Silinevica S, Jakobsone G, Urtane I.

Address for correspondence
Professor Stephen Richmond
Applied Clinical Research and Public Health
School of Dentistry, College of Biomedical and Life Sciences
Heath Park
Cardiff
CF14 4XY

Abstract:

Introduction: Several studies have highlighted differences in the facial features in a White European population. Genetics appears to have a major influence on normal facial variation and environmental factors are likely to have minor influences on face shape directly or through epigenetic mechanisms.

Aim: The aim of this longitudinal cohort study is to determine the rate of change in midline facial landmarks in 3 distinct homogenous population groups (Finnish, Latvian and Welsh) from 12.8 to 15.3 years of age. This age range covers the pubertal growth period for the majority of boys and girls.

Methods: A cohort of children aged 12 were monitored for facial growth in 3 countries (Finland (n=60), Latvia (n=107) and Wales (n=96)). 3D facial surface images were acquired (using either laser or photogrammetric methods) at regular intervals (6 to 12 months) for 4 years. Ethical approval was granted in each country. 9 midline landmarks were identified and the relative spatial position of these surface landmarks were measured relative to the mid-endocanthion (men) over a 4-year period.

Results: This study reports the children who attended 95% of all scanning sessions (Finland 48 out of 60; Latvia 104 out of 107; Wales 50 out of 96). Considerable facial variation is seen for all countries and sexes. There are clear patterns of growth that show different magnitudes at different age groups for the different country groups, sexes and facial parameters. The greatest single yearly growth rate (5.4mm) was seen for Welsh males for men-pg distance at 13.6 years of age. Males exhibit greater rates of growth compared to females. These variations in magnitude and timings are likely to be influenced by genetic ancestry as a result of population migration.

Conclusion: The midline points are a simple and valid method to assess the relative spatial positions of facial surface landmarks. This study confirms previous reports on the subtle differences in facial shapes and sizes of male and female children in different populations and also highlights the magnitudes and timings of growth for various midline landmark distances to mid-endocanthion point.

Introduction

For many years there has been lack of clarity in detailing the influences of genetics and environment on face shape. Recent Genome Wide Association Studies (GWAS) have identified over 50 genes associated with distinct facial surface features occurring in normal facial variation (1-8) and there has been an improved understanding of biological pathways affecting facial shape and size (9, 10); heritability (11), arguably cardio-metabolic risk factors (12), sex hormones such as testosterone (13) and childhood illnesses (14). These studies highlight the dominance of the genetic influences on face shape, although face shape can be significantly altered by trauma and surgery (2, 3). Reports on the influence of maternal alcohol intake on normal variation has been conflicting at one year of age and 15 years of age (15-17). The possible shared genetic influences of breathing disorders (and environmental pollutants) on facial shape currently suggest minor facial shape differences compared to control groups in a growing population (2, 3, 18, 19).

Several studies have highlighted the close relationship between the surface soft tissues and underlying skeletal structures (20-23), which supports the assessment of facial surface to evaluate growth changes. Normal facial variation in a White European population can be explained by face height (28.8%), face width (10.4%) and nose prominence (6.7%) and relative prominence of the maxilla to the mandible (5.3%) (24). Facial features differ slightly across the European population due to ancestry (25-27).

Over 10 facial growth cohort studies have been reported (mainly in the USA and Europe) with varying cohort designs, sample sizes and ethnicities, using mostly photographs and radiography (28-38). More recently three-dimensional facial surface scanning has been employed to evaluate longitudinal changes in population groups (39-42). These studies use non-invasive 3D scanning devices to capture the facial surface on repeated occasions to map growth trajectories and determine the rate of growth in the overall face and/or facial features.

The challenge for longitudinal cohort growth studies is retaining as much of the original sample as possible throughout the period of study so that the findings are not influenced by fluctuating sample sizes or the application of inappropriate missing data modelling.

It is contentious where to superimpose serial radiographs or facial scans. For example, placing serial radiographs anchored on Sella, Nasion, mid-endocanthi or centroids of Procrustes analyses will yield very different growth projections (6, 43-45). Arguably, more importantly than superimpositions/registrations and complex analyses of sequential radiographs is to simply identify the differing rates of growth between various facial midline parameters (independent of face position) in both sexes in different population cohorts as these are common, easily recognisable facial landmarks that are routinely used in orthodontic practice. In addition, recent 3D studies have shown significant antero-posterior and vertical changes in facial shape in the midline in relation to glabella, nasion, nose upper and lower lips and chin (21, 39, 46, 47).

It has been reported that different facial shapes are present within the White European population. Different population migration patterns will inevitably result in different genetic structures arguably associated with different pubertal timings and patterns in facial growth leading to differences in face shape (26, 27, 48-53)

The aim of this longitudinal cohort study is to determine the rate of change in midline facial landmarks in 3 distinct population groups (Finnish, Latvian and Welsh) from 12.8 to 15.3 years of age. This age range covers the pubertal growth period for the majority of boys and girls (39, 43, 46, 54, 55).

Type of study: Longitudinal cohort.

Subjects

Children between 12.8 and 15.3 years of age. This age group allowed the maximum numbers of individuals to be assessed in 3 growth cohorts.

Finland: Children were White European from the Oulu area (n=48; 23 males and 25 females). Ethical approval was granted by the City of Oulu (reference 7728/2006).

Latvia: Children were recruited from the Riga area (n= 104; 56 males and 48 females). Ethical approval was granted by the Ethics Committee of Riga Stradiņš University (reference 14; 28.06.2012).

Wales: Children were White European selected from the year-7 cohort in 2 large comprehensive schools in the South Wales Valleys area within Rhondda Cynon Taf (n=50; 27 males and 23 females). Ethical approval was obtained from the director of education, head teachers, school committees, and the relevant ethics committees of Bro Taf and Cardiff University (reference 04/WSE/109). Written informed consent was obtained before obtaining the 3D facial surface scans.

