Quantitative analysis of fetal facial morphology using 3D ultrasound and statistical shape modeling: a feasibility study

Andrea Dall'Asta[1,2], Silvia Schievano[3], Jan L Bruse[3], Gowrishankar Paramasivam[1], Christine Tita Kaihura[2], David Dunaway[4], Christoph C Lees[1,5,6]

1. Centre for Fetal Care, Queen Charlotte's and Chelsea Hospital, Imperial College Healthcare NHS Trust, London, United Kingdom.

2. Obstetrics and Gynaecology Unit, University of Parma, Parma, Italy.

3. University College London Institute of Child Health & Great Ormond Street Hospital for Children, London, United Kingdom.

4. Craniofacial Unit, Great Ormond Street Hospital for Children NHS Foundation Trust & University College London Hospital, London, United Kingdom

5. Institute of Reproductive and Developmental Biology, Department of Surgery and Cancer, Imperial College London, London, United Kingdom

6. Department of Development and Regeneration, KU Leuven, Leuven, Belgium.

Address for correspondence:

Dr Christoph C Lees, MD MRCOG

Centre for Fetal Care, Queen Charlotte's and Chelsea Hospital, Imperial College Healthcare NHS Trust, Du Cane Road, London, W12 0HS, United Kingdom

Email: christoph.lees@imperial.nhs.uk

Phone: +44 208383998

Word count in the abstract: 251 words

Word count in the text (references excluded): 2997 words

Conflict of interest statement

All the Authors state no conflict of interest and no financial disclosures.

Abstract

Background – The antenatal detection of facial dysmorphism using three-dimensional ultrasound may raise the suspicion of an underlying genetic conditions but infrequently leads to a definitive antenatal diagnosis. Despite advances in array and non invasive prenatal testing, not all genetic conditions can be ascertained from such testing.

Objectives – The aim of this study was to investigate the feasibility of quantitative assessment of fetal face features using prenatal three-dimensional ultrasound volumes and statistical shape modelling.

Study design – Thirteen normal and seven abnormal stored three-dimensional ultrasound fetal face volumes were analysed, at a median gestation of 29^{+4} weeks (25^{+0} – 36^{+1}). The 20 three-dimensional surface meshes generated were aligned and served as input for a statistical shape model, which computed the mean three-dimensional face shape and three-dimensional shape variations using Principal Component Analysis.

Results – Ten "shape modes" explained over 90% of the total shape variability in the population. While the first mode accounted for overall size differences, the second highlighted shape feature changes from an overall proportionate towards a more asymmetric face shape with wide prominent forehead, and undersized, posteriorly positioned chin. Analysis of the Mahalanobis distance in PCA shape space suggested differences between normal and abnormal fetuses (median and interquartile range distance values 7.31±5.54 for the normal group versus 13.27±9.82 for the abnormal group) (p=.056).

Conclusion – This feasibility study demonstrates that objective characterization and quantification of fetal facial morphology is possible from three-dimensional ultrasound. This technique has the

potential to assist in utero diagnosis particularly of rare conditions where facial dysmorphology is a feature.

Key words: 3-dimensional ultrasound, facial dysmorphism, genetic syndrome, prenatal diagnosis, principal component analysis, statistical shape modeling

Introduction

Despite advances in invasive and non-invasive prenatal diagnostic and screening tests with the advent of CGH array and non invasive prenatal testing, many genetic syndromes or chromosomal abnormalities are still diagnosed only postnatally. In many cases of well phenotypically characterized conditions, not all genetic mutations are known, for example prominent eyebrows and everted nostrils with smooth philtrum may be present in Cornelia de Lange syndrome and depressed nasal bridge and frontal bossing are common features of Apert syndrome. Hence definitive prenatal diagnosis of these conditions is precluded unless targeted testing is performed and even then, the diagnosis is not always positively made.

Recent reports have suggested that the antenatal detection of facial dysmorphisms using three-dimensional ultrasound (3D US) can contribute in the identification of highly suggestive features or lead to in utero diagnosis of abnormalities by targeting genetic tests (1-9). Dysmorphic fetuses may show highly suggestive facial features or, conversely, the facial shape may be apparently normal but associated with additional abnormal findings which may suggest a genetic diagnosis. Whatever the condition, the accurate assessment of the fetal face and the diagnosis of facial abnormalities can be challenging and remains subjective even for experienced investigators.

Due to the higher resolution and image quality of recently developed equipment, three-dimensional and four-dimensional (4D) US, both in multiplanar and rendered views, currently act as adjunct to standard 2D sonography for the visualization of fetal anatomy (10), thus representing a useful tool for parental counselling (1,6,11).

Experience of 3D volumes acquired using computed tomography (CT) scan in children (12-16) has demonstrated that it is possible to reliably and repeatedly represent volumes by landmarking

biologically homologous points. No studies have yet assessed the feasibility of a computational approach for the antenatal analysis of the fetal facial shape. The aim of our study was to evaluate the feasibility of quantitative assessment of 3D US facial volumes via Statistical Shape Modelling (SSM) (17,18). We sought to investigate the feasibility of describing normal and abnormal facial features objectively in order to distinguish dysmorphic facial fetuses from normal ones. If this were achieved it would lead to the possibility of diagnosing genetic syndromes that manifest with craniofacial dysmorphism.

Methods

Patient population and image acquisition

Patients who underwent 3D US scan for clinical purposes between January and June 2016 at two tertiary Fetal Medicine Centres (Queen Charlotte's and Chelsea Hospital, London, UK and University Hospital of Parma, Italy) were considered for the study. Volumes from fetuses over 24 weeks of gestation were considered where fetal position and lie, amniotic fluid volume, limbs attitude and intra-amniotic floating cord allowed acquisition of good quality 3D US data. This was subjectively defined on the basis of the clearness of the facial borders and absence of shadowing in the multiplanar view – unobstructed "face on" view (19,20). Static 3D acquisition of the volume on midsagittal view of the fetal face was undertaken with a low frequency probe (4-8 MHz) equipped Voluson S8-E8 (General Electric Healthcare, London, United Kingdom and General Electric Healthcare Milan, Italy); the volumes were acquired using the highest quality mode and the lowest angle sweep, which allowed the inclusion of the whole facial surface from the forehead to the chin. A 90° coronal view was obtained orienting the volumes along the x, y and z axis during the post-acquisition processing and the oriented volume was stored in the machine as per normal clinical practice. Anonymised volumes were exported in Cartesian coordinate format for further 3D image processing. At this stage, cases were excluded when missing facial parts or due to unsatisfactory definition of the facial borders secondary to shadowing or obstructed vision caused by the uterine wall / intervening limbs / cord loops. Fetuses were labelled as "normal" or "abnormal" according to the antenatal or postnatal diagnosis (Table 1).

