Study on Facial Asymmetry in Children Under 16 years from the North West of England using Three-dimensional Images (3dMD)

Thesis submitted in accordance with the requirements of the

University of Liverpool for the degree of

Doctor of Dental Science by

Orla Carty

Submitted: July 2018

Contents

Acknowledgements

I would like to extend my gratitude to my supervisors Dr Flannigan, Dr Dominguez-Gonzalez and Dr Burnside for their input and advice regarding the planning, development and completion of this research. I am also sincerely grateful to Mr C Duncan who kindly agreed to provide the images for this project, and to Emer, my colleague and friend, for her part in the completion of the reliability. I would like to thank John and all of the Medical Illustration Team at Alder Hey Hospital for allowing us to complete data collection within the department and for their kindness and assistance during the duration of this work.

I would like to thank Ross for spending endless hours discussing my research. Thank you for all of your help, advice, love and patience, all of which are greatly appreciated.

Finally, I would like to thank my parents for the many years of encouragement during my school, undergraduate and postgraduate years. Thank you for instilling in me the hard-working attitude which has got me through every step of the way. Their love and support has made all of this possible, for which I will be forever grateful.

Abstract

Study on facial asymmetry in children under 16 years from the North West of England using three-dimensional images (3dMD)

O. Carty, N.F. Flannigan, S. Dominguez-Gonzalez and G. Burnside (School of Dentistry, The University of Liverpool.

Objectives: The primary objective was to determine the extent of facial asymmetry in a standard population of children from the North West of England using a landmark based approach on 3-dimensional (3d) images. The secondary objectives were to investigate the direction and severity of nose and chin deviation and the correlation between them, to determine the relationship between gender and facial asymmetry and to investigate if asymmetry differs with age.

Design/Setting: A retrospective cross-sectional cohort study undertaken at Alder Hey Children's Hospital.

Subjects: Participants were children of either gender and under 16 years of age (mean 7.36 years) with all types of skeletal relationships who volunteered to have 3d images (using 3dMD software) of their head and face captured for research purposes.

Methods: Images of 145 children were available. Following the application of exclusion criteria and a panel assessment including 2 Orthodontic Registrars, 2 Orthodontic Consultants and a Medical Illustration Technician, to determine adequate quality, 107 images were included in the study. Reference frame analysis was completed to orientate the images in a standardised manner. A landmark based approach was used by a single examiner positioning 8 mid and 7 bilateral facial landmarks and the analysis was applied using Vultus software (version 2.5.0.1). Intra and inter-reliability of landmark positioning were completed prior to data collection.

Results: Fifty-seven males and 50 females were included with a mean age of 7.36 years (SD 3.74). The majority of the sample was White British (82.2%). The median asymmetry indices (AI) for midfacial landmarks were all <1.10mm (25^{th} and 75^{th} IQR ranged from 0.24 to 1.75) and for bilateral facial landmarks were all <2.60mm (25^{th} and 75^{th} IQR ranged from 0.63 to 3.65). Pogonion was the most asymmetric midfacial landmark (median AI 1.08mm, IQR 0.56, 1.75) and Cheilion was the most asymmetric bilateral facial landmark (median AI 2.56mm, IQR 1.69, 3.65). A statistically significant relationship between the side of nose and chin deviation (Chi-squared p=0.00) was detected and also a significant correlation between the severity of nose and chin deviation (Pearson's correlation 0.91). There was no significant difference in the asymmetry detected between males and females when the Mann-Whitney U test was applied. There was an association between increasing asymmetry and age for landmarks: Pronasale, Subnasale, Stomion, Pogonion and Christa philtri however this is

potentially influenced by the size of the patient's head and thus should be interpreted with caution.

Conclusions: In the standard population of children from the North West of England an element of facial asymmetry is present. The direction and magnitude of nose and chin deviation are highly correlated. Males and females exhibit equal levels of facial asymmetry.

Chapter 1: Introduction

Symmetry may be considered simply to be a balanced face, whereas asymmetry describes an imbalance.[1] Research suggests that asymmetry to a certain degree can be present in aesthetically pleasing faces,[2] and that facial asymmetries, although maybe not immediately obvious, are in-fact quite common.[3] In the United Kingdom (UK) it's estimated that 1 in 100 people have a significant facial defect, which may or may not result in increased facial asymmetry.[4] Individuals with facial deformities can present with facial asymmetry due to discrepancies in the in the size, morphology, or relationship of the two sides of the face.[5]

To decide upon a point where 'normal' asymmetry becomes 'abnormal' is exceptionally difficult, with asymmetry being viewed as a range rather than a specific number or ratio. This attempt to categorise our patients is influenced by the individuals perception of their own facial imbalances, but also by the clinicians sense of balance and the amount they feel the patient varies from that seen in the standard population.[5] It is thought that as the cranium is approached there is less apparent asymmetry and more dimensional stability.[6] Greater asymmetry is thought to be tolerated in the lower face whilst maintaining a 'normal' facial appearance,[7] but this finding is not exclusively reflected throughout the literature.[3] There have also been reports that the right side of the face and neurocranium tend to be larger than the left, with the left side perhaps being more attractive.[3] Gender and age although greatly investigated appear to have no significant influence on the extent of facial asymmetry.

In an ideal world, clinicians in addition to their own experiences and perceptions, would have an objective means of assessing facial asymmetry to allow them to define what constitutes 'normal' asymmetry. This would allow for definitive standards to be outlined regarding what level of deviation may be acceptable in the facial soft tissues before considering orthodontic, orthognathic, plastic or joint treatment approaches.[7] The purpose of this research project is to analyse and quantify the degree of soft tissue facial asymmetry in a standard population of children (<16 years) in the North West of England using 3-dimensional (3d) images. It's hoped that this data may be used in future to help provide a reference population to compare patients who complain of facial imbalances and to decide whether they are within or outside the 'normal' limits of asymmetry in order to establish if intervention is indicated. This would also assist in the decisions regarding which patients are in most need for treatment particularly from the perspective of public service provision. The results of this research are also planned to use as a control group to compare with children who have undergone corrective craniofacial surgery to help provide information on likely outcomes following surgical interventions relative to deviation from the appearance of the general population.

Chapter 2: Literature Review

2.1: Symmetry

2.1.1: Background

A symmetrical face is considered to be precisely balanced, with regards to corresponding size, shape and arrangement of structures on either side of a mid-sagittal plane.[6] Symmetry has been considered a sign of better health and superior genetic make-up, and it appears that there is an intimate relationship between symmetry and attractiveness.[8] Perfect bilateral symmetry is more of a theoretical concept which is actually a rare occurrence in nature[5], with an element of mild asymmetry thought to occur in all individuals.[9,10] Humans have an ability to form judgements almost immediately in relation to facial attractiveness.[11] However, what is considered 'attractive' varies between individuals and is possibly influenced by gender, age, ethnicity and culture.[8] This is a complicated area to research as the literature suggests that a certain degree of asymmetry can be present in aesthetically pleasing faces, and that facial asymmetries,[2] although perhaps not immediately obvious, are in-fact quite common.[3] Certain minor facial asymmetries are even considered to improve attractiveness, an example of this is navei which are considered to be a trait of facial beauty.[12] Interestingly, computer generated perfectly symmetrical faces (hemi-facial duplication) have not been shown to be 'maximally attractive' when compared to natural faces with 'normal' levels of asymmetry present.[8,13,14] Therefore, a mild degree of asymmetry may be desirable. This is demonstrated in Figure 1 which is taken from Wang et al. (2017).[8]

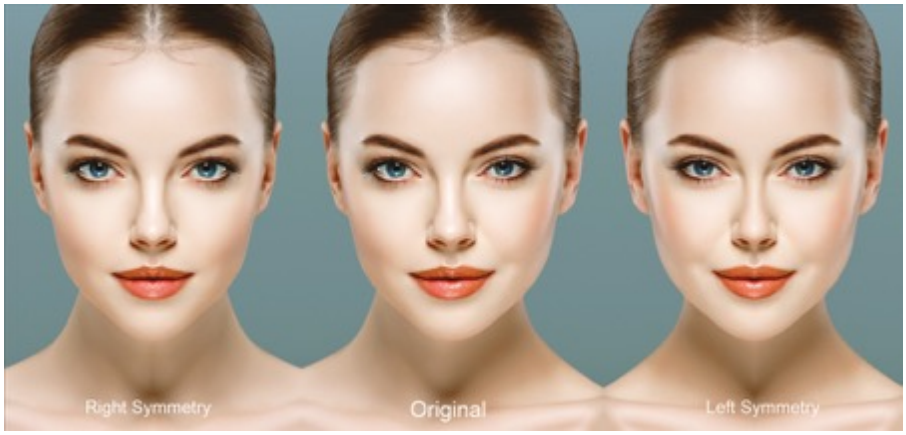


Figure 1: Taken from Wang et al. (2017).[8] Using hemi-facial duplication the left image represents perfect right sided symmetry, the middle image is the original (slightly asymmetrical image with different left and right sides), and the right image represents perfect left sided symmetry.

2.1.2: Symmetry in Different Regions of the Face

Right sided facial dominance has been reported, with the right hemi-face usually being larger than the left in both genders.[8] Reflecting on the portraits by master painters including Rembrandt, da Vinci and Vermeer, it can be concluded that the majority of these portraits depict the left side of the face (Figure 2), which is the side that has been cited to be more aesthetically pleasing. To test this hypothesis, Blackburn et al. (2012)[15] assessed images of faces in their original and mirror forms using pupil diameter confirming the hypothesis that the left side is perceived as more aesthetically pleasing. It is unclear if discrepancies on the left side are more sensitive to perception as the research is inconclusive. If this were true, then correction of left sided deformities could potentially be prioritised over those affecting the right side.

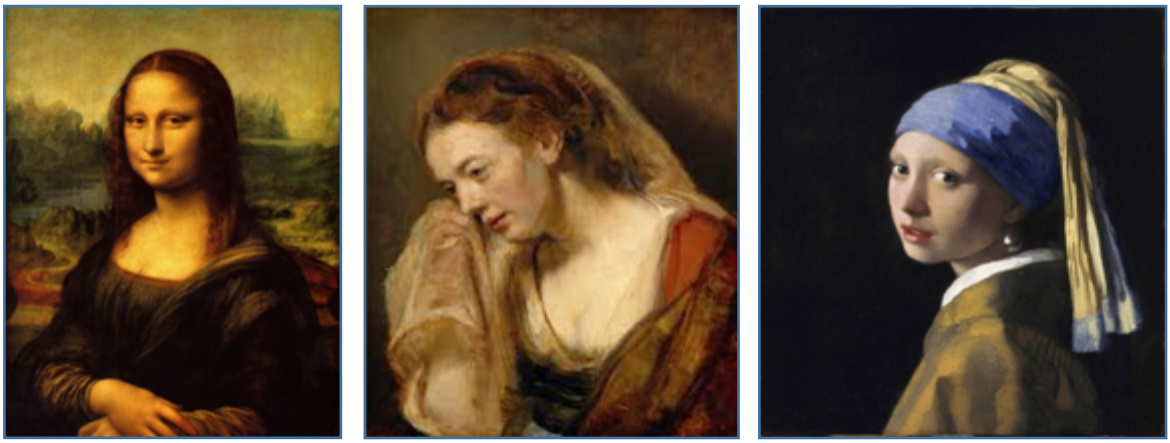


Figure 2: Illustrating the favoured left sided bias in distinguished portraits including: a) 'Mona Lisa' by Leonardo Da Vinci (early 1500's)[16], b) 'A weeping woman' by Rembrandt van Rijn (1640's)[17], and c) 'Girl with a pearl earring' by Johannes Vermeer (1670's).[18]

The occurrence of asymmetry is reported by many to increase in the lower part of the face, as you progressively move downward, away from the cranium.[6] However, there are some standard population studies, such as that of a Northern American population by Farkas and Cheung (1981)[3], which have reported asymmetries to occur most commonly in the upper third of the face using a direct anthropometry technique. In their study of 154 boys and 154 girls (each aged 6, 12 or 18 years) the average difference between the right and left sides of the face was 3mm. Facial asymmetry is often thought to be more noticeable the closer to the midline it occurs with asymmetries closer to the midface evaluated more negatively.[11]
Gender and age appear not to significantly influence on the extent of facial asymmetry. This is supported by several studies using measurements of children from of the 'standard population' in places including Canada, Finland and the United Kingdom.[2,3,19]

2.1.3: The Perception of Asymmetry

An important question to consider is at what extent does deviation from symmetry become noticeable, and furthermore, at what level is it significant enough to consider an intervention? It's not simply the presence, or absence of detected asymmetry, our perception is influenced by multiple factors including location on the face, extent, and observer. It has been reported that asymmetry detected by clinicians and lay people differ, which clinicians including orthodontists detecting relatively small discrepancies. McAvinchey et al. (2014)[9] found that the observer type influenced the identification and perception of the severity of facial asymmetry. This study had 4 groups initially: lay people, dental care professionals, dental undergraduates and orthodontists. They identified significant differences with orthodontists having the least tolerance for asymmetry and lay people the most, so it was data from these 2 groups which the discussion focused upon. The methodology included deviating the chin from 0-20mm in 2mm increments to both right and left sides using computer software on a 3d lifelike image of 2 average faces (female and male). Each image was viewed for 14 seconds with a 2 second break between images, arranged on a timed presentation in a random order. Participants in each group were asked to categorise the facial appearance into a) normal b) slightly abnormal but socially acceptable and c) abnormal appearance which merits correction. For lay people, the perceived deviation of the chin to be within 'normal' limits was a mean of 5.6mm (standard deviation (SD) ± 2.7), whereas for the orthodontist group it was 3.6mm (SD ±1.5). The orthodontist group had the lowest tolerance level for deciding the patient would benefit from surgical correction with chin deviations of 9.7mm (SD ± 3.0) being placed in this category. Lay people on the other hand had a higher threshold of 11.8mm (SD ± 4.0). The finding that orthodontists are more critical of facial discrepancies is supported by previous research in this area.[20] Interestingly the gender of the image, gender of observer, or direction of asymmetry were not found to influence the perception of asymmetry in the area of the chin in that particular study.[9] Naini et al. (2012)[20] reported chin deviations of less than 5mm to be acceptable, with deviations of 10mm more likely to favour a surgical approach, with no significant impact of observer gender. Certainly, for all observers, it seems that as the extent of the deviation increases so too does its' detection.[9,20] As the level of asymmetry increases, the judgement that intervention is appropriate to improve it also increases.[8] More severe asymmetries often make it easier to decide upon a surgical intervention, whereas borderline cases present a challenging decision. The assessment of asymmetry is often subjective, particularly as there were no clearly defined predetermined thresholds of when normality changes to deformity, prior to the systematic review provided by Wang et al. (2017).[8]

2.1.4: Acceptable Limits of Asymmetry

Wang et al. (2017)[8] have combined the available evidence on the perception of facial asymmetry, and the thresholds which may be acceptable in a standard population. Orthognathic and cephalometric studies, papers which were not published in English and those that didn't report on perception were excluded. They divided the face into the following subunits: orbital, nasal, oral commissure and chin regions and identified the areas which are most and least tolerant of asymmetries. As an individual may have several areas affected (e.g. pan facial asymmetry) it would be useful to be able to plan future corrective procedures around the most negatively affected areas (deviating most from the standard population) in order to maximise the treatment outcomes for the patient and minimise the interventions required. The aim after correction would be for the asymmetry to be less perceptible to the lay person, bringing them as close to the normally acceptable limits as possible.[8]

2.1.4.1: Orbital Region

A discrepancy between left and right sides of 3mm elevation in the eyebrow area was deemed acceptable within normal limits in a study of 2-dimensional (2d) clinically altered images of a female by Hohman et al. (2014).[21] They tested asymmetries in this area varying from 0-6mm (in 1mm increments) using an online survey of physicians and lay people. Asymmetries between 3-6mm were noted by between 92-100% of the observers. Chu et al. (2011)[22] using data from 30 lay people concluded that the discernment of eyebrow asymmetry in a digitally altered 2d image of a male model changes between 3mm and 3.5mm when viewed for 10 seconds. A 3mm discrepancy was detected by 10% of lay people (3/30), whereas a 3.5mm discrepancy was detected by 73% (22/30) ($p=<0.001$) which was statistically significant. The face was considered to require surgery when the eyebrow discrepancy reached 4mm (50% of participants).[22]

Asymmetry of the eyelid appears to be readily perceived and is thought to be the most sensitive area to static facial asymmetry perception.[8] There is a sharp increase between the detection of asymmetry of eyelids by lay people at 1mm (10%) and at 2mm (85%).[21] This follows the hypothesis that once asymmetry becomes detectable, perception increases exponentially.

2.1.4.2: Nasal Region

Perceptions of nasal asymmetry have been researched in terms of millimetres[23,24] and degrees[25] relative to a mid-sagittal plane. Four millimetres has previously been reported as the visual perception threshold for nose deviation using digitally manipulated photographs.[24] Meyer-Marcotty et al. (2011)[23] used 3 groups of observers (lay people, orthodontists and maxillofacial surgeons) to analyse this. A 4mm displacement of the nasal tip was significantly

more detectable than 2mm displacement ($p<0.001$), with left sided nasal deviations assigned a more negative rating when compared to the right. This is in contrast to their findings that chin deviations to the right side are more negatively perceived. When comparing nose and chin asymmetries, nose asymmetries were viewed as more prominent and less acceptable than the same level of deviation of the chin. This is probably not surprising considering the nose is located centrally in the face and the nasal bridge corresponds with the vertical midfacial axis.[23] Lay people were reported as being quite good at detecting facial asymmetries which is contrary to the findings of other research studies in this area.[23]

Kwak et al. (2015)[25] similarly investigated various levels of nasal deviations by constructing asymmetries of 0-5° relative to the mid-sagittal plane. These were arranged in 1° increments using simulated 2d photographs. The observers were 4 groups (total n=160) composed of lay people, dental students, general dentists and orthodontists. A deviation of 2.92° was identified as the threshold for recognition of nasal asymmetry. However, the study design was not ideal as images were of a female only, were printed and displayed in order of increasing asymmetry and there was no time limit stated for each image to be observed.

2.1.4.3: Oral Commissure Region

In terms of static position of the oral commissure, Chu et al. (2011)[22] found that 73% of observers detected an asymmetry of 3mm ($p<0.001$) on a male image, with a discrepancy >5mm perceived to require intervention.[22] This threshold of 3mm in the oral commissure region is supported by other research studies which used 2d photographs of a female model.[21]

2.1.4.4: Chin Region

Silva et al. (2013)[24] found no difference between perceived attractiveness when comparing an original 'normal' 2d photograph with a digitally manipulated photograph which had up to 6mm chin deviation to the left side. It's possible that 6mm therefore is minimally detectable in this region. This is in contrast to research using 2d models by Naini et al. (2012)[20] who suggested the threshold for chin deviation to influence attractiveness to be 5mm ($p<0.001$), with 10mm deviation judged to require an intervention ($p<0.001$). In a study using 3d models Meyer-Marcotty et al. (2011)[23] reported the perceptive threshold for chin asymmetry to be 6mm ($p<0.001$), with right side deviations being more negatively perceived. The images used in this study were more 'real life' than the 2d models used in other studies. Figures 3 and 4 provide an example illustrating the difference in appearance between these types of images.

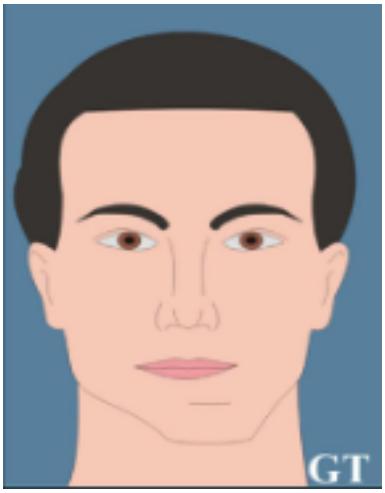


Figure 3 (Left side): Is an example of a 3d image taken from Meyer-Marcotty et al. (2011)[23] demonstrating 6mm left side chin deviation.

Figure 4 (Right side): Is an example of a 2d model taken from Naini et al. (2012)[20] demonstrating 5mm left horizontal asymmetry of the mandible/chin.

2.2: Anthropometry

2.2.1: Introduction

The word anthropometry originates from the Greek language, *Anthropos* meaning 'human' and *metron* a 'measure'. This is an area of science involving the measurement of human morphology in terms of size, shape, weight and proportions.[26]

Professor Leslie G. Farkas is well recognised as the father of modern craniofacial anthropometry who dedicated his career to researching and developing craniofacial anthropometry techniques.[27] He recognised the importance of having 'normative' data sets for various populations and parts of the world encouraging those in other countries to formulate local data.[27] In addition to researching North American craniofacial morphology, Farkas also travelled to gather craniofacial anthropometric information for Chinese, African and other populations during his career.[27,28] The literature often uses this 'gold standard' of direct anthropometry to compare modern techniques.[26–29]

2.2.2: Direct Anthropometry

The original method to assess asymmetry of the face was direct anthropometry. This involves physically measuring multiple distances, proportions and angles with the patient present.[26] It requires precise training, an understanding of landmarks and their definitions, an armamentarium of measuring tools, and both the clinician's and patient's time.

When measuring the soft tissue landmarks of the face the hard tips of the calliper are designed to touch the surface gently, but not press into the skin surface. Equally when the measuring

tape is used it should be applied carefully so as not to distort the soft tissues. When bony landmarks are being measured the blunt tips of the callipers are used (e.g. Glabella, Gonion). In order to standardise orientation, the patients' heads are positioned with Frankfort plane parallel to the true horizontal. This is identified by joining landmarks Porion (highest point of the soft tissue upper margin of the external auditory meatus) and Orbitale (lowest point on the lower margin of the orbit). To place the facial midline (vertical orientation), 3 landmarks are used: Nasion (deepest point of the nasal bridge); Subnasale (midpoint at base of columella); and Gnathion/or Menton (lower midline of mandible).[26]

There are 4 main groups of measurements taken for direct craniofacial anthropometry: horizontal, vertical, sagittal and inclinations (angles).[26] Most of the horizontal measurements and lateral measurements (taken on both sides of the face) are projective, meaning the shortest distance between the 2 landmarks is recorded. Single measurements represent the landmarks in the midline, paired measurements represent bilateral landmarks. Angles provide information on the planes relative to each other giving a more global assessment as they are made up of at least 3 landmarks and 2 planes. Farkas (1981)[26] reported that there were 174 different craniofacial measurements requiring 112 different methods of assessment. The area of the face with the most landmarks (n=55) and most measurements (n=30) were the orbits. In a landmark study of the 'normal' North American population Farkas and Cheung (1981)[3] reported the amount and prevalence of asymmetry present in children between 6-18 years of age. To describe asymmetries at different levels he divided the face into subunits known as the horizontal thirds: upper, middle and lower face. They completed their direct anthropometric technique whilst viewing the face from the lateral perspective, this is likely to be due to the fact that studies using cephalometric images were popular at that time for assessing facial asymmetry. They placed 7 landmarks to allow the assessment of 6 pairs of measurements on each side of the face, one almost perpendicular to the true horizontal, and the other 5 running in a horizontal oblique direction. These can be seen in Figure 5.[3] Black ink was used to identify the required landmarks on the face to ensure the same exact positon could be used for each measurement. A standardised method was outlined using a spreading calliper to carry out the measurements.

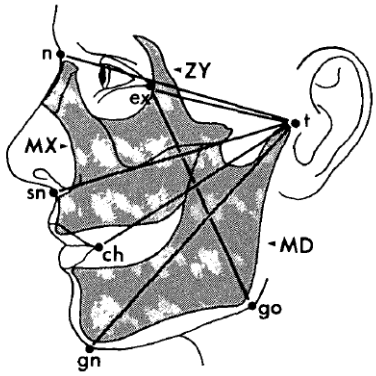


Figure 5: Taken from Farkas and Cheung (1981)[3], illustrating the 7 landmarks used: Nasion (n), Subnasale (sn), Gnathion (gn), Cheilion (ch), Exocanthion (ex), Gonion (go) and Tragion (t), and the 6 pairs of measurements: one almost perpendicular to the true horizontal and the other 5 running in a horizontal oblique direction.

2.2.3: Landmarks

Landmarks are used in many research studies investigating facial symmetry and to provide 'normative' data for the dimensions of the head and face. The definitions used are often taken or adapted from Farkas (1994).[28] They can be used as a means of comparing the measurements of right and left sides, thus assessing facial asymmetry quantitatively. In addition to being used for measurements, landmarks can also assist with image orientation. Providing a clear definition for each landmark is key in anthropometry as in order to compare datasets, different patients on different days, or even to compare the same patient at different stages of growth, the landmark position needs to be reproducible.[26] Direct anthropometry traditionally involved marking the skin with ink for each 'measuring point' (landmark). Farkas (1981)[26] outlined this step as being particularly important when the same landmark was going to be used for several measurements, ensuring it was in exactly the same place for each one.

2.2.4: Landmarks used in 3-Dimensional Images

The selection of appropriate facial anthropometric landmarks for 3d analysis is challenging. Tables 1 and 2 outline frequently positioned landmarks and their definitions which have been used in the assessment of facial asymmetry in 3d images.

Midfacial landmark	Definition
Trichion (tr)	Point located just below the hairline in the midline of the forehead
Glabella (g)	Most prominent midpoint between the eyebrows
Soft Tissue Nasion (n)	Deepest point of the nasal bridge
Pronasale (prn)	Most protruded point of the apex of the nose
Subnasale (sn)	Midpoint of the columella base at the apex of the angle where the lower part of the nasal septum and the surface of the upper lip meet.
Superior Labial Sulcus (sls)	Deepest midline point between the mouth and nose
Labial Superious (ls)	Midpoint of the upper vermillion line
Labial Inferious (li)	Midpoint of the lower vermillion line
Stomion superious (stos)	Most inferior midpoint of the vermillion border of the upper lip
Stomion (sto)	Midpoint of the mouth orifice
Stomion Inferious (stoi)	Most superior midpoint of the vermillion border of the lower lip
Lower Lip (ll)	Midway between the Cheilion and Labiale Inferious (right and left)
Sublabialis/Inferior Labiale Sulcus (sl)	Determines the lower border of the lower lip and upper border of the chin (mentolabial ridge)
Pogonion (pg)	Most anterior midpoint of the chin
Menton (me)	Most inferior point on chin

Table 1: Examples of midfacial landmarks used in the assessment of facial asymmetry and their definitions.[2,7,19,28,30–32]

Bilateral landmark	Definition
Endocanthion (en)	Point at the inner commissure of the eye fissure
Exocanthion (ex)	Point at the outer commissure of the eye fissure
Pupil (p)	Centre point of each pupil
Palpebrale Superious (ps)	Highest point in the mid-portion of the free margin of each upper eyelid
Palpebrale Inferious (pi)	Lowest point in the mid-portion of the free margin of each lower eyelid
Orbitale Superious (os)	The highest point on the lower border of each eyebrow

Orbitale (or)	Lowest point on the lower margin of each orbit
Distal end of supercilium (sc)	Point at distal end of each eyebrow
Frontaltemporale (ft)	Point of concavity on each side of the forehead above the supraorbital rim, lateral to the elevation of the linea temporalis
Zygion (zy)	Most lateral extents of the zygomatic arches
Gonion (go)	Most lateral point on the soft tissue contour of each mandibular angle located at the inter-section of the tangent lines
Tragion (t)	Point located at the most concave point (notch) of the intersection of the upper margin of the tragus
Sub-tragion (str)	Most inferior point on the anterior inferior margin of the helix attachment of the face
Superaurale (su)	Highest point on the free margin of the auricle
Preaurale (pra)	Most anterior part of the ear, located in front of the helix attachment to the head
Otobasion Inferious (otbi)	Most inferior point on the ear lobe at the attachment to the cheek
Alare (al)	Most lateral point on each alar contour where nostril starts to curve laterally
Alare Curvature (ac)	Most lateral point in the curved baseline of each ala
Inner Alare(Ali)	Inner marking level at mid-portion of the alae were the thickness of each ala is measured
Outer Alare (alo)	Outer marking level at mid-portion of the alae were the thickness of each ala is measured
SubAlare (sbal)	Point at lower limit of each alar base where it joins the skin of the upper lip
Columellar high point (c)	Highest point on columellar crest
Subnasale Inner (sni)	Midpoint of columella on each side at the bottom line where the thickness of the columella is measured
Christa Philtri (cph)	Point on each elevated margin of the philtrum just above the vermillion line of the upper lip
Cheilion (ch)	Point at each lateral labial commissure

Table 2: Examples of bilateral landmarks used in the assessment of facial asymmetry and their definitions.[2,7,19,28,30–32]

Many of these landmarks clinicians will be familiar with as the names correspond with

landmarks used in cephalometry. It is essential to note that in cephalometry these mainly refer to bony landmarks, whereas in anthropometry of the face landmarks refer to the soft tissue position.

Gonion and mention require palpation to correctly locate and in the presence of excess muscular or adipose tissue their identification can be challenging.[26] Farkas (1981)[26] recommended that Nasion be identified by using ones' fingernail to locate the slight ridge at the midline point where the nasal root and nasofrontal bones meet. Palpation is not possible on 3d images, although some prospective studies identify the landmarks on the face with black eye pencil prior to capture.[33,34] Orbitale superious can't be used if the eyebrows have been plucked or shaped, and Trichion can't be used if the hairline is receding. In addition to these limitations, if the quality of the image is not ideal, e.g. movement during image capture, hair obscuring face or missing data, some landmarks may be impossible to place directly on the 3d image if the sample is retrospective and the image can't be retaken to improve the quality.

2.3: Two-dimensional Imaging Techniques

2.3.1: Cephalometric Radiographs

Traditionally 2d radiographs, such as lateral and postero-anterior cephalograms, have been used to evaluate the degrees of facial asymmetry affecting the hard tissues. The lateral cephalogram is often readily available to an orthodontic clinician and should be used to assist in the assessment of facial asymmetries when it is available. At the time of Farkas and Cheung's 1981 publication, cephalometric radiographs were the most popular method of evaluating facial asymmetry.[3] This method has limitations, mainly that they are a 2d representation of a 3d facial structure.[7,35] Soft tissues may act to mask irregularities of the facial bones in both patients with average dimensions, and in those requiring treatment.[3,36–38] Therefore, asymmetries noted radiographically may not be a true representation of how the individual appears in real life.

2.3.2: Photographs

Routinely in Orthodontics we use facial photographs to assess soft tissue facial balance and symmetry.[39] The primary limitation with photographs is that they are a 2d representation of a 3d subject with the depth of facial form left unaccounted for.[40] Before 3d imaging systems increased in availability, 2d photographs were a popular method of capturing facial morphology. They have been used to calculate an asymmetry value for the face, usually involving landmark identification and using mathematical formulae to work out the difference between right and left sides.[12]

Fudalej et al. (2011)[41] used a hand held digital camera to record images (basal and frontal) of

cleft and non-cleft patients (at rest) to calculate and compare asymmetries and aesthetic differences. There are potential problems with using a more traditional hand-held device to obtain images in research. In order to ensure accuracy, you need the exact same perspective and the same object to camera distance for each patient, with the same camera settings to be able to compare the images.

Farrera et al. (2015)[42] also used photographs to assess facial asymmetry and aesthetics but captured the images using a standardised method with the Frankfort plane horizontal and usually a neutral facial expression. In most research, images are not likely to be accepted if they have hair obscuring the face, active muscles of facial expression (unless desired) or plucked eyebrows which may alter landmark identification. If photographs are to be used, the camera should be high quality, for example the Pentax K1000 (with 135mm AF Pentax lens), and should be positioned in a stable tripod at a set distance from the subject (e.g. 2 metres) with a scale in millimetres and standard white background.[42]

2.3.3: Measuring Facial Asymmetry using 2-Dimensional Images

Berlin et al. (2014)[13] published a comparison of methods of assessing facial asymmetry using 2d analysis. They concluded that the most recommendable techniques to assess asymmetry of 2d faces were: overall facial asymmetry (FA), asymmetry index (AI) and z-score (symmetry value accounting for horizontal and vertical differences).[13]

The FA uses the centres of bilateral facial landmarks (mi) which are placed on the image, these are the absolute values of the x-coordinates on the left (xli) and on the right (xri). The mi is calculated using the following formula[12]:

$$m_i = \left| \frac{xr_i - xl_i}{2} \right|.$$

Following this the horizontal distances between the centres may be calculated and it's the sum of these that gives the overall facial asymmetry score. An advantage of this method is that it doesn't require the construction of a reference line.[13] In a study comparing 2d and 3d methods of analysing facial asymmetry it was shown that there is no correlation between a 3d analysis method (asymmetry index) and the 2d FA method.

The asymmetry index (AI) was utilised by Nakamura et al. (2001)[43] to form a percentage to represent overall facial asymmetry in postero-anterior cephalograms and frontal view photographs. Thus, the results are quite straightforward to interpret when presenting data, as a perfectly symmetrical face would have an asymmetry index of 0%. This calculation will be outlined further in the literature review specifically related to 3d analysis methods. When 3d and 2d AI's were compared there was a significant correlation between them, suggesting that research using these 2 methods may be comparable.

2d asymmetry index formula:[12]

$$AI = \left| \frac{d_R - d_L}{d_R + d_L} \right|.$$

The z-score was also significantly correlated with the 3d AI when the 2 methods were compared. The z score accounts for the signed x and y coordinates (horizontal and vertical symmetry) for bilateral landmarks and x coordinate only for midfacial single landmarks.[12] As this is a 2d analysis method, it doesn't account for discrepancies in the third dimension.

2.4: Three-Dimensional Imaging Methods

2.4.1: Introduction

This study is focused on measuring facial soft tissues using imaging techniques which do not require any radiation exposure for the patient. Cone beam computed tomographs (CBCT) and plain radiographs have also been used in the assessment of facial asymmetry. These studies tend to investigate the facial morphology of patients requiring treatment (therefore have an indication for the exposure) for conditions such as Parry-Romberg syndrome[44], and also for orthognathic planning and assessment of post-surgical outcomes. As this thesis is based on the 'standard population,' methods involving ionising radiation exposure are not covered in detail in the literature review.

2.4.2: Laser Scanning

Laser scanners have been used to obtain 3d images in many research studies investigating facial form and symmetry.[10,39,45,46] This method commonly involves the use of 2 laser scanners (one capturing each side of face, left and right respectively), with significant overlap of the pair of scanners (known as a stereo pair) in the anterior midfacial region. The lasers used are eye safe class I lasers with a wavelength of λ =690nm at a power of 30mW. This is considered a non-invasive procedure and results in no ionising radiation exposure to the patient.[47,48]

Patients are usually positioned seated on an adjustable chair in a natural head position 1350mm away from each camera (may vary between makes and models), looking directly into a mirror located between the two scanners, prior to the image being taken (Figure 6).[19,39] The mirror may have vertical and horizontal guidelines to assist the participant in orientating their head. Participants are asked to sit with the middle of their face in line with the vertical line on the mirror and their eyes level with the horizontal plane marked, the adjustable seat may be moved (up/down, left/right) as necessary.[45,49] One adult study didn't use a mirror and instead simply asked the patients to look directly at an object positioned between the 2 cameras.[50] In a clinical setting the natural head position has previously been reported to be reproducible.[51]

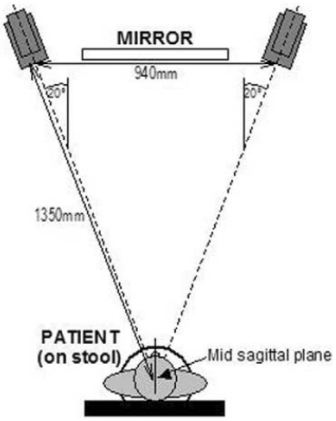


Figure 6: Participant and camera positioning prior to laser scanning, taken from Kau et al. (2005).[52]

To prepare participants, hair is tucked behind the ears or pinned up, glasses are removed, individuals are asked to wear no make-up and to be clean shaven where applicable.[48] Depending on the aims and objectives of the study they may be asked to maintain a relaxed facial form (some asking the patients to swallow immediately before[47] and to close their lips[45]) or to smile if this is of interest to the study.[30] If there is any notable movement or alteration in facial expression during acquisition the process is repeated to gain an image of acceptable quality.[45] Lenses of medium range with focal length of 14.5mm are commonly used.[45,48,50]

The images are stored on the selected software as 2 files initially. The area of overlap anteriorly makes it possible to merge or 'stitch' the 2 images together. Each scan is systematically processed to fuse the images and create a single 3d composite facial image. It is commonplace to merge the right and left scans only if there is 70% matching between the paired images in the area of overlap, with the tolerance set to +/-0.5mm.[39] Figure 7 illustrates the right and left facial shells with surface matching prior to merging.[53] Images can be obtained in 2.5 seconds but different makes and models report acquisition times of up to 8 seconds.[19,39,45]

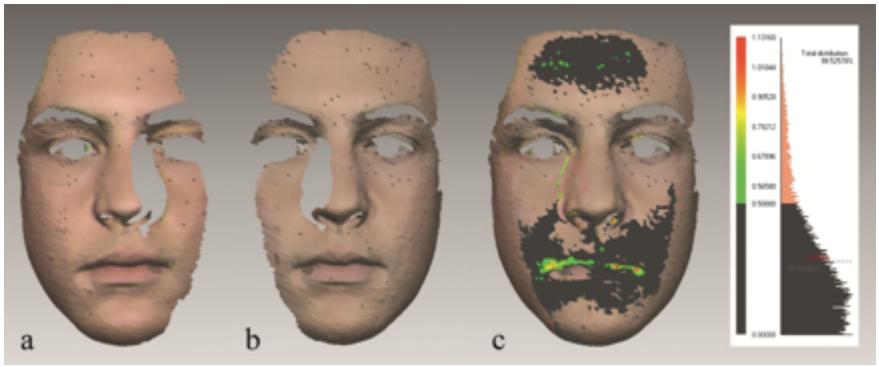


Figure 7: Taken from Djordjevic et al. (2014)[53] illustrating the a) right facial shell, b) left facial

shell and c) surface matching of overlap of the 2 shells. The dark grey indicates deviations between the shells of <0.5mm, light green 0.51-0.79mm, yellow 0.80-0.90mm and red 0.91-1.13mm deviations. Overall these shells were acceptable for merging with an average overlap of 0.28mm (SD 0.24mm) and 85.53% matching between shells.