Exclusion criteria: non-White European and individuals with a history of previous craniofacial trauma, craniofacial anomalies, facial disfigurement and clinically evident facial asymmetry.

Facial surface acquisition and image processing

The facial surfaces of the Finnish and Welsh children (in natural head posture) were acquired by two high resolution laser scanners (Konica Minolta® Vivid 900/910 (Konica Minolta is a registered trademark of Konica Minolta INC). The capture and image processing have been reported extensively and has been shown to be reliable (56-60). All the Latvian children faces were captured using the 3dMDface system scanner (3dMDFace is a registered trademark of 3DMD LIMITED). The facial scans were planned for every 6 months to monitor facial growth. The facial landmarks were recorded by 2 calibrated examiners (Examiner 1, Welsh and Finnish; Examiner 2, Latvian cohort).

Identifying soft tissue landmarks

Nine facial landmarks described by (61) (Figure 1) were identified and recorded by zooming and rotating the images to locate each specific landmark. The mid-endocanthion point (men) was constructed from the left and right inner canthi. The mid-endocanthion point is regarded as a relatively stable landmark with the inner canthi only increasing by 1.6mm and 1.9mm for males and females respectively (between ages 9 to 16 years of age) (6, 62, 63).

Statistical analyses

The intra-operator reliability test was conducted using 80 facial images which were randomly selected across all years and populations. All the facial scans were re-landmarked seven days apart. Agreement of landmark positioning was categorized according to 3 levels (≤ 0.5mm, ≤ 1.0mm, > 1.0mm) (64). Inter-examiner reliability was also conducted and the mean error between two landmarks and the percentage variance due to landmarking errors as a proportion of total mean variance will be reported.

Some of the subjects did not attend all scanning sessions. The missing data were linearly interpolated by the algorithm (see supplemental information). The data points for the rate-of-growth curves were calculated with an increment of one month (δ = 1 month = 0.083 year) for all ages involved in the study. The rate-of-growth curves were plotted using spline smoothing.

Results

This study reports the children who attended the majority of the scanning sessions (Finland 48 out of 60; Latvia 104 out of 111; Wales 50 out of 96 – Table 1). Not all the children were able to attend on each occasion due to illness, examinations, school trips and school related activities. There were no significant facial asymmetries in the sample all, facial asymmetries were less than 0.5mm.

Landmark reliability ranged from 0.053 mm to 1.593 mm for Examiner 1 (Table 2). Out of 27 coordinates, 20 were classified as highly reliable, 13 coordinates were moderately reliable, and 3 coordinates were poorly reliable. 91.7% of landmarks were found to be reproducible to at least 1.0 mm, indicating a moderate-to-high level of reliability. The reliability of z-coordinates appeared to be much higher than that of the x- and y-coordinates, with the most coordinates in the z-axis reliable to less than 1.0 mm. The y-coordinates showed the poorest reproducibility for *g*Y, *n*Y, and *pg*Y which agrees with a previous study (64). For Examiner 2 the intra-examiner error was generally less than 0.5mm except for the following coordinates (gY, 0.61; nY, 0.52; alLZ, 0.51; cphRX, 0.51 and pgY, 0.69). The inter-examiner

mean error on the distances were less than 2mm (men-g, 0.59; men-n, -0.56; men-prn, 0.25; men-sn, 0.04; men-ls, -0.29; men-li, 1.12; men-pg, 1.90).

Country specific growth rates

The country growth rates for both sexes are shown (Figure 2). The age range 12.8 to 15.3 years includes matched data for all 3 cohorts.

Males

There are clearly different patterns of growth for the individuals in each country. There are 2 peaks of growth for Welsh children 12.8 and 13.6 years of age for all measures (except men-g and men-n). After the 2nd peak, growth essentially stabilises for all parameters until 15.3 years of age. The maximum peak 5.4mm/year occurs for the Welsh at 13.6 years of age. The Welsh males are distinctly different from both Finnish and Latvian males the latter 2 populations reveal similarity both showing 2 peaks in growth at similar age ranges with the Latvians showing a longer sustained growth for men-pg (13.4 to 14.8 years), men-g and men-n show the smallest growth rates.

Females

In contrast to the males, females show greater similarity in growth of the landmarks and the growth rates are significantly less than shown in the males. There are 2 significant peaks in the Finnish and Welsh females (12.8 to 14 years of age). Again men-g and men-n show the smallest growth rates.

Combining all data (Figure 3)

Males

When the data is combined, the graphs demonstrate the averaging of growth peaks for the 3 countries with essentially a smoother and flatter curve with a maximum/minimum of 3.3/1.4mm per year at 12.8 and 14.9 years of age respectively for men-pg.

Females

The peak growth rate for men-li occurs at 13.4 years of age. The decline in growth rate is seen earlier than males at 13.6 years of age. The maximum growth rate occurs for men-li (2mm) at 13.4 years of age compared with males (2.3mm).

Discussion

Reliability

The placement of landmarks was generally reliable in respect to intra and inter-examiner error. The variance due to landmarking errors as a proportion of total mean variation was less than 3.4% (men-n, 3.42%; men-prn, 1.60%; men-sn, 3.57%; men-ls, 2.93%; men-li, 2.15%; men-pg, 3.35%). Therefore, the changes in growth velocity and patterns of growth in the 3 populations are valid.

Sampling

With the Bhatia and Leighton study (20), 202 individuals were recruited at birth and 159 were recalled at 16 years of age (80% retained although the number may be considerably lower as 736 individuals were recruited at birth and the samples between 12 and 16 years of age may reflect a cross sectional rather than a true cohort sample). This compares to 80% Finnish; 94% Latvian and 52% Welsh retention rates for all scheduled image acquisitions 12.8 to 15.3 years of age. However, the findings of this study reflect a strict cohort with all individuals attending the majority (95%) of the acquisition sessions. The samples for the 3 countries are relatively small but compares favourably to other cohort and cross-sectional studies (15, 25, 27). For this study it was important to keep the integrity of the cohorts with modelling for missing data for each session (where appropriate) to reflect the facial growth that is seen in children.