Following a query to the National Research Ethics Service ("Developing ultrasound fetal facial dysmorphology analysis based on 3D imaging datasets"), ethical approval for this technical feasibility study using anonymised routinely collected volumes was determined as not necessary.

Image post-processing

An overview of the steps performed in this study for 3D US volume post-processing is provided in Figure 1. Image elaboration included manual face segmentation and 3D reconstruction using Mimics v17 (Materialise, Leuven, Belgium) to create computational surface meshes representing the 3D face shapes. Meshes were cut in order to include the facial shape from the forehead to the chin using three points located in the middle part of the lenses bilaterally and the internal border of the upper alveolar ridge at the incisive foramen as landmarks for the sectional cutting plane.

The cut meshes were manually aligned (Rhinoceros 4.0 McNeel & Associates, Seattle, WA, USA) in a common coordinate system using as references the midline point of the internal border of the upper alveolar ridge at the level of the incisive foramen and a second midline point which was drawn at the symmetry plane of the mesh at the top of the forehead (Figure 1).

Statistical Shape Modelling and Principal Component Analysis of facial shape features

SSM is an advanced statistical method applicable to a set of computational surface meshes, to study complex 3D shape features (17,18,21,22). It allows building descriptive and predictive statistical models based on 3D shape populations and thus extraction of "shape biomarkers" (23). Using the non-parametetric *Deformetrica* SSM framework (24), we computed the 3D mean anatomical facial shape or template based on the aligned surface meshes, with no need for landmarks or point-to-point correspondence (Figure 2). The unique facial shape features of each subject are captured by computing subject-specific transformations that deform the template shape towards each subject shape (Figure 1). Each subject shape *i* is thus characterised by a multitude of deformation vectors ϕ_i rather than by its actual 3D mesh surface (23,25,26,27).

All deformation vectors together can be combined in a deformation matrix *D* which constituted the input for principal component analysis (PCA), a mathematical technique used to condense large amount of 3D shape data into a smaller number of relevant variables, the "principal components" or "shape modes", without a reduction of the data itself (13,28). Shape modes, representing the key shape features of the input population, can be visualised as shape deformation of the mean shape and allow the application of conventional statistical testing on 3D shape data. The principal contributors to facial shape variability (i.e. shape modes) in our cohort were computed in Matlab (The Mathworks, Natick, MA) (26,29,30). Total shape variability was broken down into *m* shape modes, each accounting for a certain percentage of the total variability (28).

In order to visualise dominant 3D fetal face features as derived by PCA, shape modes were visualised in ParaView (31) as deformations of the computed template shape along the mode, from -3 standard deviations (SD) to +3SD around the mean shape. Furthermore, shape modes were numerically described by so-called "shape vectors" $\{f_{i,k}\}_{k=1\ldots m}$ (23,30,32) with each shape vector entry quantifying how much of the shape features described by the respective shape mode is contained within a subject's shape.

The shape vector entries $\{f_{i,k}\}$ for each PCA shape mode *k* and subject *i* were computed. Differences in shape vector distributions between normal and abnormal subjects were visualised and related to 3D facial features described by the respective shape mode. To provide a quantitative "abnormality score" that takes into account all 3D facial features (parameterised as deformations of the template shape in *D*), we computed the subject-specific *Mahalanobis distance* $d_{Mahal,i}$ in the PCA shape space, defined as (28):

$$d_{Mahal,i} = \sum_{m=1}^{p} \frac{\{f_{i,m}\}^2}{Var(\{f_{i,m}\})} \qquad \textit{(WHICH REF?)}$$

With $Var(\{f_{i,m}\})$ being the variance of the shape vector of the m^{th} PCA shape mode. Distributional differences in $d_{Mahal,i}$ between the normal and abnormal group were assessed using non-parametric Mann-Whitney U test, considering $p<0.05$ as statistically significant. Median and interquartile ranges (IQR) are reported. All statistical tests were carried out using IBM SPSS Statistics v22 (SPSS Inc., Chicago, IL).

Results

Twenty-nine patients fulfilled the inclusion criteria of which 20 were suitable for further image processing and were included in the SSM and PCA: the median gestational age at volume collection was 29^{+4} weeks [25^{+0} – 36^{+1} weeks]. The normal group accounted for 13 fetuses (median gestation at inclusion 30^{+0} weeks [25^{+0} – 36^{+1} weeks]), including 12 structurally normal fetuses and one case of coarctation of the aorta and polyhydramnios in a fetus who had negative postnatal genetic testing. Five fetuses considered as abnormal were affected by a genetic or chromosomal abnormality, whereas the remaining abnormal cases consisted in 1 case of unilateral right-sided cleft lip, alveolar ridge and palate, and 1 case of micrognathia (median gestation at inclusion 28^{+4} weeks [26^{+3} – 33^{+0}]). Table 1 summarises patient characteristics and gestational age at volume collection of all cases whose surface meshes served as input for the SSM. From this cohort, anatomical mean face shape was computed (Figure 2).

According to PCA results, m=10 shape modes accounted for over 90% of the shape variability in the study group (Figure 3). Shape mode 1 found differences in overall facial *size* to be the most important contributor to shape variability within the population, accounting for 53.1% of total shape variability (Figure 4). Distributions of shape vector entries for shape mode 1 did not differ on visual assessment between the normal and abnormal cases, implying that abnormal cases could not be determined merely based on facial size differences.

Shape mode 2 accounted for 9.6% of shape variability (Figure 5). Within this shape mode, all but one abnormal case (case 15, Table 1) were clustered together towards low, negative shape vector values, while normal cases were predominantly grouped towards high, positive shape vector values as quantified by median shape vector values of -2.95±115.9 (median±IQR) for the normal versus -56.4±101.6 for the abnormal group.

Looking at the shape mode visualisation as deformation of the mean shape, PCA shape mode 2 described facial shape features relating to differences in proportion between the chin and the forehead (Figure 6). High positive shape vector values corresponded to large, positive deformations of the template towards +3SD, which showed a trend towards an equally prominent forehead and chin region and appeared to be more representative for the “normal” shape. On the other hand, negative values, representing template deformations towards -3SD, showed gradually more prominent disproportion between the chin (increasingly smaller) and the forehead (increasingly protruding), which were common key features of most of the “abnormal” cases.