2.4.3: Accuracy of 3-Dimensional Laser Scanning

The manufacturing accuracy of laser scanners is reported to be within 0.1mm (2 high resolution Minolta Vivid VI900 3D cameras used as a pair) (Konica Minolta, Tokyo, Japan) with 'real life' research reporting accuracy of the composite facial scans (merging the 2 images) to be clinically acceptable with 90% matching of the face within 0.85mm.[47] A prospective clinical trial was undertaken by Kau et al. (2005)[52] investigating the reliability of 3d laser scanning of the faces of school children (mean age 11.3 years). Images were captured on 3 different occasions (baseline T1, within 3 minutes T2 and 3 days later T3). The lower jaw was found to have the greatest error of reproducibility and this is thought to be due the mandible being freely moveable relative to the rest of the face. Therefore, it is hypothesised that mandibular position may be slightly altered between images taken on separate occasions. Despite the slightly reduced reproducibility of the lower face, the error did not exceed 1.35mm.[47] The overall mean deviations between superimposed images were 0.31mm (SD ±0.18) for scans taken 3 minutes after baseline, and 0.40mm (SD ±0.11) for scans taken 3 days after baseline.[47] This method was predicted to be useful for future studies investigating growth and development of children's' facial morphology.[52]

Due to their relatively quick capture time and ability to store data long term, laser scanners have become increasingly popular in research over the last 2 decades. This method lessens the time required during direct anthropometry techniques and also minimises the impact on clinical outpatient appointments as images can be taken and stored for later analysis. Their reliability and accuracy has resulted them being a popular means of obtaining 3d images for the purposes of anthropometry.[19] The use of laser scanners for young children has been questioned by some due to the time required for the patient to stay still. In an attempt to answer this, Djordjevic et al. (2014)[46] did an exploratory study of facial laser scans of 5-year-old patients with cleft lip and/or palate. In an attempt to keep the children still the investigator was positioned directly opposite the child and provided both visual and audio ques simultaneously, counting aloud numbers 1 through to 8 and holding up the relative number of fingers to help the child stay aware of the time required to remain motionless. The merged midfacial area, when right and left scans were combined, showed an average difference between right and left scans to be 0.25mm (SD ±0.06), this was true for 92.1% of the images (SD ±2.9%).[46] This was considered by the authors to be acceptable as their results were within the tolerance level suggested by Kau et al. (2004)[49] which considered left and right side scans which match in at

least 90% of the overlapping area, within an error of 0.75mm, to be acceptable. This was determined using images of both children (mean age child group 11.6 years) and adult 3d laser scanning images. It was noted that adults did show a slightly more favourable outcome of 90% of the overlapping surface within an error of 0.5mm, suggesting that images of adults are slightly more accurate.[49] The Konika Minolta manufactured high resolution laser scanners were used by several research teams in varying models: Vivid 910 [39,48] and Vivid 900 [19,45] although other systems are available.

2.4.4: Stereophotogrammetry

Stereophotogrammetry has become an increasingly popular means of non-contact 3d surface image acquisition particularly among clinicians and researchers in craniofacial units.[54] It is reported to be the most common class of 3d surface imaging system.[54] Their use is moving to replace more traditional direct anthropometry measurements which are both time consuming clinically, and demanding in relation to patient cooperation. There are 3 types of stereophotogrammetry systems commercially available: active, passive and hybrid.[55]

Active stereophotogrammetry uses the projection of structured light onto the surface which needs to be recorded, the system uses information gained from the deformation of this pattern from multiple viewpoints (2 or more cameras) to generate the 3d image using a method known as triangulation (Figure 8). The system is precisely set up to allow for this with specific camera to camera distances and camera to object distances. Passive stereophotogrammetry on the other hand does not use the projection of pattern onto the surface to be recorded but relies on surface detail such as skin texture to determine the geometry. Similarly it uses precisely positioned cameras (2 or more) but the merging and matching of information from these cameras is more challenging due the lack of projected pattern which helps to match the images in the active strategy (Figure 9).[55] The hybrid technique combines features of both active and passive stereophotogrammetry to achieve a superior quality 3d surface image.[55]

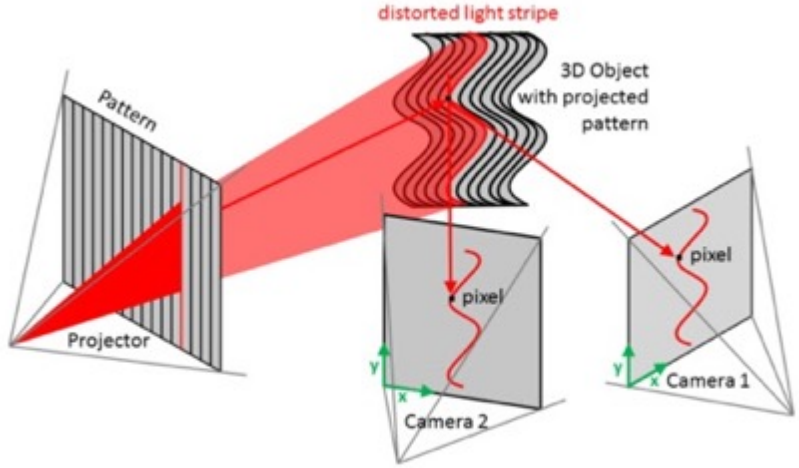


Figure 8: Active stereogrammetry diagram taken from Tzou et al. (2014).[55]

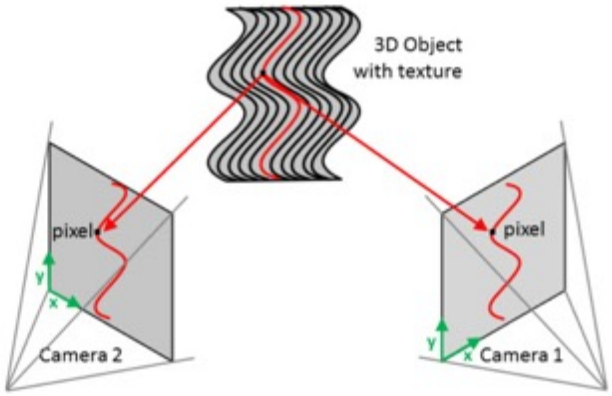


Figure 9: Passive stereogrammetry diagram taken from Tzou et al. (2014).[55]

2.4.5: 3dMD

3dMD™ is a stereophotogrammeteric system used by multiple craniofacial research units with the exact set up varying between individual models.[2,31,56,57] The Static 3dMDHead™ System (3dMD LLC, Atlanta, USA) works by active stereophotogrammetry and calculates a high quality 3d image of the surface of the participants whole head and face. It does this by using a series of photographs generated by the systems synchronised cameras and flashes which are precisely arranged around the participant to achieve 360° detail. In this technique it's the use of multiple cameras in optimum configuration that obviates the need to stitch multiple images together, which is a potential source of error experienced with the laser scanning method.[58] Five modular camera units are used, each are industrial grade machine vision cameras with external white flash units. The surface of the face and head is recorded simultaneously by these cameras and flashes which are timed to act in sync during a rapid 1.5 millisecond (ms) capture window which generates one continuous point cloud.[59,60] Firstly, 10 white light speckle projectors are activated at the same time by the 5 monochrome stereo camera pairs. Then, 0.5ms later, the external white light flash units and 5 colour cameras are activated together. These colour cameras are located in the middle of the modular camera unit. This active stereophotogrammetry technique can identify and link the features of surfaces (such as the face) which are recorded by each monochrome stereo camera pair by using complex stereo triangulation algorithms. This defines the 3d shape and contour of the image. The next stage is the generation of a colour texture map which is achieved using a different software algorithm which merges the data from the 5 colour cameras which corresponds to the shape contour information. A polygon surface mesh is formed with a master x, y, and z coordinate system.

Following acquisition, the point cloud can be immediately transferred to a linked computer. The result is a full 360° photorealistic scan of the face and head.[2] The surface texture and colour are mapped onto the geometric shape of the head creating a life-like image.[54] The process time is less than 15 seconds. The equipment is ideally operated by an experienced medical illustration photographer and needs to be calibrated on a daily basis, or anytime the system is moved.[57] Calibration time for active technique using 3dMDHead™ is <90 seconds and using 3dMDFace™ is 20 seconds. The calibration time for passive techniques include: Canfield VECTRA M3 (Canfield Imaging Systems, Fairfield, NJ) <3 minutes, and Di3D (Di4D, Glasgow, UK) 5 minutes.

It is important that the conditions under which the images are taken are standardised.[57] Lighting requirements may vary slightly between imaging systems but simple overhead fluorescent lighting is usually acceptable. If the room is too bright, for example beside a large a natural light source, it may interfere with the intricate flash mechanisms.[54] Images should be taken in rooms without windows, or in settings where blinds or curtains are available to control external light. The control of ambient light is more essential for passive than active stereophotogrammetry systems.[55]

In research investigating facial asymmetry participants are frequently instructed to maintain a relaxed facial expression with lips gently together and eyes open without straining which could affect the appearance of the soft tissues. Darby et al. (2015)[30] measured facial symmetry changes using 3d images between at rest, smiling and maximal smile, participants were given specific instructions in relation to the facial expression required before each image was taken.[30] Areas prone to data loss and poor resolution include submental and subnasale areas, around the ear, or any aspect obscured by strands of hair. In studies where a detailed view of the nostril shape is required, the patients head may be positioned by tilting the head back (10 degrees) and slightly extending the patients neck.[2,54,59] This positioning technique can be seen in studies investigating symmetry in pre and post-surgical outcomes of participants with cleft lip and palate.[59]

Other 3d vision-based non-contact stereophotogrammetry imaging devices (C3D®) can be arranged slightly differently. If the face is the only part required, then it is not necessary to position the cameras all around the participant's head. In the 2-pod system (C3D®) employed by Hajeer et al. (2004)[61] cameras act as a stereo pair to obtain 6 images which are taken over 50ms. The patient is seated 1.5 metres from the imaging pods, asked to look directly into a mirror in front of them and are positioned carefully within the yellow zone (illustrated in Figure 10) which represents the working volume area.[61] This system was developed at Glasgow University and is reported to have accuracy within 0.4mm when tested on scans of facial casts of infants with cleft lip.[62]

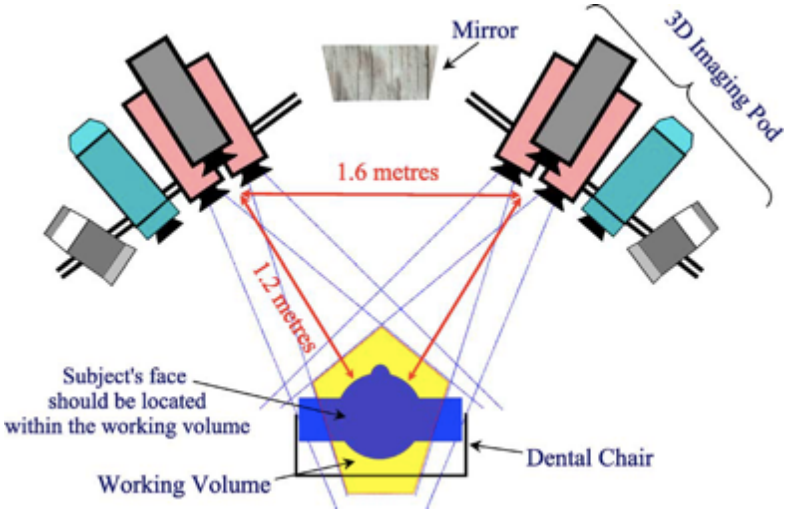


Figure 10: Illustrating the patients position for 3d imaging using a 2 pod system (C3D®) taken from Hajeer et al. (2004).[61] The yellow area is the 'working volume area' which the participant needs to be positioned within.

Three-dimensional stereophotogrammetry enables images to be captured as rapidly as 1.5-2ms (3dMD, Atlanta, USA).[59] The speed of the process may reduce the occurrence of motion artefact by lessening the potential time for the participant to alter position or facial expression during image acquisition.[59] It is the speed, accuracy and minimal invasiveness that has made stereophotogrammetry particularly useful when assessing the craniofacial features of young children. The image is stored as a digital archive which forms a record for the patient, and is a potential source of data for future research analysis depending on the consent.[54] Young children who often present great challenges for direct anthropometry can have their images obtained in less than a fraction of a second and the image can then be stored and referred to at a later date without requiring further cooperation from the patient. On the other hand, direct anthropometry requires patients to sit still for several minutes as each individual measurement is made.[58] Another advantage of stereophotogrammetry is that the process does not expose the patient to any ionising radiation.[54] The image is available almost immediately for assessment of quality. If required, the image can be retaken without incurring increased costs and without greatly increasing the time burden on the patient or the photographer.

2.4.6: Accuracy of 3-Dimensional Stereophotogrammeteric Images

The 3dMDHead™ image generated and the measurement software associated has been verified by the manufacturer to have consistent geometric accuracy with the root mean square of less than 0.2mm. Aldridge et al. (2005)[58] conducted a study using a mixed sample of children and adults (n=15) to investigate the precision, repeatability and measurement error of landmarks due to device and due to digitisation of images attained using 3dMDFace™.

They concluded that the 3d anthropometric data obtained was 'highly reliable'.[58] Areas more prone to error included the mandible, which has previously been discussed in relation to laser scanning, and the measurements crossing the labial fissure. Most of the landmarks were located with error of less than 1mm. Glabella, right and left Gonion were the landmarks associated with the most error. In order to locate Glabella and Gonion properly it's suggested that manual palpation of the area may be required as the perception of their location may change depending on the orientation of the patient or the image on the viewing screen.[58] The software does allow images to be moved freely but it is not possible to repeat the exact position. The best way to overcome this is to standardise the orientation of the patients' images both during acquisition and during data analysis to minimise any error due to variation in positioning.[48]

Weinberg et al. (2004)[63] reported that physically marking landmarks on faces prior to capturing the images improved precision of both landmarks identified on images taken using the Genex 3D imaging system, and on measurements using traditional direct anthropometry.[63] However, this changes the process from 'non-contact' to 'contact' and requires more time from both the clinician and patient, in addition to superior patient cooperation.[48] Regardless of whether landmarks were physically marked or not prior to acquisition the Genex 3D system used was found to be more precise in obtaining craniofacial measurements than direct anthropometry.[58] Weinberg et al. (2004)[63] proposed a threshold level of 2mm for landmark error as their research team felt this was the at which the error would become clinically significant.[30,63] Huang et al. (2013)[7] also used the Genex 3D system reporting its acquisition time to be within 400ms. In this study the landmarks were identified and digitised directly onto the 3d images. Their intra-reliability of landmark identification (using x, y and z coordinates) was calculated to have a mean error of 0.52mm (range 0.31-0.95 mm).[7]

De Menzes et al. (2010)[33] conducted a prospective clinical study to investigate the Vectra 3D imaging system (passive stereophotogrammetry) on a sample of 10 healthy adult volunteers. They used black liquid eyeliner to mark facial landmarks on the individuals prior to image capture. Using 2 different operators they found no significant systematic errors ($p>0.05$) and only negligible errors in facial measurements between the 2 measurements. The mean absolute distances between the values of the 2 sets of measurements ranged from 0.05mm ($p=0.22$) (inter-zygia) to 0.9mm (mouth width) ($p=0.86$), neither of which were found to be significant ($p>0.05$).[33] The accuracy of the system was not found to be influenced by facial form of the participant in the image (dolichocephalic/brachycephalic or class II/ III skeletal pattern), although it's possible the sample size may have been too low to detect a difference. Their results conformed with previous research deducing that the stereophotogrammeteric system was both repeatable, accurate and appropriate for use as a clinical analytic tool.[33]

2.5: Measuring Facial Asymmetry using 3-Dimensional Images

2.5.1: Image Orientation

The first step after the file has been saved is to follow a specific sub-routine to standardise image orientation. This should be done in consideration of all 3 coordinate planes and often requires the positioning of a specified number of anthropometric landmarks.

Primozic et al. (2012)[39], using the software package Rapidform, positioned the image using the mid-sagittal (YZ) and transverse plane (through bilateral Endocanthions) (XZ). These planes were automatically generated once 3 landmarks were located manually: Endocanthion right and left and Pogonion.[39] The mid-sagittal plane (YZ) of the original face represents the plane around which the image may be mirrored.[39]

Kaipainen et al. (2016)[31] used Maxilim software to analyse 3dMDFace™ images. They firstly orientated the faces horizontally and vertically, and then placed 22 soft tissue anthropometric landmarks manually, with 15 associated facial planes. To fabricate a mirror of the image they used Exocanthion right and left to construct a transverse plane with the pupil reconstructed point representing the half-way point between the pupils. Perpendicular to the transverse plane, the coronal plane was positioned passing through 2 midfacial landmarks: the pupil reconstructed point and Subnasale. Similarly, the sagittal plane was used as the symmetry plane to mirror the image at a later stage.

Alqattan et al. (2015)[48] described firstly placing 6 anthropometric landmarks including: Glabella, Endocanthion right and left, Pronasale, Subnasale and Pogonion. The sagittal, coronal and transverse planes were then defined using the combined original-mirrored image (Figure 11).[48] The symmetry plane was represented as the sagittal plane (YZ), the cylinder that coincided with all data points defined the transverse plane (XZ) and finally the coronal plane (XY) was positioned at 90° to the sagittal and transverse planes. The origin of the system was the point at which these 3 planes met and can be considered the origin of the 3-dimensional image coordinate system. In this example the origin can be found at the midpoint between the inner canthi, but the origin position can be modified depending on the studies protocol.[31,48] It is important to know the origin of the image as at that point the coordinates are 0,0,0 for the x, y and z coordinate planes respectively and thus it is from there which the other landmarks are measured in relation to.

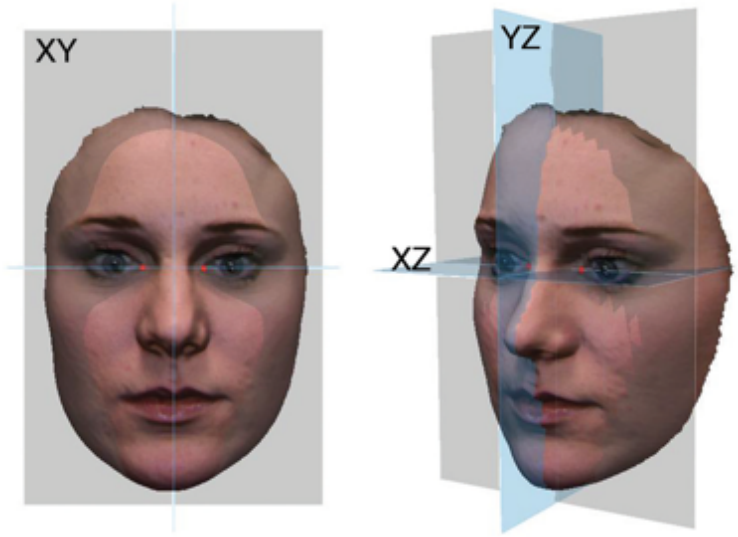


Figure 11: Illustrates the reference frames (sagittal plane (YZ), transverse plane (XZ) and coronal plane (XY)) viewed from the facial view and three-quarter view, taken form Alqattan et al. (2015).[48]

2.5.2: Mirroring

Computer software can be used to remove any extraneous unwanted data from the image including unwanted collars or areas not of interest to the particular study (e.g. hair or neck).[19,57] Once the specific area of interest is separated from the unwanted elements (e.g. face without neck), specialised software can be used to create a mirrored image. This is then superimposed on the original image and their surfaces registered together to create a distance map between the two. This process is known as the 'best fit' procedure.[64] Some research has used the Iterative Closest Point Algorithm which is a complex algorithm used to register the surfaces together. This map allows the user to visualise the distance between the original and mirrored photographs using their corresponding points. It also allows for the quantification of asymmetry which numerically is represented as an absolute mean difference in distance between the 2 images which is expressed in millimetres. The larger the distance the greater the overall asymmetry (difference between original and mirrored image). This gives an overall numerical asymmetry score for the total face known as the 'mean absolute asymmetry,' usually given with 95th percentiles.[57]

The same method can be applied using landmarks to define facial planes to separate specific areas. Kuijpers et al. (2015)[57] calculated, in addition to an overall mean facial asymmetry, values for different facial parts including the nose, cheeks, lips and chin which were separated by several clearly defined facial planes. Primozic et al. (2012)[39] reported their results in terms of the percentage of asymmetry, this represents the percentage of the face within which the original and mirrored images did not coincide within 0.5mm.

Djordjevic et al. (2013, 2014)[19,53] presented their findings in relation to facial symmetry (rather than asymmetry) as a percentage of the whole face. The tolerance level was set at 0.5mm (acceptable match between 2 images). Using this method, the higher the percentage (amount of match between the 2 images with <0.5mm difference) the more symmetrical the area assessed is (Figure 12a).[19,65] Their methodology also divided the face into upper, middle and lower facial thirds separating them using the inner canthus plane and a horizontal plane through the outer commissure of the lips (Figure 12b). These planes were also used by Primozic et al. (2012)[39] and Ovsenik et al. (2014)[64] to represent the dividing lines between facial thirds when assessing facial asymmetry.

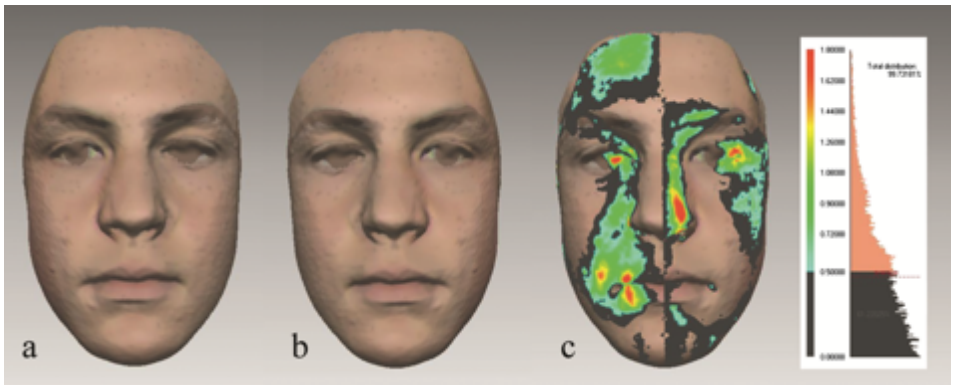


Figure 12a: Demonstrates an absolute colour map the difference in distance between the original and mirrored images. The different colours represent variation in the distances: black 0.0-0.5mm (tolerance level), blue 0.5-1.4mm, green 1.4-2.4mm, yellow 2.4-2.8mm and red 2.8-3.5mm. The overall percentage within 0.5mm tolerance in this particular image was found to be 61%. Taken from Djordjevic et al. (2014).[53]

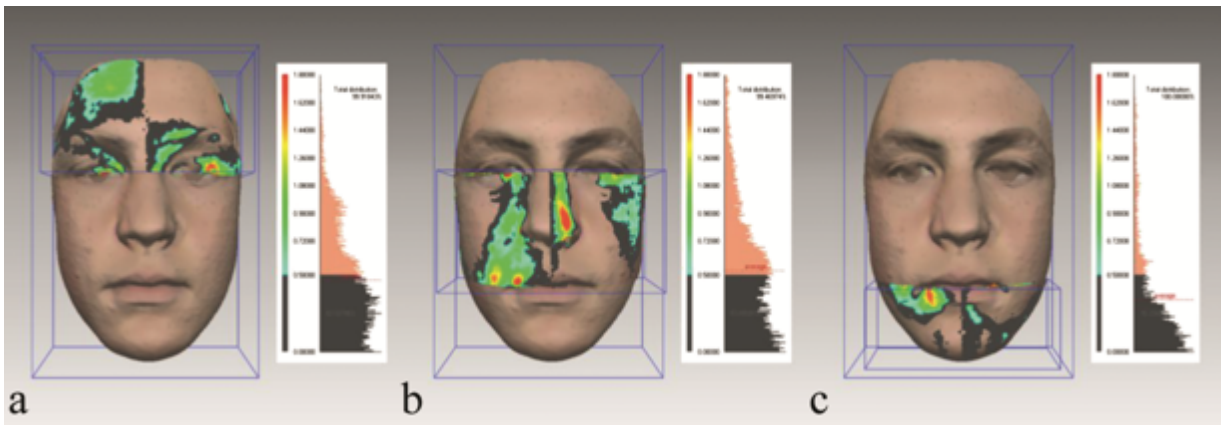


Figure 12b: Demonstrates an absolute colour map of the difference in distance between the original and mirrored images but divides the face into upper middle and lower facial thirds. The same colours are used as in Figure 12a to represent differences in distances between the images. This allows a different symmetry score to be assigned to each facial third. Here the symmetry percentages were: 62.0% upper, 53.7% middle and 78.3% lower facial third,

with the middle third in this male patient being the least symmetrical. Taken from Djordjevic et al. (2014).[53]

In a prospective clinical study, 5-year-old children with repaired cleft lip and/or palate were compared with average faces generated from a standard population non-cleft control group which were age and gender matched.[46] In this study the face was divided into 4 regions which were divided by 3 horizontal planes: 1. connecting Endocanthion to Endocanthion (inner canthus plane), 2. plane positioned horizontally through Subnasale and 3. connecting the bilateral corners of the mouth. As this was a study based on the outcomes for cleft patients the authors wanted to be able to assess the nose and upper lip separately, thus they used 4 regions as opposed to 3.[19,46]

2.5.3: Average Face Construction

Research using both laser surface scanning and stereophotogrammetry has been used to generate what are known as 'average faces.'[46,59] In studies investigating outcomes after cleft surgery, this method can be used to compare operated children with an age and gender matched 'average faces' generated from a sample of the standard population. Using this approach, it can be estimated how asymmetrical the operated patients are relative to their counterparts who never had a cleft or surgical repair. The images are scaled, rotated and translated one by one in order to optimally superimpose them. Generalised Procrustes Analysis is used to minimise the spaces between landmarks in different images. Djordjevic et al. (2014)[46] created different average faces for males and females of the standard population and used these to compare individual images of 5-year-olds with operated cleft lip and/or palate. It's not always possible to separate genders to create male and female average faces if the sample size is small, for example, if creating an average face of a child with a cleft.[59] If the aim is to create an average face of a child with a cleft it's good practice to choose a side (e.g. cleft to occur on the left) and to mirror any right side clefts (so they become left sided) prior to superimposing the images to generate the 'average' in order to increase the sample size available.[59] In Figure 13, A. is the image of the average face for a control group (mean age 10.5 years, genders combined) and B. represents the average face of the group of patients with operated unilateral cleft lip and palate (mean age 10.1 years, genders combined).[59]

Figure 13: Top row (A) shows the average face for the control group, bottom row (B) shows the average face which was constructed using images of patients with operated unilateral cleft lip and palate. Taken from Bugaighis et al. (2014).[59]

2.5.4: Anthropometric Mask Construction

The anthropometric mask (AM) has been designed as a template to fit over the areas of interest on any particular face. This has been outlined by Claes et al. (2012)[56] using information gained from 3d stereophotogrammeteric images (3dMD) of 400 healthy Australians (age range 5-25 years), as a potential alternative to the distance map created by closest point analysis. The facial surface template is defined by approximately 10,000 points (known as quasi-landmarks) equidistant from each other (~2mm apart) linked together by approximately 20,000 triangles. The authors compare the use of the AM to fitting an elastic mask over a statue by aligning the faces and deforming the geometry of the mask as required to get them to fit together. The registration method is non-rigid which allows for differences in facial shape between the face being analysed and the AM template. The authors report that this facilitates the analysis of different images in a spatially dense and consistent manner whilst avoiding the use of landmarks which may not fully represent the entire facial form. They also suggest that surface based techniques which use closest point algorithms may not be the most sensible biologic approach to assessing facial form in 3d.[56] A comparison between the AM and closest point analysis can be seen in Figure 14.

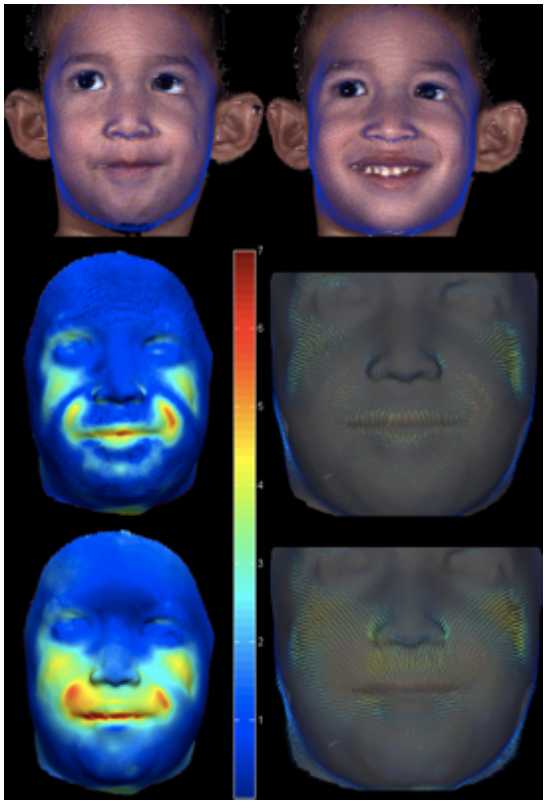


Figure 14: The top of the image shows the AM mappings applied to 2 images of a 4-year-old boy with Treacher Collins Syndrome at rest and smiling. The middle 2 images show the closest point analysis with the distance map on the left side and the vector field on the right side. They report that discontinuities in this image map represent intersections between the image (lack of continuity). On the other hand, the bottom row shows the method using the AM with quasi-landmark map on the left side and vector field on the right side which the authors report to be a more realistic outcome. Taken from Claes et al. (2012).[56]

2.5.5: Asymmetry Index

Katsumata et al. (2005)[66] first outlined the use of the Asymmetry Index (AI) for analysing asymmetry present on computed tomography (CT) images.[7] An AI can be used to quantify facial asymmetry by means of entering the x, y and z coordinates of a bilateral landmark of interest into the following formula:[7,66]

$$AI = \sqrt{(Ldx\text{-}Rdx)^2 + (Ldy\text{-}Rdy)^2 + (Ldz\text{-}Rdz)^2}$$

If there was perfect symmetry (e.g. no difference in position of paired bilateral landmarks) then the AI score would be zero. The larger the AI value, the bigger the discrepancy is in the differences between x, y and z for the particular bilateral landmark measured. In order to calculate the AI for midfacial landmarks the x coordinate alone is used as this represents

deviation from the midfacial (sagittal) plane.[7]

Huang et al. (2013)[7] investigated asymmetry in an adult Chinese population (n=60, ages 20-35 years) and reported the mean AI to be 0.76-2.82mm (SD 0.42 – 2.50) for midline and bilateral landmarks combined. The AI's were larger for both midfacial and bilateral landmarks in the lower aspect of the face (e.g. Menton and Cheilion when compared to Glabella and Endocanthion) suggesting asymmetry increases as you move away from the cranium. Katsumata et al. (2005)[66] investigated facial AI's using CT and reported that asymmetry increased in the lower part of the face in a standard population. They considered, for each landmark, a deviation of more than 2 standard deviations to be 'marked asymmetry.' Cheilion was the landmark which demonstrated the most asymmetry with a mean AI of 2.82mm (SD ±1.42). Interestingly, the patients had been selected on the basis of an Angles Class I dental malocclusion with 'face regarded as normal symmetry' by a team of assessors including a plastic surgeon and an orthodontist. The fact that the bilateral Cheilion discrepancy mean AI was 2.82mm but this patient was still included and regarded as generally facially symmetrical, supports a publication by Farkas and Chung (1981)[3] which claims that 3mm of asymmetry is acceptable or unnoticeable in faces of the standard population.[3,7]

Katsumata et al. (2005)[66] represented their AI findings (using cephalometric based landmarks on CT scans) on a graph with colours green, yellow and red depicting symmetrical, asymmetrical and marked asymmetrical landmarks (Figure 15a). The line between green and yellow represents 1 SD from the mean AI value for that landmark, the line between yellow and red represents 2 SD from the mean AI value. The second image (Figure 15b) shows how this graph can be used to compare an individual to the 'normative' data. It's clear that the landmarks located in the 'red' region (lower 1st incisor, Menton, Condyle, lower 1st molar, Coronoid process and Gonion) were significantly asymmetrical for the particular individual being presented. A limitation of these graphs is that it doesn't give information regarding the direction of the asymmetry.

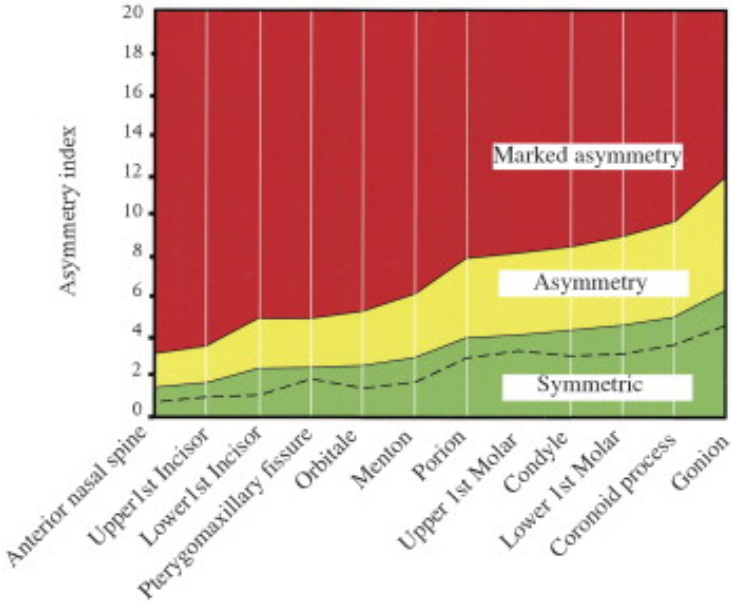


Figure 15a: Graph outlining the asymmetry index (AI) (y axis) for each landmark (x axis). The line between green and yellow represents 1 standard deviation (SD) from the mean AI value (above this considered asymmetrical), the line between yellow and red represents 2 SD from the mean AI (above this considered marked asymmetry). Taken from Katsumata et al. (2005).[66]

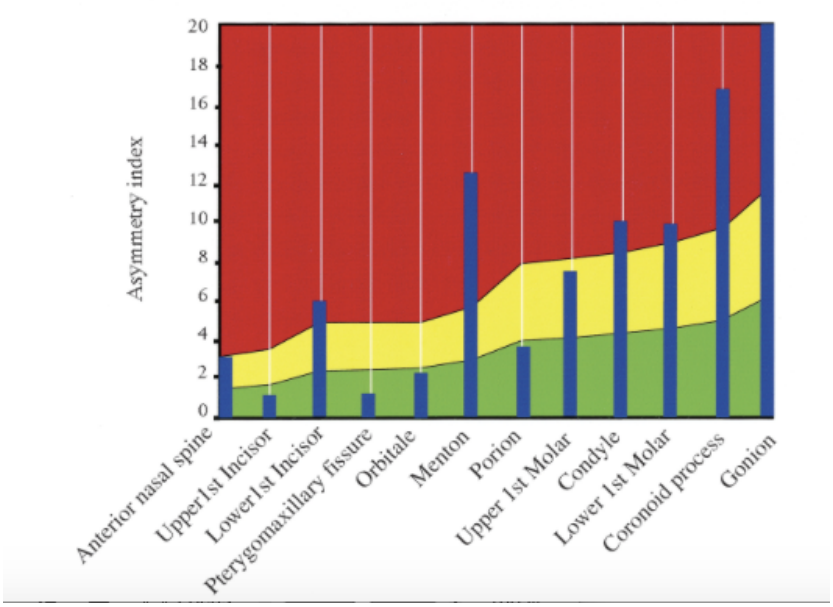


Figure 15b: Bar chart superimposed on graph from Figure 15a shows the specific AI for each landmark from a single patient to compare their asymmetry to the 'normal' sample. Colours on graph are divided by 1 and 2 SD of mean AI as in Figure 15a above. Taken from Katsumata et al. (2005).[66]

Huang et al. (2013)[7] also represented their findings visually on a facial asymmetry diagram to illustrate the deviation from perfect symmetry (Figures 16a and b). Similarly, to the graph used

by Katsumata et al. (2005)[66] this is a useful visual tool to demonstrate the location and severity of asymmetry, it also has the added value of demonstrating in what direction the deviation is in.[7]

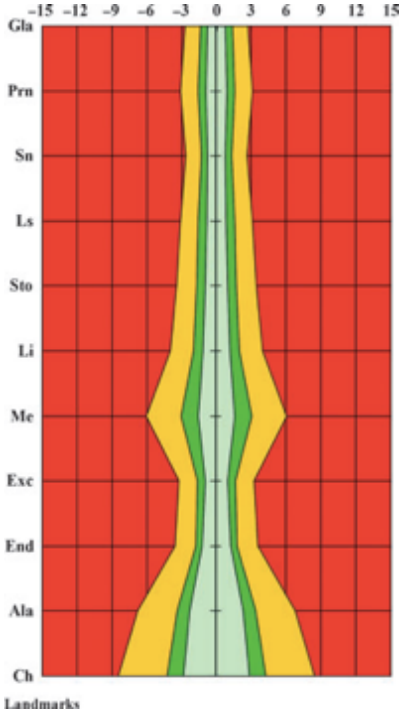


Figure 16a

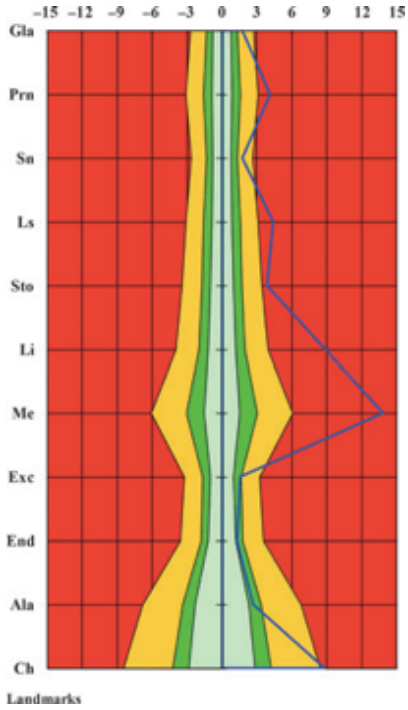


Figure 16b

Figures 16a & b: Asymmetry is illustrated in the graphs as: light green colour represents symmetry within the mean values, darker green is symmetry with deviation <1 SD, yellow is symmetry with deviation within 1-2 SD and areas coloured in red represent AI values deviating greater than 2 SDs away from the mean, considered to represent significant facial asymmetry.
Figure 16a: Illustrates these values.
Figure 16b: Shows the AI values for an individual relative to the 'normal' data, seen as a blue line. As the values are mostly positive, this indicates the patients' landmarks are deviated to the left of the midfacial plane (as formula is calculated left minus right). For this particular individual, it may be concluded that they have significant asymmetry (>2 SD away from the mean) for the following landmarks: Pronasale, Labrale superious, Stomion, Menton, and Cheilion. Taken from Huang et al. (2013).[7]

Alqattan et al. (2015)[48] conducted a UK based study used a sample of 85 'normal' adults (age range 19-54 years) to represent the standard population for a study comparing landmark versus surface 3d analysis techniques of facial asymmetry quantification. They calculated asymmetry in millimetres separately in each of the 3 coordinate planes for bilateral landmarks and in the x direction only for midfacial landmarks, using a total of 21 anthropometric landmarks. The values for 3 planes were then combined to present an overall AI for each

bilateral landmark using the same formula as Huang et al (2013).[7,48] The data was presented in terms of median, 25th and 75th interquartile (IQR) values for each of the landmarks. For midfacial points the highest AI was measured for Pogonion, and the lowest AI for Pronasale. For bilateral landmarks the highest AI was recorded for Cheilion (as was found by Huang et al. (2013)[7] and the lowest AI was for Christa philtri. This was true for males and females with no statistically significant differences in AI scores according to gender when Mann-Whitney U test was applied ($p>0.05$). These median values and IQRs for males and females are outlined in Table 3 below representing the highest and lowest values of asymmetry for midfacial and bilateral landmarks in this study by Alqattan et al. (2015).[48]

Landmark	Median AI males (IQR)	Median AI females (IQR)
Pogonion	1.5 mm (0.7, 3.0)	1.8 mm (0.7, 2.5)
Pronasale	0.1 mm (0.0, 0.3)	0.2 mm (0.1, 0.3)
Cheilion	3.2 mm (2.1, 4.1)	3.5 mm (2.4, 5.0)
Christa philtri	2.2 mm (1.6, 3.0)	1.6 mm (0.9, 3.5)

Table 3: Outlines the asymmetry indices for landmarks: Pogonion, Pronasale, Cheilion and Christa philtri in a sample of adult males and females which has been adapted from Alqattan et al. 2015.[48]

In some papers, the term asymmetry index is used for a different calculation. Berssenbrügge et al. (2014)[12] calculated an AI by firstly mirroring an image, superimposing the mirrored image with the original, and then registering the two together using the iterative closest point algorithm as previously discussed regarding the mirroring technique. In this study, the 3d AI refers to the mean difference between the original and mirrored images when registered together. Similarly, the higher the value, the more facial asymmetry present. To compensate for the possibility of face size influencing the results they apply the diagonal of the bounding box using the image in frontal view and divide the mean distance by this. As the resultant number is quite small it is multiplied by factor of 1000 which does not impact the results.[13]
The 2d AI is calculated using the relative distance of a landmark to the midfacial plane and the difference between the right and left measurements.[12,13,67] This can be done for each landmark using the formula below when: 'R' is right distance and 'L' is left distance to the medial line of the face. This method was reportedly introduced by Nakamura et al. (2001)[43] and is given as a percentage, with a higher percentage representing more asymmetry.[13,67]

$$\textbf{2D AI} = |(R - L) / (R + L)| \times 100\ (\%)\ ^{67}$$

2.6: Reliability Testing

Aldridge et al. (2005)[58] investigated the precision and error of measurements obtained using images captured by the 3dMD stereophotogrammeteric system. In order to investigate this, 2 images were taken of each subject (15 subjects), which included a sample of 'normal' adults (n=7), 2 children with Downs syndrome, 5 children who had corrective surgery for unicoronal synostosis and known craniofacial disorders, and one non-syndromic child. A single observer identified 27 anthropometric landmarks (6 midfacial, 7 bilateral) as defined by Farkas (2004)[28] on both images of each subject. They defined precision as the mean difference between 2 repeated measurements of the same image. Precision was considered to be highly precise for mean differences <1mm, precise if it was between 1-2mm and less precise for mean differences >2mm. Fourteen landmarks were found to be highly precise with an error sub-millimetre in terms of x, y, and z coordinates of both facial scans, this is comparable to results found using other imaging techniques.[58,63] Three landmarks had precision between 1-2mm (Nasion and bilateral Tragion), and 3 landmarks showed error >2mm (Glabella and bilateral Gonion). The overall median error for all landmarks combined in all 3 planes of space was 0.44mm, ranging from 0.17mm (right Endocanthion) to 4.1mm (right Gonion). Nasion error was <1mm for x (0.34mm) and z coordinates (0.71mm) but 1.35mm in the y axis (superior-inferior). Glabella also shows increased error in the y axis (4.03mm) suggesting it is challenging to locate in a superior-inferior direction reliably. Gonion showed a poor level of precision in all 3 planes of space with bilateral landmarks demonstrating errors between 1.45 – 4.10mm. In relation to error due to digitisation there was 0.9% variance when all of the linear dimensions were averaged, suggesting that digitisation of a subject doesn't greatly contribute to overall error. Error due to the imaging system was reported as 1.5% (mean) which again is a small proportion of total error. There were 11 linear distances (out of total 190) which had an error due to imaging of greater than 5%, six of these involved the landmarks Otobasion inferious, Gonion or Tragion. The authors reported that positioning of these landmarks can be influenced by missing data, hair covering skin, or shadows cast particularly in the ear region.