Missing data

Missing data is inevitable and are a challenge to manage in cohort studies. Data can be categorised as missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) (65, 66). The data in this study can be classified as MCAR and will yield asymptomatic unbiased estimates. In a systematic review of 84 reported epidemiological studies 26% missing data was reported (67). In this present study the overall missing data are considerably lower at 5.5% (Finnish 14.3%, Latvian 4.2% and Welsh 0%).

There are many techniques to model missing data and each method can have a significant effect on the outcome especially when a high percentage of missing data occurs at specific

time points (68),(69). In this study the missing data were relatively small across the common age interval and a simple extrapolation was undertaken. The difference in curves produced by the current study and that reported by Bhatia and Leighton (20) with only 10 and 15 complete data sets for individuals between the ages of 12 and 15 highlight over averaging when data is missing essentially producing gentle curves as opposed to fluctuating changes identified with the present study.

Facial landmarking

Facial anatomical landmarks are not finite structures and are subject to developmental changes. Most cephalometric landmark definitions relate to the deepest/most prominent concavity/convexity and therefore subject to developmental changes as a direct result from deposition and resorption (70), genetic influences (1), hormonal changes (13) and mandibular rotation (36). These developmental changes may alter the relative spatial position of landmarks resulting in small negative increments during the growth period (which also may be influenced by the error in landmarking particularly in the y-axis). However, the negative increments follow patterns. For instance, there were negative changes (0.8 to 1mm) in sn-pg for both Finnish and Welsh males (13 and 13.6 years of age respectively) and can be explained by a forward mandibular growth rotation in these 2 populations (36) (Figure 3).

Use of mid-endocanthion point

The mid-endocanthion point is regarded as a relatively stable landmark with the inner canthi only increasing by 1.6mm and 1.9mm for males and females respectively (age 9 to 16 years of age) (6, 55, 61-63, 71-73).

The mid-endocanthion constructed landmark and distance to soft tissue nasion has been used repeatedly in Genome Wide Association studies (GWAS) and is associated with the *PAX3* gene (1, 6-8). It is known that the nasion moves forward and upwards during growth and if there was substantial movement of the mid-endocanthion it would be highly unlikely that a genetic association would be found.

Validation (comparison with Bhatia and Leighton study (20))

There are very few studies that provide the rate of facial growth in children. The growth rates for the 3 country groups pooled data are compared with those provided by Bhatia and Leighton (20) (Figure 4). The distances used by Bhatia and Leighton were soft tissue nasion to subnasale (n-sn), soft tissue nasion to pronasale (n-prn) and subnasale to soft tissue menton (sn-me). The Bhatia and Leighton sample can be considered as a cross-sectional longitudinal sample with the rate of growth curves averaged and smoothed compared to a distinct cohort followed 12 to 16 years of age.

n-sn: Males: The patterns of growth for the 3 country groups were similar with greater rates of growth between 13 and 14 years of age and the Bhatia and Leighton sample shows essentially a steady growth rate. Females: There is considerable variation with the Wales sample showing several surges of growth at 12.8, 13.6 and 14.2 years of age. The Latvian and Finnish children also show different peaks of growth at different ages but less dramatic compared to the Welsh.

n-prn: Males: The patterns of growth for all groups are similar for all groups although there appears to be a late marginal surge in the Latvian children at 15.2 years of age. Females: The rate of growth is less than in the males with considerable variability in peaks and timings across the 3 populations.

sn-pg: This parameter was recorded for the facial surface capture systems but Bhatia and Leighton used soft tissue menton instead. There is considerable variation in all country groups. Males: The Finnish cohort show 3 periods of rapid growth and this differs in timing compared to Welsh and Latvian children. The greatest growth rate was shown by the Welsh cohort at 13.6 years of age (3.8mm/year) although there was an earlier period at 12.8 years of age. The Latvian children also showed 2 sustained peaks at 13 years, from 13.8 to 14.8 and also a late surge at 15 years of age. Females: 2 peaks of growth are seen for the Finnish and Welsh children and 3 peaks for the Latvian children. The average pattern for the 3 countries is similar to the Bhatia and Leighton study (20) for both males and females, although the latter rate is on average slightly higher due to the difference in the landmarks measured (pg and men).

For males, the men to glabella distance also follows a different growth pattern compared to other parameters with an increased rate of growth 13.1 to 13.7 years with an increase again from 15 years of age. The men-pg distance is greatest at 12.8 years of age (3.3mm/yr) and 2 additional peaks at 13.8 and 14.2 years of age. As with the females the men-nasion distance shows relatively steady growth throughout, with an average of 0.4mm/yr for females and 0.7mm/yr for males. This is a relatively small rate of growth compared to other nasion measures.

Looking at the growth rates for all females (n=95), generally there appears to be 2 definite periods of rapid growth 12.8 to 14 and 14.4 to 14.9 years of age. However, mid-endocanthion to glabella tends to follow a slightly different pattern with periods of growth 12.8 to 13.5 and 13.6 to 13.9 years of age.

Differences in growth variation

The differences in growth variation between the countries are likely to be due to the different genetic ancestries which will influence subtle differences in timings in puberty and patterns of facial shape (48, 49, 74, 75). Shared genetics has also been reported in respect to tooth eruption, height and craniofacial distances (5). To explore the differences between the 3 population groups the centroid size was calculated with 95% confidence limits (Figure 5). The Welsh males shows a different rate of growth pattern compared to the Finnish and Latvian males which may emphasise the differing genetic ancestry. There is little difference between the female population groups in terms of centroid size. The male growth rate is substantially higher compared to females.