No such clustering of shape vector entries between the two groups could be found in the remaining 8 PCA modes.

Regarding the abnormality score $d_{Mahal,i}$, abnormal cases generally showed higher distances in PCA shape space as defined by m=10 PCA shape modes (Figure 7), with $d_{Mahal,i}$ being 7.31±5.54 (median±IQR) for the normal group and 13.27±9.82 for the abnormal group ($p=.056$).

Discussion

In this paper we describe quantitative analysis of fetal facial morphology based on Cartesian volumes acquired via 3D ultrasound, and our findings suggest that quantitative differences based on PCA may differentiate between normal and abnormal facial morphology. Antenatal computed evaluation of fetal facial shape is feasible and SSM visually captured and quantified facial shape features that may help to distinguish between normal and abnormal fetuses.

The role of 3D US in the assessment of the fetal face has been previously described (1,6,11) and its advantages over 2D US include the capability to offline reconstruct the acquired volume for further analysis (9). Dysmorphisms have been associated with neurodevelopmental impairment in the context of different genetic syndromes such as Apert, Cornelia de Lange, Costello, Noonan and Goldenhar (4-8) though this is not invariably so, for example in the case of Treacher Collins and Stickler Syndromes – both conditions having well characterised facial appearances. Where fetuses are considered to be at risk they may be studied using 3D US however, this requires considerable skill and experience, and most importantly is dependent on subjective assessment.

A standardised and objective approach using computed assessment of US fetal facial volumes would allow a greater degree of confidence in antenatal diagnosis of genetic syndromes or chromosomal abnormalities. Ideally, such an approach allowing quantitation of the degree of normality or abnormality of the face, or even recognizing specific patterns associated with different syndromes, could ultimately lead to non-invasive diagnosis based on the recognition of associated dysmorphic features. We have not tested this latter hypothesis in this study.

Postnatal experiences in craniofacial imaging have seen computational representations of the human face starting from CT-acquired volumes, which were analysed using geometric

morphometrics (14-16). However, CT is not suitable for prenatal assessment, and previous morphometric techniques were based on landmarking, dependent on individual manual labelling of the biological structures. In former studies on antenatal 3D US volumes, a semi-automated technique, whose acronym was VOCAL ("Virtual Organ Computer-aided AnaLysis"), was described for the estimation of the volume of the gestational sac in early pregnancy. Such obtained volumes could be converted into shape vectors and computed by means of PCA as we have done in our study (LEE, DETER). Differently from our experience, VOCAL technique lead to the rendering of volume estimations volumes, whereas in our setting we aimed to reconstruct actual facial shapes. Our approach, using a non-parametric SSM that does not require anatomical landmarks is independent from subjective face feature evaluation and thus more reproducible.

In our study, the application of unsupervised PCA yielded promising results, which suggest that disproportion in the facial areas in the vertical plane can distinguish normal and abnormal face shapes. Normal cases were more likely to show more symmetric proportions between the forehead and the chin, whereas the abnormal faces showed a wider, prominent forehead and undersized and posteriorly positioned chin. Numerical shape vector results showed a clear clustering of normal and abnormal cases with one outlying case (case 15, Table 1) within the abnormal group. This subject presented with a rounded initial volume, which resulted in a small, round overall face shape with shortened and less prominent forehead, but a larger chin region after performing cutting operations. In fact, this face was so far outlying (i.e. very different from all the other abnormal cases and also different from the normal cases) that we suspect this subject to mark the beginning of another, abnormal population, which simply lacks members due to the small cohort size. This may be due to the fact that we labelled cases as "abnormal", regardless of their specific diagnosis. One could imagine that subgroups within the "abnormal" group might exist that cluster together facial features

associated with one specific defect. Such a group may indeed be placed very far away from both the normal and the abnormal cases contained in our patient cohort.

This was a feasibility study to describe the technique and establish whether it was possible to quantify facial features from prenatal ultrasound. We did not intend to derive an algorithm for differentiating normal from abnormal faces. Nevertheless, despite the small sample size and the fact that our abnormal group was made up by fetuses of different and varied abnormalities, we found a trend towards negative shape vector values compared to normal cases. It would be interesting to test normal versus abnormal fetuses carrying all the same genetic abnormality in a larger cohort of fetuses as we believe that condition specific vector patterns are likely to emerge.

Further, we computed the Mahalanobis distance $d_{Mahal,i}$ in order to define one quantitative "abnormality score" as a distance of one subject face shape in multi-dimensional PCA shape space with regard to the rest of the population. *D*ifferences of borderline statistical significance were found between normal and abnormal cases in terms of $d_{Mahal,i}$, with abnormal cases generally showing higher values, implying greater distance from the normal face shape population. Hence we show that fetal face features derived from 3D US assessment, can indeed be quantified and thus compared between groups. Trends in both shape vectors and $d_{Mahal,i}$ reveal potential of detecting abnormal facial characteristics using these methods.

We are aware of the limitations of 3D US imaging in terms of reproducibility and quality of the imaging. Fetal position, reduced amniotic fluid volume, maternal habitus and unfavourable conditions may negatively affect the collection of the data. However, about two-thirds of the routinely collected volumes in our cohort were suitable for data analysis and would be close to 100% if repeat visits for ultrasound imaging were scheduled. We have not included volumes acquired from fetuses aged less than 24 weeks given that volume acquisition and quality of the multiplanar and

the rendered images may be negatively affected by the smaller size of the fetus and active fetal movements. Furthermore, we do not envisage this technique to be suitable on a routine basis. Indeed, given the wide variation of facial shape features even within normal individuals (e.g. depending upon ethnicity), it is likely that an unrestricted application of SSM and PCA would provide a great amount of false positive results. Clinical use of the technique we have described would be appropriate only in cases where a genetic syndrome is suspected but the actual etiology cannot be diagnosed or suspected by additional tests, including non-targeted CGH array. Realistically, it seems unlikely that such differential diagnosis process would be completed before 24 weeks of gestation.

Another limitation of this technique is the time required for post-processing, currently mostly manually performed. In first instance, adjustment for size prior to processing could allow implementing a strategy of automated alignment and would further standardise the technique. Over the last decade, learning databases capable to automatically reconstruct complex shapes starting from anatomical landmarks have been described (SCHEIRER, GHESU, CARNEIRO, SOFKA). Nevertheless we believe that the use of automated techniques in such settings could lead to inaccurate reconstruction of the actual 3D shape and, ultimately, “normalize” abnormal faces.