Without physically palpating Orbitale, locating the Frankfort plane is challenging. A possible solution is to place landmarks prior to image acquisition, such that was completed by Weinberg et al. (2004)[63], or to use the natural head position as an alternative.[63] The correct positioning of Gonion requires palpation, and with that in mind Aldridge et al. (2005)[58] suggested that if Gonion is required for a study it may be appropriate to palpate this and mark physically on the subject's face prior to acquisition. Despite some minimal limitations, overall 3dMD images are repeatable and can be used to measure facial morphology in a highly reliable manner.[58]

Chapter 3: Aims and Study Objectives

3.1: Aim

To analyse and quantify the magnitude of soft tissue facial asymmetry in a standard population of children (<16 years old) in the North West of England using three-dimensional images.

3.2: Objectives and Research Question

3.2.1: Primary Study Objective

1. To determine the extent of facial asymmetry in a standard population of children (<16 years old) using a landmark based approach.

3.2.2: Secondary Study Objectives

1. To investigate the direction of nose and chin deviation relative to a midfacial plane, and to investigate if there is a relationship between:
    - The magnitude of deviation of the nose and chin.
    - The direction of nose and chin deviation.
2. To compare the results of 2 methods of measuring facial asymmetry:
    a. Linear measurements
    b. Surface measurements
3. To investigate the correlation between extent of facial asymmetry and gender.
4. To investigate if the extent of facial asymmetry differs with age.

3.2.3: Research Question

Is there facial asymmetry present in a sample of the standard population of children under 16 years old in the North West of England?

Chapter 4: Methods and Participants

4.1: Design

This was a retrospective cross-sectional cohort study investigating the extent of facial asymmetry in a standard population of children under 16 years of age in the North West of England.

4.2: Sample

4.2.1: Image Selection

The sample was retrospective and included 3d images of participants faces and heads (360˚) taken using Static 3dMDHead™ System (2011 model) (3dMD LLC, Atlanta, USA) technology. These had been captured in Liverpool (2013) and consisted of standard people who volunteered for their images to be taken for research purposes. They formed part of the "Headspace" project led by Consultant Plastic and Reconstructive Surgeon at Alder Hey Hospital (AHH) Mr. C. Duncan, to develop values for a standard shaped cranium, and to develop future craniofacial research and surgical planning. For this research project on measuring facial asymmetry, Mr Duncan kindly provided details of images of 172 participants from the standard population which were available for use and stored at AHH.

4.2.2: Inclusion Criteria

Images of participants were included if they:

- Were under 16 years old.
- Had any skeletal relationship including Class I, II or III.
- Had no obvious craniofacial dysmorphology.
- Had adequate quality 3d image available which was determined by a panel assessment.
- Had parent or guardians informed consent to have their image captured and stored for research purposes.

4.2.3: Exclusion Criteria

Images of participants were excluded if they:

- Had active muscles of facial expression including smiling or crying.
- Had eyes closed.
- Had mouth wide open.
- Were diagnosed or suspected to have craniofacial syndromes, cleft lip and/or palate or any condition which might influence craniofacial development.

- Had a history of craniofacial complications.
- Had a facial deformity due to trauma.
- Missing data which would result in inaccurate landmark positioning and image analysis.
- Presence of clothing/other which directly impedes landmark positioning or image analysis.

4.2.4: Consent

Informed consent was obtained from the parent or guardian on the day the images were acquired. All participants had volunteered to have their 3d images taken for research purposes, however some patients didn't consent for their images to be published. The image numbers of these patients were given to the primary researcher (OC) to ensure those specific images would not be published.

4.2.5: Sample Size

No sample size calculation was completed as research was limited to include a sample of images retrospectively. The research aimed to include as many appropriate images as possible to increase the sample size and improve the generalisability of the results.

4.2.6: Panel Assessment

A panel assessment was organised to assess the suitability of the images to be included. Prior to this any images of adults, missing images or duplicates were excluded from the sample. One hundred and forty-five images were assessed using Vultus (2.5.0.1) software by a panel of 5 members consisting of: 2 Orthodontic Registrars (OC and EB), 2 Orthodontic Consultants (NF and SDG) and a Medical Imaging Technician (JO).

Members of the panel had copies of the proposed methodology sent to them in advance of the panel assessment date. On the day of the assessment, prior to the images being examined there was a team meeting in which the landmarks and definitions which had been agreed were discussed again to ensure all members were familiar with the specific areas of particular relevance. The inclusion and exclusion criteria were reiterated to all members to ensure clarity. Images were examined by each of the 5 members individually on separate computers in different locations within AHH to ensure that there was no discussion between assessors. Each member was given a hard copy of a table to complete with either a 'Y' (yes) to include or 'N' (no) to exclude for each image (Appendix 1) and an option to explain the exclusion if the assessor desired. Data was uploaded to excel by primary researcher (OC) to record the data. It was agreed in advance by the panel that if 3 members deemed the image acceptable it would be included. Therefore, if an image was assigned 3 'Y's' and 2 'N's' it was included. Whereas, if an image was assigned 2 'Y's' and 3 'N's' it was to be excluded.

4.3: Method

Included images were orientated into the natural head position initially by visual assessment and manual manipulation using Vultus software. Subsequently 5 bilateral landmarks were identified: Endocanthion, Exocanthion, Pupil, Pre-aurale, and Supra-aurale, which are defined in Table 4. The natural head position was selected and reference frame analysis (provided by 3dMD™) was applied which triggers the software to orientate the images in a standardised manner. This was agreed as the method of registration for all images to ensure the origins of each were the same. The origin was known as the Pupil reconstructed point (prp) which was positioned at (0,0,0) in terms of x, y and z coordinate positions respectively. The prp may be defined as the point in the midline of the nose at the level of the inter-pupillary line.[68] Figures 17a and b illustrate an image of a 10-year-old male with the reference frame applied.

Landmark	Definition
Endocanthion	Inner commissure of the eye fissure [7,28]
Exocanthion	Outer commissure of the eye fissure [7,28]
Pupil	Centre point of each pupil [28]
Pre-aurale	Most anterior part of the ear, located in front of the helix attachment to the head [28]
Super-aurale	Highest point on the free margin of the auricle [31]

Table 4: Definitions the 5 bilateral landmarks used for the reference frame analysis.

Figure 17a: Frontal image of a 10-year-old male illustrating the bilateral landmarks (shown in green) used for the reference frame analysis: Endocanthion, Exocanthion, Pupil, Pre-aurale, and Supra-aurale. The pupil reconstructed point is shown in red. The scale on the x and y axis is in millimetres.

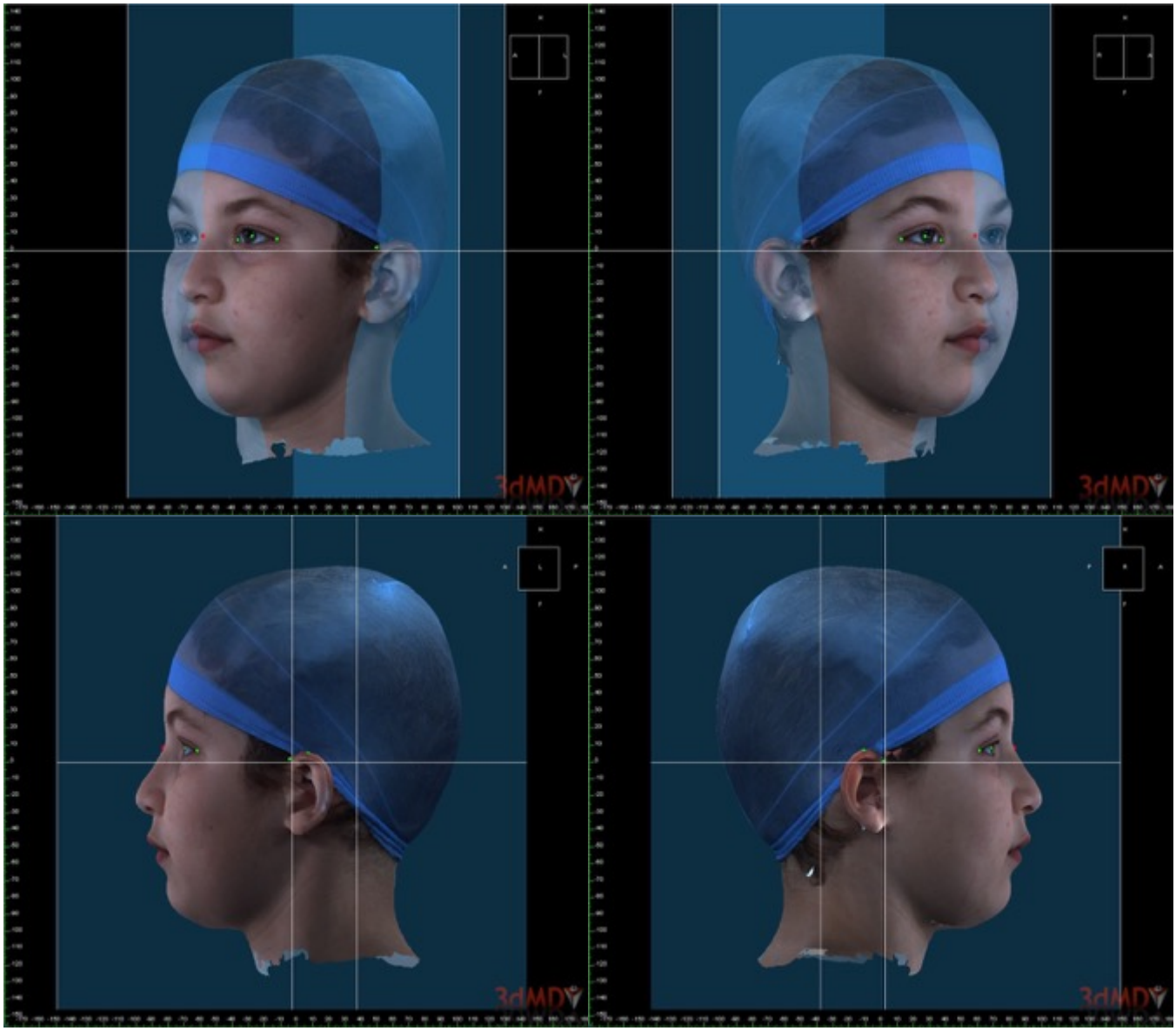


Figure 17b: Images of a 10-year-old male illustrating the reference frame analysis in right and left three-quarter and profile view.

Landmark identification (22 landmarks: 8 midfacial and 7 bilateral) was then completed on each 3d image manually using Vultus software by the primary researcher (OC) within the Medical Illustration Department at AHH. Certain landmarks which have been reported to be unreliable without palpation were excluded e.g. Gonion.[69] The anthropometric landmarks and definitions were selected following literature review and agreement from the research team (OC, EB, NF and SDG), these are commonly used in research investigating facial asymmetry and are mainly based on the work of Farkas (1994).[28] Table 5 defines the midfacial and bilateral landmarks used. Figures 18a and b illustrate an image of a 10-year-old image with 22 landmarks in positon.

Midfacial landmarks	
Landmark	**Definition**
Glabella (g)	Most prominent midpoint between eyebrows [28,48]
Nasion (n)	Deepest point of nasal bridge [7,28]
Pronasale (prn)	Most protruded point of the apex nasi [7,28,48]
Subnasale (sn)	The midpoint of the angle at the columella base where the lower border of the nasal septum and the surface of the upper lip meet [28,48]
Labrale Superious (ls)	The midpoint of the upper vermillion line [7,28,48,70]
Stomion (sto)	Midpoint of the mouth orifice [7,28]
Labrale Inferious (li)	The midpoint of the lower vermillion line [7,28,48,70]
Pogonion (pg)	The most prominent midpoint of the chin [28,48]
Bilateral landmarks	
Landmark	**Definition**
Endocanthion (en)	Inner commissure of the eye fissure [7,28]
Exocanthion (ex)	Outer commissure of the eye fissure[7,28]
Palpebrale Superious (ps)	The highest point in the mid-portion of the free margin of each upper eyelid [28,48]
Palpebrale Inferious (pi)	The lowest point in the mid-portion of the free margin of each lower eyelid [28,48]
Alare (al)	The most lateral point on the alar contour [28,48,71]
Christa Philtri (cph)	Point on each elevated margin of the philtrum just above the vermillion line of the upper lip [30]
Cheilion (ch)	Point at the corner of the mouth at the labial commissure [71]

Table 5: Midfacial and bilateral landmark titles and definitions used in this research project for the purpose of analysing of facial asymmetry.

Figure 18a: Frontal image of a 10-year-old male illustrating the 22 landmarks which were identified on each image (8 midfacial and 7 bilateral), which are listed and defined in Table 5. The scale on x and y axis is in millimeters.

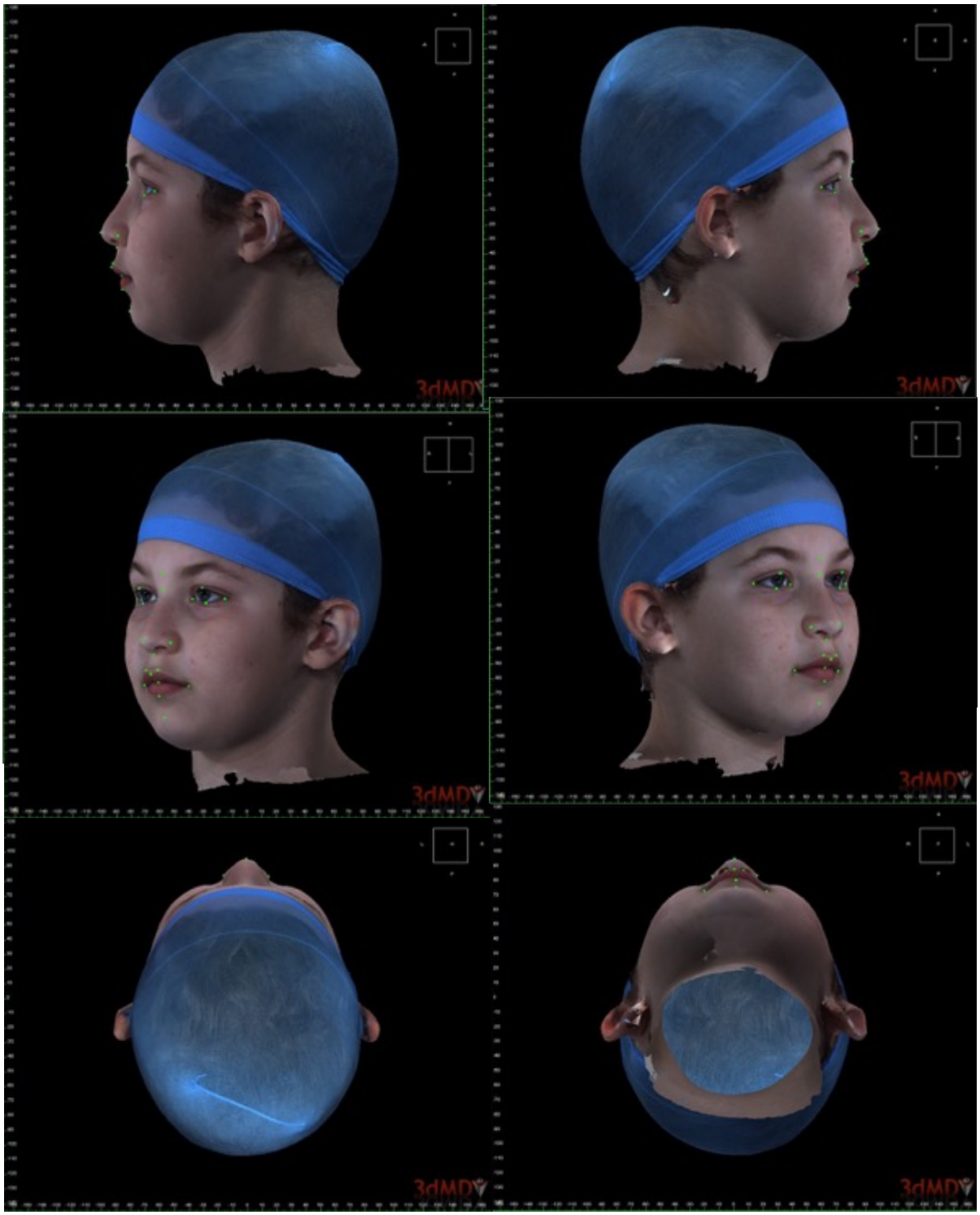


Figure 18b: Images of a 10-year-old male illustrating the 22 landmarks which were identified on each image (8 midfacial and 7 bilateral), which are listed and defined in Table 5.

Once landmarks were correctly positioned they were analysed using Vultus software. This analysis generated the following information in an Excel document for each patient:

- x, y, and z coordinates for all landmarks.
- Euclidean (linear) distances between all landmarks in millimeters.

- Distances along the surface of the soft tissues in millimeters (surface distances) between Nasion point and the 7 bilateral landmark pairs: Endocanthion, Exocanthion, Palpebrale superious, Palpebrale inferious, Alare, Christa philtri and Cheilion. These values are not affected by the positioning of the patients' head.
- The angle (degrees) from Pronasale (Figure 19) and Pogonion (Figure 20) to the true vertical. In order to determine the side of deviation, a positive x coordinate of Pronasale/ Pogonion indicated there was deviation of the point to the left side, whereas a negative x coordinate indicated the point was on the right side of the face. This was possible to infer as the origins (prp) of the images were standardised. Alterations in head position could influence the angles and therefore the images were orientated in a standardised manner as previously outlined at the beginning of this section (4.3). These 2 angles were constructed using:
    - Landmarks: Nasion and Pronasale (prn) or Pogonion (pg)
    - Projected Nasion point (p_prn) or (p_pg) which is a point 5mm below the Nasion point on the true vertical that passes through Nasion point.

Figure 19: Image of a 7-month old female demonstrating the angle formed between landmark Pronasale and the midfacial plane. Pronasale is located 1.2° to the right side. The scale on the x and y axis is in millimeters.

Figure 20: Image of a 7-month old female demonstrating the angle formed between landmark Pogonion and the midfacial plane, with Pogonion located only marginally to the right by 0.7°. The scale on the x and y axis is in millimeters.

The Excel document for each image was named according to the original image number and the photo of the participant which was present in the generated Excel document was removed. No patient identifiable details were on the spreadsheets once they were saved for analysis purposes. A master Excel was created to include coordinates, measurements (linear and surface) and angles for all included images. The data was subsequently analysed using StatsDirect (V3) and the Statistical Package for the Social Sciences (SPSS V24).

4.4: Statistics

4.4.1: Panel Assessment

Cohen's Kappa was used to analyse the level of agreement between the 5 panel members.

4.4.2: Landmark Identification Method Error

4.4.2.1: Intra-observer Agreement

In order to assess intra-reliability, 12 images from the sample of the 'standard population' children (>10% of the included images) were selected at random by statistician (GB). These images were orientated according to agreed standards and 22 landmarks (8 midfacial and 7 bilateral) were positioned by the primary researcher (OC) generating x, y and z coordinates for each landmark and linear measurements between the points in millimetres. A minimum of two weeks later the process was repeated by the same operator to test the intra-reliability. The measurements used for comparison between the 2 episodes included:

- Linear measurements from Nasion to the following landmarks (7 bilateral and 7 midfacial):
    - Bilateral landmarks: Endocanthion, Exocanthion, Palpebrale superious, Palpebrale inferious, Alare, Christa philtri and Cheilion.
    - Midfacial landmarks: Glabella, Pronasale, Subnasale, Labrale superious, Stomion, Labrale inferious and Pogonion.
- x, y and z coordinates for Nasion.

4.4.2.2: Inter-observer Agreement

Inter-reliability was calculated using 12 images from the sample of 'standard population' children (>10% of the included images) and 12 images of patients from a different sample with operated non-syndromic unicoronal synostosis. These 24 images were randomly selected by statistician (GB). Two operators (OC and EB) orientated 24 images and placed landmarks to generate x, y and z coordinates and linear distances between these points. Clear landmark definitions and practice using the software was required to obtain acceptable reliability. When inter-reliability testing was tested for the second time 24 new images were selected at random. The measurements analysed were the same as those used for intra-reliability outlined in section 4.4.2.1.

To assess both intra and inter-reliability for each landmark and for Nasion x, y and z coordinates separately, intra-class correlation coefficients were used. Bland-Altman plots were constructed to visually present the reliability using means and upper and lower 95% limits of agreement and to assess for systematic error.

4.4.3: Assessing Facial Asymmetry

4.4.3.1: Asymmetry of Midfacial Landmarks

The asymmetry of midfacial landmarks was calculated using the x coordinate only as this determines the distance from the origin (prp) and reflects the landmarks location in relation to the midface. The Mann-Whitney U Test was used to assess for differences in midfacial asymmetry between genders. Pearson's correlation coefficient was used to determine if there was a relationship between age and asymmetry.

4.4.3.2: Asymmetry of Bilateral Landmarks using x, y and z Coordinates

Asymmetry of bilateral landmarks was firstly assessed using the differences in x, y and z coordinates individually. These values were measured relative to the origin of the image (point 0,0,0). For example, to calculate the asymmetry in millimeters in terms of the x coordinate for landmark Exocanthion the following formula was used:

$$\boldsymbol{Asymmetry\ in\ mm's = ABS\ ((ExoRx) - (ExoLx))}$$

- Asymmetry in mm's = difference between left and right sides
- ABS = absolute value
- ExoRx = Exocanthion right side x coordinate
- ExoLx = Exocanthion left side x coordinate

4.4.3.3: Asymmetry Index for Bilateral Landmarks

An asymmetry index (AI) was calculated for each of the bilateral landmarks. This encompasses the differences in all 3 planes of space (x, y and z coordinates) to give an overall value of asymmetry for the particular landmark. The following formula was used:

$$\boldsymbol{AI = \sqrt{(Ldx\text{-}Rdx)^2 + (Ldy\text{-}Rdy)^2 + (Ldz\text{-}Rdz)^2}}$$

4.4.3.4: Linear and Surface Measurements of Asymmetry

Linear (closest distance) and surface measurements from bilateral and midline landmarks to Nasion were generated using Vultus software. For the bilateral landmarks this allowed differences between right and left sides to be assessed for both linear and surface measurements. Following this the differences between the amount of asymmetry indicated by the two measurement methods (linear and surface) were compared. A paired t-test was undertaken comparing right and left linear measurements to establish if there was a significant difference in mean values between sides, thus representing asymmetry. This was also completed for the surface measurements.

4.4.3.5: Deviation of Nasal Tip (Pronasale) and Chin (Pogonion)

The angles formed by Pronasale and Pogonion relative to the true vertical were used to represent the amount of nose and chin deviation from the midline. The relationship between extent of deviation (degrees) and side of deviation (right or left) was assessed for both the nose and chin using the Mann-Whitney U Test. To investigate if there was an association between the side of deviation of the tip of the nose (Pronasale) and the tip of the chin (Pogonion) a chi-squared test was used.

4.5: Ethics

Ethical approval was not required for this project. Please see Appendix 4 letter from Director of Research at Alder Hey Hospital (AHH) to supervisor Dr Dominguez-Gonzalez.

4.6: Access to Source Data and Participant Information

The names of participants were not required and therefore were never obtained by the primary researcher. The images used had been anonymised and assigned numbers for identification purposes. The date the image was taken could be identified by viewing when the file was created. Patient information including date of birth, ethnicity, place of birth, first language and history of craniofacial conditions were kindly provided by Mr C. Duncan.

4.6.1: Data Handling

Data collection was completed on site within the Medical Illustration Department at AHH. This is a secure area requiring staff keycard access. The laptop was formally approved for storing patient data and had the appropriate security provided by the in-house information technology team. The laptop was never brought off site during this period.

4.7: Funding

The DDSc research fund (Orthodontic Department) at the University of Liverpool was utilised to purchase a laptop for the purpose of data collection. As it was located in AHH it was arranged through their information technology department to ensure it complied with local safety protocols and regulations.

Chapter 5: Results

5.1: Panel Assessment

A quality assessment of 145 images was undertaken by 5 panel members to determine which were appropriate to be included in the study. All images were assessed by each rater, with no missing responses. Eighty-three (57.2%) images were approved for inclusion by all 5 panel members, 23 (15.9%) by 4 members and 7 (4.8%) by 3 panel members. Consequently, there were a total of 113 images (77.9%), which received approval from 3 or more panel members, remaining for analysis.

Thirty-two images (22.1%) were excluded which had met less than 3 out of 5 panel members' approval. The full results of the assessment can be seen in Table 6.

Panel assessment of 145 images		
Number of panel members (0-5) that voted to include image	Number of images (percentage)	Total included or excluded
5	83 (57.2%)	**Included: n=113 (77.9%)**
4	23 (15.9%)	
3	7 (4.8%)	
2	10 (6.9%)	**Excluded: n=32 (22.1%)**
1	8 (5.5%)	
0	14 (9.7%)	

Table 6: Outlining the results of the panel assessment in terms of number of images (and percentage of total images n=145), which were approved for inclusion by 5, 4, 3, 2, 1 or no panel members. The right-hand column outlines the overall number and percentage of images included and excluded by the panel.

Table 7 lists the number and percentage of included images for each examiner. Assessor A included the most images n=123 (84.8%) with assessor E being more critical including the fewest images n=100 (69%).

Cohens Kappa (Fleiss-Cuzik extension) was used to evaluate the overall agreement between the raters for the panel assessment. Unweighted kappa was selected as data was categorical rather than ordinal. The result was determined using software StatsDirect (V3). The agreement was found to be moderate between the 5 assessors with kappa=0.564 (95% CI 0.513 to 0.616) ($p<0.0001$), according to the standards outlined by Landis and Koch (1977).[72]

Six of the 113 images included by panel assessment were subsequently excluded due to missing data or medical history involving a craniofacial condition. This resulted in a total of 107 images being available for final data analysis.

Images included by each assessor out of total n=145		
Panel members	Selected number of images included (%)	Selected number of images excluded (%)
A	123 (84.8%)	22 (15.2%)
B	117 (80.7%)	28 (19.3%)
C	103 (71.0%)	42 (29.0%)
D	113 (77.9%)	32 (22.1%)
E	100 (69.0%)	45 (31.0%)

Table 7: Outlining the number of images (and percentage of total images n=145) each individual assessor included or excluded.

5.2: Reliability

Ten percent of the sample was deemed an appropriate proportion of images to use for reliability testing (inter and intra) under the guidance of an experienced statistician. Twelve of the 113 initially included standard population 3dMD™ images (prior to the further exclusion of 6 images), and 12 of the 36 unicoronal synostosis images were selected to be used for intra and inter-reliability testing. Images were selected at random by statistician (GB) to reduce the potential for selection bias which is good practice when undergoing reliability testing.

5.2.1: Intra-observer Reliability

5.2.1.1: Intra-observer Reliability for all Landmarks to Nasion

The intra-class correlation coefficients (ICC) (one-way random effects) for intra-observer agreement of the distance between each landmark (n=21) to point Nasion were all >0.97 which is classed as excellent agreement. The ICC's, mean differences and 95% limits of agreement are listed for each landmark individually in Table 8. Landmarks; Labrale superious, Stomion, Labrale inferious, Pogonion, Christa philtri right and left and Cheilion right and left all demonstrated ICC's of 0.999.

Bland-Altman plots were constructed for each landmark to help assess for the possibility of systematic errors, and to illustrate the likely size of measurement errors. The 95% limits of agreement were used to judge whether the range of errors was acceptable for each landmark in terms of absolute difference <1mm. Figures 21a and b illustrate the agreement for Pronasale and Pogonion, these landmarks were selected as examples as they were used to evaluate nose and chin deviation by constructing angles. Pronasale positioning did not show meaningful signs of systematic error, with the mean difference of -0.18mm between the 2 values (upper and lower 95% limits of agreement: -0.86 to 0.50). Pogonion location had a slight tendency towards systematic error with the second set of measurements between

Pogonion and Nasion generally measuring as larger than the first (numbers tending to be of negative sign), the mean difference was -0.31mm (95% limits of agreement: -0.00 to 0.39). Agreement plots for intra-reliability for all landmarks can be found in Appendix 2. The results of the intra-observer agreement were discussed with the research team and deemed acceptable prior to the commencement of data collection.

Intra-observer reliability using liner distances between landmarks and Nasion			
Landmark relative to Nasion	Intra-class correlation coefficient (ICC)	Mean difference (mm) (SD)	95% limits of agreement (mm)
Midline landmarks			
Glabella	0.99	0.13 (0.30)	-0.46 to 0.71
Pronasale	0.99	-0.18 (0.35)	-0.86 to 0.50
Subnasale	0.99	-0.19 (0.29)	-0.76 to 0.38
Labrale Superious	0.99	0.08 (0.25)	-0.41 to 0.56
Stomion	0.99	0.06 (0.26)	-0.45 to 0.57
Labrale Inferious	0.99	-0.13 (0.42)	-0.96 to 0.70
Pogonion	0.99	-0.31 (0.35)	-0.99 to 0.39
Bilateral landmarks			
Endocanthion Right	0.98	-0.23 (0.43)	-1.06 to 0.61
Endocanthion Left	0.98	-0.07 (0.34)	-0.74 to 0.60
Exocanthion Right	0.99	-0.31 (0.34)	-0.98 to 0.36
Exocanthion Left	0.99	-0.13 (0.37)	-0.79 to 0.53
Palpebrale Superious Right	0.99	0.00 (0.41)	-0.80 to 0.80
Palpebrale Superious Left	0.99	0.22 (0.25)	-0.28 to 0.71
Palpebrale Inferious Right	0.99	-0.02 (0.41)	-0.83 to 0.80
Palpebrale Inferious Left	0.99	0.12 (0.30)	-0.48 to 0.71
Alare Right	0.99	-0.46 (0.22)	-0.90 to -0.02
Alare Left	0.99	-0.37 (0.30)	-0.97 to 0.22
Christa Philtri Right	0.99	-0.09 (0.36)	-0.80 to 0.61
Christa Philtri Left	0.99	0.02 (0.26)	-0.50 to 0.54
Cheilion Right	0.99	0.00 (0.39)	-0.77 to 0.77
Cheilion Left	0.99	0.08 (0.14)	-0.20 to 0.36

Table 8: Intra-observer reliability using liner distances between landmarks and Nasion.

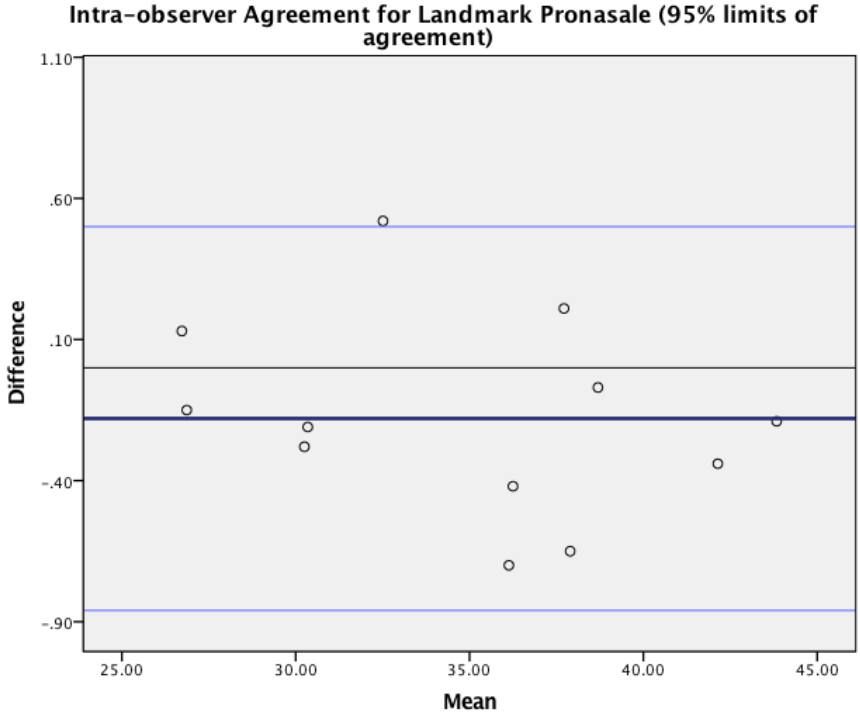


Figure 21a: Bland-Altman plot for intra-observer agreement for Pronasale; Mean difference: -0.18mm, Limits of agreement: -0.86 to 0.50mm.

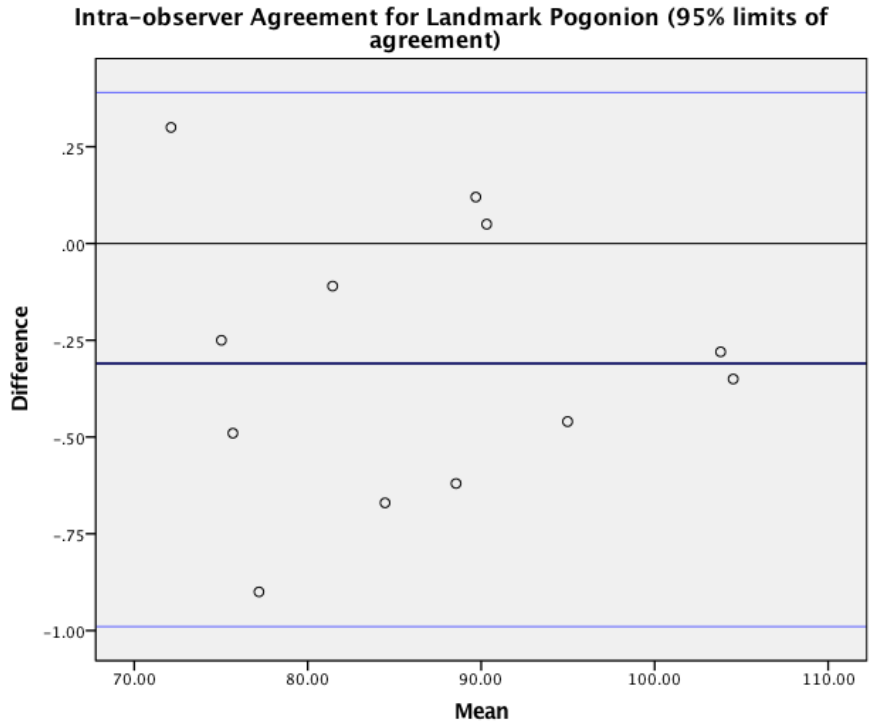


Figure 21b: Bland-Altman plot for intra-observer agreement for Pogonion; Mean difference: -0.31mm, Limits of agreement: -0.99 to 0.39.

5.2.1.2: Intra-observer Reliability for Landmark Nasion x, y and z Coordinates

The intra-class correlation coefficients (ICC) (one-way random effects) for intra-observer agreement of landmark Nasion positioning in terms of x, y and z coordinates were: 0.979, 0.998 and 0.999 respectively. This indicates excellent agreement and is supported by the low values for mean differences and the narrow confidence intervals shown in Table 9. Bland-Altman plots were constructed and can be seen in Figures 22a-c for x, y and z coordinates. The mean differences are close to zero (0.01, 0.02 and -0.01) indicating no suggestion of systematic error between the first and second measurements and thus excellent intra-observer reliability.

Intra-observer reliability of Nasion coordinates (x, y and z)			
Coordinate	Intra-class correlation coefficient (ICC)	Mean differences mm's (SD)	95% limits of agreement
X	0.979	0.01 (0.13)	-0.25 to 0.28
Y	0.998	0.02 (0.16)	-0.29 to 0.34
Z	0.999	-0.01 (0.02)	-0.05 to 0.03

Table 9: Intra-observer reliability of Nasion coordinates x, y and z.

Figures 22a-c: Bland-Altman Plots for intra-observer agreement for Nasion x, y and z coordinates.

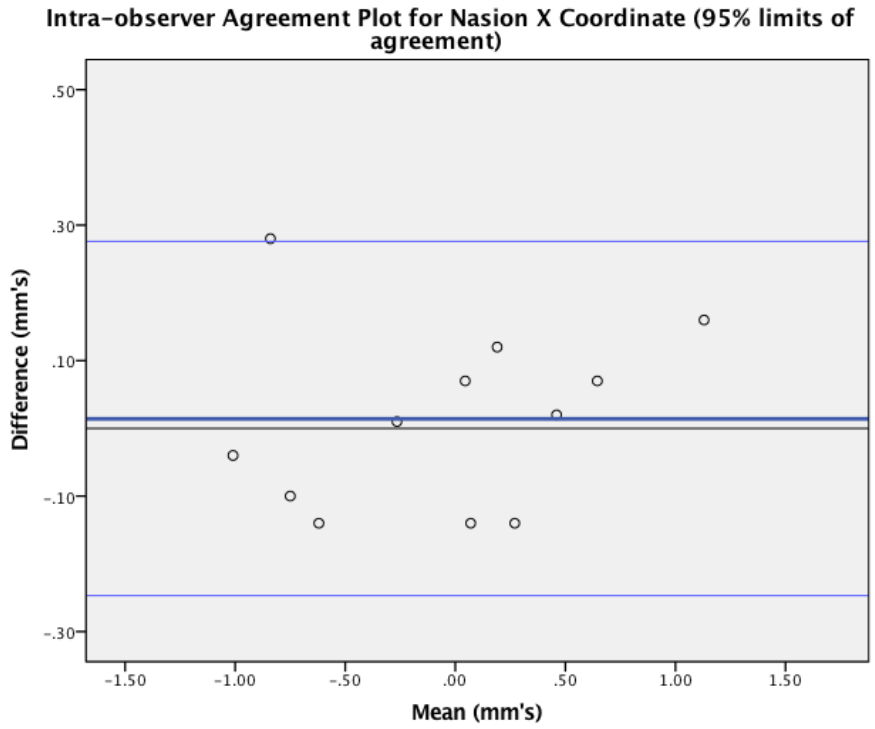


Figure 22a: Bland-Altman Plot for intra-observer agreement for Nasion x coordinate; Mean difference: 0.1mm, Limits of agreement: -0.25 to 0.28.