Orthodontic treatment

Orthodontic treatment was undertaken in a proportion of the cohorts (Finland 30%; Latvia 15%; Wales 20%). The majority of treatment was undertaken using fixed and less than 5% functional appliances. The lip morphologies of Finnish and Welsh children tend to be different (47, 76). However, the strength of association between men-pg and men-ls/li for males were similarly high (R^2=0.72) and marginally higher for females (men-pg and men-ls, R^2=0.82 and men-pg and men-li, R^2=0.94). The rate of growth in men-pg is 3.6 times greater in males compared to females, with men-ls 1.89 and men-li 1.72 and these growth

variations are reflected in the various strengths of association. The effects of the treatment on midline points used in this study are likely to be minimal due to the inherited nature of facial features such as face height, inter-ocular distance and relative prominence of the lips (h^2=0.8) (2, 3, 11, 77, 78) and the distinct pubertal growth patterns shown in the 3 cohorts.

Strengths and limitations of the study

The strength of this study is it maps simple midline facial landmarks in 3 cohorts in 3 countries (Finland, Latvia and Wales) over the same time period. The number of individuals followed in each cohort are similar to those reported in other longitudinal growth studies (20, 79). Repeated measures of simple midline points have enabled sequential facial assessment in 3 cohorts. Comparisons over the same time period has highlighted similarities and differences between the cohorts. The greatest differences were observed in males, with the Welsh males exhibiting different growth patterns compared to the Finnish and Latvian children.

Although the samples studied are similar to previous cohorts the sample size is insufficient to explore the influence of different orthodontic treatment modalities on the midline points. This study follows rate of growth from 12.8 to 15.3 years of age. There appears to be significant growth before the study period judged by the rate of reduction in mm/year (12.8 years) and there may be a suggestion of minor later growth surges after 15 years of age. Future studies should capture growth from at least 10 to 18 years of age at 6 monthly intervals.

When making comparisons in terms of growth and possible treatment changes, researchers should be aware that the patterns of growth may be altered in different population cohorts. With the recent advances in facial genetics, the genes associated with distinct facial features and shared genetics should also be taken into account when considering facial growth (1, 5, 53, 80).

Conclusion

The midline points are a simple and valid method to assess the relative spatial positions of facial surface landmarks. This study confirms previous reports on the subtle differences in

facial shapes and sizes of male and female children in different populations and also highlights the magnitudes and timings of growth rates for various midline landmark distance to mid-endocanthion point. These variations of magnitude and timings are likely to be influenced by genetic ancestry as a result of initial population migration and the biological basis of facial shape and growth will need to be validated in larger scale population studies.

Acknowledgements

We would like to thank the support of the schools and children who enthusiastically participated in this longitudinal growth study. We also wish to thank the affiliated universities for their support.

Conflict of Interest

The authors do not have any financial or any other interests or connections directly or indirectly, that have influenced this research study. The authors are completely impartial.

References

1. Claes P, Roosenboom J, White JD, Swigut T, Sero D, Li J, et al. Genome-wide mapping of global-to-local genetic effects on human facial shape. Nat Genet. 2018;50(3):414-23.
2. Richmond S, Howe LJ, Lewis S, Stergiakouli E, Zhurov A. Facial Genetics: A Brief Overview. Front Genet. 2018;9:462.
3. Richmond S, Wilson-Nagrani C, Zhurov A, Farnell D, Galloway J, Mohd Ali A, et al. Factors influencing facial shape. 2018.
4. Xiong Z, Dankova G, Howe LJ, Lee MK, Hysi PG, de Jong MA, et al. Novel genetic loci affecting facial shape variation in humans. bioRxiv. 2019:693002.
5. Fatemifar G, Hoggart CJ, Paternoster L, Kemp JP, Prokopenko I, Horikoshi M, et al. Genome-wide association study of primary tooth eruption identifies pleiotropic loci associated with height and craniofacial distances. 2013;22(18):3807-17.
6. Paternoster L, Zhurov AI, Toma AM, Kemp JP, St Pourcain B, Timpson NJ, et al. Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position. Am J Hum Genet. 2012;90.
7. Adhikari K, Fontanil T, Cal S, Mendoza-Revilla J, Fuentes-Guajardo M, Chacon-Duque JC, et al. A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. Nat Commun. 2016;7:10815.
8. Liu F, van der Lijn F, Schurmann C, Zhu G, Chakravarty MM, Hysi PG, et al. A genome-wide association study identifies five loci influencing facial morphology in Europeans. PLoS Genet. 2012;8(9):e1002932.
9. Wilderman A, VanOudenhove J, Kron J, Noonan JP, Cotney J. High-Resolution Epigenomic Atlas of Human Embryonic Craniofacial Development. Cell Rep. 2018;23(5):1581-97.
10. Sun Z, da Fontoura CSG, Moreno M, Holton NE, Sweat M, Sweat Y, et al. FoxO6 regulates Hippo signaling and growth of the craniofacial complex. PLoS Genet. 2018;14(10):e1007675.