Additionally, other landmark points for referencing the cutting plane might perform better. Indeed, shadowing from the alveolar ridge or the nasal septum may impair the labelling of the lower reference point placed on the midline on the internal border of the upper alveolar ridge at the level of the incisive foramen. Realistically, it is difficult to identify reproducible upper referencing points other than the lenses of the eyes, whereas a new landmarking site, placed for example at the external border of the lower alveolar ridge, might be more feasible.

Once processing steps such as segmentation, reconstruction, cutting and alignment become more automated and robust, studies on larger cohorts may provide further basis for antenatal diagnosis of genetic syndromes starting from 3D facial volumes.

Conventional genetic testing with karyotyping has been used in prenatal diagnosis since the 1960s and allows the identification of aneuploidies as well as many structural rearrangements (33) but cannot detect microscopic deletions or duplications. More recently introduced CGH array provides higher resolution and capability to provide more diagnostic information however the clinical significance of abnormal array results may be unknown and false negatives are still possible unless targeted testing is undertaken (2). Another disadvantage of invasive genetic testing is represented by the failure to obtain a result, which is more likely at advanced gestation (34). Computed assessment of the fetal facial morphology could overcome the limitations of currently available tests and represent a link between abnormal facial morphology and genetic testing.

In summary, this is the first study describing a combined approach using 3D US, SSM and PCA for the antenatal assessment of biologic structures and for the quantitative assessment of fetal facial morphology. Further improvement in the automation of the post-processing and testing on larger cohorts of fetuses with specific conditions manifesting in dysmorphisms is required before any practical implementation of such a technique. Nevertheless, results from this study are promising and 3D US-based quantitative assessment of facial shape using SSM and PCA has the potential to contribute to antenatal diagnosis of genetic and chromosomal abnormalities where facial dysmorphism is a feature.

References

1) De Jong-Pleij EA, Ribbert LS, Tromp E, Bilardo CM. Three-dimensional multiplanar ultrasound is a valuable tool in the study of the fetal profile in the second trimester of pregnancy. Ultrasound Obstet Gynecol. 2010 Feb;35(2):195-200.

2) Chen CP SY, Hsu CY, Tsai FJ et al.. Abnormally flat facial profile on two- and three-dimensional ultrasound and array comparative genomic hybridization for the diagnosis of Pallister-Killian syndrome. Taiwan J Obstet Gynecol 2010 Mar;49(1):124-8.

3) Clark DM SI, Deardorff MA, Byrne JL et al.. Identification of a prenatal profile of Cornelia de Lange syndrome (CdLS): a review of 53 CdLS pregnancies. Am J Med Genet A 2012; 158A: 1848–1856.

4) Mahieu-Caputo D SP, Amiel J, Simon I et al.. Prenatal diagnosis of sporadic Apert syndrome: a sequential diagnostic approach combining three-dimensional computed tomography and molecular biology. Fetal Diagn Ther 2001 Jan-Feb;16(1):10-2.

5) Thellier E, Levaillant JM, Roume J, Quarello E, Bault JP. Cornelia de Lange syndrome: Specific features for a prenatal diagnosis. Ultrasound Obstet Gynecol 2015 Oct 20.

6) David AL, Turnbull C, Scott R et al.. Diagnosis of Apert syndrome in the second-trimester using 2D and 3D ultrasound. Prenatal Diagnosis. 2007;27(7):629-32.

7) Levaillant J-M, Gérard-Blanluet M, Holder-Espinasse M et al.. Prenatal phenotypic overlap of Costello syndrome and severe Noonan syndrome by tri-dimensional ultrasonography. Prenatal Diagnosis. 2006;26(4):340-4.

8) Guzelmansur I CG, Ceylaner S, Ceylan N, Daplan T. Prenatal diagnosis of Goldenhar syndrome with unusual features by 3D ultrasonography. Genet Couns 2013;24(3):319-25.

9) Conner SN, Longman RE, Cahill AG. The role of ultrasound in the diagnosis of fetal genetic syndromes. Best Pract Res Clin Obstet Gynaecol 2014 Apr;28(3):417-28.

10) Rubesova E, Barth RA. Advances in fetal imaging. Am J Perinatol. 2014 Aug;31(7):567-76.

11) Picone O, Levaillant JM, Hirt R, Frydman R, Boulvain M, Senat MV. Correlation between referral ultrasound with suspected foetal anomalies and autopsy examination in two prenatal diagnosis centres. Impact of the routine use of 3D/4D scan. Prenat Diagn. 2008 Mar;28(3):191-6.

12) Ponniah AJT, Witherow H, Evans R, Dunaway D, Richards R, Ruff C. Planning reconstruction for facial asymmetry. Int J Simulation 2006;7(5):32-9.

13) Pluijmers BI, Ponniah AJ, Ruff C, Dunaway D. Using principal component analysis to describe the Apert skull deformity and simulate its correction. J Plast Reconstr Aesthet Surg. 2012 Dec;65(12):1750-2.

14) Crombag GA, Verdoorn MH, Nikkhah D, Ponniah AJ, Ruff C, Dunaway D. Assessing the corrective effects of facial bipartition distraction in Apert syndrome using geometric morphometrics. J Plast Reconstr Aesthet Surg. 2014 Jun;67(6):e151-61.

15) Nikkhah D, Ponniah A, Ruff C, Dunaway D. Planning surgical reconstruction in Treacher-Collins syndrome using virtual simulation. Plast Reconstr Surg. 2013 Nov;132(5):790e-805e.

16) Staal FC, Ponniah AJ, Angullia F, Ruff C, Koudstaal MJ, Dunaway D. Describing Crouzon and Pfeiffer syndrome based on principal component analysis. J Craniomaxillofac Surg. 2015 May;43(4):528-36.

17) Cootes T, Hill A, Taylor C, Haslam J. Use of active shape models for locating structures in medical images. Image Vis. Comput. 1994;12:355–65.

18) Young AA, Frangi AF. Computational cardiac atlases: from patient to population and back. Exp. Physiol. 2009;94:578–96.

19) Hassan WA, Lees CC. Facial cleft detected: is the palate normal? Best Pract Res Clin Obstet Gynaecol. 2014 Apr;28(3):379-89.

20) Offerdal K, Blaas HG, Eik-Nes SH. Prenatal ultrasound detection of facial clefts: a prospective study of 49,314 deliveries in a non- selected population in Norway. Ultrasound Obstet Gynecol 2008;31:639–46.

21) Heimann T, Meinzer H-P. Statistical shape models for 3D medical image segmentation: A review. Med. Image Anal. 2009;13:543–63.