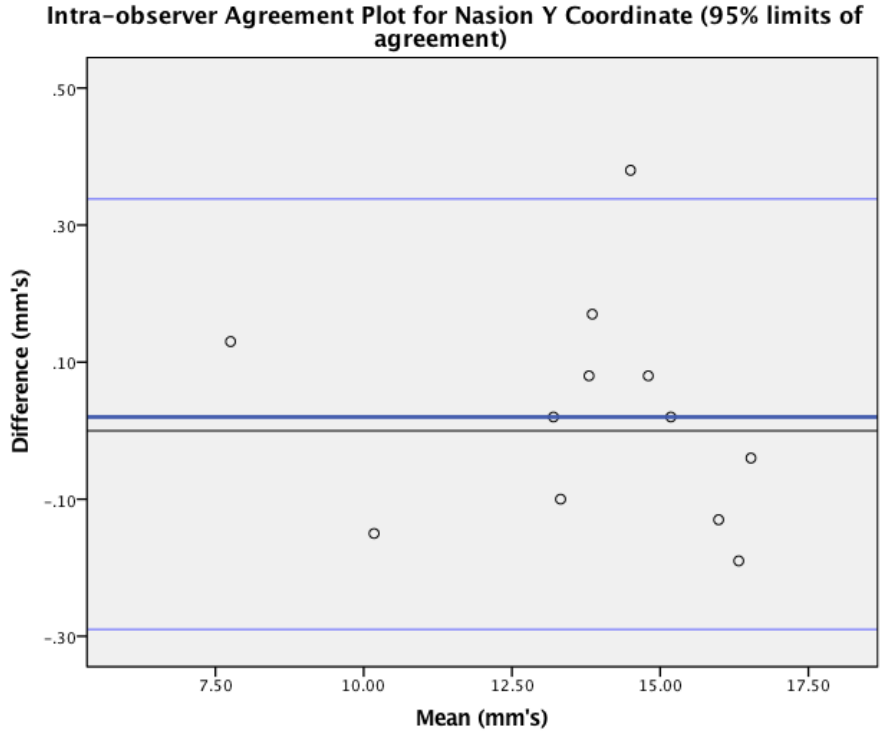


Figure 22b: Bland-Altman Plot for intra-observer agreement for Nasion y coordinate; Mean difference: 0.02mm, Limits of agreement: -0.29 to 0.34mm.

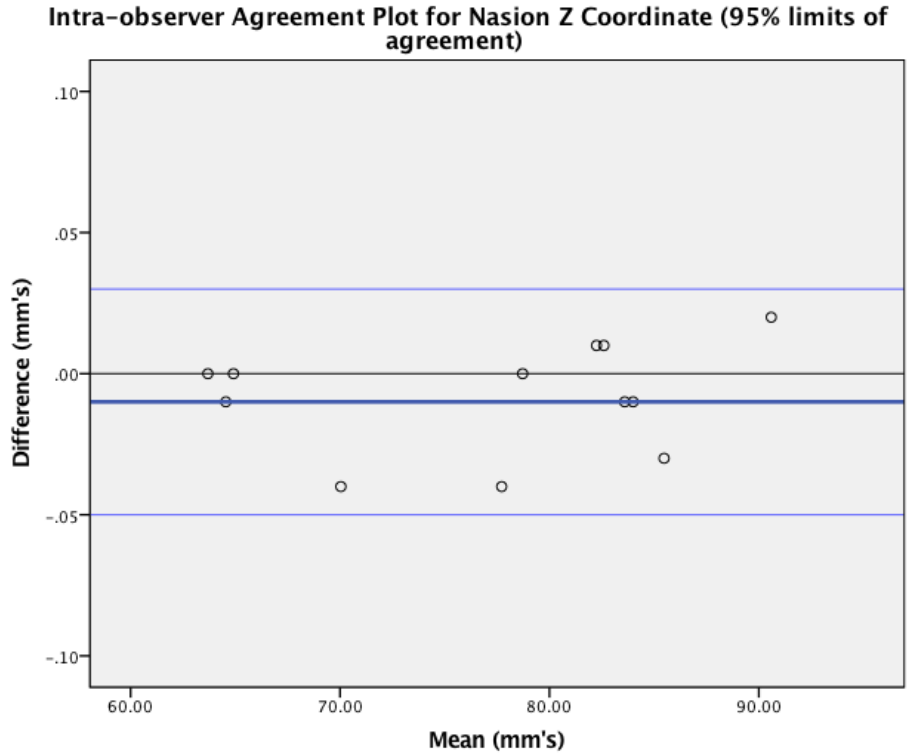


Figure 22c: Bland-Altman Plot for intra-observer agreement for Nasion z coordinate; Mean difference: -0.01mm, Limits of agreement: -0.05 to 0.03mm.

5.2.2: Inter-observer Reliability

5.2.2.1: Inter-observer Reliability for all Landmarks to Nasion

The initial inter-reliability assessment results can be seen in Table 10 for the 2 observers (OC and EB) analysing 24 images. Following further experience and training using the software, inter-reliability testing was repeated using another randomly selected 24 images to try to improve the agreement. The final ICCs, mean differences and 95% limits of agreement for inter-reliability are shown in Table 11. Bland-Altman plots were constructed for each landmark to help assess for the possibility of systematic errors, and to illustrate the likely size of measurement errors. The 95% limits of agreement were used to judge whether the range of inter-observer error was acceptable for each landmark in terms of absolute difference <2mm. The intra-class correlation coefficients ranged from 0.96 for Glabella, to ≥0.99 for landmarks: Pronasale, Subnasale, Labrale superious, Stomion, Labrale inferious, Pogonion, Alare right, Christa philtri right, Christa philtri left, Cheilion right and Cheilion left. Inter-observer agreement was therefore considered excellent for all landmarks. The mean differences for all landmarks were ≤0.52mm with standard deviations ranging from 0.38 mm (Endocanthion left) to 0.85mm (Pogonion). Pogonion and Pronasale, shown in Figures 23a and b, have again been selected as examples. All final Bland-Altman plots for inter-observer agreement may be found in Appendix 3.

First inter-observer reliability using liner distances between landmarks and Nasion			
Landmark relative to Nasion	Intra-class correlation coefficient (ICC)	Mean difference	95% limits of agreement
Midfacial landmarks			
Glabella	0.90	-0.56	-1.96 to 0.85
Pronasale	0.98	-0.05	-2.25 to 2.14
Subnasale	0.99	0.34	-1.59 to 2.27
Labrale Superious	0.99	0.18	-1.45 to 1.80
Stomion	0.99	0.18	-1.54 to 1.89
Labrale Inferious	0.99	0.12	-1.87 to 2.10
Pogonion	0.99	-0.67	-4.11 to 2.77
Bilateral landmarks			
Endocanthion Right	0.91	-0.46	-2.00 to 1.09
Endocanthion Left	0.87	-0.22	-1.86 to 1.43
Exocanthion Right	0.96	0.00	-1.58 to 1.58
Exocanthion Left	0.87	0.91	-1.35 to 3.18
Palpebrale Superious Right	0.89	-0.59	-2.36 to 1.19
Palpebrale Superious Left	0.74	1.10	-1.28 to 3.47
Palpebrale Inferious Right	0.70	-0.15	-2.10 to 1.80
Palpebrale Inferious Left	0.86	0.63	-1.63 to 2.90
Alare Right	0.97	-0.81	-2.91 to 1.29
Alare Left	0.98	-0.25	-2.32 to 1.81
Christa Philtri Right	0.99	-0.14	-1.90 to 1.62
Christa Philtri Left	0.99	-0.18	-1.88 to 1.51
Cheilion Right	0.99	-0.22	-1.81 to 1.37
Cheilion Left	0.99	0.03	-2.00 to 2.06

Table 10: First inter-observer reliability using to linear distances between landmarks and Nasion.

Second inter-observer reliability using liner distances between landmarks and Nasion			
Landmark relative to Nasion	Intra-class correlation coefficient (ICC)	Mean difference mm's (SD)	95% limits of agreement
Midfacial landmarks			
Glabella	0.96	0.19 (0.77)	-1.34 to 1.71
Pronasale	0.99	0.22 (0.60)	-0.95 to 1.39
Subnasale	0.99	0.06 (0.55)	-1.01 to 1.14
Labrale Superious	0.99	0.27 (0.47)	-0.66 to 1.19
Stomion	0.99	0.52 (0.42)	-0.31 to 1.34
Labrale Inferious	0.99	0.38 (0.56)	-0.72 to 1.48
Pogonion	0.99	-0.04 (0.85)	-1.70 to 1.62
Bilateral landmarks			
Endocanthion Right	0.98	0.25 (0.40)	-0.53 to 1.04
Endocanthion Left	0.97	0.43 (0.38)	-0.32 to 1.18
Exocanthion Right	0.99	0.08 (0.54)	-0.98 to 1.08
Exocanthion Left	0.99	0.13 (0.58)	-1.01 to 1.28
Palpebrale Superious Right	0.96	-0.5 (0.49)	-1.47 to 0.47
Palpebrale Superious Left	0.97	0.15 (0.58)	-0.99 to 1.30
Palpebrale Inferious Right	0.98	-0.04 (0.57)	-1.15 to 1.07
Palpebrale Inferious Left	0.98	0.27 (0.50)	-0.71 to 1.25
Alare Right	0.99	-0.22 (0.70)	-1.60 to 1.16
Alare Left	0.99	0.13 (0.74)	-1.31 to 1.57
Christa Philtri Right	0.99	0.04 (0.59)	-1.11 to 1.19
Christa Philtri Left	0.99	0.01 (0.51)	-0.99 to 1.02
Cheilion Right	0.99	0.05 (0.45)	-0.83 to 0.94
Cheilion Left	0.99	0.20 (0.47)	-0.72 to 1.11

Table 11: Second inter-observer reliability using to linear distances between landmarks and Nasion.

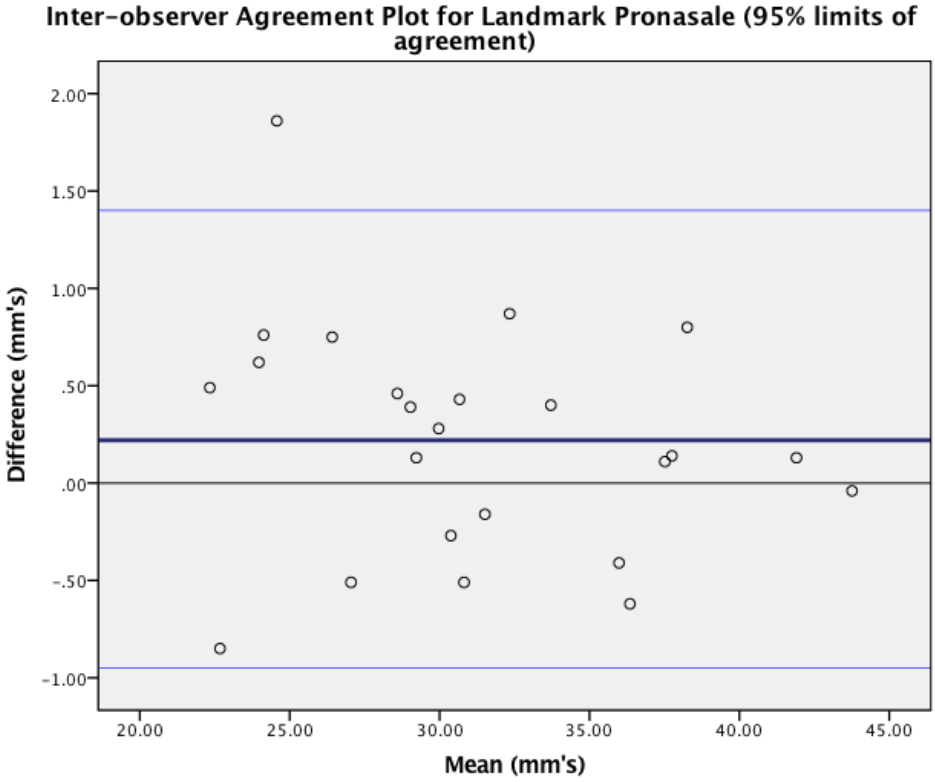


Figure 23a. Bland-Altman Plot for inter-observer agreement for Pronasale; Mean difference: 0.22mm, Limits of agreement: -0.95 to 1.39mm.

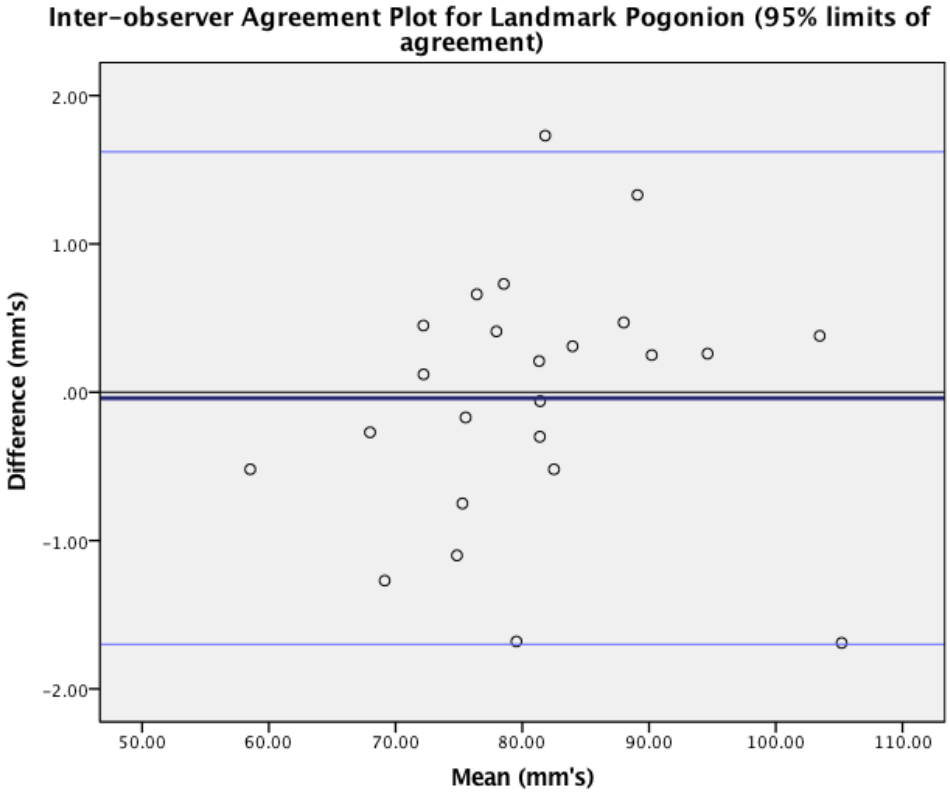


Figure 23b: Bland-Altman Plot for inter-observer agreement for Pogonion; Mean difference: -0.04mm, Limits of agreement: -1.70 to 1.62 mm.

5.2.2.2: Inter-observer Reliability for Landmark Nasion x, y and z Coordinates

As for intra-reliability, the inter-observer reliability was also tested for the x, y and z coordinates of Nasion individually. The inter-observer reliability for each coordinate of Nasion also demonstrated excellent agreement with intra-class correlation coefficients (one-way random effects) of 0.77, 0.99 and 0.99 (Table 12). The ICC was the least favourable for Nasion x coordinate (left to right) 0.77 with 95% levels of agreement -1.08 to 0.68 (Figure 24a). This is still considered to be excellent. The accuracy in the z (antero-posterior) direction was highly precise and is illustrated on the Bland-Altman plot with narrow 95% limits of agreement (Figure 24c). The mean differences were close to zero for all 3 coordinate planes which suggests that there were no apparent trends for systematic errors between the assessors. Bland-Altman plots were constructed to illustrate this and can be seen in Figures 24a-c for coordinates x, y and z respectively.

Inter-reliability of Nasion coordinates (x, y and z)			
Coordinate	Intra-class correlation coefficient (ICC)	Mean difference mm (SD)	95 % limits of agreement
X	0.77	-0.20 (0.45)	-1.08 to 0.68
Y	0.99	0.19 (0.42)	-0.64 to 1.02
Z	0.99	-0.03 (0.14)	-0.30 to 0.24

Table 12: Inter-observer reliability for Nasion x, y and z coordinates.

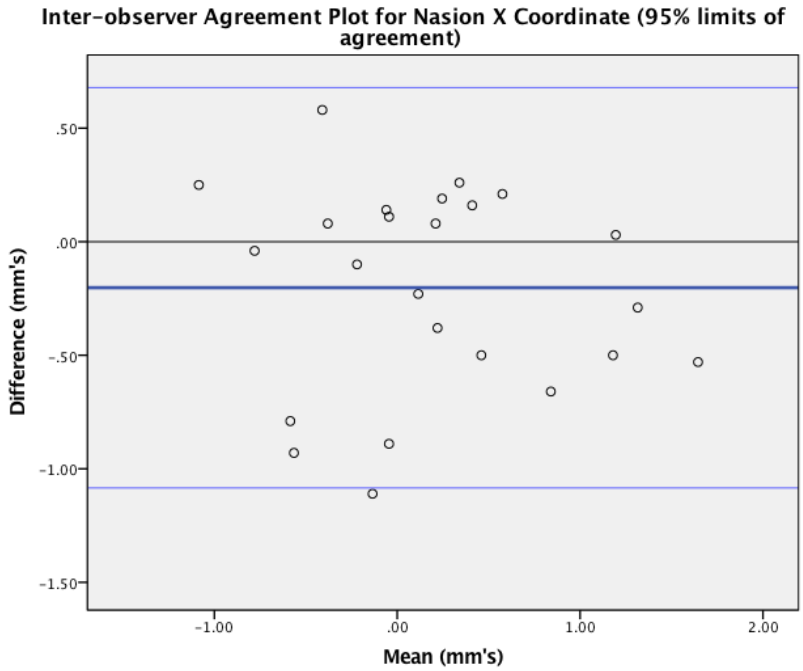


Figure 24a: Bland-Altman Plot for OC and EB inter-observer agreement for Nasion x coordinate; Mean difference: -0.20mm, Limits of agreement: -1.08 to 0.68mm.

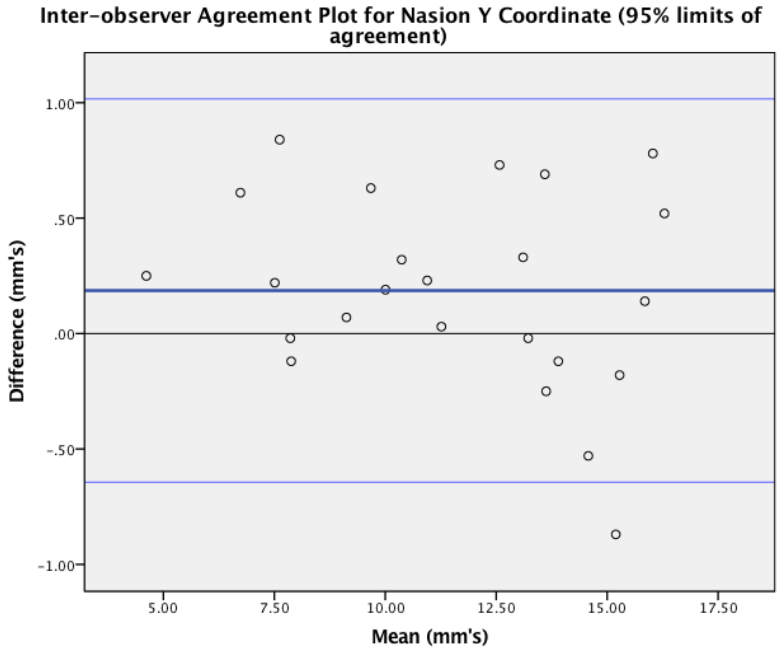


Figure 24b: Bland-Altman Plot for OC and EB inter-observer agreement for Nasion y coordinate; Mean difference: 0.19mm, Limits of agreement: -0.64 to 1.02mm.

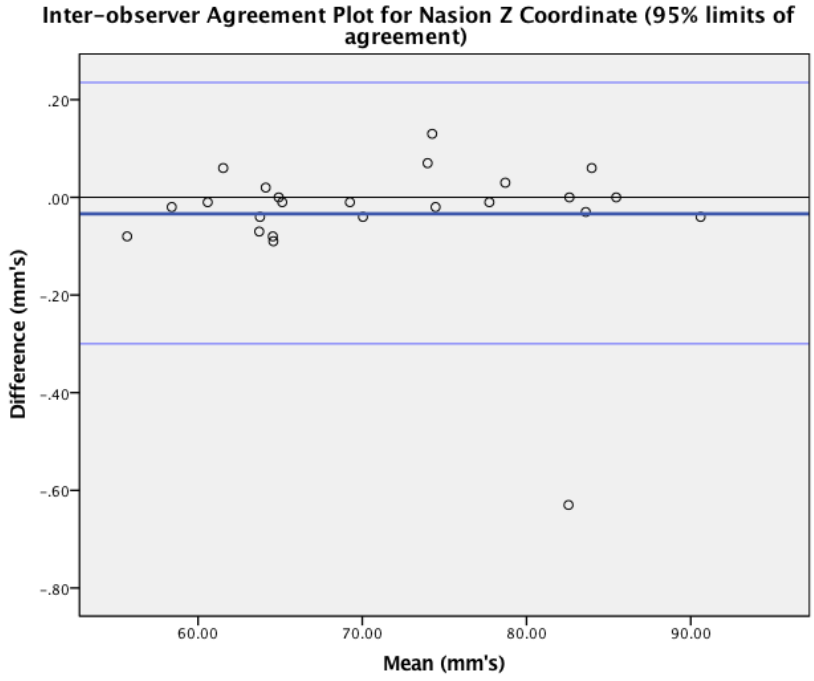


Figure 24c: Bland-Altman Plot for OC and EB inter-observer agreement for Nasion z coordinate; Mean difference: -0.03 mm, Limits of agreement: -0.30 to 0.24mm.

5.3: Descriptive Statistics of Sample

There were images of 154 individuals available for use in this study. After the initial exclusion of patients older than 16 years, missing images and those with apparent craniofacial

anomalies, 145 patients were assessed for inclusion by the panel. Following the panel assessment 113 images were included, with 32 images excluded due to reasons including active muscles of facial expression, mouth open, eyes closed and areas of missing data. A further 6 images were subsequently excluded due identification of a craniofacial condition or patient data not being fully accessible. This resulted in a final number of 107 images of different children to be analysed for this study. The flow of patients is illustrated in Figure 25.

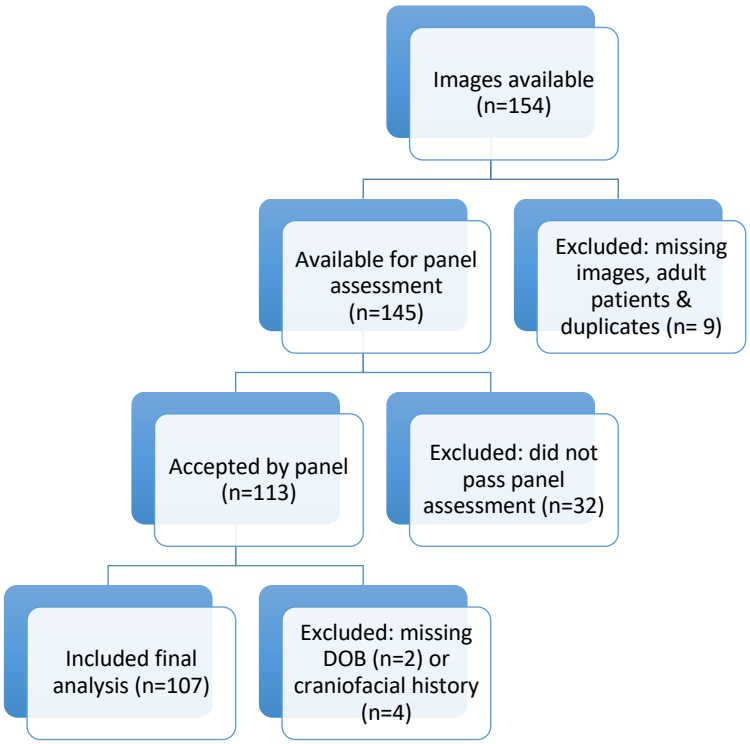


Figure 25: Flow diagram of images of individuals in the study

There were 57 males (53.3%) and 50 females (46.7%) included in the final sample (Table 13). The mean age was 7.36 years (SD 3.74 years). The ages ranged from 7 months to 14 years old (Figure 26). The mean age of males was 7.96 years (SD 3.51 years) (95% confidence intervals: 7.03, 8.90), which was higher than the mean age of females 6.68 years (SD 3.91 years) (95% confidence intervals: 5.57, 7.79).

Gender	Number of Participants	Mean Age (years)	Standard Deviation (years)	Minimum (years)	Maximum (years)	95% Confidence Interval (years)	
						Lower bound	Upper bound
Males	57	7.96	3.51	0	14	7.03	8.90
Females	50	6.68	3.91	0	14	5.57	7.79
Total	107	7.36	3.74	0	14	6.65	8.08

Table 13: Table to show breakdown of males and females included in the study and their age distribution.

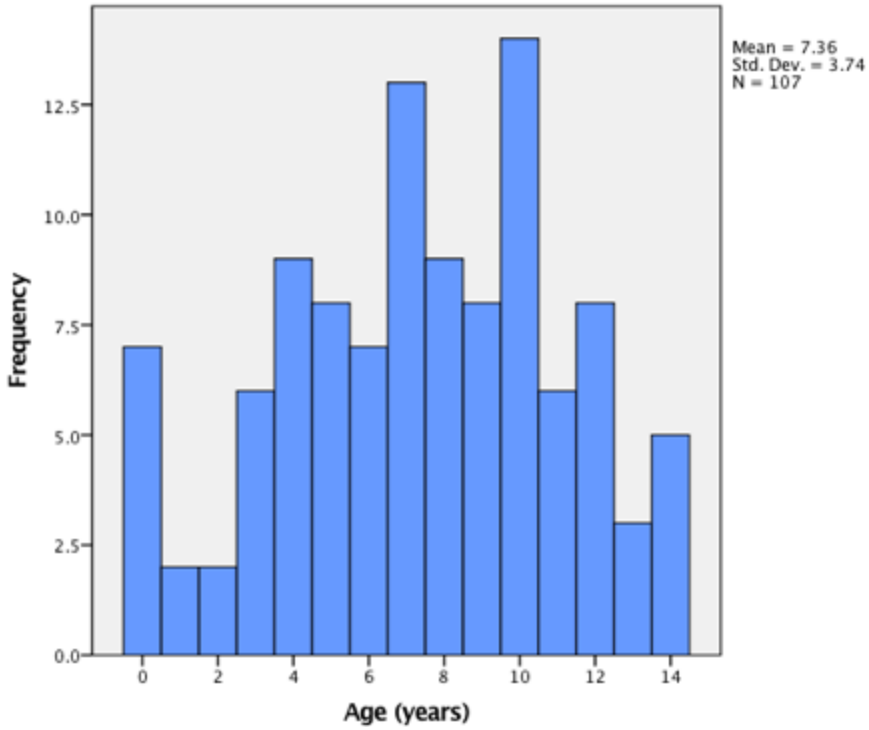


Figure 26: Age distribution of participants included in the study (n=107).

The number of children at each age (years 0-14) can be seen in Table 14. The greatest number of participants were age 7 years (n=13, 12.1%) and 10 years (n=14, 13.1%) when the images were taken. The ages with the least images were 1 and 2 years of age with each of these having 2 participants, representing 1.9% of the total sample each respectively.

Age (years)	Number of children in age group (out of total 107 images)	Percentage (%)	Cumulative Percent (%)
0	7	6.5%	6.5%
1	2	1.9%	8.3%
2	2	1.9%	10.2%
3	6	5.6%	15.7%
4	9	8.4%	24.1%
5	8	7.5%	31.5%
6	7	6.5%	38.0%
7	13	12.1%	50.0%
8	9	8.4%	59.3%
9	8	7.5%	66.7%
10	14	13.1%	79.6%
11	6	5.6%	85.2%
12	8	7.5%	92.6%

13	3	2.8%	95.4%
14	5	4.7%	100.0%

Table 14: Number of participants of each age (0-14 years) and percentage of final sample (n=107).

The ethnic grouping system used for the UK 2001 census was used to categorise the racial backgrounds of the included patients for this study (Table 15). The majority of children were White British (n=88, 82.2%) with small numbers representing other ethnicities including: Asian or Asian British, Black or Black British, other mixed racial backgrounds and White Irish.

Ethnicity	Frequency	Percent	Cumulative
White British (A)	88	82.2%	White British or Irish n=93 (86.1%)
White Irish (B)	4	3.7%	
Mixed White and Black Caribbean (D)	1	0.9%	Mixed backgrounds n=9 (8.3%)
Mixed White and Black African (E)	2	1.9%	
Mixed White and Asian (F)	3	2.8%	
Any other mixed background (G)	3	2.8%	
Asian or Asian British (Indian) (H)	2	1.9%	Asian or Asian British (Indian or other) n=4 (3.7%)
Asian or Asian British (any other Asian background) (L)	2	1.9%	
Black or Black British (any other black background) (P)	1	0.9%	Black or Black British n=1 (0.9%)
Unknown (X)	1	0.9%	Unknown (0.9%)

Table 15: Table to show the breakdown of ethnicities for final sample (n=107).

5.4: Asymmetry Index for Midfacial Anthropometric Landmarks (x Coordinate)

The asymmetry index (AI) for midline landmarks (Glabella, Nasion, Pronasale, Subnasale, Labrale superious, Stomion, Labrale inferious and Pogonion) was measured using differences in the x direction (left to right) as determined relative to the origin of the image. This directly indicated the amount of asymmetry in millimeters. Histograms were constructed, and the Shapiro-Wilk test applied to test for normality. As there was evidence of non-normality the data is presented in Table 16 as median values with interquartile ranges (IQR) (25^{th} and 75^{th} percentile) (in brackets), minimum and maximum values. The overall medial landmark asymmetry ranged between 0.60mm (IQR 0.24, 1.04) and 1.08mm (IQR 0.56, 1.75), with landmark Pogonion demonstrating the highest degree of asymmetry, and Nasion the least.

Differences in gender were investigated using the Mann-Whitney U test and are also shown in Table 16. Males were significantly more asymmetric than females in terms of landmarks: Glabella ($p=0.015$), Labrale inferious ($p=0.05$) and Pogonion ($p=0.041$). However, once the Bonferroni correction was applied to account for multiple hypothesis testing these differences no longer remained significant as they had $p>0.006$. This corrected significance value was calculated by dividing 0.05 by the number of tests used (8 midfacial). The highest maximum value for asymmetry was seen for Pogonion in one male (3.94mm).

Landmarks		X Coordinate		
		Median (IQR)	Minimum	Maximum
Glabella	All	.62 (.27, 1.16)	.01	2.28
	Males	.77 (.35, 1.37)	.01	2.28
	Females	.43 (.22, .96)	.01	2.27
	Difference (p)	.015*		
Nasion	All	.60 (.24, 1.04)	.00	2.28
	Males	.63 (.25, 1.19)	.01	2.28
	Females	.49 (.24, .90)	.00	1.98
	Difference (p)	.160		
Pronasale	All	.74 (.36, 1.27)	.00	3.07
	Males	.72 (.35, 1.31)	.01	2.85
	Females	.75 (.33, 1.28)	.00	3.07
	Difference (p)	.881		
Subnasale	All	.82 (.36, 1.37)	.00	2.81
	Males	.83 (.33, 1.38)	.01	2.81
	Females	.78 (.36, 1.34)	.00	2.72
	Difference (p)	.988		
Labrale Superious	All	.86 (.47, 1.59)	.01	3.89
	Males	1.15 (.50, 1.75)	.02	3.89
	Females	.84 (.39, 1.35)	.01	3.26
	Difference (p)	.189		
Labrale Inferious	All	.98 (.42, 1.70)	.02	3.6
	Males	1.33 (.32, 2.25)	.02	3.6
	Females	.79 (.46, 1.33)	.03	2.47
	Difference (p)	.050*		
Stomion	All	.90 (.40, 1.60)	.00	3.25
	Males	1.20 (.35, 1.80)	.02	3.25
	Females	.85 (.45, 1.23)	.00	2.37
	Difference (p)	.163		
Pogonion	All	1.08 (0.56, 1.75)	.00	3.94
	Males	1.18 (.64, 2.25)	.03	3.94
	Females	.99 (.44, 1.37)	.00	3.16
	Difference (p)	.041*		

Table 16: Asymmetry in the x coordinate for midfacial landmarks (Glabella, Nasion,

Pronasale, Subnasale, Labrale superious, Labrale inferious and Pogonion) in terms of median with inter-quartile range (IQR), maximum and minimum vales in millimetres. These are given for whole sample (All) and also for males and females separately. Gender difference was calculated using the Mann-Whitney U test, * indicates statistically significant prior to the Bonferroni correction.

To assess the relationship between age and asymmetry index of midfacial landmarks Pearson's correlation was used (Table 17). The absolute value of the x coordinate was used for each of the midfacial landmarks to prevent direction of asymmetry influencing the result. Six out of 8 of the landmarks demonstrated a statistically significant relationship between age and asymmetry: Pronasale 0.298 (p=0.002), Subnasale 0.317 (p=0.001), Labrale superious 0.242 (p=0.012), Stomion 0.297 (p=0.002), Labrale inferious 0.256 (p=0.008) and Pogonion 0.273 (p=0.004). However, after the Bonferroni correction was applied the correlation only remained significant for 4 midfacial landmarks: Pronasale (p=0.016), Subnasale (p=0.008), Stomion (p=0.016), and Pogonion (p=0.032). These correlations are illustrated using scatter plots in Figures 27a-d.

Correlation between midfacial landmark asymmetry indices and age		
Landmark	**Pearson's correlation**	**2-tailed significance**
Glabella	-0.075	0.443
Nasion	-0.005	0.956
Pronasale	0.298*	0.002
Subnasale	0.317**	0.001
Labrale superious	0.242*	0.012
Stomion	0.297**	0.002
Labrale inferious	0.256**	0.008
Pogonion	0.273*	0.004

Table 17: Outlining the relationship between asymmetry indices for midfacial landmarks and age in terms of: Pearson's correlation and 2 tailed significances. * = Correlation is significant at the 0.05 level (2-tailed). ** = Correlation is significant at the 0.01 level (2-tailed).

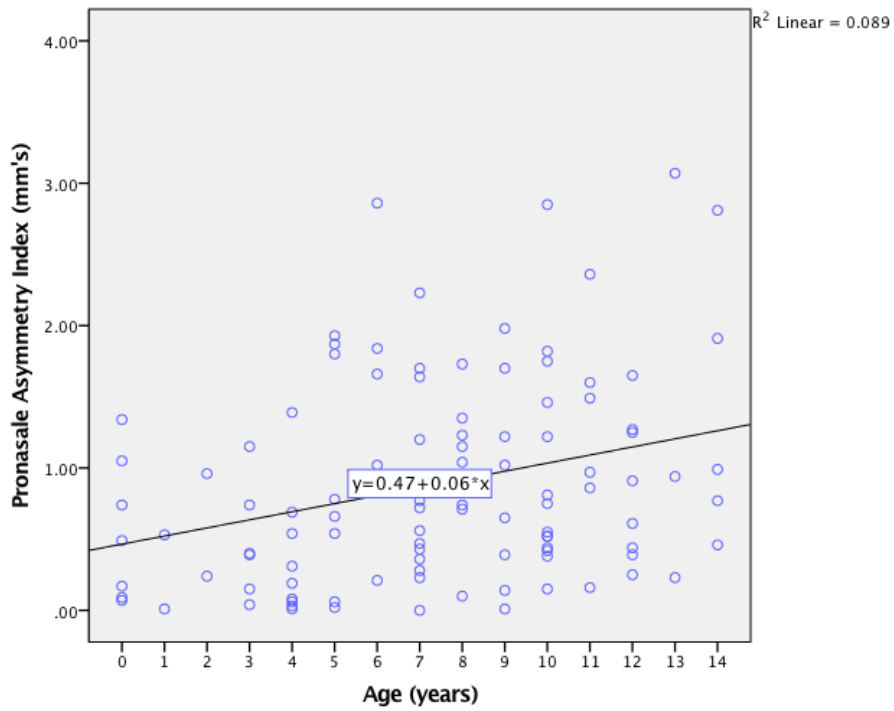


Figure 27a: Scatter plot illustrating the correlation between increasing age and increasing asymmetry index for landmark Pronasale (R^2 Linear = 0.089).

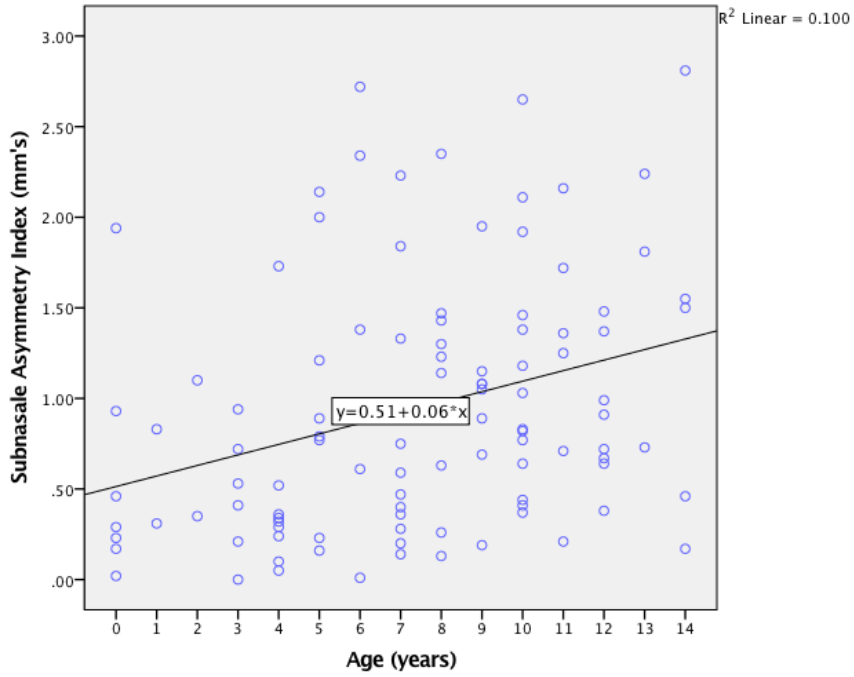


Figure 27b: Scatter plot illustrating the correlation between increasing age and increasing asymmetry index for landmark Subnasale (R^2 Linear = 0.100).