11. Hoskens H, Li J, Indencleef K, Gors D, Larmuseau MH, Richmond S, et al. Spatially Dense 3D Facial Heritability and Modules of Co-heritability in a Father-Offspring Design. 2018;9.
12. Djordjevic J, Lawlor DA, Zhurov AI, Toma AM, Playle R, Richmond SJBo. A population-based cross-sectional study of the association between facial morphology and cardiometabolic risk factors in adolescence. 2013;3(5):e002910.
13. Roosenboom J, Indencleef K, Lee MK, Hoskens H, White JD, Liu D, et al. SNPs Associated With Testosterone Levels Influence Human Facial Morphology. Front Genet. 2018;9:497.
14. Pound N, Lawson DW, Toma AM, Richmond S, Zhurov AI, Penton-Voak ISJPotRSBBS. Facial fluctuating asymmetry is not associated with childhood ill-health in a large British cohort study. 2014;281(1792):20141639.
15. Muggli E, Matthews H, Penington A, Claes P, O'Leary C, Forster D, et al. Association Between Prenatal Alcohol Exposure and Craniofacial Shape of Children at 12 Months of Age. JAMA Pediatr. 2017;171(8):771-80.
16. Howe LJ, Sharp GC, Hemani G, Zuccolo L, Richmond S, Lewis SJ. Prenatal alcohol exposure and facial morphology in a UK cohort. Drug Alcohol Depend. 2019;197:42-7.
17. Mamluk L, Edwards HB, Savovic J, Leach V, Jones T, Moore THM, et al. Low alcohol consumption and pregnancy and childhood outcomes: time to change guidelines indicating apparently 'safe' levels of alcohol during pregnancy? A systematic review and meta-analyses. BMJ Open. 2017;7(7):e015410.
18. Al Ali A, Richmond S, Popat H, Toma AM, Playle R, Pickles T, et al. A three-dimensional analysis of the effect of atopy on face shape. 2013;36(5):506-11.
19. Al Ali A, Richmond S, Popat H, Playle R, Pickles T, Zhurov AI, et al. The influence of snoring, mouth breathing and apnoea on facial morphology in late childhood: a three-dimensional study. BMJ Open. 2015;5(9):e009027.
20. Bhatia SN, & Leighton, B. C. Manual of facial growth: A computer analysis of longitudinal cephalometric growth data: Oxford University Press; 1993.
21. Al Ali A, Richmond S, Popat H, Toma AM, Playle R, Zhurov AI, et al. The influence of asthma on face shape: a three-dimensional study. Eur J Orthod. 2014;36(4):373-80.
22. Shrimpton S, Daniels K, de Greef S, Tilotta F, Willems G, Vandermeulen D, et al. A spatially-dense regression study of facial form and tissue depth: towards an interactive tool for craniofacial reconstruction. Forensic Sci Int. 2014;234:103-10.
23. Young NM, Sherathiya K, Gutierrez L, Nguyen E, Bekmezian S, Huang JC, et al. Facial surface morphology predicts variation in internal skeletal shape. Am J Orthod Dentofacial Orthop. 2016;149(4):501-8.
24. Toma AM, Zhurov AI, Playle R, Marshall D, Rosin PL, Richmond S. The assessment of facial variation in 4747 British school children. Eur J Orthod. 2012;34(6):655-64.
25. Galloway J, Farnell D, Zhurov A, Richmond S, Katić V, Lahdesmaki R, et al., editors. The Influence of Ethnicity, Sex and Subject on Facial Shape. 2018 IADR/PER General Session & Exhibition; 2018.
26. Kau CH, Richmond S, Zhurov A, Ovsenik M, Tawfik W, Borbely P, et al. Use of 3-dimensional surface acquisition to study facial morphology in 5 populations. Am J Orthod Dentofacial Orthop. 2010;137(4 Suppl):S56.e1-9; discussion S-7.
27. Hopman SM, Merks JH, Suttie M, Hennekam RC, Hammond P. Face shape differs in phylogenetically related populations. Eur J Hum Genet. 2014;22(11):1268-71.

28. Saunders SR, Popovich F, Thompson GW. A family study of craniofacial dimensions in the Burlington Growth Centre sample. American journal of orthodontics. 1980;78(4):394-403.
29. Riolo ML. Moyers RE. McNamara JA. HW. An Atlas of Craniofacial Growth (Craniofacial Growth Series, Monograph Number 2)1974.
30. Blanchette ME, Nanda RS, Currier GF, Ghosh J, Nanda SK. A longitudinal cephalometric study of the soft tissue profile of short- and long-face syndromes from 7 to 17 years. American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics. 1996;109(2):116-31.
31. Bishara SE, Jakobsen JR, Hession TJ, Treder JE. Soft tissue profile changes from 5 to 45 years of age. American Journal of Orthodontics and Dentofacial Orthopedics. 1998;114(6):698-706.
32. Moorrees CF, Efstratiadis SS, Kent RL, Jr. The mesh diagram for analysis of facial growth. Proceedings of the Finnish Dental Society Suomen Hammaslaakariseuran toimituksia. 1991;87(1):33-41.
33. Richardson E. Atlas of craniofacial growth in Americans of African descent. Univ of Michigan Center for (January 1, 1992): Univ of Michigan Center for (January 1, 1992); 1992.
34. Krogman WM, Mazaheri M, Harding RL, Ishiguro K, Bariana G, Meier J, et al. A longitudinal study of the craniofacial growth pattern in children with clefts as compared to normal, birth to six years. The Cleft palate journal. 1975;12(00):59-84.
35. Chaconas SJ, Bartroff JD. Prediction of normal soft tissue facial changes. Angle Orthod. 1975;45(1):12-25.
36. Skieller V, Bjork A, Linde-Hansen T. Prediction of mandibular growth rotation evaluated from a longitudinal implant sample. Am J Orthod. 1984;86(5):359-70.
37. Bhatia SN, Leighton BC. Manual of facial growth: A computer analysis of longitudinal cephalometric growth data: Oxford University Press; 1993. 560 p.
38. Prahl-Andersen B, Kowalski CJ. Analysis of cohort effects in mixed longitudinal data sets. International journal of sports medicine. 1997;18 Suppl 3:S186-90.
39. Matthews H, Penington A, Clement J, Kilpatrick N, Fan Y, Claes P. Estimating age and synthesising growth in children and adolescents using 3D facial prototypes. Forensic Sci Int. 2018;286:61-9.
40. Wen YF, Wong HM, McGrath CP. A longitudinal study of facial growth of Southern Chinese in Hong Kong: Comprehensive photogrammetric analyses. PLoS One. 2017;12(10):e0186598.
41. Kau CH, Richmond S. Three-dimensional analysis of facial morphology surface changes in untreated children from 12 to 14 years of age. Am J Orthod Dentofacial Orthop. 2008;134(6):751-60.
42. Matthews HS, Penington AJ, Hardiman R, Fan Y, Clement JG, Kilpatrick NM, et al. Modelling 3D craniofacial growth trajectories for population comparison and classification illustrated using sex-differences. Sci Rep. 2018;8(1):4771.
43. Richmond S, Toma AM, Zhurov AI. [New perspectives on craniofacial growth]. Orthod Fr. 2009;80(4):359-69.
44. Houston WJ, Lee RT. Accuracy of different methods of radiographic superimposition on cranial base structures. Eur J Orthod. 1985;7(2):127-35.
45. Richmond S, Wilson-Nagrani C, Zhurov A, Farnell D, Galloway J, Mohd Ali A, et al. Factors influencing facial shape. Evidenced-Based Orthodontics. 2018.