22) Cootes TF, Taylor CJ. “Statistical Models of Appearance for Medical Image Analysis and Computer Vision,” in In Proc. SPIE Medical Imaging, 2001, pp. 236–248.

23) Bruse JL, McLeod K, Biglino G et al. A statistical shape modelling framework to extract 3D shape biomarkers from medical imaging data: assessing arch morphology of repaired coarctation of the aorta. BMC Med. Imaging. 2016;16:40.

24) Durrleman S, Prastawa M, Charon N et al.. Morphometry of anatomical shape complexes with dense deformations and sparse parameters. NeuroImage. 2014;101:35–49.

25) Durrleman S, Pennec X, Trouvé A, Ayache N. Statistical models of sets of curves and surfaces based on currents. Med. Image Anal. 2009;13:793–808.

26) Leonardi B, Taylor AM, Mansi T et al.. Computational modelling of the right ventricle in repaired tetralogy of Fallot: can it provide insight into patient treatment? Eur. Heart J. Cardiovasc. Imaging. 2013;14:381–6.

27) Vaillant M, Glaunès J. Surface Matching via Currents. In Information Processing in Medical Imaging. Edited by Christensen GE, Sonka M. Springer Berlin Heidelberg; 2005:381–392.

28) Jolliffe, I.T. Principal Component Analysis. 2nd ed. Springer-Verlag New York, Inc.; 2002.

29) Mansi T, Durrleman S, Bernhardt B et al.. A Statistical Model of Right Ventricle in Tetralogy of Fallot for Prediction of Remodelling and Therapy Planning. Med Image Comput Comput Assist Interv. 2009;12(Pt 1):214-21.

30) Mansi T, Voigt I, Leonardi B et al.. A statistical model for quantification and prediction of cardiac remodelling: application totetralogy of Fallot. IEEE Trans Med Imaging. 2011 Sep;30(9):1605-16.

31) Ahrens J, Geveci B, Law C. 36 ParaView: An End-User Tool for Large-Data Visualization. Vis. Handb. 2005;717.

32) Bruse JL, McLeod K, Biglino G et al.. A Non-parametric Statistical Shape Model for Assessment of the Surgically Repaired Aortic Arch in Coarctation of the Aorta: How Normal is Abnormal? O Camara Al Eds Stat. Atlases Comput. Models Heart 2015. Munich: Springer International Publishing Switzerland 2016; 2016. p. 21–9.

33) Mansi T, Durrleman S, Bernhardt B, et al. A statistical model of right ventricle in tetralogy of Fallot for prediction of remodelling and therapy planning. In: Yang G-Z, Hawkes D, Rueckert D, Noble A, Taylor C, eds. Medical image computing and computer-assisted intervention—MICCAI 2009 (lecture notes in computer science). Berlin, Heidelberg: Springer; 2009:214-21.

34) Bruse JL, Cervi E, McLeod K, et al. Looks do matter! Aortic arch shape following hypoplastic left heart syndrome palliation correlates with cavopulmonary outcomes. Ann Thorac Surg 2017;103:645-54.

35) Bruse JL, Khushnood A, McLeod K, et al. How successful is successful? Aortic arch shape following successful aortic coarctation repair correlates with left ventricular function. J Thorac Cardiovasc Surg 2017;153:418-27.

36) De Maesschalck R, Jouan-Rimbaud D, Massart DL. The Mahalanobis distance. Chemom Intell Lab Syst 2000;50:1-18.

37) Elfadaly FG, Garthwaite PH, Crawford JR. On point estimation of the abnormality of a Mahalanobis index. Comput Stat Data Anal 2016;99:115-30.

38) Lee W, Deter RL, McNie B, et al. Quantitative and morphological assessment of early gestational sacs using three-dimensional ultrasonography. Ultrasound Obstet Gynecol 2006;28: 255-60.

39) Deter RL, Li J, Lee W, Liu S, Romero R. Quantitative assessment of gestational sac shape: the gestational sac shape score. Ultrasound Obstet Gynecol 2007;29:574-82.

40) Carneiro G, Amat F, Georgescu B, Good S, Comaniciu D. Semantic-based indexing of fetal anatomies from 3-D ultrasound data using global/semi-local context and sequential sampling. In Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2008.

41) Sofka M, Zhang J, Zhou SK, Comaniciu D. Multiple object detection by sequential Monte Carlo and hierarchical detection network. In Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2010.

42) Scheirer WJ, de Rezende Rocha A, Sapkota A, Boult TE. Toward open set recognition. IEEE Trans Pattern Anal Mach Intell 2013;35:1757-72.

43) Ghesu FC, Krubasik E, Georgescu B, et al. marginal space deep learning: efficient architecture for volumetric image parsing. IEEE Trans Med Imaging 2016;35:1217-28.

44) Kirchhoff MRH, Lundsteen C. High resolution comparative genomic hybridisation in clinical cytogenetics. J Med Genet 2001;38:740-4.

45) Lawin O'Brien A, Dall'Asta A, Tapon D, et al. Gestation related karyotype, QF-PCR and CGHarray failure rates in diagnostic amniocentesis. Prenat Diagn 2016;36:708-13.

Figures

Figure 1 – Overview of processing chain from reconstructed 3D US volumes over the manual registration (alignment via defined landmarks) and cutting, mean shape (blue) and transformation (ϕ) computation towards individual cases, principal component analysis (PCA) and final visualization and quantitative assessment of shape modelling results. *Shape modes* thereby describe key 3D shape features contributing to the overall facial shape variability, while *shape vectors* are their numerical equivalents, describing shape features associated with a certain shape mode with respect to an individual's face shape.