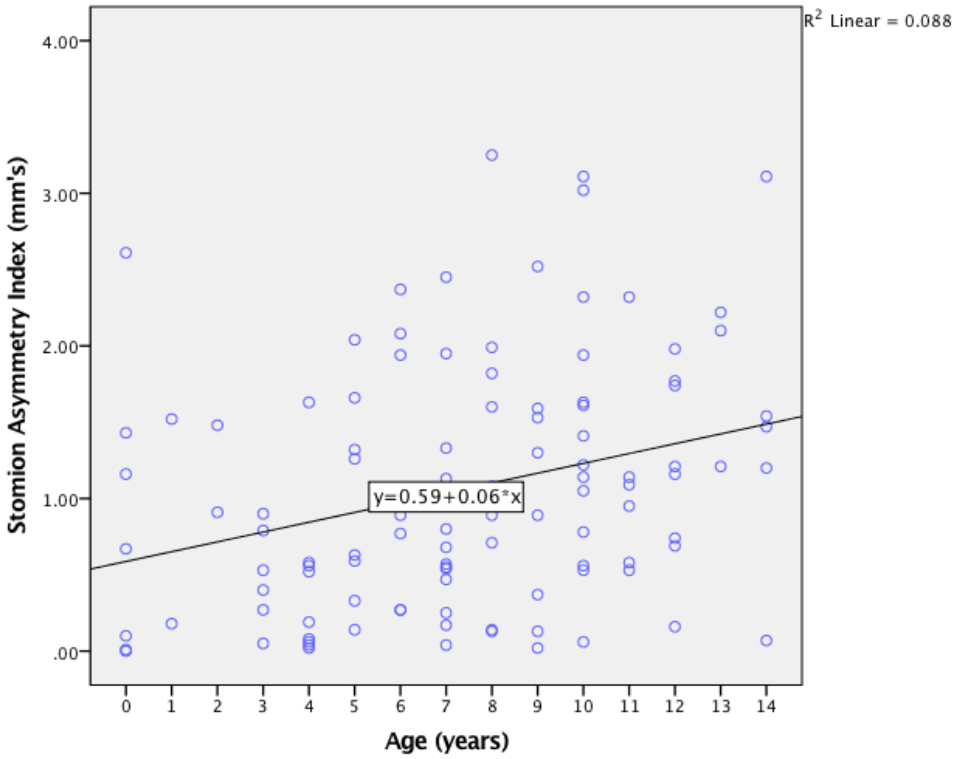


Figure 27c: Scatter plot illustrating the correlation between increasing age and increasing asymmetry index for landmark Stomion (R^2 Linear = 0.088).

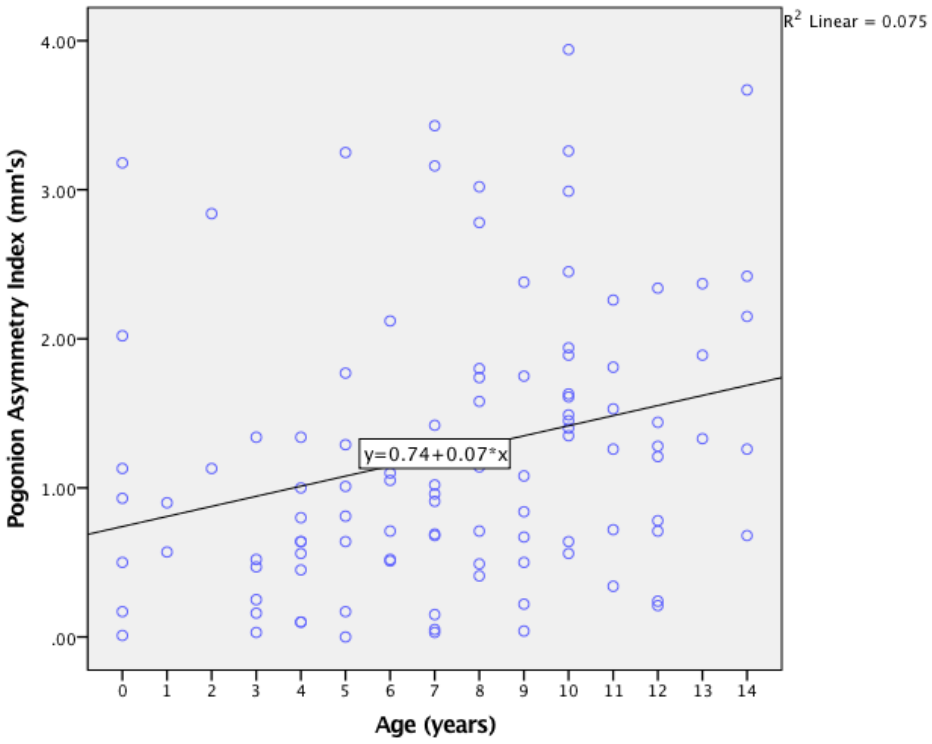


Figure 27d: Scatter plot illustrating the correlation between increasing age and increasing asymmetry index for landmark Pogonion (R^2 Linear = 0.075).

5.5: Asymmetry for Bilateral Anthropometric Landmarks using Coordinates: x, y and z

For the bilateral anthropometric landmarks (Endocanthion, Exocanthion, Palpebrale superious, Palpebrale inferious, Alare, Christa philtri and Cheilion), the asymmetry was measured in each of the 3 coordinate planes using the absolute values of the differences in x, y and z coordinates between the left and the right sides. The results are presented in Table 18 as medians with interquartile ranges (IQR) (25^{th} and 75^{th} percentile), minimum and maximum values as the data showed evidence of non-normality when assessed using histograms and the Shapiro-Wilk test.

The median values for all bilateral landmarks in the 3 coordinate planes were <2mm except for Christa philtri in the x direction which had a median of 2.02mm (IQR 0.83, 2.39). Overall, in the x direction the asymmetry ranged from 0.56mm (IQR 0.29, 1.16) to 2.02mm (IQR 0.83, 3.36). Christa philtri had the highest maximum difference in the x-coordinate plane of 6.49mm difference in one 8-year-old male. In the y direction (superior-inferior) the asymmetry ranged between 0.20mm (IQR 0.09, 0.36) and 0.59mm (IQR 0.24, 0.98). Each landmark in the y coordinate had a minimum difference of 0.00mm indicating no difference between the vertical landmark positioning in at least one male and one female. The only maximum value above 2mm in the superior inferior direction was 2.83mm for Cheilion in one 10-year-old male. In the z direction, the asymmetry ranged from 0.28mm (IQR 0.14, 0.51) for Christa-philtri to 1.45mm (IQR 0.65, 3.15) for Palpebrale superious. The median differences were ≤1mm for landmarks: Endocanthion. Palpebrale inferious, Alare, Christa philtri and Cheilion. Exocanthion and Palpebrale superious had median differences of 1.41mm (IQR 0.54, 2.19) and 1.45mm (IQR 0.65, 3.15) respectively. The maximum values for Palpebrale superious were very high for both males and females (23.99mm males and 26.53mm females) in the antero-posterior direction. Figure 28 illustrates the differences in the z direction for Palpebrale superious in a box and whisker plot for the 107 images (genders combined) with multiple outliers evident outside the 25^{th} and 75^{th} percentiles.

The Mann-Whitney U test was used to assess for significant differences in asymmetry between genders in all three coordinate planes. Males had more asymmetry in the y coordinate of 3 landmarks: Exocanthion (p=0.048), Palpebrale inferious (p=0.044) and Cheilion (p=0.017). These gender differences were no longer significant once the Bonferroni correction was applied. Therefore, overall there was no significant difference of asymmetry between genders for bilateral landmarks in terms of x, y and z coordinates in the population of children included.

Asymmetry for bilateral landmarks in 3 coordinate planes (mm's)							
Landmarks		X		Y		Z	
		Median (IQR)	Minimum, maximum	Median (IQR)	Minimum, maximum	Median (IQR)	Minimum, maximum
Endocanthion	All	.77 (.53, 1.30)	.04, 3.55	.35 (.16, .64)	.00, 1,43	.96 (.53, 1.5)	.01, 3.87
	Males	.75 (.36, 1.31)	.06, 4.01	.35 (.17, .68)	.00, 1.38	.93 (.61, 1.60)	.01, 3.87
	Females	.93 (.56,1.31)	.13, 3.55	.35 (.13, .63)	.00, 1.43	1.00 (.47, 1.45)	.01, 3.21
	Difference (p)	.352		.617		.532	
Exocanthion	All	1.08 (.40, 1.94)	.02, 4.91	.29 (.11, .53)	.00, 1.76	1.41 (.54, 2.19)	.00, 5.66
	Males	.95 (.37, 1.87)	.06, 4.01	.38 (.13, .61)	.00, 1.76	1.54 (.58, 2.56)	.01, 5.66
	Females	1.14 (.63, 2.06)	.02, 4.91	.23 (.03, .44)	.00, 1.01	1.09 (.48, 2.15)	.00, 5.62
	Difference (p)	.389		.048*		.327	
Palpebrale Superious	All	.56 (.29, 1.16)	.01, 4.13	.38 (.15, .70)	.00, 1.72	1.45 (.65, 3.15)	.04, 26.53
	Males	.56 (.28, .1.05)	.01, 2.26	.45 (.17, .91)	.00, 1.72	1.32 (.60, 2.77)	.04, 23.99
	Females	.56 (.29, 1.35)	.06, 4.13	.35 (.14, .57)	.00, 1.52	1.60 (.73, 4.24)	.10, 26.53
	Difference (p)	.290		.077		.422	
Palpebrale Inferious	All	.63 (.26, 1.2)	.02, 4.15	.32 (.17, .62)	.00, 1.52	.93 (.44, 1.67)	.02, 3.98
	Males	.53 (.26, 1.14)	.02, 2.70	.44 (.15, .66)	.00, 1.40	.92 (.44, 1.67)	.02, 3.98
	Females	.72 (.24, 1.21)	.02, 4.15	.25 (.19, .45)	.00, 1.52	.95 (.44, 1.74)	.05, 3.78
	Difference (p)	.494		.044*		.908	
Alare	All	1.49 (.65, 2.39)	.01, 5.01	.58 (.29, .95)	.00, .1.75	.68 (.33, 1.34)	.01, 2.47
	Males	1.49 (.78, 2.36)	.02, 5.01	.61 (.31, 1.00)	.00, 1.63	.67 (.34, 1.37)	.07, 2.35
	Females	1.42 (.59, 2.50)	.01, 4.39	.48 (.26, .92)	.00, 1.75	.68 (.31, 1.35)	.01, 2.47
	Difference (p)	.793		.339		.703	
Christa Philtri	All	2.02 (.83, 3.36)	.02, 6.49	.20 (.09, .36)	.00,.79	.28 (.14, .51)	.00, 1.83
	Males	2.48 (.78, 3.59)	.03, 6.49	.22 (.10, .38)	.00, .73	.30 (.14, .52)	.00, 1.83
	Females	1.89 (.84, 3.16)	.02, 5.40	.19 (.08, .36)	.00, 0.79	.28 (.14, .50)	.02, 1.08
	Difference (p)	.410		.392		.596	
Cheilion	All	1.78 (.81, 3.03)	.02, 5.97	.59 (.24, .98)	.00, 2.83	.90 (.50, 1.74)	.05, 4.15
	Males	2.05 (.65, 3.04)	.07, 5.97	.71 (.30, 1.12)	.00, 2.83	.99 (.51, 1.75)	.09, 4.15
	Females	1.76 (.93, 2.74)	.02, 4.99	.47 (.17, .83)	.00, 1.49	.89 (.50, 1.73)	.05, 3.54
	Difference (p)	.722		.017*		.776	

Table 18: Asymmetry in the x, y and z coordinates for bilateral landmarks (Endocanthion, Exocanthion, Palpebrale superious, Palpebrale inferious, Alare, Christa philtri, and Cheilion) in terms of median with inter-quartile range (IQR), maximum and minimum values in millimetres. These are given for whole sample 'All' and also for males and female separately. Gender difference was calculated using the Mann-Whitney U test, * indicates statistically significant prior to Bonferroni correction.

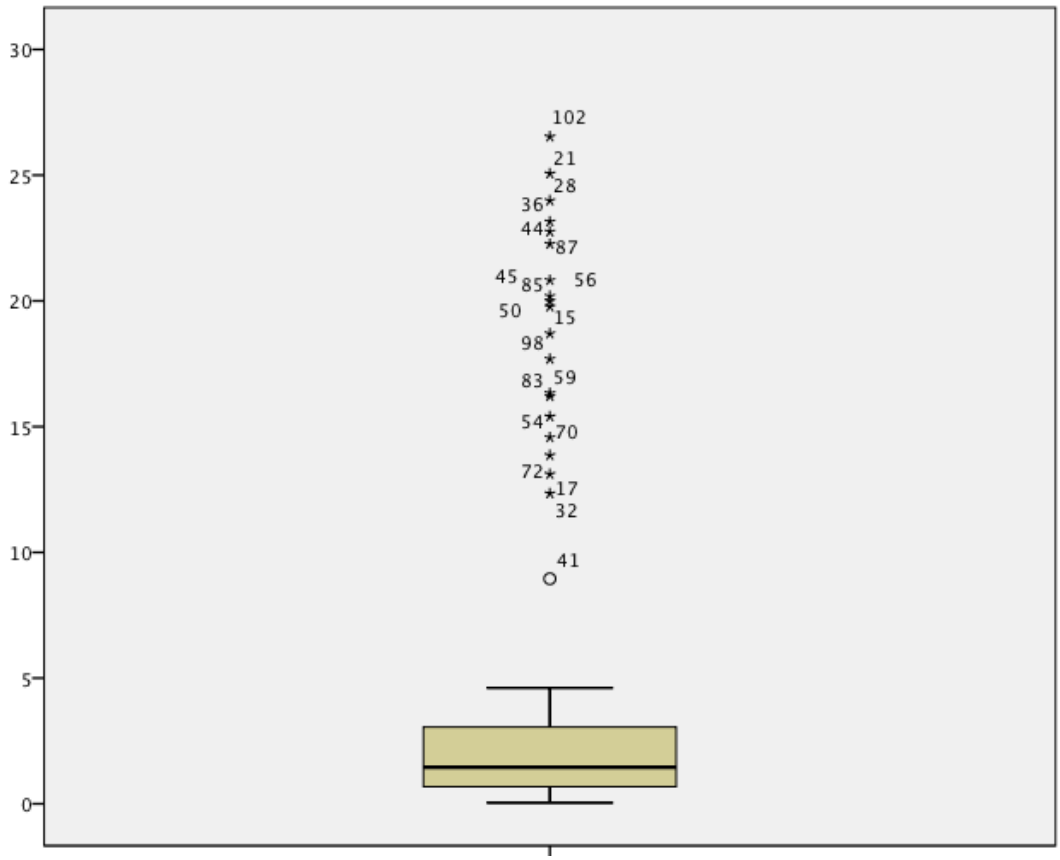


Figure 28: Box and whisker plot illustrating the differences in the z direction for landmark Palpebrale superious. Median difference: 1.45mm (IQR 0.63, 3.15), minimum 0.04, maximum 26.53mm with multiple outliers evident.

5.6: Asymmetry Index for Bilateral Anthropometric Landmarks

The asymmetry index (AI) provides an overall value (in mm's) of asymmetry for each bilateral landmark, taking into account discrepancies in all 3 planes of space. The results showed evidence of non-normality when assessed using histograms and Shapiro-Wilk tests and are presented in Table 19 as median values with interquartile ranges (25^{th} and 75^{th} percentiles). The bilateral landmarks with the least asymmetry were Endocanthion 0.56mm (IQR 1.15, 2.42) and Palpebrale inferious 1.56mm (IQR 1.06, 2.22). The highest median value for asymmetry was identified for Cheilion 2.56mm (1.69, 3.65). Minimum indices ranged from 0.25mm (Cheilion in females) to 0.65 (Alare in females) and maximum values ranged from 4.02mm (Endocanthion in females) to 6.72mm (Christa philtri in males). The application of the Mann-Whitney U test did not reveal any significant differences in the asymmetry indices between genders ($p>0.05$).

Landmarks		Asymmetry Index (mm)		
		Median (IQR)	Minimum	Maximum
Endocanthion	All	1.56 (1.15, 2.42)	.60	4.18
	Males	1.44 (1.13, 2.53)	.62	4.18
	Females	1.71 (1.26, 2.11)	.60	4.02
	Difference (p)	.685		
Exocanthion	All	2.16 (1.41, 3.34)	.41	5.78
	Males	2.16 (1.49, 3.33)	.41	5.78
	Females	2.16 (1.24, 3.35)	.50	5.62
	Difference (p)	.808		
Palpebrale Superious	All	1.68 (1.23, 3.29)	.35	4.71
	Males	1.73 (1.24, 2.28)	.44	4.71
	Females	1.65 (1.20, 2.29)	.35	4.54
	Difference (p)	.793		
Palpebrale Inferious	All	1.56 (1.06, 2.22)	.20	4.23
	Males	1.36 (1.03, 2.05)	.20	4.00
	Females	1.61 (1.11, 2.01)	.51	4.23
	Difference (p)	.617		
Alare	All	1.94 (1.27, 2.73)	.59	5.05
	Males	2.04 (1.40, 2.75)	.59	5.05
	Females	1.83 (1.18, 2.71)	.65	4.66
	Difference (p)	.392		
Christa Philtri	All	2.09 (.63, 3.46)	.32	6.72
	Males	2.51 (1.07, 3.70)	.32	6.72
	Females	1.92 (.88, 3.17)	.32	5.48
	Difference (p)	.288		
Cheilion	All	2.56 (1.69, 3.65)	.25	6.68
	Males	2.59 (1.73, 3.77)	.66	6.68
	Females	2.36 (1.64, 3.30)	.25	5.05
	Difference (p)	.318		

Table 19: Asymmetry index for whole sample 'All,' and for males and females separately, presented in terms of median with inter-quartile range (IQR), maximum and minimum values in millimetres. Gender difference was calculated using the Mann-Whitney U test.

To assess the relationship between age and asymmetry index for bilateral landmarks Pearson's Correlation coefficient was used (Table 20). Three out of the 7 bilateral landmarks demonstrated a statistically significant relationship between asymmetry and age: Alare 0.242 (p=0.012), Christa philtri 0.270 (p=0.005) and Cheilion 0.236 (p=0.015), with older patients

demonstrating more asymmetry. However, once the Bonferroni correction was applied, Christa philtri was the only bilateral landmark to show a significant positive correlation with age (p=0.035 after correction), which is illustrated using a scatter plot in Figure 29.

Correlation between bilateral landmark asymmetry indices and age		
Landmark	**Pearson's correlation**	**2-tailed significance**
Endocanthion	0.055	0.577
Exocanthion	0.056	0.565
Palpebrale Superious	-0.007	0.941
Palpebrale Inferious	0.014	0.883
Alare	0.242*	0.012
Christa Philtri	0.270**	0.005
Cheilion	0.236*	0.015

Table 20: Outlining the relationship between asymmetry indices for bilateral landmarks and age in terms of: Pearson's correlation and 2 tailed significances. * = Correlation is significant at the 0.05 level (2-tailed). ** = Correlation is significant at the 0.01 level (2-tailed).

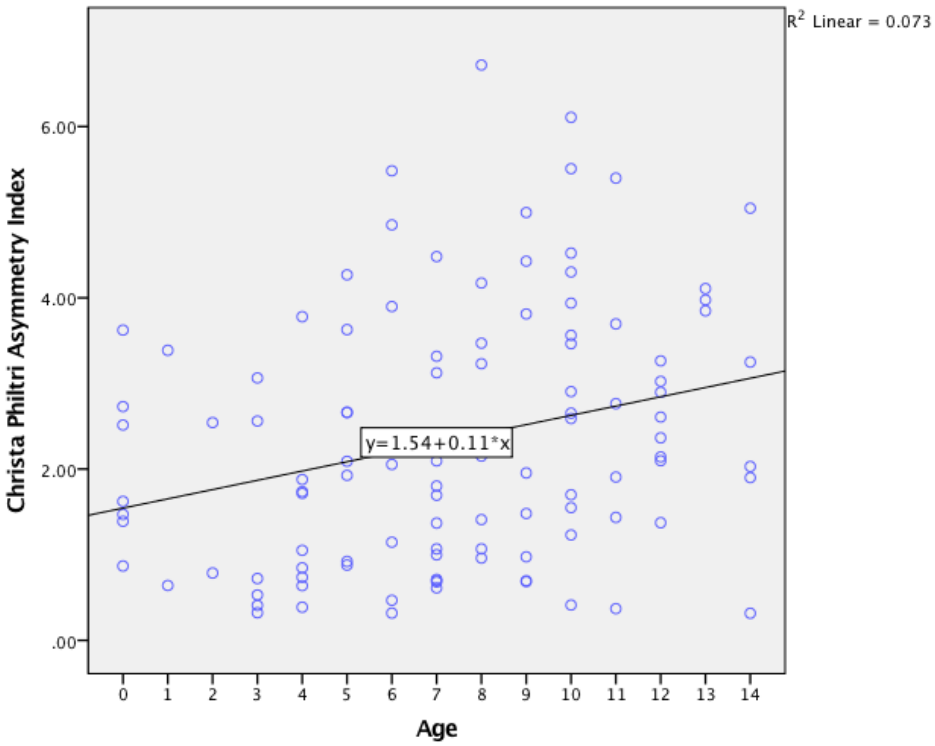


Figure 29: Scatter plot illustrating the positive correlation between increasing age and increasing asymmetry index for landmark Christa philtri (p=0.005) (R^2 Linear = 0.073).

5.7: Linear and Surface Measurements of Asymmetry between Bilateral Landmarks

The distances between each bilateral landmark and Nasion were measured using both linear and surface measurement tools (Vultus software). The differences between the right and left sides were then calculated for both methods to signify the asymmetry between bilateral landmark pairs in millimetres. The results are presented in Table 21 as medians with interquartile ranges (IQR) (25^{th} and 75^{th} percentile), minimum and maximum values as the differences between right and left sides did not follow normal distribution.

The absolute differences between bilateral landmark pairs using linear measurements ranged from 0.20mm (IQR 0.07, 0.37) for Christa philtri to 1.33mm (IQR 0.53, 1.97) for Exocanthion. The absolute differences for surface measurements ranged from 0.51mm (IQR 0.24, 0.97) for Alare to 2.60mm (IQR 1.16, 4.59) for Cheilion. There were larger maximum discrepancies between landmarks using the surface method with maximum values ranging from 3.37 to 18.35mm, compared to maximum linear values of 0.87mm to 4.97mm.

Paired t-tests were used to compare mean values between right and left sides for both linear and surface measurement methods, as the actual distances (rather than the differences) followed normal distribution. After applying the Bonferroni correction, there were significant differences between bilateral linear measurements for landmarks: Endocanthion ($p=0.001$), Exocanthion ($p<0.001$), Christa philtri ($p=0.006$) and Cheilion ($p=0.001$). There were also significant differences in bilateral surface measurements after the Bonferroni correction was applied for landmarks: Exocanthion ($p=0.001$) and Cheilion ($p=0.001$).

Overall differences were larger for surface measurements than for linear. This was expected due to the method of assessment. Paired t-test also shows a highly significant difference between means of linear and surface measurements for all landmarks ($p=<0.001$). The differences (mm's) between linear and surface measurements are presented in Table 22. There was no statistically significant difference between genders for bilateral landmarks using either linear or surface measurements when the Mann-Whitney U test was applied ($p>0.05$).

Landmark	Measurement type (linear or surface)	Median (IQR) difference between left and right sides (mm)	Minimum difference (mm)	Maximum difference (mm)	Mean difference (SD)	Paired t-test comparing right and left sides (Sig 2-tailed)
Endocanthion	Linear	.71 (.39, 1.33)	.02	3.10	0.38 (1.12)	0.001**
	Surface	.94 (.42, 1.61)	.00	3.37	0.28 (1.34)	0.030*
Exocanthion	Linear	1.33 (.53, 1.97)	.02	4.97	1.09 (1.33)	0.000**
	Surface	1.41 (.41, 2.56)	.01	9.10	0.73 (2.13)	0.001**
Palpebrale Superious	Linear	.76 (.37, 1.70)	.00	4.38	0.26 (1.47)	0.066
	Surface	1.67 (.86, 2.85)	.01	18.35	0.01 (3.68)	0.976
Palpebrale Inferious	Linear	.99 (.44, 1.74)	.02	4.42	0.14 (1.47)	0.325
	Surface	1.06 (.51, 1.96)	.00	4.13	0.30 (1.55)	0.047*
Alare	Linear	.47 (.24, .82)	.01	2.15	0.12 (0.73)	0.096
	Surface	.51 (.24, .97)	.01	4.12	0.09 (0.96)	0.361
Christa Philtri	Linear	.20 (.07, .37)	.01	.87	0.08 (0.31)	0.006**
	Surface	.75 (.256, 1.35)	.00	10.11	0.22 (1.98)	0.258
Cheilion	Linear	.73 (.47, 1.11)	.01	2.58	0.35 (0.90)	0.000**
	Surface	2.60 (1.16, 4.59)	.03	11.61	1.42 (4.25)	0.001**

Table 21: Differences in bilateral landmark pairs using both linear and surface measurement tools. Results are presented as median differences with inter-quartile ranges (IQR), maximum and minimum values in millimetres. Paired t-test in the right-side column shows significance of differences between bilateral landmarks; *=significant prior to Bonferroni correction and **=significant following the Bonferroni correction.

Landmark	Median (IQR) differences between linear and surface asymmetry measurements (mm)
Endocanthion	0.19 (0.08, 0.39)
Exocanthion	0.59 (0.30, 1.25)
Palpebrale Superious	0.58 (0.24, 1.62)
Palpebrale Inferious	0.51 (0.34, 0.93)
Alare	0.13 (0.06, 0.33)
Christa Philtri	0.48 (0.17, 0.99)
Cheilion	1.75 (0.75, 3.95)

Table 22: Median differences between linear and surface measurements with interquartile ranges (mm).

5.8: Pronasale and Pogonion

The angles formed by the deviation of the tip of the nose (Pronasale) and chin point (Pogonion) from the midfacial plane were analysed (Table 23). Histograms for Pronasale and Pogonion are illustrated in Figures 30a and b, demonstrating that the data for angles is skewed towards zero (symmetry). As the data did not follow normal distribution it is reported as medians with interquartile ranges (IQR) (25^{th} and 75^{th} percentile), minimum and maximum values. The median deviation of the nasal tip was 1.77° (IQR 0.84°, 3.12°) with 57% (n=61) deviating to the right side 42.1% (n=45) to the left and 0.9% (n=1) coincident with the midface. The median deviation of the chin point was 1.12° (IQR 0.52°, 2.51°). Fifty-four percent (n=54) of chin deviations were to the right side, with 48.6% to the left (n=52) and 0.9% (n=1) coincident with the midface.

Two participants Pronasale or Pogonion didn't deviate from the midface, these 2 individuals were excluded from further analysis of the extent (degrees) and side (right or left) of nose and chin deviation leaving 105 participants out of the total 107. There was a statistically significant correlation between the extent of nasal tip and chin deviation from the midfacial plane with a Pearson's correlation coefficient of 0.910 (2 tailed significance p<0.001) (Figure 30c).

Landmark	Angle (degrees)			Pearson's correlation between deviation of Pronasale and Pogonion (2-tailed significance) (n=105)
	Median (IQR)	Minimum	Maximum	
Pronasale Right n=61 (57.0%) Left n=45 (42.1%) None n=1 (0.9%)	1.77 (0.84, 3.12)	0.10	11.08	0.910* (p<0.001)
Pogonion Right n=54 (50.5%) Left n=52 (48.6%) None n=1 (0.9%)	1.12 (0.52, 2.51)	0.00	12.99	

Table 23: Outlines the angles of Pronasale and Pogonion relative to the midfacial plane in terms of medians with interquartile ranges (IQR) (25^{th} and 75^{th} percentile) (in brackets), minimum and maximum values. The Pearson's correlation coefficient 0.910* (p=0.00) indicates a high correlation between the extent of Pronasale and Pogonion deviation, *= statistically significant.

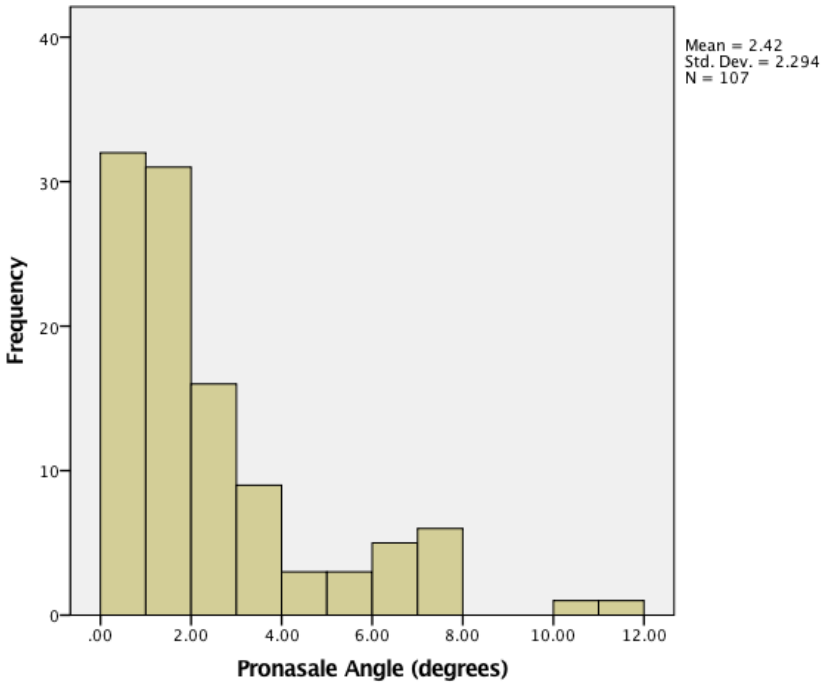


Figure 30a: Histogram illustrating the skewed distribution of the extent of the angle formed by Pronasale relative to the midface.

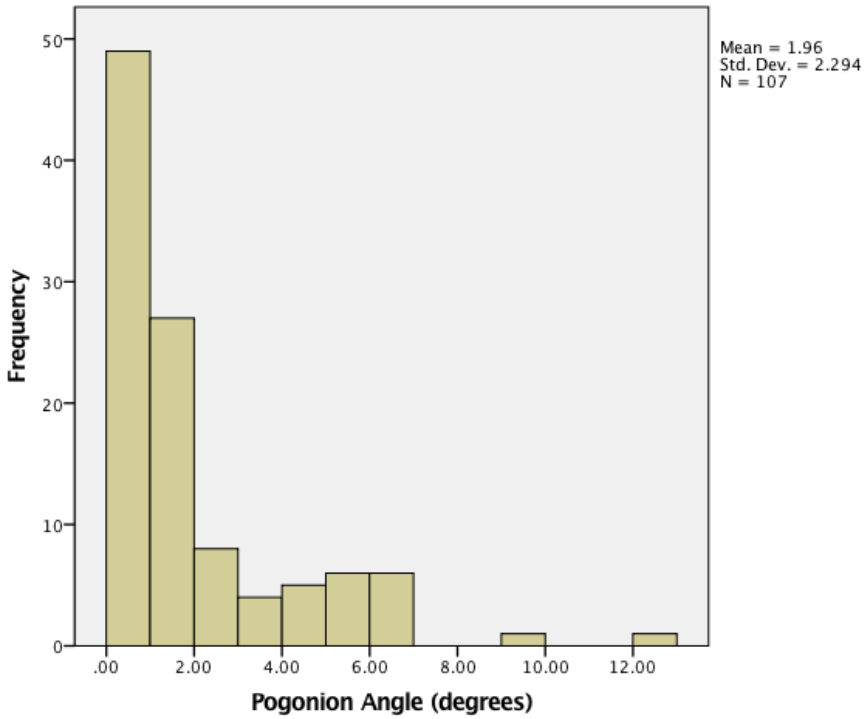


Figure 30b: Histogram illustrating the skewed distribution of the extent of the angle formed by Pogonion relative to the midface.

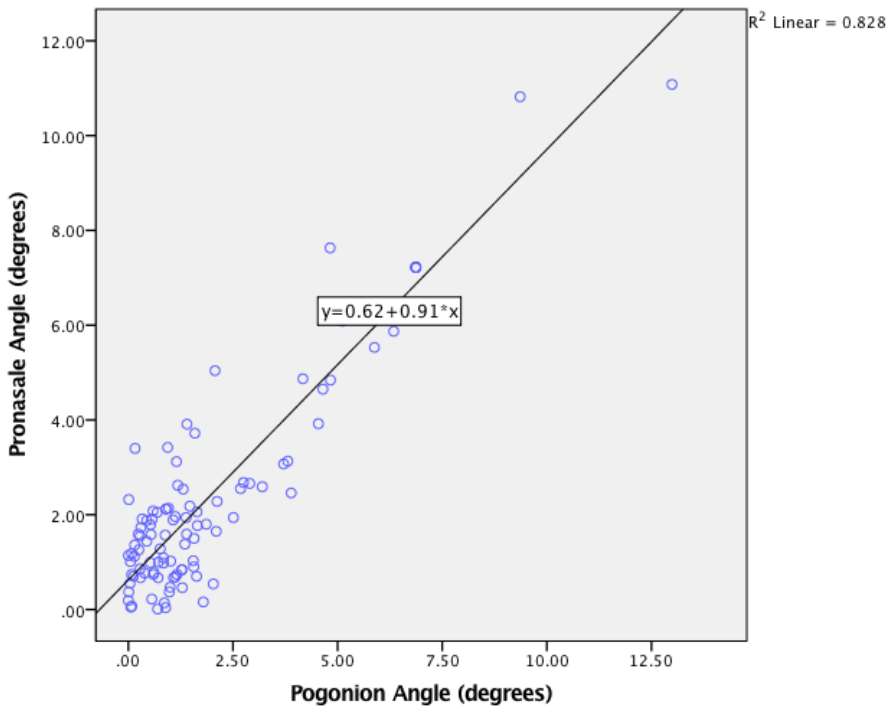


Figure 30c: Scatter plot illustrating the correlation between the extent of Pronasale and Pogonion deviation (degrees) (R2 linear = 0.828). Pearson's correlation 0.910 (2 tailed significance 0.00).

There was no significant difference observed in relation to the extent of deviation of Pronasale (p=0.11) or Pogonion (p=0.29) between genders when the Mann-Whitney U test was applied (n=105). There was no significant correlation between the age of participants and the extent of deviation of Pronasale or Pogonion with Pearson's correlations of 0.163 for Pronasale (2-tailed significance p=0.096), and 0.090 for Pogonion (2-tailed significance p=0.362).

A Mann-Whitney U test showed there was a relationship between side (right or left) of Pronasale deviation (n=105) and extent (number of degrees) (p=0.022). Left sided nasal deviations tended to be more severe (Table 24).

Side of Pronasale deviation	Median (degrees) (IQR)	Maximum and minimum values (degrees)
Left side Pronasale deviation (n=45)	2.05 (1.41, 3.27)	0.04, 11.08
Right side Pronasale deviation (n=60)	1.27 (0.71, 2.97)	0.01, 7.63

Table 24: Direction and severity of nasal tip (Pronasale) deviations outlined as medians (degrees) with 25th and 75th interquartile ranges.

The Mann-Whitney U test was also applied to assess the relationship between the side and extent of deviation for Pogonion, but no significant association was found (p=0.61) (Table 25). There was no significant relationship between extent of Pronasale deviation and side of

Pogonion deviation (Mann-Whitney U test 0.635). There was no significant relationship between the extent of Pogonion deviation and the side of Pronasale deviation (Mann-Whitney U test 0.087).

Side of Pogonion deviation	Median (degrees) (IQR)	Maximum and minimum values (degrees)
Left side Pogonion deviation (n=51)	1.29 (0.54, 2.51)	0.00, 12.99
Right side Pogonion deviation (n=54)	0.99 (0.53, 2.99)	0.00, 6.87

Table 25: Direction and severity of chin (Pogonion) deviations outlined as medians (degrees) with 25th and 75th interquartile ranges (IQR).

There was statistically significant association between the side (right or left) of Pronasale and Pogonion deviation, which was demonstrated using a Chi-square test (p=<0.001). The cross tabulation for side of nose and chin deviation can be seen in Table 26.

		Pronasale side (number of participants and %)		
		Left	Right	Total
Pogonion side (number of participants and %)	Left	37 (35%)	8 (8%)	45 (43%)
	Right	14 (13%)	46 (44%)	60 (57%)
	Total	51 (48%)	54 (52%)	105 (100%)

Table 26: Cross tabulation of the side of deviation of Pronasale and Pogonion

Chapter 6: Discussion

6.1: Limitations of the Study

6.1.1: Sample

The sample was taken from a group of individuals from the 'Headspace' project, who volunteered to have 3dMD™ images captured for research purposes. It could be hypothesised that due to potential volunteer bias these participants may not show the same level of asymmetry as a random sample obtained from the general population. However, this is unlikely as asymmetry was not the focus of the Headspace project, it was mainly promoted as a means of creating standard cranial values for the population to aid and develop future treatment methods. In addition to this, patients were enrolled regardless of sex, skeletal classification or dental malocclusion which should have reduced bias. Patients were excluded if they had a history of craniofacial syndromes or surgery but were not excluded based on the presence of facial asymmetry as the aim of this study was to quantify asymmetry in the standard population.

In contrast to other studies in this area,[7,73] this sample included children of all ethnicities in order to optimally represent the local population. The majority of the sample (85.9%) were Caucasian (White British 82.2% or White Irish 3.7%). Due to a lack of diversity, following advice from an experienced statistician, statistical tests were not thought to be appropriate to investigate potential differences between ethnic groups. It is unlikely that variations in ethnic origin would have influenced the level of asymmetry detected significantly as the non-Caucasian individuals were few in number.

6.1.2: Quality and Number of Images Available

There were images of 154 individuals available for use in this study, as previously illustrated in the flow diagram in section 5.3 (Figure 25). Nine of these were firstly removed due to missing images, adult patients and duplicates. After initial screening, it was apparent that the ear region was of limited quality in a number of images. Following discussion between all members of the research team it was agreed that ear landmarks, such as Tragion, would not be included in the final analysis of asymmetry so that as many images were eligible for inclusion as possible. A panel assessment was then undertaken to determine if the quality of images which were acceptable for inclusion.

In order to measure inter-rater agreement between panel assessors, Cohens weighted kappa (Fleiss-Cuzik extension) was used. This method was selected as data was categorical and there were more than 2 raters. Unlike a percentage of agreement, kappa accounts for the percentage agreement which has occurred by chance. The scale used to define the results

were those outlined by Landis and Koch (1977)[72] (see Table 27), which describe agreement ranging from 'no agreement' to 'almost perfect' based on the numerical value. A kappa of above 0.8 indicates very good agreement beyond chance.

The agreement between the 5 panel members was moderate with Cohens kappa=0.564 (95% CI 0.513 to 0.616), and the upper and lower 95% confidence intervals may be described between moderate and substantial agreement. There was a difference in the numbers of images included by assessors with one member accepting 123 images (84.4%) and another accepting only 100 (69%) out of the possible 147 images for inclusion in the study. This difference suggests that some members were more critical of image quality than others.

In particular when using a retrospective sample, the quality assessment of data available is imperative as images cannot be repeated. In an ideal research situation, there would be 'almost perfect' agreement between assessors to ensure the inclusion of a maximum number of high quality images. The literature demonstrates that studies which collect images prospectively are able to control quality by repeating images as required and applying exclusion criteria prior to image capture (for example those outside the age or ethnic criteria, or those with obvious facial asymmetry).[7,19,74] Other studies state that they have excluded images due to poor quality but do not outline the method of determining this conclusion, for example Hatch et al. (2017)[75] excluded 12 (3.7%) out of 325 3d facial images but did not report how this 'poor image quality' was assessed.[75]

There is a balance between quality and quantity as being hyper critical can lead to exclusion of many images, reducing the sample size, whereas on the other hand, accepting low quality images would lead to unreliable landmark positioning and inaccurate data. The use of 5 examiners and the resultant 'moderate' level of overall agreement suggests that the images selected for inclusion were of an acceptable quality for the purposes of this study. Overall 113 images were accepted following the panel assessment.