46. Koudelova J, Hoffmannova E, Dupej J, Veleminska J. Simulation of facial growth based on longitudinal data: Age progression and age regression between 7 and 17 years of age using 3D surface data. PLoS One. 2019;14(2):e0212618.
47. Hamilton SM. Measuring change in lip shape from 12 to 17 years of age in two different ethnic populations. Cardiff: Cardiff University; 2014.
48. Leslie S, Winney B, Hellenthal G, Davison D, Boumertit A, Day T, et al. The fine-scale genetic structure of the British population. Nature. 2015;519(7543):309-14.
49. Nelis M, Esko T, Magi R, Zimprich F, Zimprich A, Toncheva D, et al. Genetic structure of Europeans: a view from the North-East. PLoS One. 2009;4(5):e5472.
50. Persyn E, Redon R, Bellanger L, Dina C. The impact of a fine-scale population stratification on rare variant association test results. PLoS One. 2018;13(12):e0207677.
51. Day FR, Thompson DJ, Helgason H, Chasman DI, Finucane H, Sulem P, et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. Nat Genet. 2017;49(6):834-41.
52. Cousminer DL, Stergiakouli E, Berry DJ, Ang W, Groen-Blokhuis MM, Korner A, et al. Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. Hum Mol Genet. 2014;23(16):4452-64.
53. Cousminer DL, Widen E, Palmert MR. The genetics of pubertal timing in the general population: recent advances and evidence for sex-specificity. Curr Opin Endocrinol Diabetes Obes. 2016;23(1):57-65.
54. Tanner JM, Whitehouse RH, Takaishi M. Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. II. Arch Dis Child. 1966;41(220):613-35.
55. Nute SJ, Moss JP. Three-dimensional facial growth studied by optical surface scanning. J Orthod. 2000;27(1):31-8.
56. Kau C, Zhurov A, Scheer R, Bouwman S, Richmond S. The feasibility of measuring three-dimensional facial morphology in children. Orthodontics & Craniofacial Research. 2004;7(4):198-204.
57. Kau CH, Richmond S. Three-dimensional analysis of facial morphology surface changes in untreated children from 12 to 14 years of age. American Journal of Orthodontics and Dentofacial Orthopedics. 2008;134(6):751-60.
58. Kau CH, Zhurov A, Richmond S, Bibb R, Sugar A, Knox J, et al. The 3-dimensional construction of the average 11-year-old child face: A clinical evaluation and application. Journal of Oral and Maxillofacial Surgery. 2006;64(7):1086-92.
59. Kau CH, Richmond S, Zhurov AI, Knox J, Chestnutt I, Hartles F, et al. Reliability of measuring facial morphology with a 3-dimensional laser scanning system. American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics. 2005;128(4):424-30.
60. Zhurov AI. Computer methods for measuring 3D facial morphology. In: Middleton J, Shrive MG, Jones ML, editors. Computer methods in biomechanics and biomedical engineering. 5. [London] :2005. p. 1.
61. Farkas LG. Anthropometry of the head and face. 2nd ed. New York: Raven Press; 1994.
62. Feingold M, Bossert WH. Normal values for selected physical parameters: an aid to syndrome delineation. Birth Defects Orig Artic Ser. 1974;10(13):1-16.