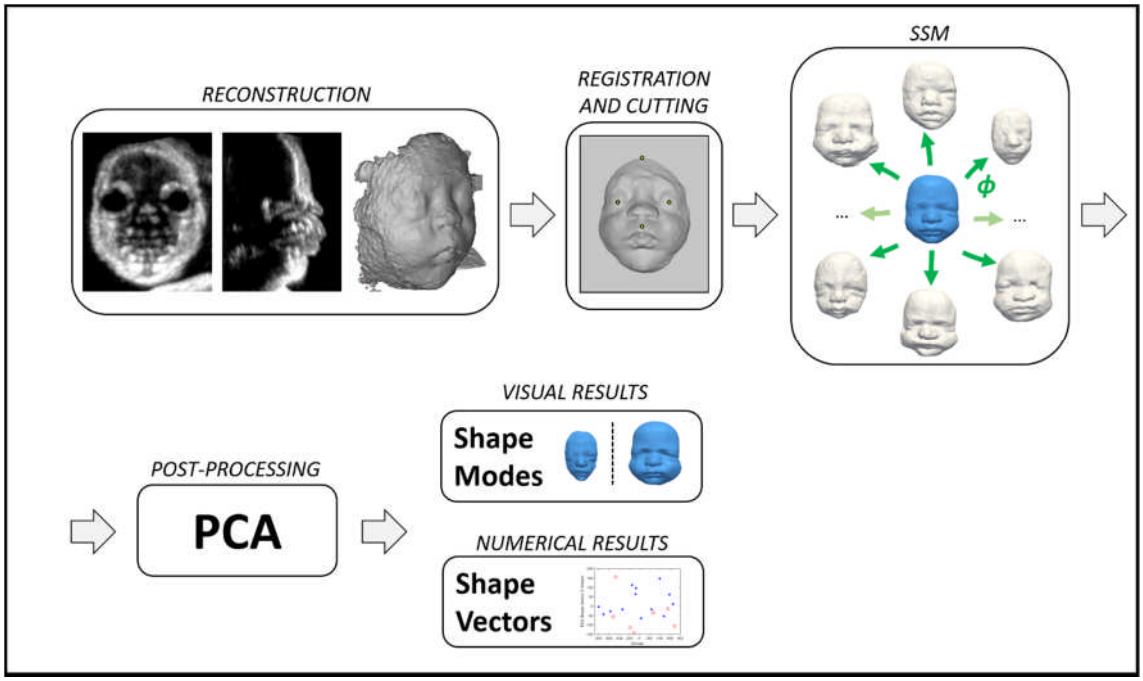


Figure 2 –Computed mean (i.e. *template*) fetal face shape, based on all 20 faces reconstructed from 3D US image data from different viewing angles. The 3D average face of normal and abnormal cases shows a relatively symmetric and even face shape without any overly prominent regions. By deforming the mean shape towards each individual's face shape, all 3D facial features are parameterised coherently in one common mathematical framework.

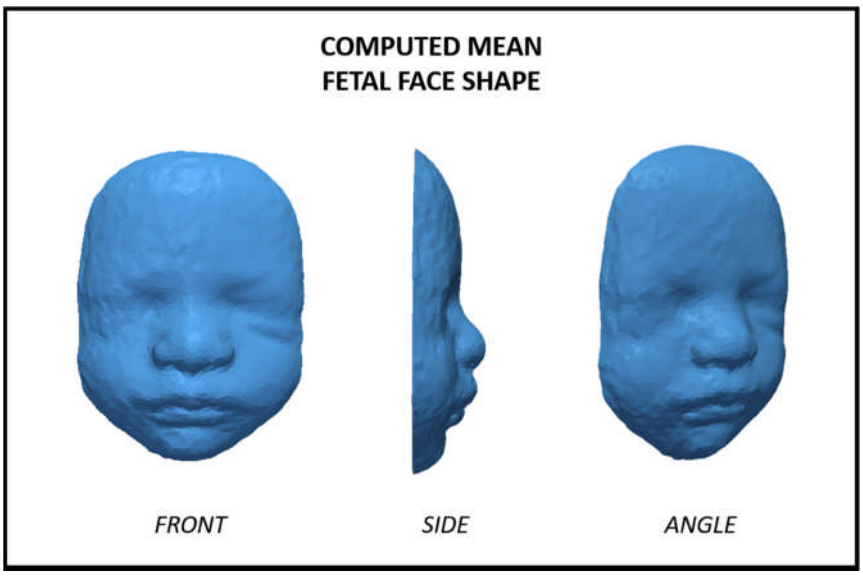


Figure 3 – Decomposition of the shape features and shape modes in the included population according to Principal Component Analysis (PCA). The first shape mode accounted for 53.1% of total shape variability; the second mode accounted for 9.6%. Ten modes accounted for over 90% of shape variability.

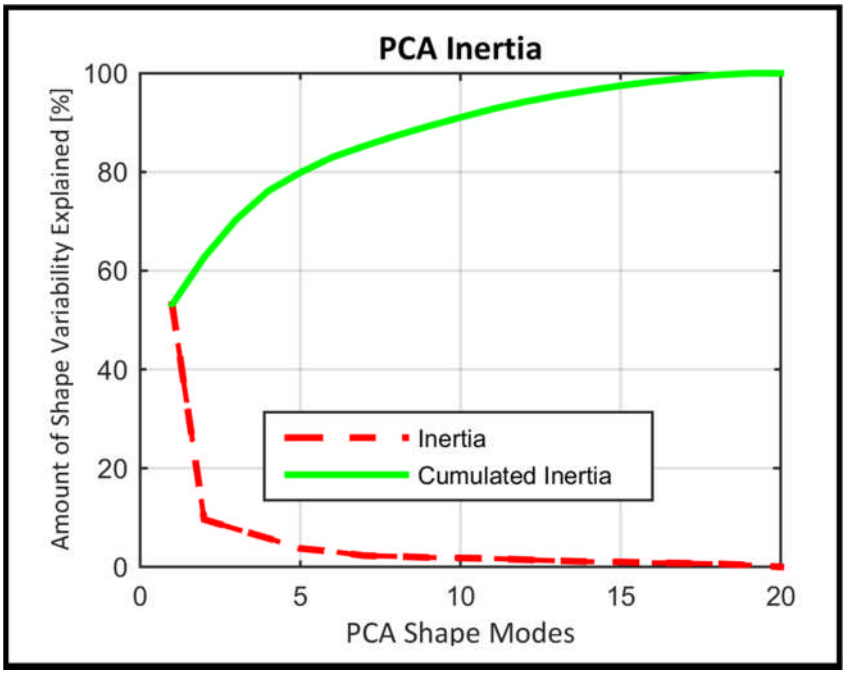


Figure 4 – PCA shape mode 1 (a), visualized as deformation of the computed mean face shape along the mode from -3 to +3 standard deviations (SD). Shape mode 1 thereby accounted mainly for size differences between the included subjects. Respective subject-specific shape vector entries are shown in (b), with large positive values relating to face features going towards +3SD and low negative values going towards -3SD as shown in (a) and indicated at the y-axis. Both groups showed similar distributions of shape vector entries, implying no substantial facial size differences between the two groups.