Kappa statistic	Strength of agreement
<0.00	Poor
0.00-0.020	Slight
0.021-0.040	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost perfect

Table 27: Agreement measurements for categorical data as defined by Landis and Koch (1977).[72]

6.1.3: Patients not Meeting Inclusion Criteria

Four of the 113 images accepted by the panel were subsequently excluded due to a history of craniofacial anomaly or surgery. A further 2 images were excluded as the dates of birth were unavailable. This left 107 3d facial images for inclusion in this research project.

6.1.4: Method of Measuring Facial Asymmetry

There was much discussion between the primary researcher, supervisors and 3dMD technician regarding how best to measure facial asymmetry. Various methods available which were outlined in the literature review (section 2.5) were considered.

Although landmark based 3d analysis of facial asymmetry has been reported to be reliable and reproducible, it has also been hypothesised that surface based 3d analysis may be less prone to measurement error as it doesn't require physical landmark positioning.[48]

6.2: Reliability

Reliability may be defined as "the degree to which the result of a measurement, calculation or specification can be depended on to be accurate."[76] As an essential consideration in research, reliability represents the reproducibility of a measurement when repeated at random, which in turn has implications when comparing results to both historical and future data. Using the results of reliability testing, the sensitivity, specificity and power of a statistical test employing the measurement may be determined. As reliability improves, so too does the sensitivity and specificity of the method. In order to calculate reliability, a quality control program should be employed with duplication of measurements in a randomly selected proportion of a sample obtained at different time points. The variation between time points that are found not to be due to measurement error is a proportion which is represented by the coefficient of reliability.[77] Intra, in addition to inter-reliability were assessed in this study to present the consistency of landmark identification and measurements, and also to facilitate future comparison of this 'standard population' with data from other research projects including a parallel DDSc project comprising of participants with operated non-syndromic unicoronal synostosis. For intra-reliability 12 images were considered to be appropriate as it represented >10% of the sample population, which was the same proportion selected by Darby et al. (2015)[30]. However, their sample size was 40 resulting in just 4 images being measured for intra-reliability testing. Inter-reliability was completed by OC and a fellow DDSc student (EB) using randomly selected 12 'standard population' images and 12 images of patients with operated unicoronal synostosis. Other research including Huang et al. (2013)[7] and Hajeer et al. (2004)[61] randomly selected 10 images for reliability testing out of their total sample sizes of 60 and 44 respectively. Intra-reliability is commonly tested by completing facial landmark positioning on

2 separate occasions at a set time apart, although 3 occasions have also been reported in the literature.[61] The 'wash out' periods used have mainly included 1 week[61] and 2 week intervals.[7,45,73] A minimum of 2 weeks was viewed as sufficient wash out without unduly delaying the progress of the research.

6.2.1: Intra-reliability

The intra-reliability was excellent for all midfacial and bilateral landmarks with intra-class correlation coefficients ranging from 0.98 to 0.99. It was decided to use the linear measurements initially and if there was an issue with reliability the x, y and z coordinates for each landmark would have been examined separately. This was agreed following statistical advice. If there were errors in landmark identification this would have been identified by discrepancies in linear measurements. The intra-reliability for x, y and z coordinates of Nasion demonstrated excellent agreement with ICCs of 0.979, 0.998 and 0.999 respectively.

According to Aldridge et al. (2005)[58], whom comparably used a landmark based approach on 3dMD stereophotogrammetric images, mean differences between 2 repeated measurements of the same image may be considered highly precise for means <1mm, and precise for means between 1-2mm. Taking this into consideration the intra-reliability for all landmarks was highly precise with mean differences <1mm for all midfacial and bilateral landmarks and Nasion x, y and z coordinates, except for midline landmark Glabella which demonstrated a mean difference of 1.13mm (95% limits of agreement -0.46 to 0.71mm), which remains to be deemed precise. Interestingly the research of Aldridge et al. (2005)[58] showed Glabella to be less precise, with a mean error of up to 2mm. Toma et al. (2009)[73] also noted poor intra and inter-examiner reproducibility of Glabella. Therefore, the findings of this research are in agreement with previous literature.

In addition to the ICCs, similarly to the methods used by Toma et al. (2009)[73], Bland-Altman plots were constructed to aid in the assessment and illustration of intra-reliability and to identify possible potential systematic errors between the first and second measurements for each image. The Bland-Altman plot for Alare right can be seen in Figure 31. This landmark had the highest intra-reliability mean difference of -0.46mm had also a tendency towards systematic error with values on the second occasion measuring slightly larger than on the first occasion. All Bland-Altman plots for intra-reliability may be found in Appendix 2.

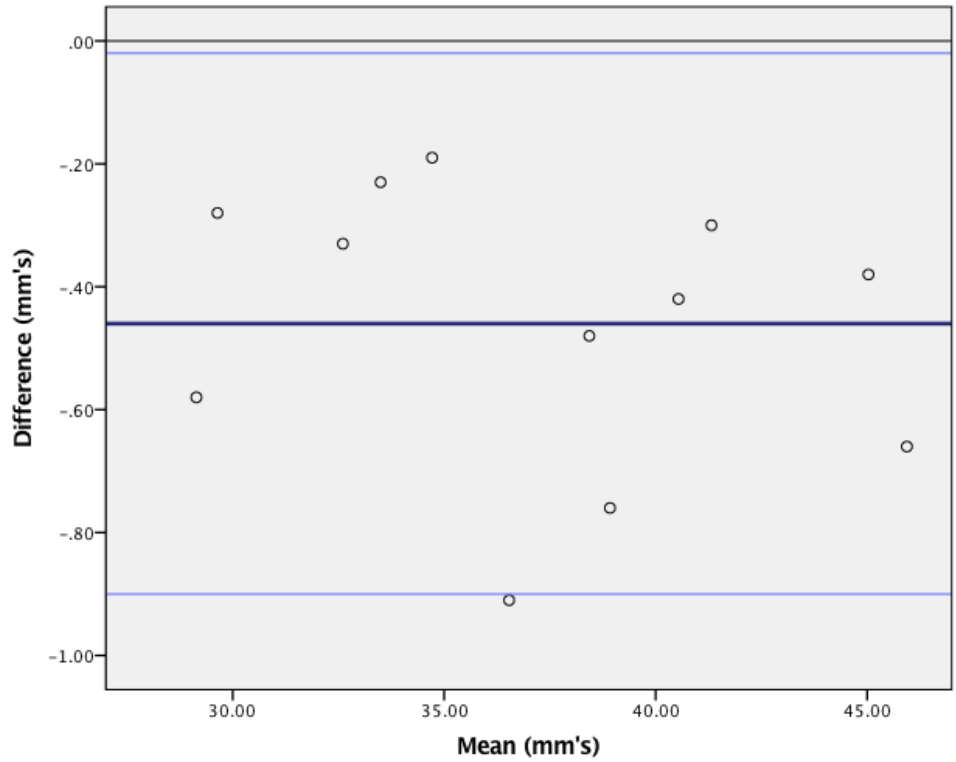


Figure 31: Bland-Altman plot for intra-observer agreement for Alare right. Mean difference: -0.46mm, Limits of agreement -0.90 to -0.02mm. The reliability is considered as moderate (<1mm).[73]

6.2.2: Inter-reliability

Intra-class correlation coefficients and Bland-Altman agreement plots were similarly used to assess inter-observer reliability. The majority of studies analysing facial asymmetry using the landmark based approach on 3d images have one investigator and thus do not report inter-reliability.[7,74,78] Research that has compared intra and inter-reliability have reported poorer reproducibility of the inter-examiner assessment.[73] In order to compare results with future studies including different examiners, it is important to be aware of the reproducibility of the technique used in terms of inter-reliability. For midfacial and bilateral landmarks the ICCs ranged from 0.96 (Glabella) to 0.99, signifying excellent agreement.[79] There was also excellent agreement for the y (super–inferior) and z coordinates (anterior-posterior) of Nasion with ICCs of 0.99. Nasion in the x direction (left to right) had an ICC of 0.77 (95% limits of agreement -1.08 to 0.68mm) which suggests good reliability.[79] The mean difference in the x coordinate was -0.20mm, y coordinate was 0.19mm and z coordinate -0.03mm. In contrast to this, Aldridge et al. (2005)[58] reported reliability to be poorest in the y direction for landmark Nasion with average mean differences of 0.43, 1.35 and 0.71mm in the x, y and z directions respectively which is slightly higher than the mean differences detected in this research. The reduced reliability for Nasion in the y direction was also described by Toma et al. (2009)[73]

using a sample of British Caucasian 15-year old children. The inter-reliability Bland-Altman plots for Nasion may be seen in Figures 24a-c in section 5.2.2.2 of the results.

Overall the intra and inter-reliability results indicate that the landmark identification applied to 3dMD images in this study was highly repeatable and precise. In general, the mean error associated for individual landmarks was <1mm which is comparable to other studies in this area and is considered moderate.[58,63,73]

6.3: Asymmetry Index of Midfacial Landmarks

The asymmetry index (AI) for midfacial landmarks was calculated using the x coordinates. As the data did not follow normal distribution the AI is presented in terms of median values and interquartile ranges (25th and 75th percentile) as this is reported to be less sensitive to the influence of outliers.[80] The median facial asymmetry was <1mm for all midfacial landmarks, except for Pogonion which had a median asymmetry of 1.08mm (IQR 0.56, 1.75mm). The results are compared with that of Alqattan et al. (2015)[48], who used a similar method to quantify facial asymmetry in a population sample which included 85 Caucasian adults whom were recruited in Cardiff University. These results are compared in Table 28 in which their data for males and females have been reported separately.

6.3.1: Nasal Region

The children in this study had a median asymmetry of 0.72mm (IQR 0.35, 1.31) in males and 0.75mm (IQR 0.33, 1.28) in females for Pronasale (nasal tip) which is slightly higher than the median asymmetry of 0.1mm (IQR 0.1, 0.3) for males and 0.2mm (IQR 0.1, 0.3) for females documented by Alqattan et al. (2015).[48] This difference in populations is further highlighted by the variability of the interquartile ranges between the results. This is also higher than the asymmetry observed in a child population (8-12 year olds) reported by Bugaighis et al. (2013)[2] who reported the mean asymmetry in males to be 0.36mm (SD 0.31mm) and in females 0.51mm (SD 0.49mm). However they used a different method of asymmetry assessment including reflection of the image and Procrustes analysis to orientate and best fit the images together, followed by the measurement of the distance between the landmarks and their corresponding reflected point to gain a value of asymmetry.[2] They had strict criteria, including only Caucasian children with a class 1 dental occlusion, competent lips and 'harmonious balanced faces' which is quite a specific cohort rather than a true reflection of the general population. Therefore, comparison between their results and the results of this study should be interpreted with caution.

Regardless of the differences between this study and some of the previous literature, it must be considered if the level of asymmetry measured is clinically or socially relevant. The current evidence regarding the discriminative thresholds of facial asymmetry have been reported in a

systematic review by Wang et al. (2017)[8] as discussed previously in the literature review (section 2.1.4).

The visual perception threshold for nasal asymmetry has been reported as 4mm, nonetheless it also has been demonstrated that 4mm nasal deviation is significantly more detectable than 2mm.[23] Bearing this in mind, in the present study the overall median asymmetry for Pronasale was 0.74mm (males and females combined) (IQR 0.36, 1.27mm) with no statistical difference identified between genders (p=0.881) using the Mann-Whitney U test. None of the sample population had nasal tip deviation of >4mm relative to the midface with the maximum asymmetry observed being 3.07mm in a 13-year-old male. Therefore, it may be concluded that although this sample had slightly more asymmetry of Pronasale than previous studies on differing populations often with more stringent exclusion criteria, there was no clinically significant asymmetry of the nasal tip present in the children of this study.

6.3.2: Chin Region

When examining Pogonion, the results indicate the sample population has slightly less asymmetry in the chin region with median asymmetry of 1.08 mm (IQR 0.56, 1.75mm), when compared to previous studies using a similar technique for adult populations which reported chin asymmetry as 1.15mm (IQR 0.7, 3.0mm) for males and 1.8mm (IQR 0.7, 2.5mm) for females (see Table 28).[48] The results exhibit slightly more asymmetry in children in England than that recorded by Bugaighis et al. (2013)[2] who reported mean asymmetry of landmark Pogonion of 0.36mm (SD 0.31) for males and 0.51mm for females (SD 0.49). Unlike our sample, these patients were recruited based on 'harmonious balanced' thus may not truly represent the standard UK population.[2]

The highest maximum value for asymmetry was 3.94mm in Pogonion for a 10-year-old male, which is much less than the maximum of 7.6mm asymmetry observed in one adult female in the Cardiff study.[48] Huang et al. (2013)[7] used Menton instead of Pogonion to quantify chin symmetry in an adult Chinese population (n=60) with Class I dental occlusion and faces with 'normal symmetry'. Correspondingly to the findings of this study, the authors concluded that the chin area had the most asymmetry when compared to other areas of the midface with a mean asymmetry index of 1.54mm (SD 1.50mm) and a maximum value of 6.40mm observed.

To consider the potential implications of our findings it is pertinent to discuss the existing literature which suggests that asymmetries in the region of the chin of up to 5mm[20] and 6mm[23,24] are minimally detectable and unlikely to influence attractiveness. Therefore, considering a median asymmetry value of Pogonion of 1.08 mm (IQR 0.56, 1.75mm), there was not clinically significant asymmetry of the chin point present in the sample population.

Landmarks		X Coordinate (n=107)			X coordinate Alqattan et al. (2015) (n=85)		
		Median (IQR)	Minimum	Maximum	Median (IQR)	Minimum	Maximum
Glabella	All	.62 (.27,1.16)	.01	2.28			
	Males	.77 (.35, 1.37)	.01	2.28	0.9 (0.6, 1.5)	0.1	4.4
	Females	.43 (.22, .96)	.01	2.27	1.0 (0.5, 2.2)	0.1	4.1
	Difference (p)	.015*			0.880		
Nasion	All	.60 (.24, 1.04)	.00	2.28			
	Males	.63 (.25, 1.19)	.01	2.28	0.8 (0.4, 1.4)	0.0	3.9
	Females	.49 (.24, .90)	.00	1.98	1.2 (0.4, 2.0)	0.0	3.9
	Difference (p)	.160			0.759		
Pronasale	All	.74 (.36, 1.27)	.00	3.07			
	Males	.72 (.35, 1.31)	.01	2.85	0.1 (0.0, 0.3)	0.0	1.1
	Females	.75 (.33, 1.28)	.00	3.07	0.2 (0.1, 0.3)	0.0	1.8
	Difference (p)	.881			0.543		
Subnasale	All	.82 (.36, 1.37)	.00	2.81			
	Males	.83 (.33, 1.38)	.01	2.81	0.8 (0.4, 1.4)	0.0	2.1
	Females	.78 (.36, 1.34)	.00	2.72	0.7 (0.3, 1.1)	0.0	2.0
	Difference (p)	.988			0.759		
Labrale Superious	All	.86 (.47, 1.59)	.01	3.89			
	Males	1.15(.50, 1.75)	.02	3.89	1.1 (0.7, 1.5)	0.2	4.0
	Females	.84 (.39, 1.35)	.01	3.26	1.2 (0.3, 1.7)	0.0	3.7
	Difference (p)	.189			0.759		
Labrale Inferious	All	.98 (.42, 1.70)	.02	3.6			
	Males	1.33 (.32, 2.25)	.02	3.6	1.2 (0.4, 1.8)	0.0	4.3
	Females	.79 (.46, 1.33)	.03	2.47	1.3 (0.4, 2.2)	0.0	5.3
	Difference (p)	.050*			0.759		
Pogonion	All	1.08 (0.56, 1.75)	.00	3.94			
	Males	1.18 (.64, 2.25)	.03	3.94	1.5 (0.7, 3.0)	0.0	5.2
	Females	.99 (.44, 1.37)	.00	3.16	1.8 (0.7, 2.5)	0.0	7.6
	Difference (p)	.041*			0.759		

Table 28: Asymmetry indices for midfacial landmarks in present research in comparison with results reported by Alqattan et al. 2015[48], * = Statistically significant (p=≤0.05).

6.3.3: Patterns of Asymmetry

Asymmetry has been suggested previously to increase as you move away from the cranium.[7] There does appear to be a pattern of increasing asymmetry in our results with median values increasing gradually from Nasion through to Pogonion. This is illustrated in Figure 32.

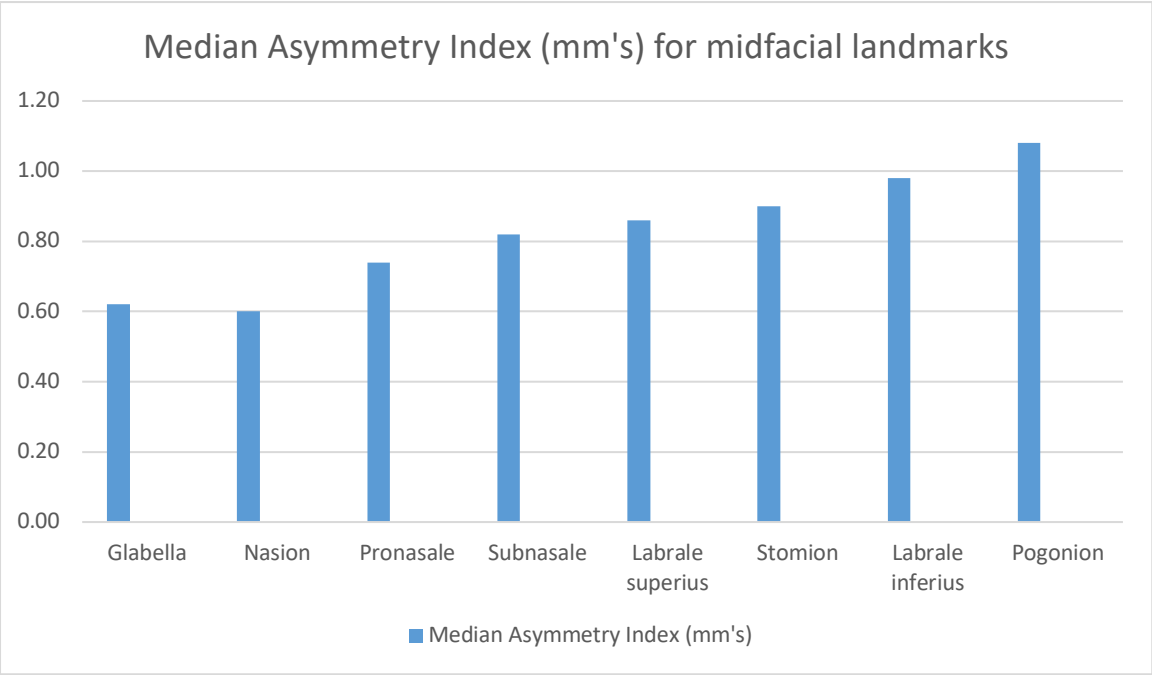


Figure 32: Bar chart illustrating the tendency for asymmetry index to increase as you move inferiorly away from the cranium.

6.3.4: Gender Differences in Midfacial Landmarks

Contrary to the results of Alqattan et al. (2015)[48] which identified no difference in gender using the Mann-Whitney u test, and Bugaighis et al. (2013)[2] whom reported no significant difference in facial asymmetry between genders (p=0.120), in this research 3 midfacial landmarks showed a significant difference between genders including: Glabella (p=0.015), Labrale inferious (p=0.050) and Pogonion (p=0.041) with males demonstrating higher values for asymmetry. However, this was no longer significant after the Bonferroni correction was applied to account for multiple hypothesis testing. It is possible that this method of correction is slightly harsh and underestimates the potential differences in asymmetry depending on gender.

Overall the findings suggest that in the standard population of children in the North West of England there is generally no significant midfacial asymmetry present. As this study is cross-sectional, it cannot provide information regarding changes overtime. Therefore, it's possible that facial asymmetry may fluctuate within individuals over time.

6.4: Asymmetry of Bilateral Landmarks

Asymmetry was initially assessed in terms of the 3 coordinate planes separately: right to left (x), superior-inferior (y) and antero-posterior (z) as recommended by previous research.[48] It is important to firstly assess asymmetry of these planes separately prior to combining the information as low levels of asymmetry in 1 or 2 coordinate planes may mask a high level of asymmetry in the third plane.[48] Following this, to give an overall value for asymmetry (all 3

planes of space combined), the asymmetry index was calculated for each of the 7 bilateral landmark pairs: Endocanthion, Exocanthion, Palpebrale superious, Palpebrale inferious, Alare, Christa philtri and Cheilion. This method has been used to assess facial asymmetry by several authors previously.[7,48,81] As it uses the absolute values it represents the overall asymmetry rather than its' direction. Comparably to the findings of previous research using a landmark based approach, the asymmetry index was higher for bilateral landmarks than medial landmarks, likely because the latter account for asymmetry in 3 coordinate planes whereas midfacial asymmetry is reflected solely in the x direction.

In order to discuss the results of our assessment of facial asymmetry using bilateral landmarks in depth the results will be outlined in terms of 3 main facial areas: orbital, nasal and oral regions. The results regarding the 3 coordinate planes separately will be compared to a recent study on Caucasian adults by Alqattan et al. (2015)[48] and the asymmetry index results will also be compared to the same Caucasian adult sample and to a sample of Chinese adults studied by Huang et al. (2013).[7,48]. The literature regarding the suggested thresholds for facial asymmetry will also be referenced in order to discuss the clinical and social significance of the findings.

6.4.1: Orbital Region

As previously outlined in the literature review, the eyelids are thought to be the most sensitive facial area to static asymmetry perception with exponential increases of detection between lay people from 1mm to 2mm of asymmetry (superior inferior (y) direction), with a rise in percentage of 10 to 85% in detection respectively.[21]

Exocanthion, Endocanthion, Palpebrale superious and Palpebrale inferious were used to assess asymmetry of the orbital region in this study. Endocanthion exhibited less asymmetry in the x, y and z direction than previously reported in the literature, for both males and females, with median values for asymmetry for both genders combined <1mm (see Table 18, section 5.5).[48] Exocanthion similarly exhibited less asymmetry than previously reported, except for males in the z direction which had a median asymmetry of 1.54mm (IQR 0.58, 2.56mm) which is slightly (0.04mm) greater than that reported by Alqattan et al. (2005)[48] (1.50mm, IQR 0.5, 3.0mm), although the IQR was smaller in this studies sample suggesting the range of asymmetry in this population was less extensive.

In the x and y direction both Palpebrale superious and inferious had median asymmetry values of <1mm (see Table 18), which according to the pre-existing literature is not likely to be significant clinically or socially.[21] This asymmetry was less than the median values of asymmetry previously reported in the literature.[48] The results in the y direction, based on the research by Hohman et al. (2014)[21], would indicate no significant asymmetry on average in the sample population in terms of eyelid position.

In the z direction asymmetry was <1mm (median 0.93mm) for Palpebrale inferious (IQR 0.44, 1.67mm) and >1mm, (median 1.45mm, IQR 25th and 75th percentiles: 0.65, 3.15mm) for Palpebrale superious. Looking at Palpebrale superious in more detail, when the z direction results for males are considered separately their asymmetry is 1.32 (IQR 0.60, 2.77mm) which is more severe than reported by Alqattan et al. (2015)[48]: 1.0mm (0.4, 2.1mm).[48] For females the asymmetry results for Palpebrale superious in this study 1.60 (0.73, 4.24mm) have the same median value than that of Alqattan et al. (2015)[48] 1.60mm but have a more limited IQR (0.7, 2.2mm). Our results suggest generally detectable asymmetry of the upper eyelid in the sample population, however there is a paucity of research regarding the perception of eyelid asymmetry specifically in the z direction, making it challenging to conclude regarding the potential significance of this finding.

When the results for Palpebrale superious in the z direction are carefully examined it can be observed that the maximum values of asymmetry between the right and left sides for Palpebrale superious are very high for both males and females (23.99mm males and 26.53mm females) in the z coordinate plane. The data in the area of this landmark can be influenced by the hair of the eyelashes which in turn can lead to areas of missing data and challenges locating landmarks particularly in the anterior-posterior direction. These outlier values are potentially due to difficulties in relation to image capture and subsequently landmark identification. Figure 28 in section 5.5 illustrates the differences in the z direction for Palpebrale superious in a box and whisker plot for all 107 images (genders combined) with multiple outliers evident outside the 25th and 75th percentiles. Considering the number of outliers and the range of values it is suggested by the authors that the results of this landmark may not truly represent the asymmetry of the upper eyelid in the sample population and thus we should be careful when drawing conclusions based on the data obtained.

The asymmetry index was lower in our sample population than that of Alqattan et al. (2015)[48] for all 4 orbital region landmarks, the results for males and females are considered separately for ease of comparison (Table 29 below). The mean AI values for Endocanthion (1.21mm, SD 0.57mm) and Exocanthion (1.00mm, SD 0.62mm) reported by Huang et al. (2013)[7] were lower for both genders than the median values for Endocanthion (1.56mm, IQR 1.15, 2.42) and Exocanthion (2.16mm, IQR 1.41, 3.34), however they selected their population based on an Angles Class 1 dental occlusion with 'normal symmetry' when assessed by an orthodontist, a plastic surgeon and a nurse so it is likely that their selected population was more symmetrical than the 'standard population' in this study, or in that of Alqattan et al. (2015)[48], which did not apply such stringent exclusion criteria. Using the asymmetry index which combines the values of all 3 coordinates it appears that the sample population in this study does not exhibit facial asymmetry in the orbital region outside that which is reported in other standard populations.

Orbital Region Bilateral Landmarks		Asymmetry Index (mm)		
		Median (IQR) (n=107) Children	Median (IQR) Alqattan et al (n=83) Caucasian adults	Mean (SD) Huang et al (N=60) Chinese adults
Endocanthion	All	1.56 (1.15, 2.42)		1.21 (.57)
	Males	1.44 (1.13, 2.53)	2.8 (1.9, 3.4)	
	Females	1.71 (1.26, 2.11)	3.1 (2.2, 4.8)	
	Difference	.685		
Exocanthion	All	2.16 (1.41, 3.34)		1.00 (0.62)
	Males	2.16 (1.49, 3.33)	3.1 (2.2, 4.8)	
	Females	2.16 (1.24, 3.35)	3.0 (2.2, 4.8)	
	Difference	.808		
Palpebrale Superious	All	1.68 (1.23, 3.29)		
	Males	1.73 (1.24, 2.28)	2.7 (1.7, 4.8)	
	Females	1.65 (1.20, 2.29)	3.1 (1.9, 4.9)	
	Difference	.793		
Palpebrale Inferious	All	1.56 (1.06, 2.22)		
	Males	1.36 (1.03, 2.05)	2.9 (1.6, 4.7)	
	Females	1.61 (1.11, 2.01)	2.8 (1.7, 4.5)	
	Difference	.617		

Table 29: Outlining the asymmetry index for orbital regions of the whole sample, and for males and females separately, presented in terms of median with inter-quartile range (IQR), maximum and minimum values in millimeters. Gender difference was calculated using the Mann-Whitney U test. The results are compared to the findings of Alqattan et al. (2015)[48] and Huang et al. (2013)[7] who presented their findings in median and mean form respectively.

Regarding the differences in gender, using the Mann-Whitney U test there was a significant difference between males and females for Exocanthion (p=0.048) and Palpebrale inferious (p=0.044) in the superior-inferior (y) direction only, with males being more asymmetric than females. Once the Bonferroni correction was applied for multiple hypothesis testing this gender difference was no longer significant and it may be concluded that there was no significant difference in orbital region asymmetry between genders in the sample population.
The eyebrows are not commonly measured or referred to in research on facial asymmetry [2,7,48], although some have investigated the perception of progressive asymmetric elevation using digitally manipulated images.[21,22] As our sample was retrospective there was a limited number of participants available. It was decided not to exclude patients if they appeared to have groomed their eyebrows in order to maintain an adequate sample size, it

was also agreed not to position landmarks on the eyebrow area as this is not a commonly used landmark and by omitting it one avoids the potential for grooming influencing the asymmetry results.

6.4.2: Nasal Region

The asymmetry of the bilateral nasal landmark (Alare) was <1.5mm in the x, y and z coordinate planes for both males and females. These were lower than the values for asymmetry reported in previous research, except for in the x direction in males which had a median of 1.49mm (IQR 0.78, 2.36) which is slightly higher than in the previous research: median 0.9mm (IQR 0.4, 2.4mm).[48] Using the AI there was less asymmetry of Alare detected than documented in the literature with a median of 1.94mm (IQR 1.27, 2.73mm) (Table 30). Most of the research regarding the perception of nasal asymmetry has focused on nasal tip deviation which has already been discussed in section 6.3.1 on midfacial landmarks. It is unlikely that median asymmetry of <1.5mm would be significant considering nasal tip deviation usually begins to be perceivable at 2mm deviation.[23] Therefore, it may be concluded that there wasn't significant asymmetry of Alare in the sample population and that there was no significant difference in the asymmetry index of Alare between males and females (Mann-Whitney U test 0.392).

Nasal Region Bilateral Landmark	**Asymmetry Index (mm)**			
	Median (IQR) (n=107) Children of multiple ethnicities		**Median (IQR) Alqattan et al. (2015) (n=83) Caucasian adults**	**Mean (SD) Huang et al. (2013) (n=60) Chinese adults**
Alare	All	1.94 (1.27, 2.73)		2.33 (1.07)
	Males	2.04 (1.40, 2.75)	2.8(1.9, 4.4)	
	Females	1.83 (1.18, 2.71)	2.8 (1.9, 4.1)	
	Difference	.392		

Table 30: Outlining the Asymmetry Index for Alare of the whole sample, and for males and females separately, presented in terms of median with inter-quartile range (IQR), maximum and minimum values in millimeters. Gender difference was calculated using the Mann-Whitney U test. The results are compared to the findings of Alqattan et al. (2015)[48] and Huang et al. (2013)[7] who presented their findings in median and mean form respectively.

6.4.3: Oral Region

The median asymmetry index for Cheilion was 2.56mm (IQR 1.69, 3.65mm), which is less asymmetry than observed by Huang et al. (2013)[7] (mean AI 2.82mm, SD 1.42mm). Three millimetres of asymmetry in the Cheilion area is reported to be noticed by 73% of people

(p<0.001) with 5mm discrepancy being perceived to require intervention.[22]

The median asymmetry index for Christa philtri was 2.09mm (IQR 0.63, 3.46mm) which is slightly less than the mean asymmetry index of 2.82mm (SD 1.42mm) reported by Huang et al. (2013).[7] The asymmetry indices for bilateral landmarks in the oral region are listed in Table 31.

Oral Region Bilateral Landmarks		Asymmetry Index (mm)		
		Median (IQR) (n=107) Children	Median (IQR) Alqattan et al (n=83) Caucasian adults	Mean (SD) Huang et al (n=60) Chinese adults
Christa Philtri	All	2.09 (.63, 3.46)		
	Males	2.51 (1.07, 3.70)	2.2 (1.6, 3.0)	
	Females	1.92 (.88, 3.17)	1.6 (0.9, 3.5)	
	Difference	.288		
Cheilion	All	2.56 (1.69, 3.65)		2.82 (1.42)
	Males	2.59 (1.73, 3.77)	3.2 (2.1, 4.1)	
	Females	2.36 (1.64, 3.30)	3.5 (2.4, 5.0)	
	Difference	.318		

Table 31: Outlining the asymmetry index for oral regions of the whole sample, and for males and females separately, presented in terms of median with inter-quartile range (IQR), maximum and minimum values in millimeters. Gender difference was calculated using the Mann-Whitney U test. The results are compared to the findings of Alqattan et al. (2015)[48] and Huang et al. (2013)[7] who presented their findings in median and mean form respectively.

6.4.4: Linear and Surface Measurements of Asymmetry

Bilateral landmarks Exocanthion and Cheilion showed significant differences between right and left sides when both linear and surface measuring tools were used, and after correcting for multiple hypothesis testing using Bonferroni correction, with p=0.001. Endocanthion and Christa philtri also showed significant discrepancies between sides using linear measurements (p=0.001 and p=0.060 respectively). This suggests that these areas of the face have statistically significant facial asymmetry, but the question is if this equates to clinical significance. The maximum median value for asymmetry using linear measurements was 1.33mm (IQR 0.53, 1.97) for Exocanthion which is not likely to be clinically significant, even though there is a statistically significant difference between the means of the right and left sides when a paired t-test is used (p=0.001). The maximum median asymmetry for surface measurements was 2.60mm for Cheilion (IQR 1.16, 4.59) which is likely to be evident clinically

but not likely to be considered to require intervention as previously discussed (as <5mm).[22] Comparing the asymmetry measured by linear and surface tools it's apparent that there are some discrepancies. The surface measurements were larger than the linear ones, which was expected. However, there were substantial differences in the asymmetry calculated by the 2 methods with as high as a 1.75mm difference in the asymmetry noted for one landmark (Cheilion). This is a large difference in results which could move a patient from a category of within 'normal limits' to one classified by asymmetrical facial appearance depending on which test is used. The primary researcher noted during data collection that the surface measurements did not always follow a consistent path across the face, and for example, could proceed around the nostril on one side, and over it on the other side, thus giving very different measurements for left/right sides. This introduces great potential for error using this system. This was discussed with the 3dMD technicians to see if the issue could be addressed but unfortunately was not possible. The tendency for error using the surface measurement tool is increased by any areas of missing data, and to attempt to rectify this, images would need to be obtained prospectively. Therefore, currently the authors would not recommend use of the surface measurement tool in its current form for the analysis of facial asymmetry.

Cheilion using linear measurements demonstrated a significant difference between genders ($p=0.031$) using the Mann-Whitney U test but this was no longer significant after accounted multiple hypothesis testing was accounted for (Bonferroni correction). Overall there were no significant differences in asymmetry observed between males and females using either the linear or surface measuring tools which is in support of the other findings of this research project ($p \leq 0.05$).

6.5: Deviation of Nasal Tip (Pronasale) from the Midface

The median deviation of the nasal tip relative to the midface was 1.77° (IQR 0.84°, 3.12°). Previous research incorporating the perspectives of lay people, orthodontists, general dentists and dental students has shown that a deviation of 2.92° of Pronasale relative to the midfacial plane is the threshold for the detection for asymmetry (mean 2.92°, SD 1.40°).[25] As in our sample the 75th percentile is 3.12°, it appears that there is clinically significant nasal tip deviation present in a proportion of the sample population, but the mean nasal tip deviation was 2.44° which is thought to be below the threshold of recognition (95% CIs 1.99°, 2.88°). It should be noted that although the patients were orientated in a standardised manner (outlined in section 4.3), it is still possible that participant positioning could have influenced the angular measurements.

The degree of nasal deviation may be seen in the pie chart below (Figure 33) with deviation illustrated firstly in 2 degree increments for the included 105 patients. Sixty-two participants (59%) had nasal tip deviation of <2degrees. The second pie chart (Figure 34) exhibits the

number of patients (n=27, 25.7%) who had nasal deviation which would be considered to be significant relative to previous literature (ie >2.92°). There was a high maximum value of 11.08° for one 2-year-old suggesting that high levels of nasal deviation do occur within individuals of the standard population.

There was clinically significant nasal deviation in 25.7% (n=27) of the sample population, a finding that was not apparent when landmark Pronasale was assessed in terms of the x coordinate. This is possibly due to the fact that the x coordinates consider simply the sagittal (left to right asymmetry) whereas the angle formed by the deviation of the landmark is also influenced by distances to the origin.

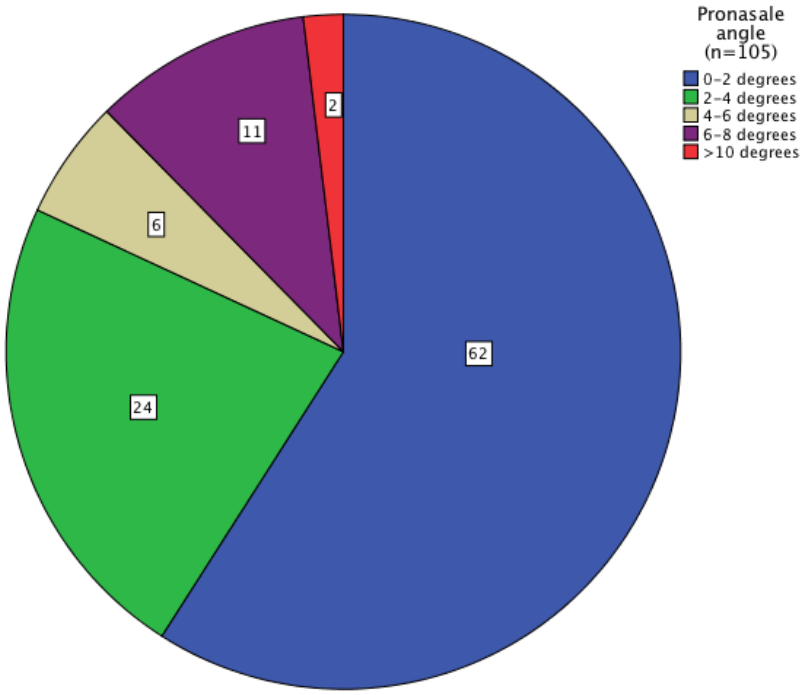


Figure 33: Pie chart illustrating the deviation of nasal tip (Pronasale) from the midface in 2-degree increments (n=105).

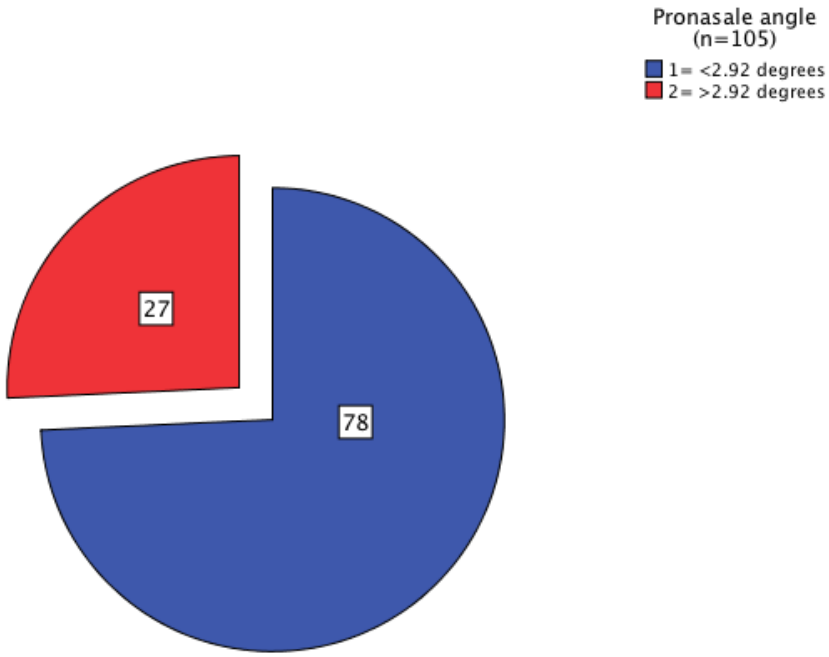


Figure 34: Pie chart illustrating the proportions of participants which based on previous literature would be considered to have significant deviation of the nasal tip (>2.92°) (n=105).