63. Zhurov AI, Richmond S, Kau CH, Toma AM. Averaging facial images. In: Kau CH, Richmond S, editors.: Wiley-Blackwell; 2010.
64. Toma AM, Zhurov A, Playle R, Ong E, Richmond S. Reproducibility of facial soft tissue landmarks on 3D laser-scanned facial images. Orthodontics & Craniofacial Research. 2009;12(1):33--42.
65. Rubin DB. Inference and missing data. Biometrika. 1976;63(3):581-92.
66. DB R. Multiple imputation for nonresponse in surveys: John Wiley & Sons; 1987 1987.
67. Eekhout I, de Boer RM, Twisk JW, de Vet HC, Heymans MW. Missing data: a systematic review of how they are reported and handled. Epidemiology. 2012;23(5):729-32.
68. Perkins NJ, Cole SR, Harel O, Tchetgen Tchetgen EJ, Sun B, Mitchell EM, et al. Principled Approaches to Missing Data in Epidemiologic Studies. Am J Epidemiol. 2018;187(3):568-75.
69. Stavseth MR, Clausen T, Roislien J. How handling missing data may impact conclusions: A comparison of six different imputation methods for categorical questionnaire data. SAGE Open Med. 2019;7:2050312118822912.
70. Enlow DH, Hans MG. Essentials of facial growth. 2nd ed. Ann Arbor, MI: Distributed by Needham Press; 2008. xvii, 507 p. p.
71. Park DH, Choi WS, Yoon SH, Song CH. Anthropometry of Asian Eyelids by Age. Plastic and Reconstructive Surgery. 2008;121(4):1405-13.
72. Beden Ü, Beltram M. Periorbital anthropometric measurements. Handbook of Anthropometry: Physical Measures of Human Form in Health and Disease. New York: Springer New York; 2012. p. 641-54.
73. Afrand M, Ling CP, Khosrotehrani S, Flores-Mir C, Lagravere-Vich MO. Anterior cranial-base time-related changes: A systematic review. American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics. 2014;146(1):21-32.e6.
74. Kerminen S, Havulinna AS, Hellenthal G, Martin AR, Sarin AP, Perola M, et al. Fine-Scale Genetic Structure in Finland. G3 (Bethesda). 2017;7(10):3459-68.
75. Al-Asadi H, Petkova D, Stephens M, Novembre J. Estimating recent migration and population-size surfaces. PLoS Genet. 2019;15(1):e1007908.
76. Davis MG. Determining Lip Phenotype and Genotype in Finnish Adults Cardiff: Cardiff University; 2018.
77. Djordjevic J, Zhurov AI, Richmond S, Consortium VJPo. Genetic and environmental contributions to facial morphological variation: a 3D population-based twin study. 2016;11(9):e0162250.
78. Thiruvenkatachari B, Harrison JE, Worthington HV, O'Brien KD. Orthodontic treatment for prominent upper front teeth (Class II malocclusion) in children. Cochrane Database Syst Rev. 2013(11):Cd003452.
79. Riolo ML, Moyer R. E., McNamara, J. A., Hunter W. S. Atlas of Craniofacial Growth University of Michigan, Ann Habor: Center for Human Growth and Development 1974.
80. Richmond S, Howe LJ, Lewis S, Stergiakouli E, Zhurov A. Facial Genetics: A Brief Overview. Frontiers in genetics. 2018;9.
81. Hazewinkel M. Linear interpolation, Encyclopedia of Mathematics: Springer Science + Business Media B.V. / Kluwer Academic; 2001.
82. Atkinson K. An Introduction to Numerical Analysis: John Wiley and Sons; 1988.

83. Shepard D, editor A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1986 ACM National Conference. 1968.
84. Levin D. The approximation power of moving least squares. Mathematics of Coputation,. 1998;67:1517-31.
85. Leemis L. Learning Base R: Lightening Source; 2016.
86. Kohl M. Introduction to statistical data analysis with R: Bookboon.com London; 2015.
87. Daróczi G. Data Analysis with R: Packt Publishing; 2015.

Supplemental information

Evaluating errors due to landmarking

Let n denote the sample size and let x_i denote the measurement of a quantity ξ (e.g., the distance between two landmarks) in the ith subject, $i = 1, \dots, n$. Let ξ_i be the true value of the quantity in subject i and let ϵ_i be the error of measurement due to different reasons (e.g., error of landmarking). Then

$$x_i = \xi_i + \epsilon_i$$

Let y_i denote another measurement of the same quantity and let ε_i be the error of measurement. Then

$$y_i = \xi_i + \varepsilon_i$$

Assuming that the true measurements and the errors are independent (uncorrelated), we can estimate the variances of x, y and $x - y$ as

$$\sigma_x^2 = \sigma_\xi^2 + \sigma_\epsilon^2, \quad \sigma_y^2 = \sigma_\xi^2 + \sigma_\varepsilon^2, \quad \sigma_{x-y}^2 = \sigma_\epsilon^2 + \sigma_\varepsilon^2$$

The variances σ_x^2, σ_y^2 and σ_{x-y}^2 are all known, since they can be calculate them from the measurements x_i and y_i ($i = 1, \dots, n$). The other three variances are unknown.

By solving the above three equations, we find that

$$\sigma_\xi^2 = \frac{\sigma_x^2 + \sigma_y^2 - \sigma_{x-y}^2}{2}, \quad \sigma_\epsilon^2 = \frac{\sigma_x^2 - \sigma_y^2 + \sigma_{x-y}^2}{2}, \quad \sigma_\varepsilon^2 = \frac{\sigma_y^2 - \sigma_x^2 + \sigma_{x-y}^2}{2}$$

These formulas can be used to estimate the variation of the quantity ξ across the population as well as the contributions due to errors.

To make the calculations more accurate, we look at the mean measurement

$$\mu_i = \frac{x_i + y_i}{2} = \xi_i + \frac{\epsilon_i + \varepsilon_i}{2}$$

The variance is evaluated as

$$\sigma_\mu^2 = \sigma_\xi^2 + \frac{\sigma_\epsilon^2 + \sigma_\varepsilon^2}{2}$$

It follows that

$$\sigma_\xi^2 = \sigma_\mu^2 - \frac{\sigma_\epsilon^2 + \sigma_\varepsilon^2}{2} = \sigma_\mu^2 - \frac{\sigma_{x-y}^2}{2}$$

The variance of the mean error is equal to $\sigma_{\text{err}}^2 = \sigma_{x-y}^2/2$.

We use this approach to estimate the contribution of landmarking errors based on the intra-examiner reliability study of 40 randomly selected facial images. The faces were landmarked twice with an interval of seven days. The table below shows the following variances for the seven distances involved in the study: σ_x^2, σ_y^2, σ_μ^2, σ_{x-y}^2, σ_ξ^2, and σ_{err}^2.

	men-g	men-n	men-prn	men-sn	men-ls	men-li	men-pg
Total variance of measurement 1	7.10	4.94	15.73	10.79	16.66	19.77	29.35
Total variance of measurement 2	7.13	5.15	16.40	10.77	16.96	20.39	30.76
Variance of mean measurement	7.05	4.98	16.02	10.74	16.76	20.03	30.00
Variance of difference between measurements 1 & 2	0.26	0.25	0.17	0.18	0.20	0.21	0.24
Total subject variance	6.99	4.92	15.98	10.69	16.71	19.97	29.94
Variance due to landmarking errors	0.13	0.12	0.08	0.09	0.10	0.11	0.12
Variance due to landmarking errors, percentage	1.82%	2.48%	0.53%	0.85%	0.59%	0.53%	0.40%

As one can see, the contribution of the landmarking errors is quite small (0.4% to 2.5%). Other sources of error such as those due to scanning, image processing and landmarking bias are also quite small. The contribution from all sources error can be estimated as only a few percent, which is unlikely to affect the main results of the study.