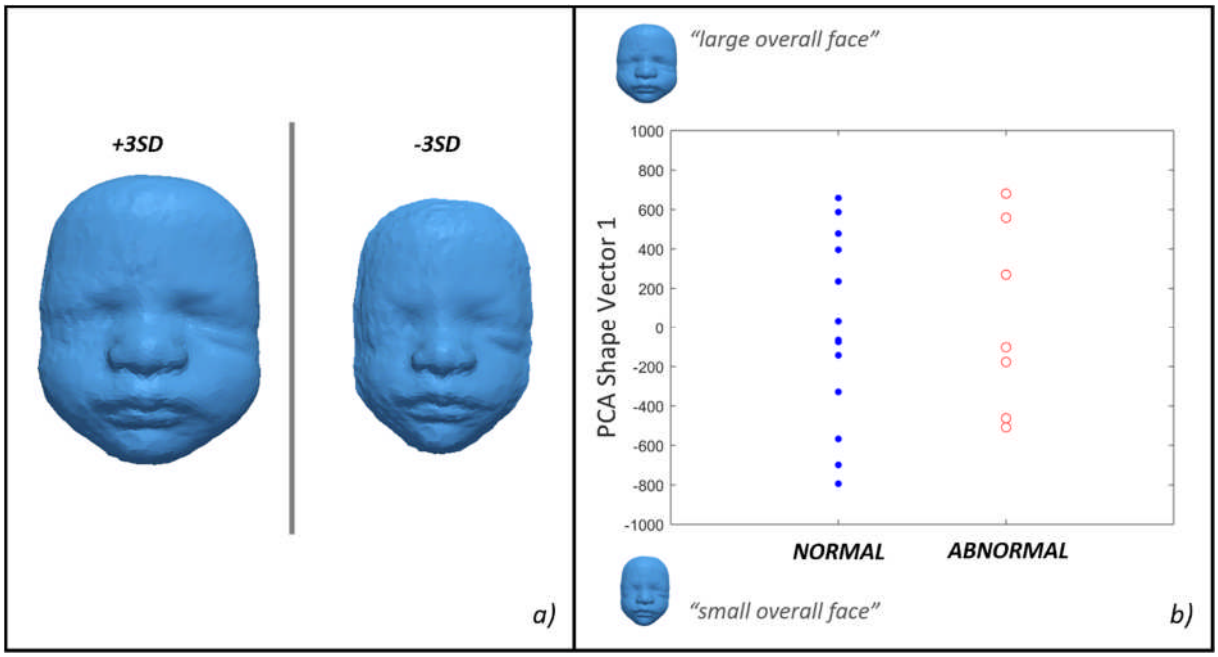


Figure 5 – Shape vector distributions of PCA shape mode 2 in normal (blue, filled circles) and abnormal (red, hollow circles) cases showing a clear clustering of normal cases towards larger, positive shape vector values and of abnormal cases towards lower, negative shape vector values. Blue face shapes indicate the extremes of PCA shape mode 2, thus showing face features associated with positive and negative shape vector values, respectively. One face shape within the abnormal group was clearly outlying (case number 15) as marked and shown.

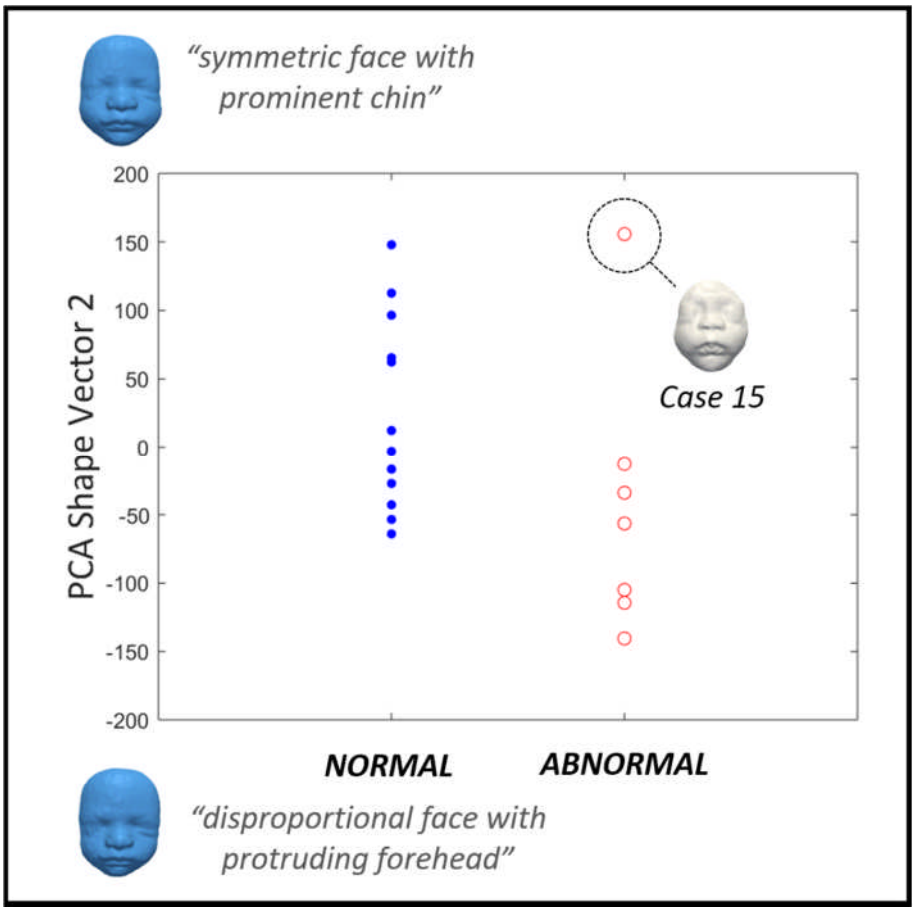


Figure 6 – Three-dimensional visualization of PCA shape mode 2 from different viewing angles (a), deforming the computed mean face shape towards shape features describing a more symmetric and even overall face shape for normal cases (associated with higher, positive shape vector values towards +3SD, Figure 5) and disproportion between the chin and a wider, protruding forehead for the abnormal group (associated with lower, negative values towards -3SD). Surface distance maps (b) quantifying local surface distances between the mean shape and the extremes of shape mode 2 as shown in (a) clarify, which facial regions are mostly affected by the shape mode 2 deformations. Red, negative surface distances thereby describe outward deformation from the mean shape, confirming protrusion and prominence of the forehead for abnormal cases.