There were more patients with their nasal tip deviated to the right (n=61, 57% of sample) than to the left (n=45, 42.1% of the sample), which is in line with previous theories of right sided facial dominance.[8] One patient's landmark Pronasale was coincident with the midface. A Mann-Whitney U test showed an association between the side of Pronasale deviation and its' extent (p=0.022), with left sided asymmetries tending to be more severe. This is interesting considering previous research has reported left sided nasal asymmetries to be more negatively perceived.[23]

6.6: Deviation of the Chin (Pogonion) from the Midface

The median deviation of the chin point was 1.12° (IQR 0.52°, 2.51°). It is unlikely that such a small deviation would be clinically significant, but there is a paucity of data looking at the angle formed between the chin and the midface so our results do not have a baseline to compare to. Considering asymmetry has formerly been reported to increase as you move downwards away from the cranium one would expect that if nasal deviation of up to 2.92° is negligible then a median of 1.12° for Pogonion deviation would not be readily detectable, and thus is not likely to be a significant asymmetry. There were some outliers with a maximum value of 12.99° observed in one 2-year-old patient (image 377). Interestingly this was the same patient that had the maximal nasal tip deviation which suggests that some patients within the standard population do have substantial deviations of both the nasal tip and chin.

There were 54 participants with chin to the right side (50.5%) and 52 (48.6%) with chin to the

left side. There was one participant with the landmark Pogonion coincident with the midface. There was no correlation between the side of chin deviation (left or right) and the severity (Mann-Whitney U test p=0.61). The potential relationship between the severity of nasal tip and chin deviation are discussed in the next section.

6.7: Relationship between Pronasale and Pogonion

There is a statistically significant correlation between the extent of nasal tip deviation (Pronasale) and the extent of chin (Pogonion) deviation with a Pearson's correlation of 0.91. As the severity of nasal deviation increases, so too does that of the chin, and vice versa. This positive correlation is clearly illustrated by the scatter plot in Figure 30c section 5.8. There is also a statistically significant relationship between the side of deviation (right or left) of Pronasale and Pogonion (chi-square test p=0.00), meaning for example that if the nose is to the left of the midfacial axis, it is likely that the chin will also be to the left of the midfacial axis. The cross tabulation outlining this is seen in Table 26 section 5.8.

From a clinical perspective, it's important to be aware of these apparent associations between asymmetry of the nose and chin. Patients who complain of specific areas of facial asymmetry should carefully be made aware of other facial asymmetries and discrepancies present prior to treatment addressing the main feature of their concern. This is because, for example, if a patient has a severe chin asymmetry complaint which is subsequently corrected, any pre-existing nasal asymmetry may become more apparent and most importantly may become detectable by the patient once the chin is in line with the midface rather than in sync with the nose. This potential for altering the perspective of other areas of the face needs to be discussed, and potential solutions outlined prior to any intervention as part of the informed consent process.

6.8: Relationship between Asymmetry and Age

There was a correlation between increasing age and increasing asymmetry of the midface. This was significant for 6 out of the 8 midfacial landmarks, with 4 remaining significant after the Bonferroni correction was applied: Subnasale, Stomion, Labrale inferious and Pogonion (see scatter plots Figures 27a-d section 5.4). Interestingly these are all in the lower third of the face. For the bilateral facial landmarks there was a relationship between increasing age and increasing asymmetry index of 3 landmarks (Alare, Christa philtri and Cheilion) with only Christa philtri remaining significant after the Bonferroni correction (corrected p=0.035). When the angles formed by Pronasale and Pogonion were examined there was no significant correlation between increasing age and extent of Pronasale deviation (Pearson's = 0.163, p=0.096) or Pogonion deviation (Pearson's=0.090, p=0.362).

It's empirical to note that the sample included children ranging from a few months old to 14

years of age. Older children would have tended to have higher values for measurements and thus this may be falsely represented as increasing asymmetry with age using the AI, when it's possible that the proportions of asymmetry are staying the same. Previous research using surface based 3d facial analysis of laser scans suggests that asymmetry is apparent at an early age and tends to remain through the pre-pubertal period without increasing or decreasing significantly.[39] This was shown in a longitudinal study of 5-year-olds whom were followed up for 54 months. Djordjevic et al. (2014)[53] also studied asymmetry changes in a longitudinal sample of Caucasian Finnish children (30 male and 30 female) with no clinically evident facial asymmetry, examining their images initially between 10-13 years of age (mean 11.5 years) (T1), then at T2 and T3 (2.5 and 4.5 years after T1 respectively) to assess changes through the pubertal growth spurt. Their results suggest that growth in healthy adolescents is largely symmetrical, but that a larger randomised sample would be required to confirm this. Farkas and Cheung (1981)[3] in a now historical sample of 308 children (6,12,18 years) reported that the prevalence and extent of facial asymmetries was not significantly influenced by age. In order to establish fluctuations in asymmetry with age and growth a longitudinal sample would be required to measure patients' relative to themselves and their cohort as they mature. Some previous research has tried to avoid the issues of face size affecting asymmetry by selecting patients in a narrower age group or by using average face construction.[7,39,53]

6.9: Clinical Implications

It has been demonstrated that facial asymmetry is poorly assessed subjectively, with Taylor et al. (2014)[82] reporting that two independent observers failed to agree regarding the degree of facial asymmetry present (mild, moderate or severe asymmetry) (r=0.56). The majority of participants were classed as having moderate asymmetry which the authors felt could be a 'catchall bin for those that were difficult to assess'.[82] This is why a reproducible objective means of assessment is greatly useful when discussing patients concerns, treatment planning and assessing surgical outcomes.

With an ever-increasing awareness of the risks of ionising radiation and being mindful of the ALARA (as low as reasonably achievable) principle, an assessment method without exposure to radiation is highly desirable. The speed at which stereophotogrammetry images may be acquired (1.5 ms) means indirect anthropometry is particularly useful in young children who may not tolerate direct measurements.

With an evidence-based approach for clinicians being regarded as essential, the fact that these images may be stored long term increases the potential for the maximum amount of information to be gained from each sample of willing participants. An example of this is that the same population used in this research study was also used in a study investigating cranial shape. This benefit optimises the use of research resources.

When people attend complaining of facial asymmetry, it is not always considered significant by the clinician but still may be very significant in the eyes of the patient. Previous research has shown that what one observer considers as mild asymmetry may be considered by another to be moderate, or even severe.[82] It can be challenging for the clinician to quantify the presence or severity of the problem and to place it within or outside the range of asymmetry in the standard population. Information regarding this would allow for baseline levels of asymmetry to be set at which surgical or other interventions may be considered depending on the patient's desires. By conducting research such as this study we take one more step forward in defining a normative level of asymmetry. This is the first study of its kind to assess facial asymmetry in the North West of England population. It is hoped the findings can be used to help define symmetry within normal limits; to compare patient's presenting asymmetry with normative data, to communicate with patients and parents the severity of their presenting complaint and to assess surgical outcomes of children, for example, with craniofacial syndromes relative to their un-operated peers which in turn would assist in the process of informed consent prior to the surgical intervention. As Sir Harold Gillies, a world renowned pioneer of plastic surgery famously advised we must "first diagnose, then treat", and with this in mind, 3d photogrammetry allows us to follow this key principle by objectively diagnosing prior to intervention.[82,83]

Chapter 7: Conclusions

- In response to the research question; yes, there is facial asymmetry present in the sample population of children under 16 years old in the North west of England.
- The median facial asymmetry detected is generally at a level which is minimal and not beyond the thresholds of perception based on the available evidence.[8]
- Using the Asymmetry Index the most asymmetric landmarks were detected in the lower third of the face. Cheilion was the most asymmetric bilateral landmark, and Pogonion the most asymmetric midfacial landmark.
- No significant differences in facial asymmetry could be demonstrated between genders.
- There were some significant differences between ages, with asymmetry tending to increase with age for landmarks: Pronasale, Subnasale, Stomion, Pogonion and Christa philtri. However, this correlation should be interpreted with caution as it could be influenced by the size of the participant's head.
- The severity (Pearsons correlation coefficient = 0.91, $p<0.001$) and direction (Chi-squared $p<0.001$) of nasal and chin deviation are highly correlated.
- Three-dimensional images (3dMD) are a useful diagnostic tool for the assessment of facial soft tissue asymmetry in a child population.

References

1. Fischer B. Asymmetries of the Dentofacial Complex. *Angle Orthod.* 1954;24(4):179–92.
2. Bugaighis I, Mattick CR, Tiddeman B, Hobson R. Three-dimensional gender differences in facial form of children in the North East of England. *Eur J Orthod.* 2013;35(3):295–304.
3. Farkas LG, Cheung G. Facial asymmetry in healthy North American Caucasians. An anthropometrical study. *Angle Orthod.* 1981;51(1):70–7.
4. Ryan FS, Moles DR, Shute JT, Clarke A, Cunningham SJ. Social anxiety in orthognathic patients. *Int J Oral Maxillofac Surg.* 2016;45(1):19–25.
5. Bishara SE, Burkey PS, Kharouf JG. Dental and facial asymmetries: a review. *Angle Orthod.* 1994;64(2):89–98.
6. Peck S, Peck L, Kataja M. Skeletal asymmetry in esthetically pleasing faces. *Angle Orthod*. 1991;61(1):43–8.
7. Huang CS, Liu XQ, Chen YR. Facial asymmetry index in normal young adults. *Orthod Craniofac Res.* 2013;16(2):97–104.
8. Wang TT, Wessels L, Hussain G, Merten S. Discriminative Thresholds in Facial Asymmetry: A Review of the Literature. *Aesthetic Surg J.* 2017;37(4):375–85.
9. McAvinchey G, Maxim F, Nix B, Djordjevic J, Linklater R, Landini G. The perception of facial asymmetry using 3-dimensional simulated images. *Angle Orthod.* 2014;84(6):957–65.
10. Primozic J, Perinetti G, Richmond S, Ovsenik M. Three-dimensional evaluation of facial asymmetry in association with unilateral functional crossbite in the primary, early, and late mixed dentition phases. *Angle Orthod.* 2013;83(2):253–8.
11. Meyer-Marcotty P, Alpers GW, Gerdes ABM, Stellzig-Eisenhauer A. Impact of facial asymmetry in visual perception: A 3-dimensional data analysis. *Am J Orthod Dentofac Orthop.* 2010;137(2):168.e1-168.e8.
12. Berssenbrügge P, Berlin NF, Kebeck G, Runte C, Jung S, Kleinheinz J, Dirksen D. 2D and 3D analysis methods of facial asymmetry in comparison. *J Craniomaxillofac Surg. 2014*;42(6):e327-34.
13. Berlin NF, Berssenbrügge P, Runte C, Wermker K, Jung S, Kleinheinz J, Dirksen D. Quantification of facial asymmetry by 2D analysis - A comparison of recent approaches. *J Cranio-Maxillofacial Surg.* 2014;42(3):265–71.
14. Zaidel DW, Deblieck C. Attractiveness of natural faces compared to computer constructed perfectly symmetrical faces. *Int J Neurosci.* 2007;117(4):423–31.
15. Blackburn K, Schirillo J. Emotive hemispheric differences measured in real-life portraits using pupil diameter and subjective aesthetic preferences. *Exp Brain Res.* 2012;219(4):447–55.
16. LeonardoDaVinci.net. The Mona Lisa - by Leonardo Da Vinci [Internet]. 2011 [cited 2017 Jul 10]. Available from: https://www.leonardodavinci.net/the-mona-lisa.jsp
17. Rembrandt van Rijn | A Weeping Woman | Detroit Institute of Arts | Buy Prints Online [Internet]. [cited 2017 Jul 10]. Available from: https://www.1000museums.com/art_works/rembrandt-van-rijn-a-weeping-woman?from=artists
18. Janson J. Girl with a pearl earring - by Johannes Vermeer [Internet]. [cited 2017 Jul 10]. Available from: http://www.essentialvermeer.com/catalogue/girl_with_a_pearl_earring.html#.WWPqWhiZM9c
19. Djordjevic J, Pirttiniemi P, Harila V, Heikkinen T, Toma AM, Zhurov AI, Richmond S. Three-dimensional longitudinal assessment of facial symmetry in adolescents. *Eur J Orthod.* 2013;35(2):143–51.
20. Naini FB, Donaldson ANA, McDonald F, Cobourne MT. Assessing the Influence of Asymmeftry Affecting the Mandible and Chin Point on Perceived Attractiveness in the

Orthognathic Patient, Clinician, and Layperson. *J Oral Maxillofac Surg.* 2012;70(1):192–206.
21. Hohman MH, Kim SW, Heller ES, Frigerio A, Heaton JT, Hadlock TA. Determining the threshold for asymmetry detection in facial expressions. *Laryngoscope.* 2014;124(4):860–5.
22. Chu EA, Farrag TY, Ishii LE, Byrne PJ. Threshold of Visual Perception of Facial Asymmetry in a Facial Paralysis Model. *Arch Facial Plast Surg*. 2011 1;13(1):14–9.
23. Meyer-Marcotty P, Stellzig-Eisenhauer A, Bareis U, Hartmann J, Kochel J. Three-dimensional perception of facial asymmetry. *Eur J Orthod.* 2011;33(6):647–53.
24. Silva BP, Jiménez-Castellanos E, Martinez-de-Fuentes R, Greenberg JR, Chu S. Laypersons' perception of facial and dental asymmetries. *Int J Periodontics Restorative Dent.* 2013;33(6):e162-71.
25. Kwak K-H, Kim Y-I, Nam H-J, Kim S-S, Park S-B, Son W-S. Differences Among Deviations, Genders, and Observers in the Perception of Eye and Nose Asymmetry. *J Oral Maxillofac Surg.* 2015;73(8):1606–14.
26. Farkas L. Anthropometry of the Head and Face in Medicine. 1st ed. Molyneux, J., Heller, S., Kudlak V, editor. Elsevier North Holland; 1981.
27. Naini FB. Leslie G. Farkas - Pioneer of Modern Craniofacial Anthropometry. *Arch Facial Plast Surg.* 2010;12(3):599–616.
28. Farkas LG. Anthropometry of the head and face. 2nd edn. Vol. 82. New York: Raven Press ltd, New York, NY; 1994. 438 p.
29. Naini FB. Leslie G. Farkas, 1915-2008. *Am J Orthod Dentofac Orthop.* 2009;136(4):614.
30. Darby LJ, Millett DT, Kelly N, McIntyre GT, Cronin MS. The effect of smiling on facial asymmetry in adults: a 3D evaluation. *Aust Orthod J.* 2015;31(2):132–7.
31. Kaipainen AE, Sieber KR, Nada RM, Maal TJ, Katsaros C, Fudalej PS. Regional facial asymmetries and attractiveness of the face. *Eur J Orthod.* 2016;38(6):602–8.
32. Deutsch CK, Shell AR, Francis RW, Bird BD. The Farkas System of Craniofacial Anthropometry: Methodology and Normative Databases. In: Handbook of Anthropometry. New York, NY: Springer New York; 2012. p. 561–73.
33. de Menezes M, Rosati R, Ferrario VF, Sforza C. Accuracy and Reproducibility of a 3-Dimensional Stereophotogrammetric Imaging System. *J Oral Maxillofac Surg.* 2010;68(9):2129–35.
34. Ferrario VF, Sforza C, Miani A, Serrao G. A three-dimensional evaluation of human facial asymmetry. *J Anat.* 1995;186(Pt 1):103–10.
35. Gateno J, Xia JJ, Teichgraeber JF. Effect of facial asymmetry on 2-dimensional and 3-dimensional cephalometric measurements. *J Oral Maxillofac Surg.* 2011;69(3):655–62.
36. Shah SM, Joshi MR. An assessment of asymmetry in the normal craniofacial complex. *Angle Orthod.* 1978;48(2):141–8.
37. Farkas LG, James JS. Anthropometry of the face in lateral facial dysplasia: the unilateral form. *Cleft Palate J.* 1977;14(3):193–9.
38. Farkas LG, Ross RB, James JS. Anthropometry of the face in lateral facial dysplasia: the bilateral form. *Cleft Palate J.* 1977;14(1):41–51.
39. Primozic J, Perinetti G, Zhurov A, Richmond S, Ovsenik M. Assessment of facial asymmetry in growing subjects with a three-dimensional laser scanning system. *Orthod Craniofacial Res.* 2012;15(4):237–44.
40. Da Silveira AC, Daw JL, Kusnoto B, Evans C, Cohen M. Craniofacial applications of three-dimensional laser surface scanning. *J Craniofac Surg.* 2003;14(4):449–56.
41. Fudalej P, Katsaros C, Hozyasz K, Borstlap WA, Kuijpers-Jagtman AM. Nasolabial symmetry and aesthetics in children with complete unilateral cleft lip and palate. *Br J Oral Maxillofac Surg.* 2012;50(7):621–5.
42. Farrera A, Villanueva M, Quinto-Sánchez M, González-José R. The relationship between facial shape asymmetry and attractiveness in Mexican students. *Am J Hum Biol.* 2015 6;27(3):387–96.

43. Nakamura T, Okamoto K, Maruyama T. Facial asymmetry in patients with cervicobrachial pain and headache. *J Oral Rehabil.* 2001;28(11):1009–14.
44. Damstra J, Oosterkamp BCM, Jansma J, Ren Y. Combined 3-dimensional and mirror-image analysis for the diagnosis of asymmetry. *Am J Orthod Dentofac Orthop.* 2011;140(6):886–94.
45. Djordjevic J, Jadallah M, Zhurov AI, Toma AM, Richmond S. Three-dimensional analysis of facial shape and symmetry in twins using laser surface scanning. *Orthod Craniofac Res.* 2013;16(3):146–60.
46. Djordjevic J, Lewis BM, Donaghy CE, Zhurov AI, Knox J, Hunter L, Richmond S. Facial shape and asymmetry in 5-year-old children with repaired unilateral cleft lip and/or palate: an exploratory study using laser scanning. *Eur J Orthod.* 2014;36(5):497–505.
47. Kau CH, Richmond S, Zhurov AI, Knox J, Chestnutt I, Hartles F, Playle R. Reliability of measuring facial morphology with a 3-dimensional laser scanning system. *Am J Orthod Dentofacial Orthop.* 2005;128(4):424-30.
48. Alqattan M, Djordjevic J, Zhurov AI, Richmond S. Comparison between landmark and surface-based three-dimensional analyses of facial asymmetry in adults. *Eur J Orthod.* 2015;37(1):1–12.
49. Kau CH, Zhurov A, Scheer R, Bouwman S, Richmond S. The feasibility of measuring three-dimensional facial morphology in children. *Orthod Craniofac Res.* 2004;7(4):198–204.
50. Božič M, Kau CH, Richmond S, Ovsenik M, Hren NI. Novel method of 3-dimensional soft-tissue analysis for Class III patients. *Am J Orthod Dentofacial Orthop.* 2010;138(6):758–69.
51. Chiu CSW, Clark RKF. Reproducibility of natural head position. *J Dent.* 1991;19(2):130–1.
52. Kau CH, Richmond S, Zhurov AI, Knox J, Chestnutt I, Hartles F, Playle R. Reliability of measuring facial morphology with a 3-dimensional laser scanning system. *Am J Orthod Dentofac Orthop.* 2005;128(4):424–30.
53. Djordjevic J, Toma AM, Zhurov AI, Richmond S. Three-dimensional quantification of facial symmetry in adolescents using laser surface scanning. *Eur J Orthod.* 2014;36(2):125–32.
54. Heike CL, Upson K, Stuhaug E, Weinberg SM. 3D digital stereophotogrammetry: a practical guide to facial image acquisition. *Head Face Med.* 2010 28;6:18.
55. Tzou CHJ, Artner NM, Pona I, Hold A, Placheta E, Kropatsch WG, Frey M. Comparison of three-dimensional surface-imaging systems. *Journal of Plastic, Reconstructive and Aesthetic Surgery.* 2014;67(4):489–97.
56. Claes P, Walters M, Clement J. Improved facial outcome assessment using a 3D anthropometric mask. *Int J Oral Maxillofac Surg.* 2012;41(3):324–30.
57. Kuijpers MA, Desmedt DJ, Nada RM, Bergé SJ, Fudalej PS, Maal TJ. Regional facial asymmetries in unilateral orofacial clefts. *Eur J Orthod.* 2015;37(6):636–42.
58. Aldridge K, Boyadjiev SA, Capone GT, DeLeon VB, Richtsmeier JT. Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. *Am J Med Genet.* 2005;138A(3):247–53.
59. Bugaighis I, Tiddeman B, Mattick CR, Hobson R. 3D comparison of average faces in subjects with oral clefts. *Eur J Orthod.* 2014;36(4):365–72.
60. Bugaighis I, Mattick CR, Tiddeman B, Hobson R. 3D asymmetry of operated children with oral clefts. *Orthod Craniofac Res.* 2014;17(1):27–37.
61. Hajeer MY, Ayoub AF, Millett DT. Three-dimensional assessment of facial soft-tissue asymmetry before and after orthognathic surgery. *Br J Oral Maxillofac Surg.* 2004;42(5):396–404.
62. Ayoub A, Garrahy A, Hood C, White J, Bock M, Siebert JP, Spencer R, Ray A. Validation of a Vision-Based, Three-Dimensional Facial Imaging System. *Cleft Palate-Craniofacial J.* 2003;40(5):523–9.
63. Weinberg SM, Scott NM, Neiswanger K, Brandon CA, Marazita ML. Digital Three-

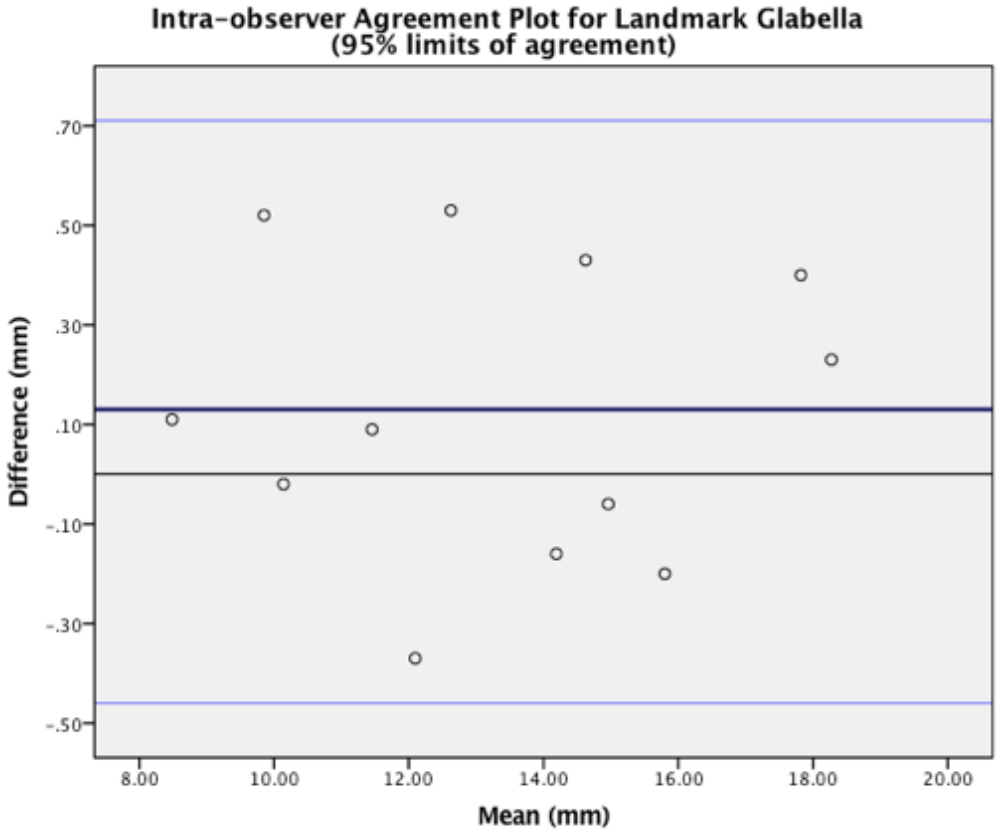
Dimensional Photogrammetry: Evaluation of Anthropometric Precision and Accuracy Using a Genex 3D Camera System. Cleft *Palate-Craniofacial J.* 2004;41(5):507–18.
64. Ovsenik M, Perinetti G, Zhurov A, Richmond S, Primozic J. Three-dimensional assessment of facial asymmetry among pre-pubertal class III subjects: A controlled study. *Eur J Orthod.* 2014;36(4):431–5.
65. Djordjevic J, Toma AM, Zhurov AI, Richmond S. Three-dimensional quantification of facial symmetry in adolescents using laser surface scanning. *Eur J Orthod.* 2014;36(2):125–32.
66. Katsumata A, Fujishita M, Maeda M, Ariji Y, Ariji E, Langlais RP. 3D-CT evaluation of facial asymmetry. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2005;99(2):212–20.
67. Nakamura T, Okamoto K, Maruyama T. Facial asymmetry in patients with cervicobrachial pain and headache. *J Oral Rehabil.* 2008;28(11):1009–14.
68. Brons S, van Beusichem ME, Maal TJJ, Plooij JM, Bronkhorst EM, Bergé SJ, et al. Development and reproducibility of a 3D stereophotogrammetric reference frame for facial soft tissue growth of babies and young children with and without orofacial clefts. *Int J Oral Maxillofac Surg.* 2013;42(1):2–8.
69. Aldridge K, Boyadjiev SA, Capone GT, DeLeon VB, Richtsmeier JT. Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. *Am J Med Genet.* 2005;138A(3):247–53.
70. Bugaighis I, O'Higgins P, Tiddeman B, Mattick C, Ben Ali O, Hobson R. Three-dimensional geometric morphometrics applied to the study of children with cleft lip and/or palate from the North East of England. *Eur J Orthod.* 2010;32(5):514–21.
71. Miller SF, Weinberg SM, Nidey NL, Defay DK, Marazita ML, Wehby GL, Morbe Uribe LM. Exploratory genotype-phenotype correlations of facial form and asymmetry in unaffected relatives of children with non-syndromic cleft lip and/or palate. *J Anat.* 2014;224(6):688–709.
72. Landis JR, Koch GG. Agreement measures for categorical data. *Biometrics.* 1977;33(1):159–74.
73. Toma AM, Zhurov A, Playle R, Ong E, Richmond S. Reproducibility of facial soft tissue landmarks on 3D laser-scanned facial images. *Orthod Craniofac Res.* 2009;12(1):33–42.
74. Darby LJ, Millett DT, Kelly N, McIntyre GT, Cronin MS. The effect of smiling on facial asymmetry in adults: a 3D evaluation. *Aust Orthod J.* 2015;31(2):132–7.
75. Hatch CD, Wehby GL, Nidey NL, Moreno Uribe LM. Effects of Objective 3-Dimensional Measures of Facial Shape and Symmetry on Perceptions of Facial Attractiveness. *J Oral Maxillofac Surg.* 2017;75(9):1958–70.
76. Oxford Dictionaries. reliability | Definition of reliability in English by Oxford Dictionaries [Internet]. [cited 2017 Sep 17]. Available from: https://en.oxforddictionaries.com/definition/reliability
77. Lachin JM. The role of measurement reliability in clinical trials. *Clin Trials.* 2004;1(6):553–66.
78. Alqattan M, Djordjevic J, Zhurov AI, Richmond S. Comparison between landmark and surface-based three-dimensional analyses of facial asymmetry in adults. *Eur J Orthod.* 2015;37(1):1–12.
79. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016;15(2):155–63.
80. McCluskey A, Lalkhen AG. Statistics II: Central tendency and spread of data. *Contin Educ Anaesthesia, Crit Care Pain.* 2007;7(4):127–30.
81. Katsumata A, Fujishita M, Maeda M, Ariji Y, Ariji E, Langlais RP. 3D-CT evaluation of facial asymmetry. *Oral Surgery, Oral Med Oral Pathol Oral Radiol Endodontology.* 2005;99(2):212–20.
82. Taylor HO, Morrison CS, Linden O, Phillips B, Chang J, Byrne ME, Sullivan SR, Forrest CR. Quantitative Facial Asymmetry. *J Craniofac Surg.* 2014;25(1):124–8.
83. Bamji A. Sir Harold Gillies: surgical pioneer. *Trauma*. 2006;8:143–56.

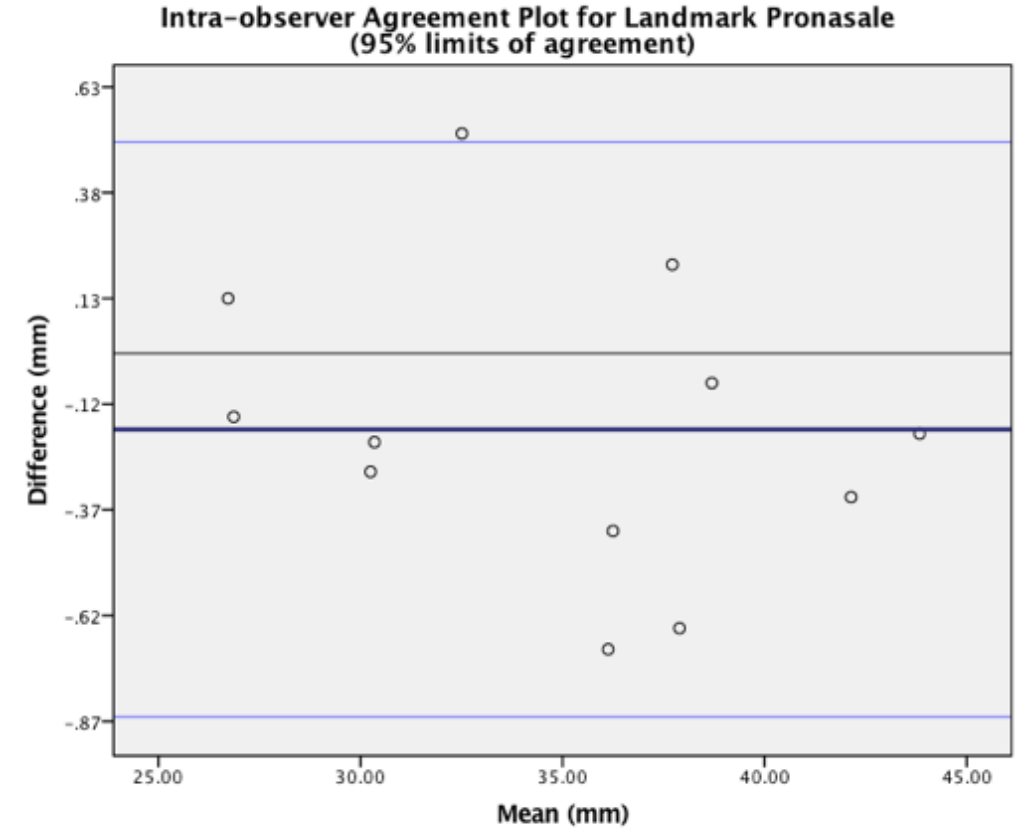
Appendices

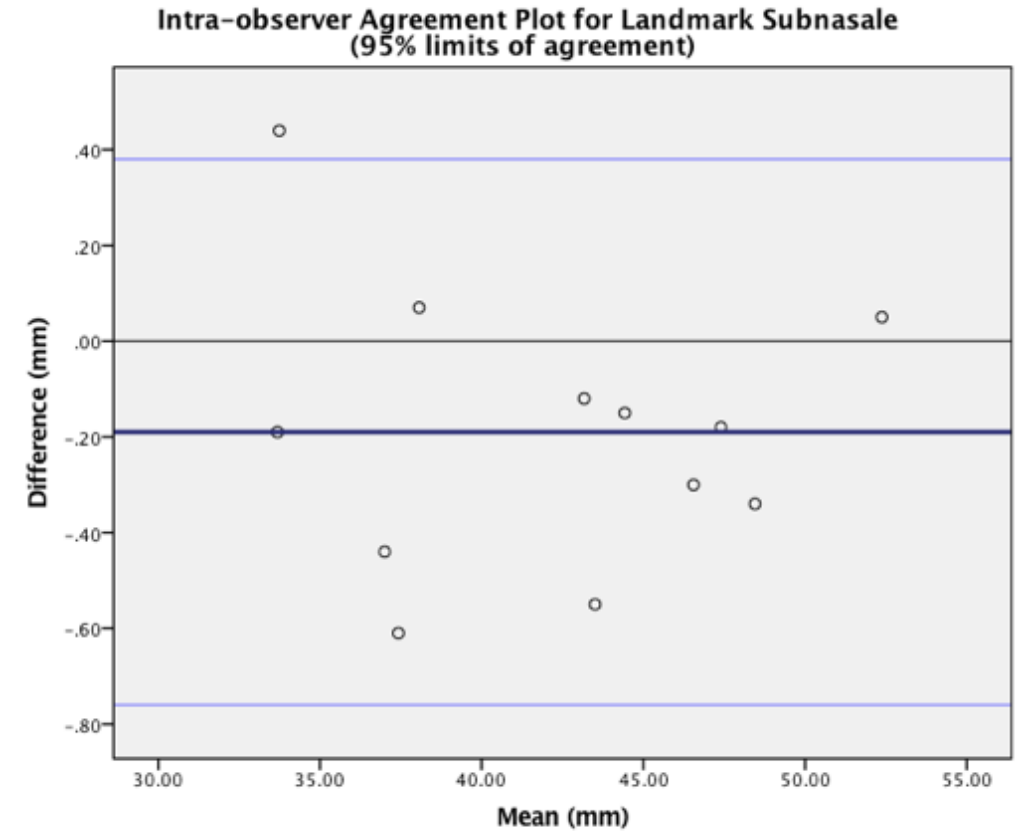
Appendix 1: Sample of Panel Assessment Form

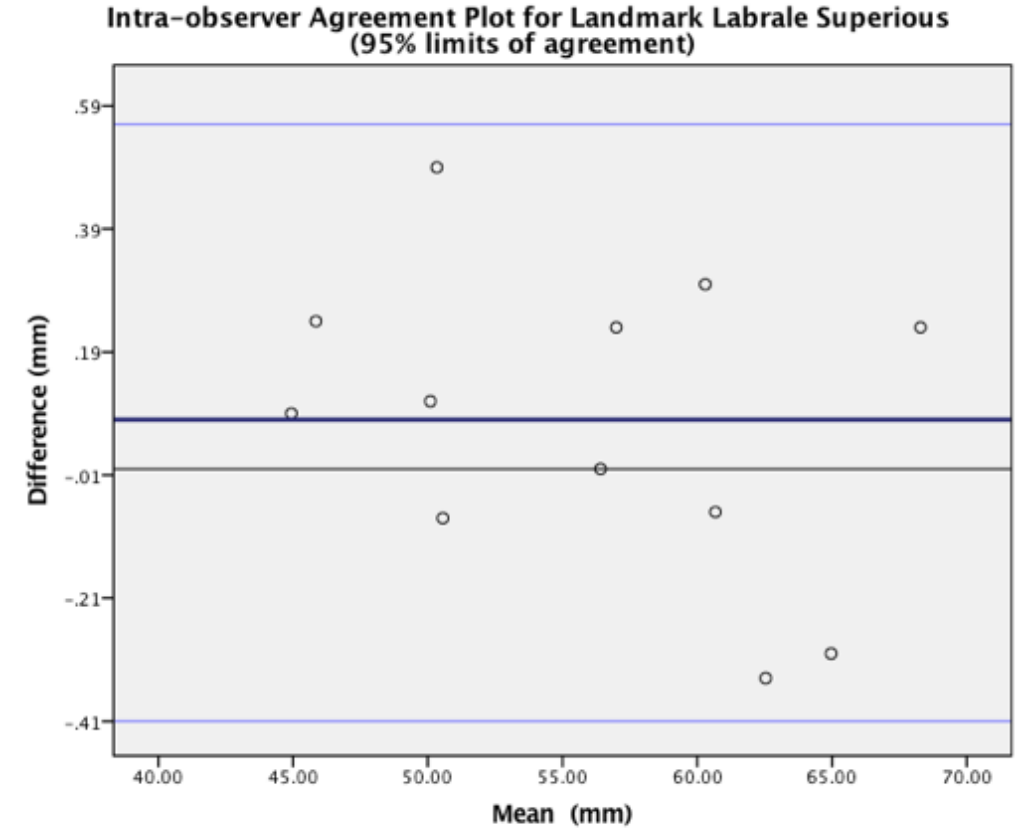
	Patient Number	Yes or No Y / N	Why excluded
1	6		
2	11		
3	16		
4	21		
5	75		
6	76		
7	77		
8	83		
9	87		
10	94		
11	95		
12	112		
13	124		
14	125		
15	133		
16	137		
17	138		
18	143		
19	145		

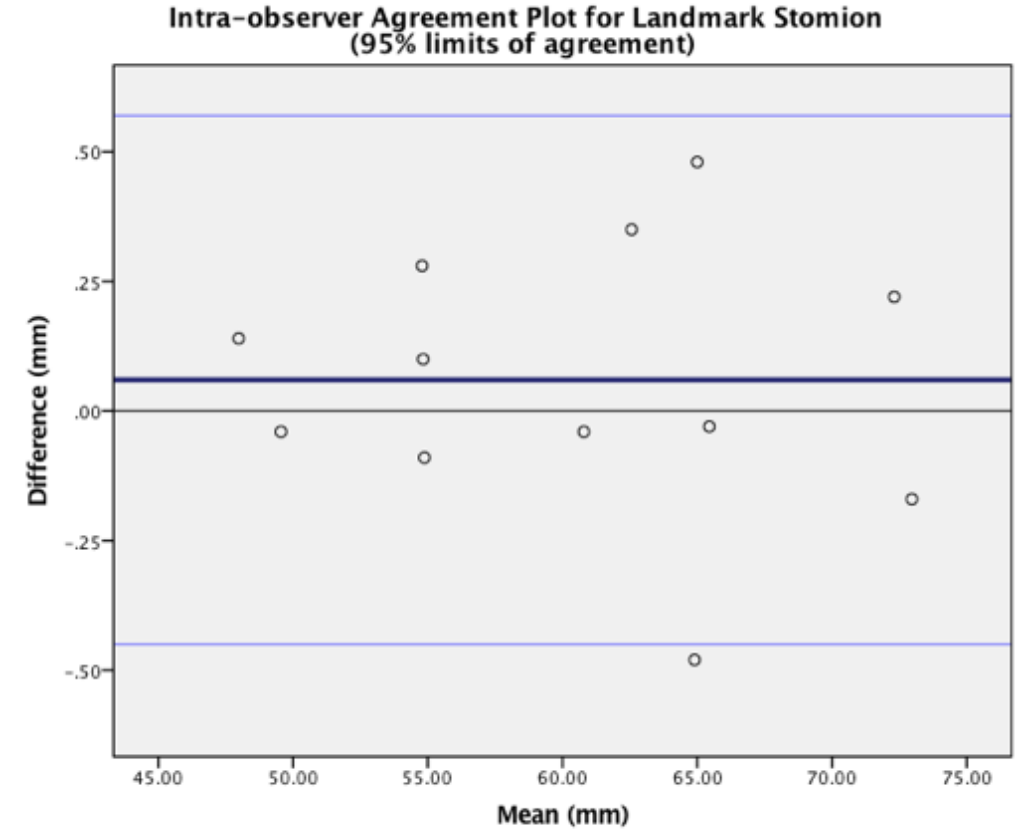
Appendix 2: Bland-Altman Plots for Intra-reliability