Methodology to obtain the rate-of-growth curves

Suppose that a sample consists of M subjects scanned at N sessions. Let a_{ij} denote the age of subject i $(i = 1, \dots, M)$ at scanning session j $(j = 1, \dots, N)$ and let d_{ij} denote a measurement d (in this study, distance between two landmarks) taken at a_{ij}, for subject i from the image acquired at session j. We intend to plot the average change in the measurement d over an age interval $[a_{\text{min}}, a_{\text{max}}]$ as well as the average rate of change of this measurement, d', over the same age interval. The former will be referred to as a growth curve and the latter as a rate-of-growth (or growth velocity) curve.

The age interval $[a_{\text{min}}, a_{\text{max}}]$ will be defined as the common age range among all subjects, which suggests that

$$a_{\min} = \max_{i=1,\dots,M} a_{i1}, \quad a_{\max} = \min_{i=1,\dots,M} a_{iN}$$

This ensures that the measurements from all subjects are involved in the calculation of the average quantities, d and d', between $a_{\min}$ and $a_{\max}$.

The age interval will be divided into K equal subinterval of length δ and $K+1$ age points, a_0 to a_K, will be used to calculate the average quantities, so that

$$a_{\min} = a_0 < a_1 < \cdots < a_{K-1} < a_K = a_{\max}, \quad a_k = a_{\min} + k\delta, \quad \delta = \frac{a_{\max} - a_{\min}}{K}$$

The age points at which the actual measurements d_{ij} were taken do not generally coincide with a_k (k denotes a would-be visit to a scanning session as though all subjects were scanned at the same ages from a_0 to a_K). Therefore, we have to interpolate the data. Let D_{ik} ($i = 1, \dots, M$ and $k = 0, \dots, K$) denote the data points obtained from d_{ij} using linear interpolation. This can be symbolically written as

$$D_{ik} = \text{interpolate}\big(d_{ij}, a_{ij}, a_k\big), \quad i = 1, \dots, M, \quad j = 1, \dots, N, \quad k = 0, \dots, K$$

Specifically, this means that

$$D_{ik} = d_{ij_0} + (a_k - a_{ij_0}) \frac{d_{ij_1} - d_{ij_0}}{a_{ij_1} - a_{ij_0}}$$

where j_0 and j_1 indicate two consecutive scanning sessions, such that $1 \le j_0 < j_1 \le N$ and $a_{ij_0} \le a_k < a_{ij_1}$, which subject i attended. The interpolation formula for D_{ik} suggests that all data d_{ij} and a_{ij} is used for each i, with j running from 1 to N. The session numbers j_0 and j_1 are calculated (using logical operators) from all ages a_j ($j = 1, \dots, N$) and selected age a_k considering possible missed visits. So this approach covers the situation when the subject missed a scanning session.

We choose the simplest method of linear interpolation (81) to be consistent with other cited studies. For example, the article by Bhatia & Leighton (20) uses this method to evaluate rates of growth. Potentially, there are a number of suitable methods to choose from, such as polynomial, spline, trigonometric, inverse distance weighting, moving least squares and many other interpolation methods (82-84). As there are relatively few missing data across the course of the study, it makes little sense to employ advanced methods, as they would not provide much improvement in accuracy. This is especially true because there are other sources of error (due to scanning, image processing and landmarking), which are greater than interpolation errors.

To plot the average measurement d against age a, we need to determine the values d_k and at the age points a_k, which are calculated as

$$d_k = \text{mean}_i\, D_{ik} = \frac{D_{1k} + \cdots + D_{Mk}}{M}, \quad k = 0, \ldots, K$$

To plot the average rate of change of d against age a, we calculate the values $d'_{k+1/2}$ and at the age points $a_{k+1/2} = a_{\text{min}} + \left(k + \frac{1}{2}\right)\delta$ with $k = 0, \ldots, K-1$. First, we approximate the rates using the difference quotient (a common way of approximating derivatives numerically):

$$D'_{i,k+1/2} = \frac{D_{i,k+1} - D_{ik}}{\delta}, \quad i = 1, \ldots, M, \quad k = 0, \ldots, K-1$$

Then, we calculate the average rates as

$$d'_{k+1/2} = \text{mean}_i\, D'_{i,k+1/2}, \quad k = 0, \ldots, K-1$$

The averaging is done across all subjects in the sample, $i = 1, \ldots, M$.

The growth and rate-of-growth curves for measurement d will be plotted using the evaluated data sets:

$$d_0, d_1, \ldots, d_K \text{ versus } a_0, a_1, \ldots, a_K \quad \text{(growth data)}$$

$$d'_{1/2}, d'_{3/2}, \ldots, d'_{K-1/2} \text{ versus } a_{1/2}, a_{3/2}, \ldots, a_{K-1/2} \quad \text{(rate-of-growth data)}$$

Apart from the average data, 95% confidence intervals are calculated at each age point. To this end, resampling techniques (bootstrapping) are used with 10,000 permutations. The resulting points produce upper and lower 95% confidence curves.

The algorithm outlined above was implemented as a set of in-house subroutines in the R language (85-87). The input data was prepared as a set of Excel® CSV files (Excel is a Registered trademark of Microsoft Corporation). The output data was saved in the same format.