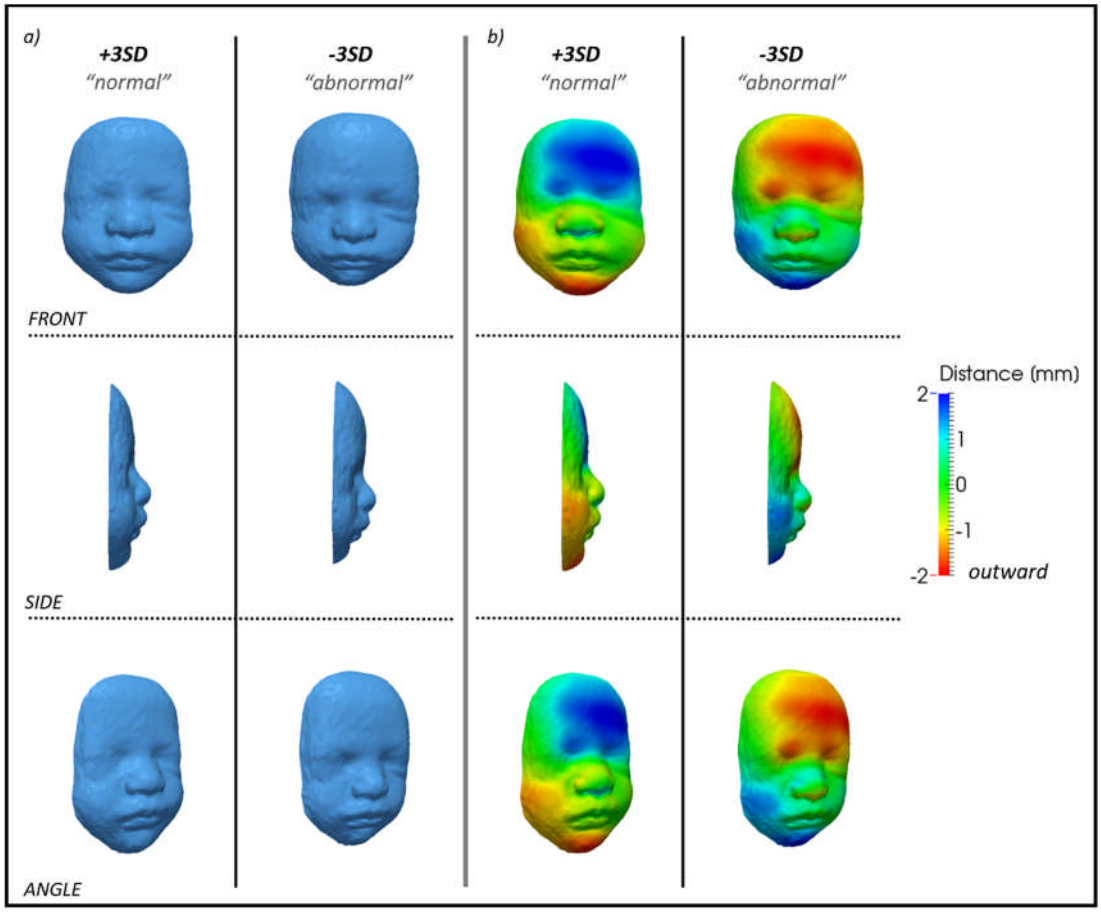
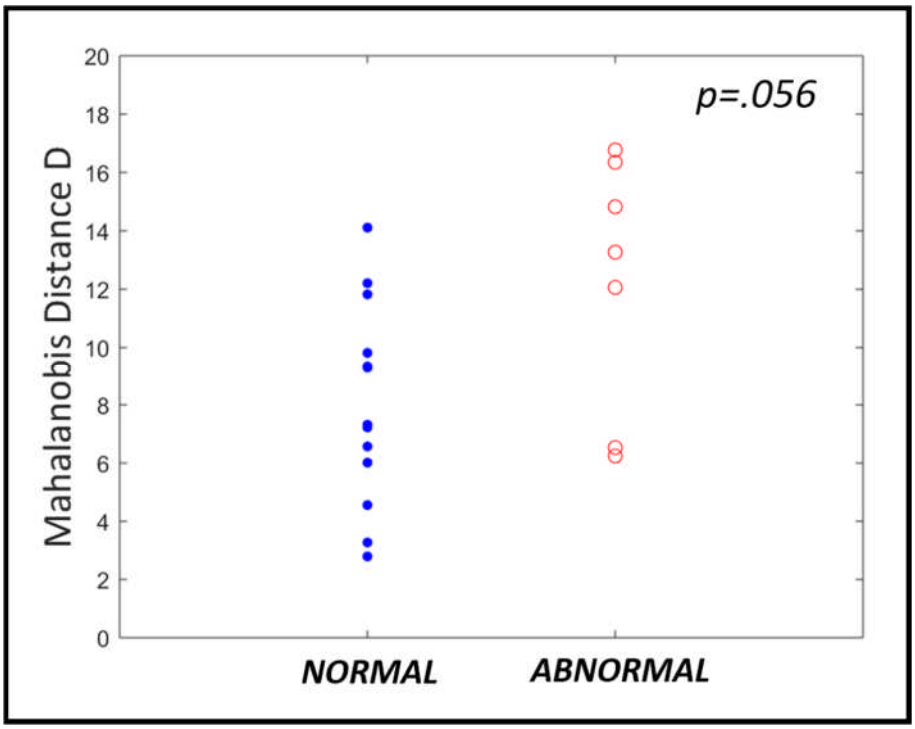


Figure 7 – Mahalanobis distance was computed in order to quantify how distant each subject 3D face shape is from the entire shape population. Of borderline statistical significance, median distance for abnormal cases was higher, implying higher facial abnormality.



Tables

Table 1 – Gestational ages at time of 3D US scan and outcome diagnosis of the included cases.

Fetus	**Gestation [weeks^{+days}]**	**Outcome**
1	25^{+0}	Normal
2	26^{+6}	Normal
3	27^{+1}	Normal
4	27^{+1}	Normal
5	28^{+4}	Normal
6	29^{+0}	Normal
7	30^{+0}	Normal
8	31^{+1}	Normal
9	32^{+5}	Normal
10	33^{+0}	Normal
11	34^{+0}	Normal
12	35^{+0}	Normal
13	36^{+1}	Normal – Coarctation of the aorta, polyhydramnios
14	26^{+3}	Abnormal (unilateral right-sided cleft lip, alveolar ridge and palate)
15	27^{+2}	Abnormal (consanguineous multiple malformations)
16	28^{+0}	Abnormal (consanguineous multiple malformations)
17	29^{+1}	Abnormal - Micrognathia
18	30^{+0}	Abnormal - Noonan Syndrome
19	33^{+0}	Abnormal - Trisomy 18
20	33^{+0}	Abnormal - Trisomy 21