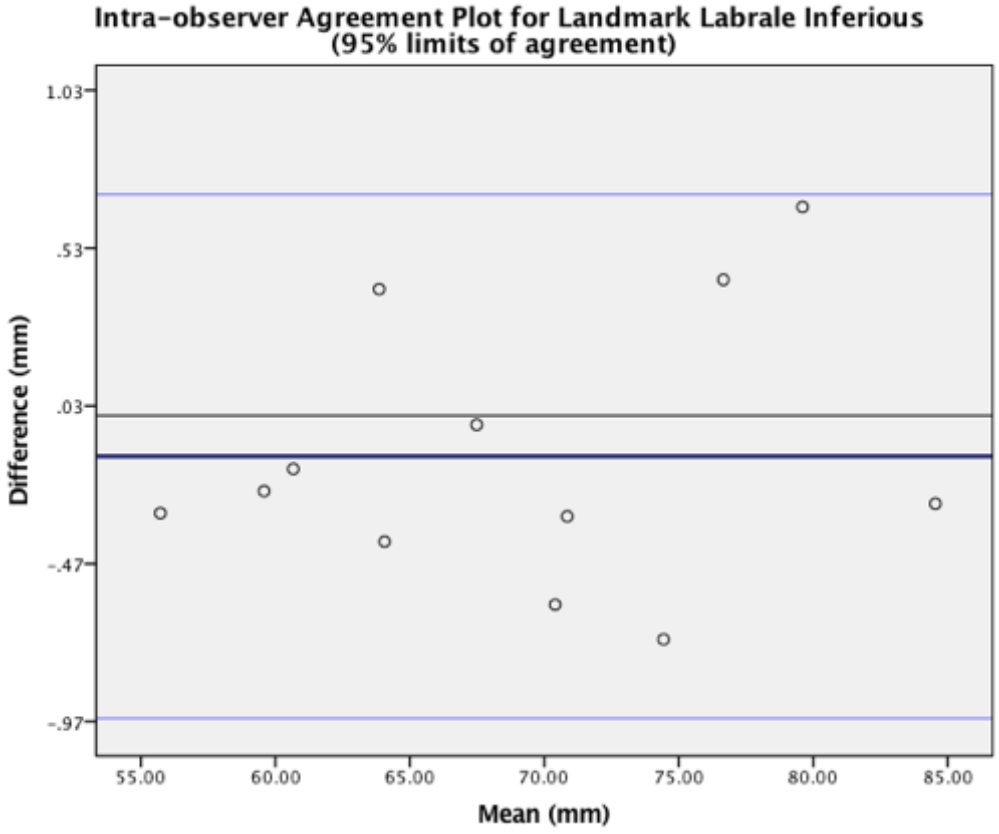


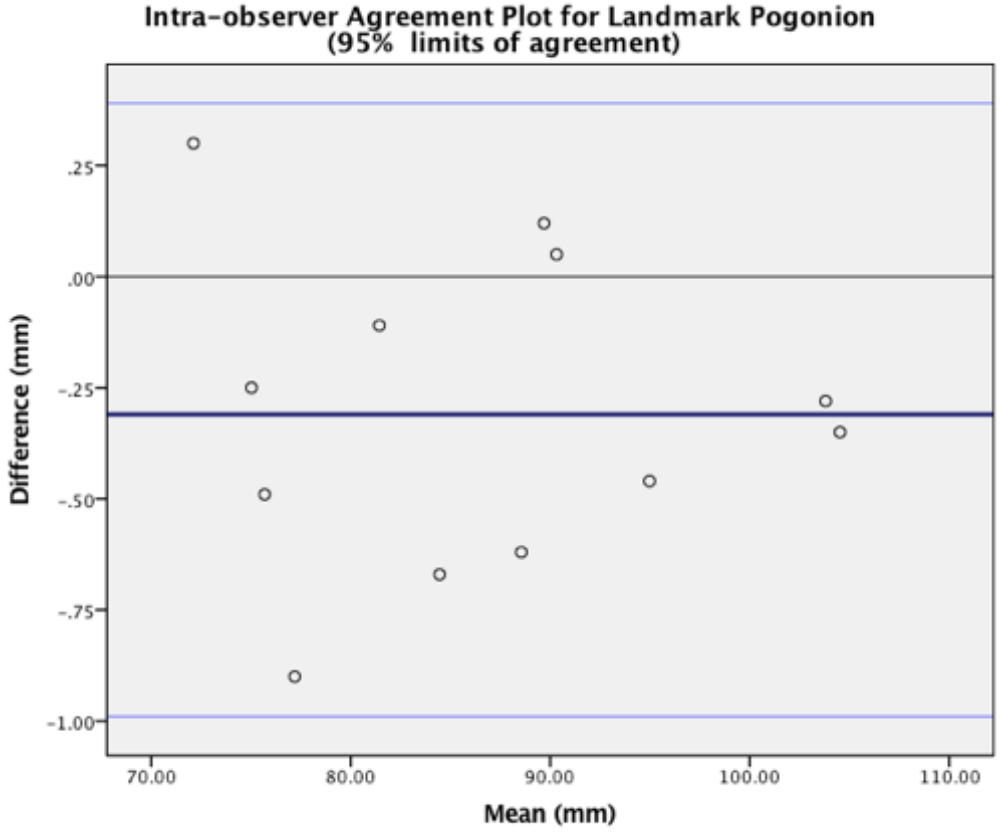

Intra-observer Agreement Plot for Landmark Pronasale
(95% limits of agreement)
Difference (mm)
.63
.38
.13
-.12
-.37
-.62
-.87
25.00
30.00
35.00
40.00
45.00
Mean (mm)

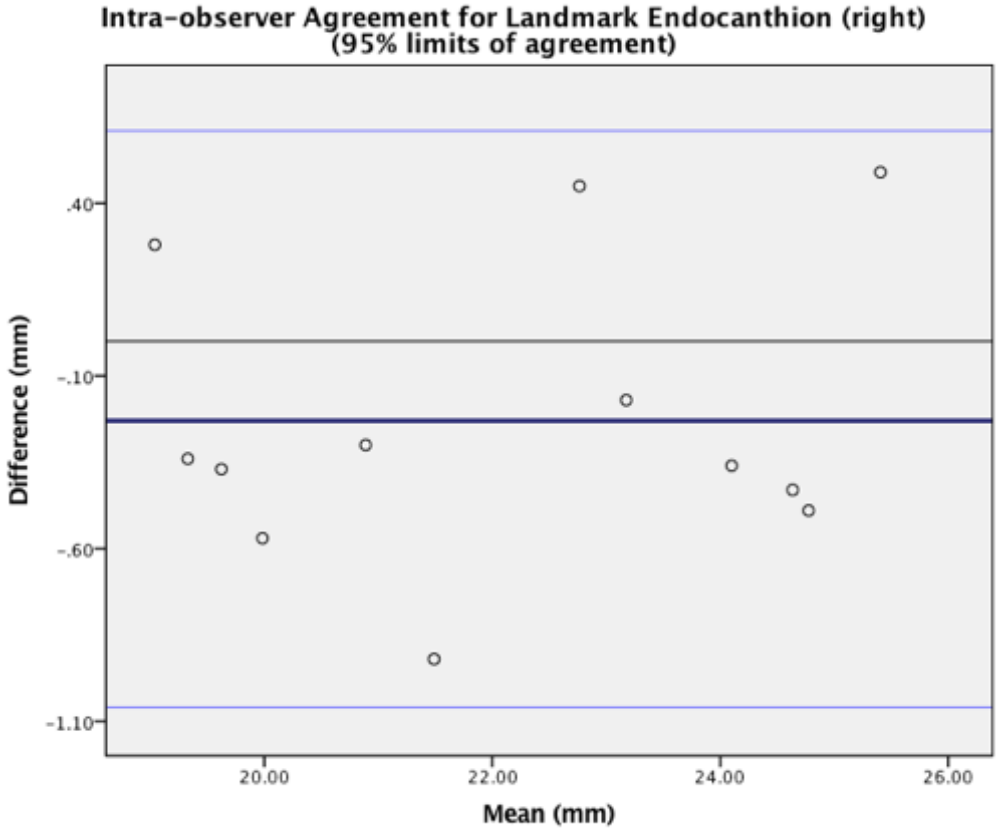

Intra-observer Agreement Plot for Landmark Subnasale
(95% limits of agreement)
Difference (mm)
.40
.20
.00
-.20
-.40
-.60
-.80
30.00
35.00
40.00
45.00
50.00
55.00
Mean (mm)

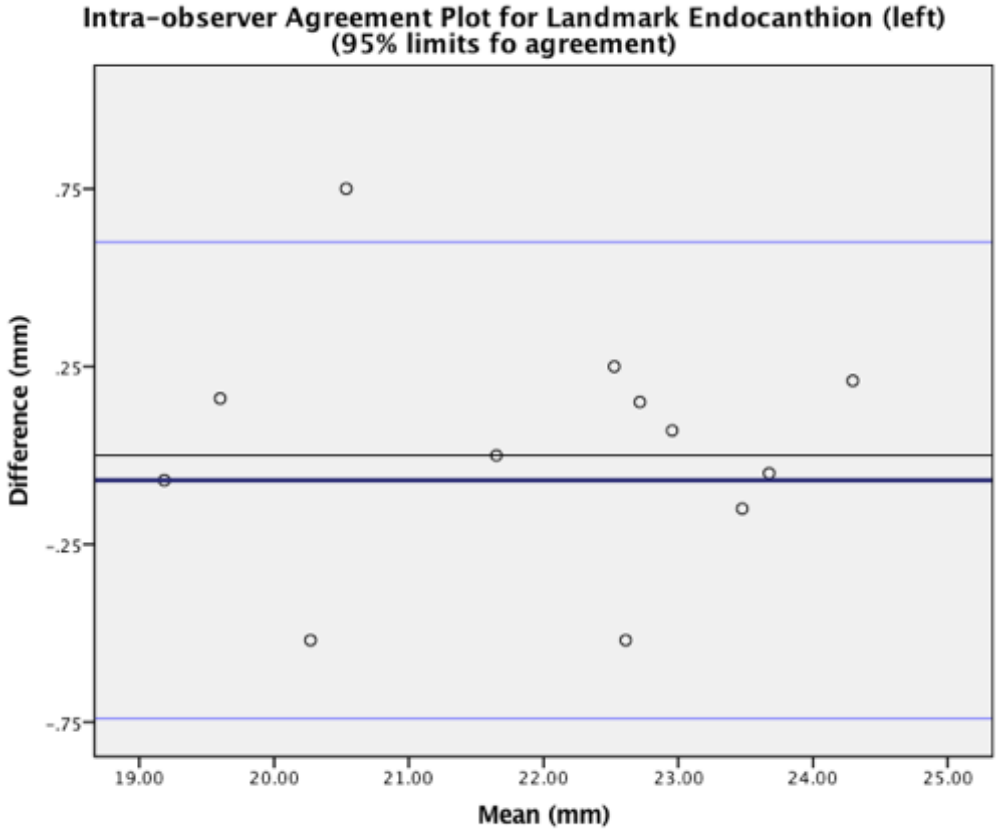

Intra-observer Agreement Plot for Landmark Labrale Superious
(95% limits of agreement)
Difference (mm)
.59
.39
.19
-.01
-.21
-.41
40.00
45.00
50.00
55.00
60.00
65.00
70.00
Mean (mm)

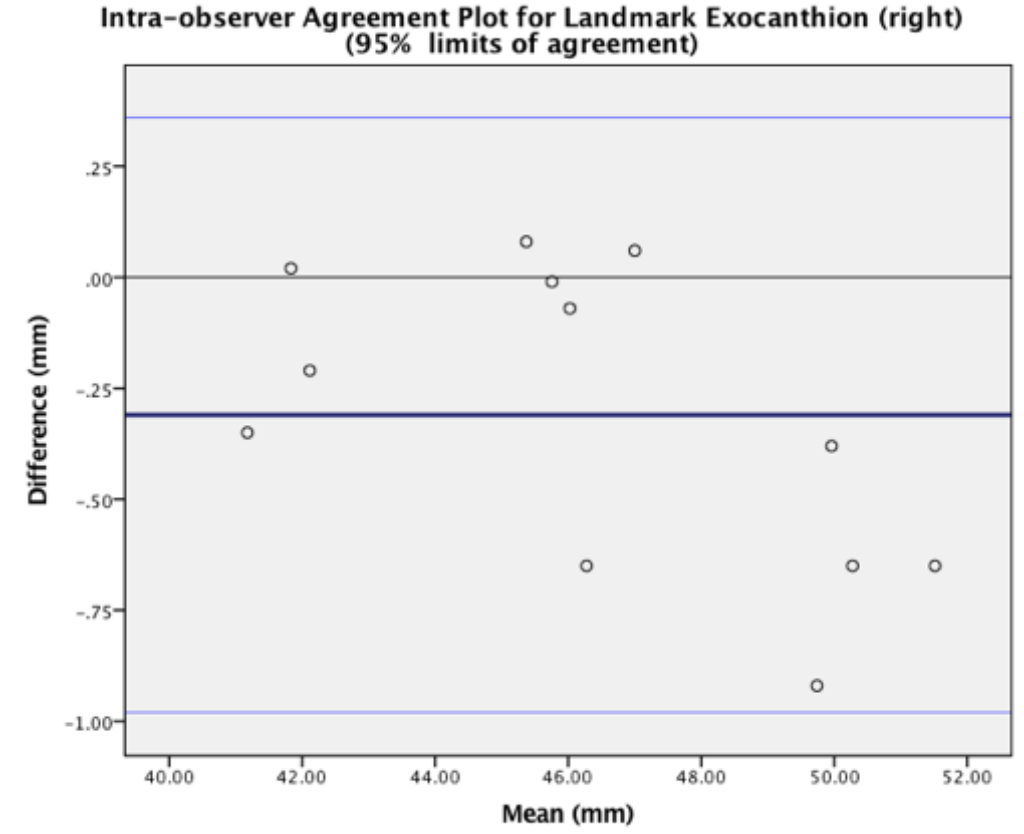

Intra-observer Agreement Plot for Landmark Stomion
(95% limits of agreement)
Difference (mm)
.50
.25
.00
-.25
-.50
45.00
50.00
55.00
60.00
65.00
70.00
75.00
Mean (mm)

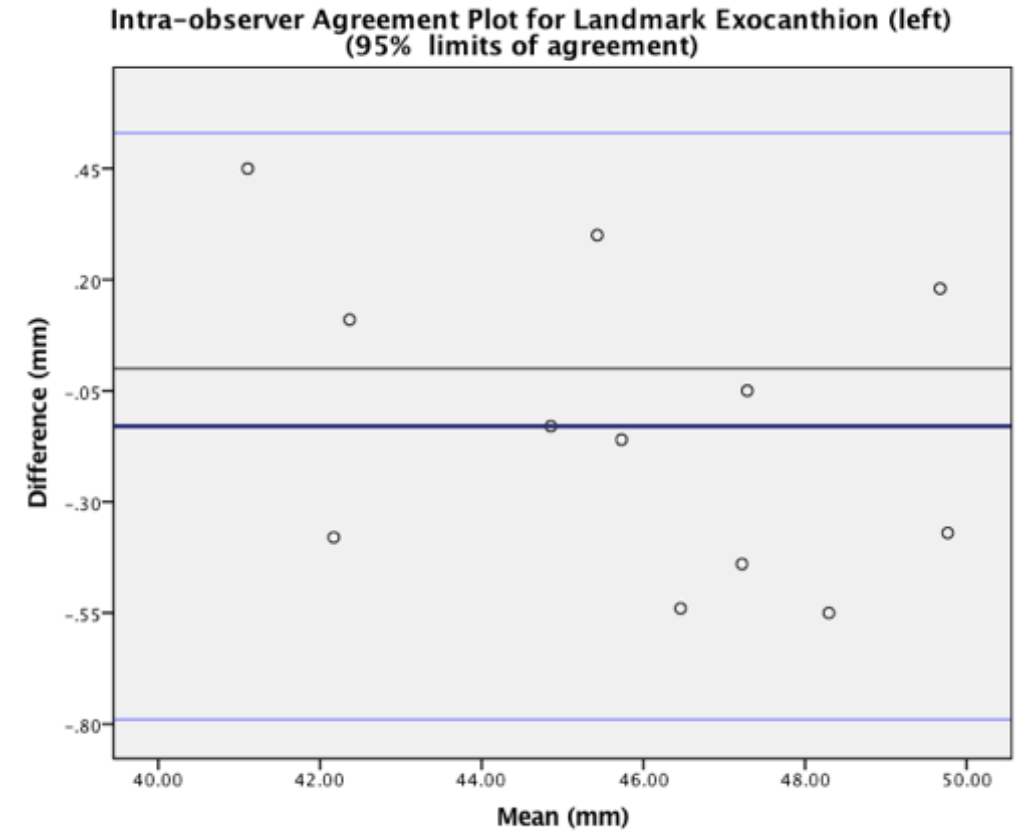

Intra-observer Agreement Plot for Landmark Labrale Inferious
(95% limits of agreement)
Difference (mm)
1.03
.53
.03
-.47
-.97
55.00
60.00
65.00
70.00
75.00
80.00
85.00
Mean (mm)

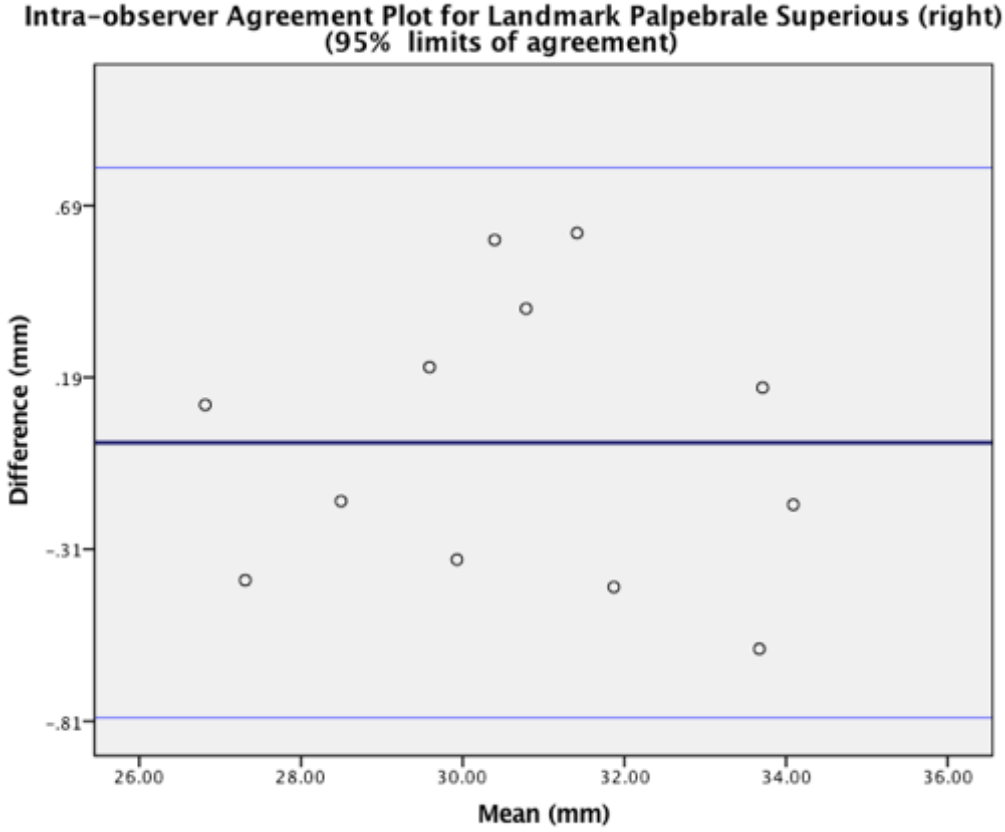

Intra-observer Agreement Plot for Landmark Pogonion
(95% limits of agreement)
Difference (mm)
.25
.00
-.25
-.50
-.75
-1.00
70.00
80.00
90.00
100.00
110.00
Mean (mm)

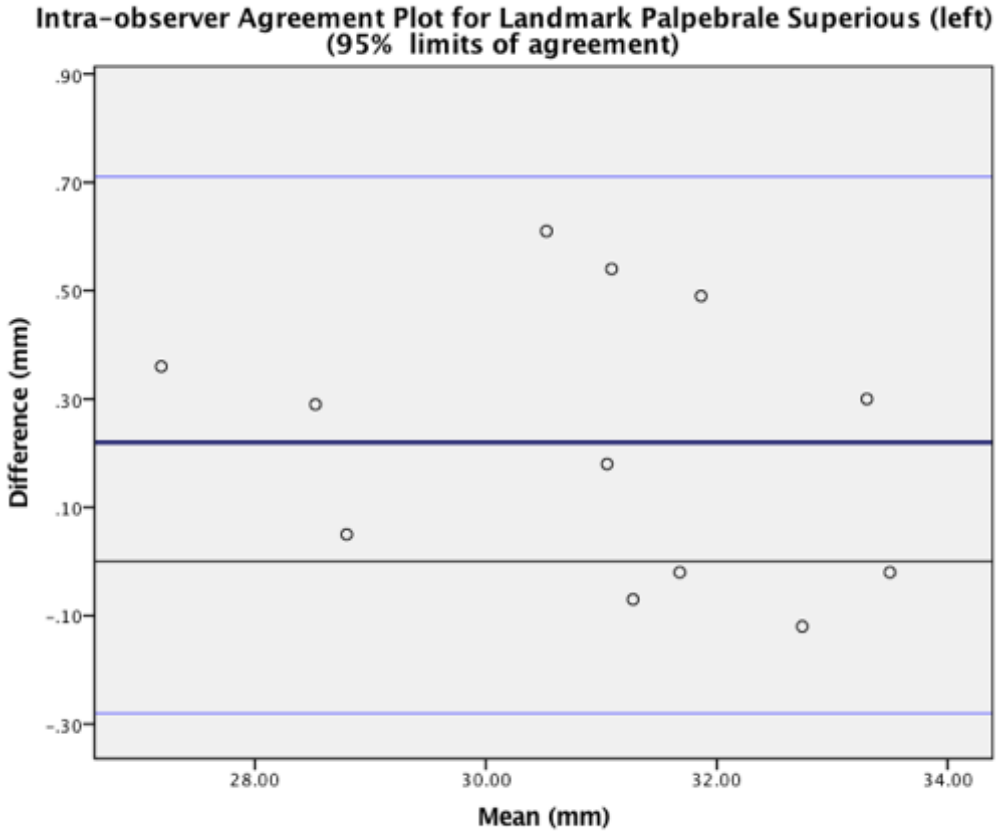

Intra-observer Agreement for Landmark Endocanthion (right)
(95% limits of agreement)
Difference (mm)
.40
-.10
-.60
-1.10
20.00
22.00
24.00
26.00
Mean (mm)

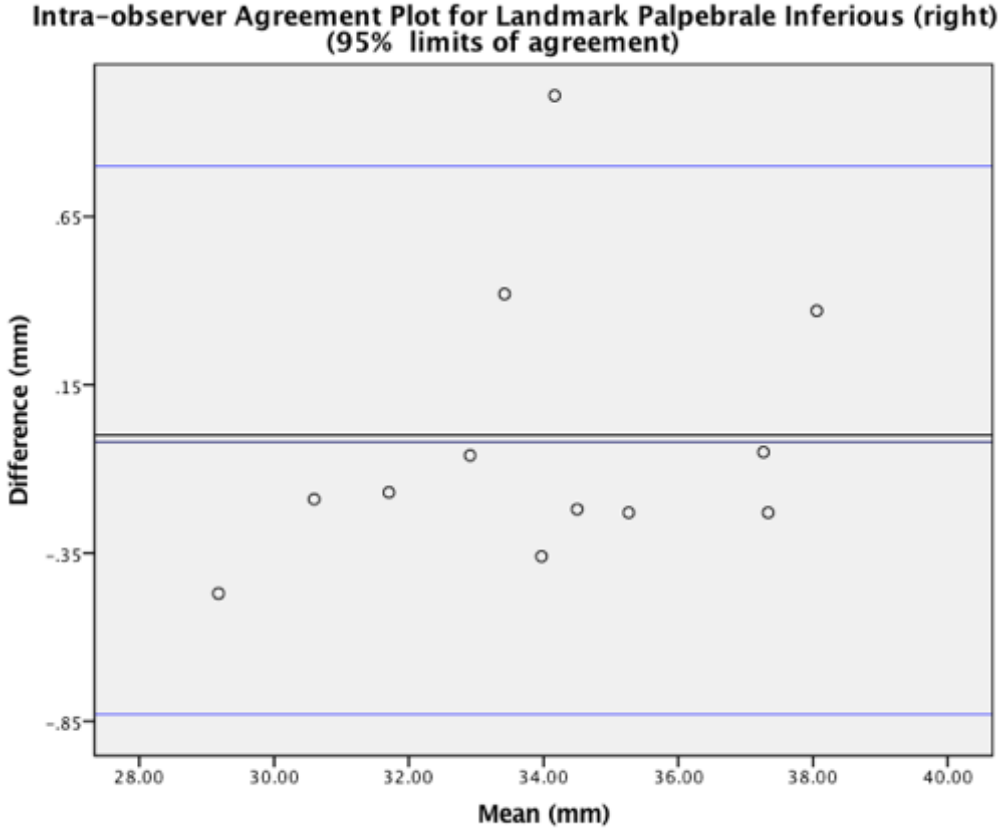

Intra-observer Agreement Plot for Landmark Endocanthion (left)
(95% limits fo agreement)
Difference (mm)
.75
.25
-.25
-.75
19.00
20.00
21.00
22.00
23.00
24.00
25.00
Mean (mm)

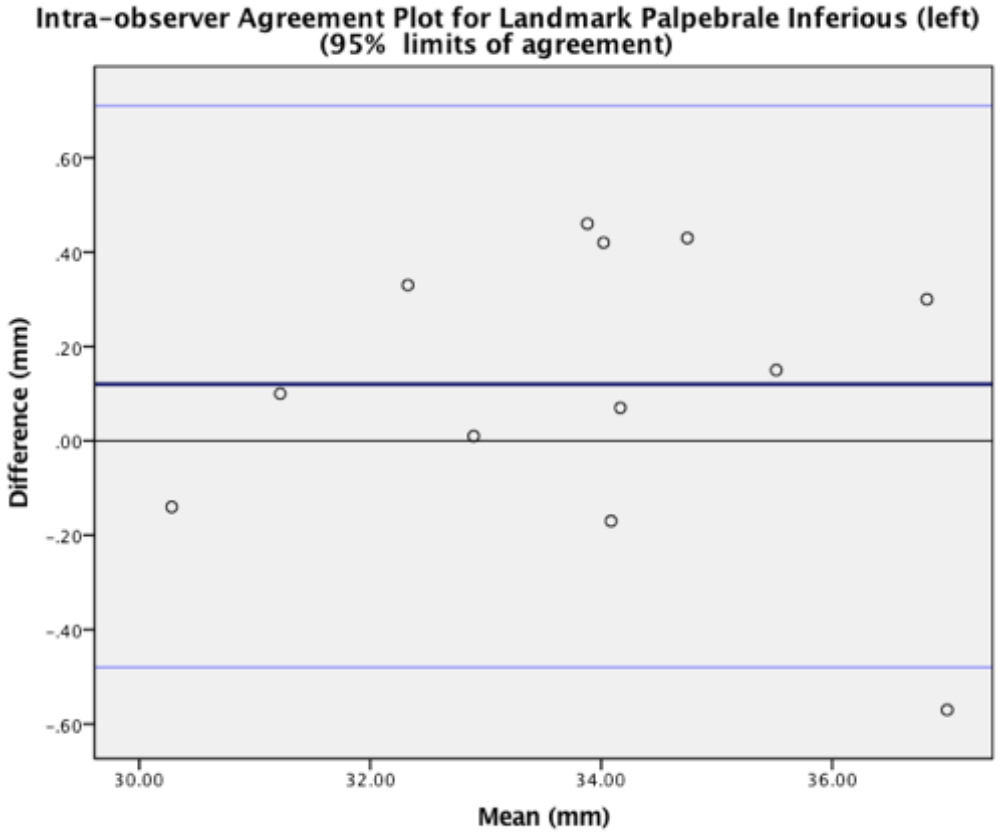

Intra-observer Agreement Plot for Landmark Exocanthion (right)
(95% limits of agreement)
Difference (mm)
.25
.00
-.25
-.50
-.75
-1.00
40.00
42.00
44.00
46.00
48.00
50.00
52.00
Mean (mm)

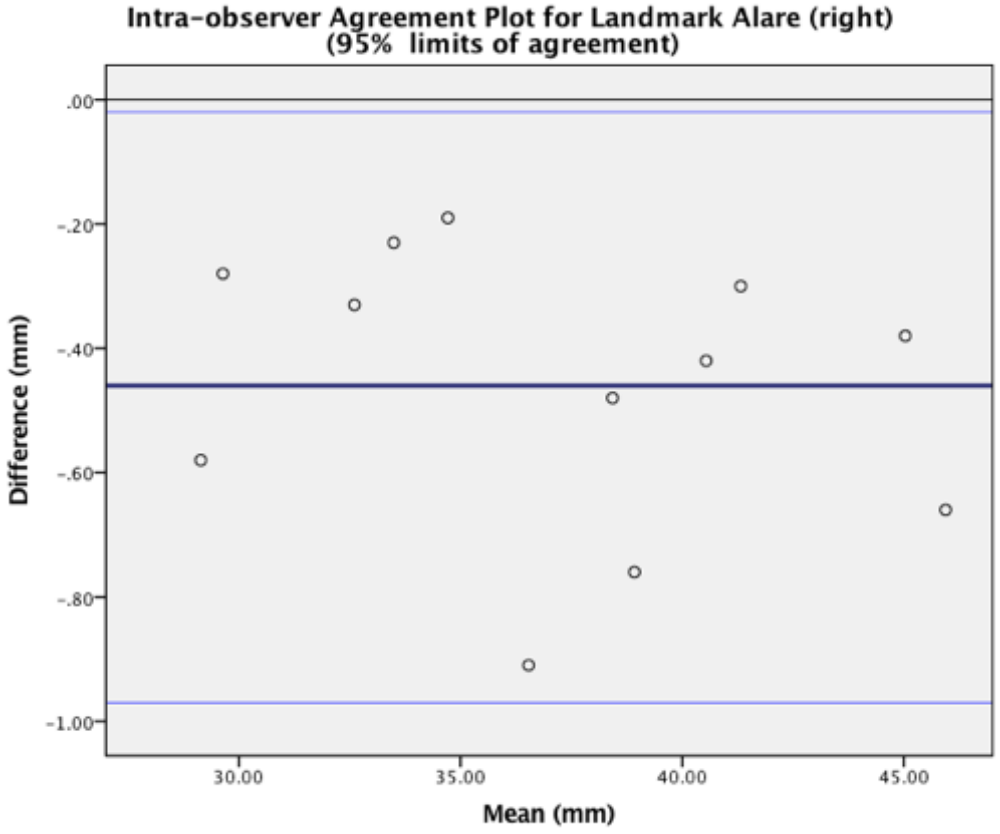

Intra-observer Agreement Plot for Landmark Exocanthion (left)
(95% limits of agreement)
Difference (mm)
.45
.20
-.05
-.30
-.55
-.80
40.00
42.00
44.00
46.00
48.00
50.00
Mean (mm)

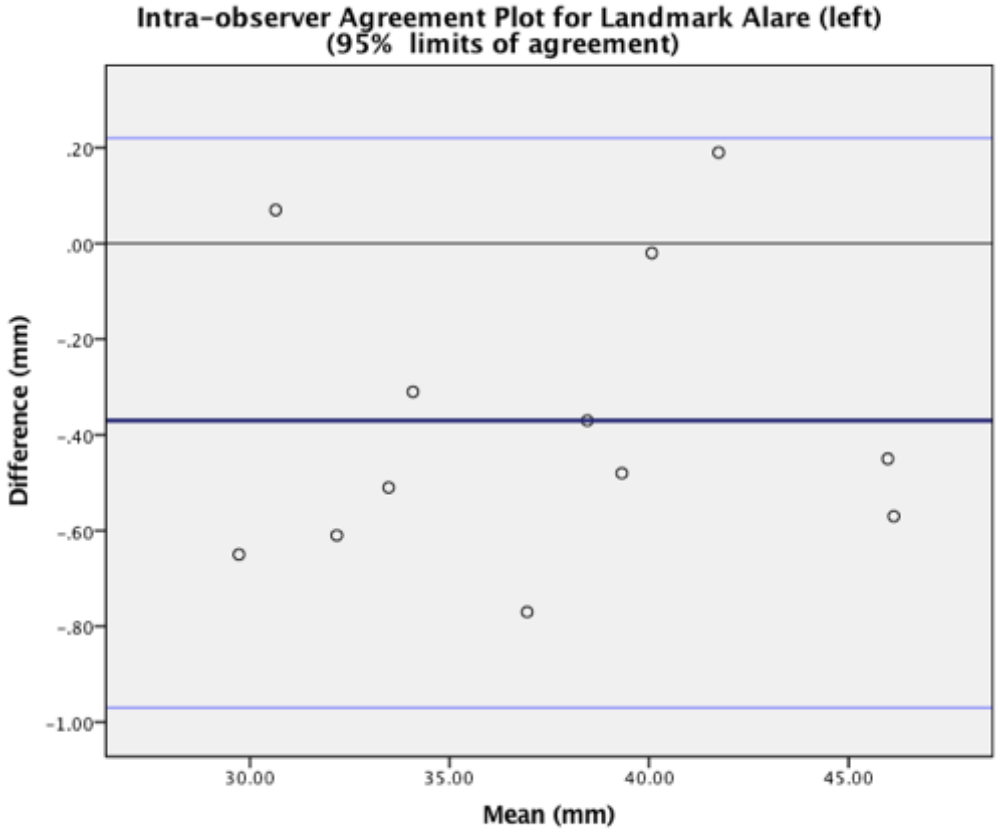

Intra-observer Agreement Plot for Landmark Palpebrale Superious (right)
(95% limits of agreement)
Difference (mm)
.69
.19
-.31
-.81
26.00
28.00
30.00
32.00
34.00
36.00
Mean (mm)

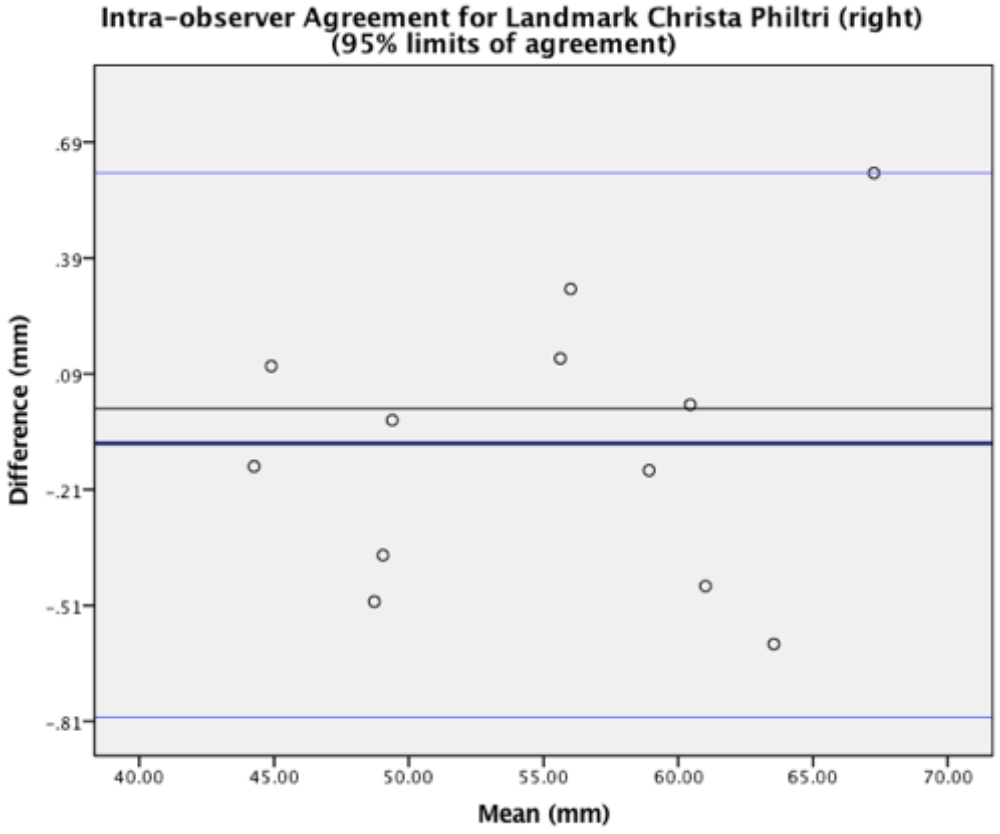

Intra-observer Agreement Plot for Landmark Palpebrale Superious (left)
(95% limits of agreement)
Difference (mm)
.90
.70
.50
.30
.10
-.10
-.30
28.00
30.00
32.00
34.00
Mean (mm)

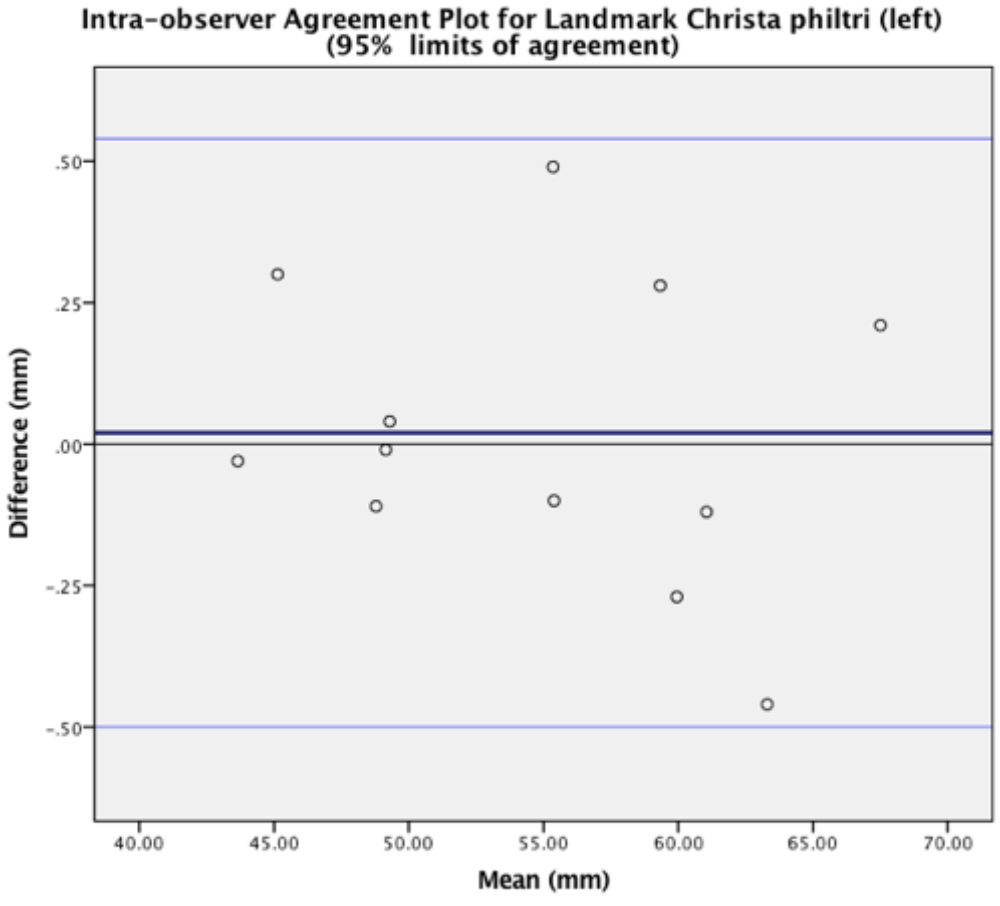

Intra-observer Agreement Plot for Landmark Palpebrale Inferious (right)
(95% limits of agreement)
Difference (mm)
.65
.15
-.35
-.85
28.00
30.00
32.00
34.00
36.00
38.00
40.00
Mean (mm)

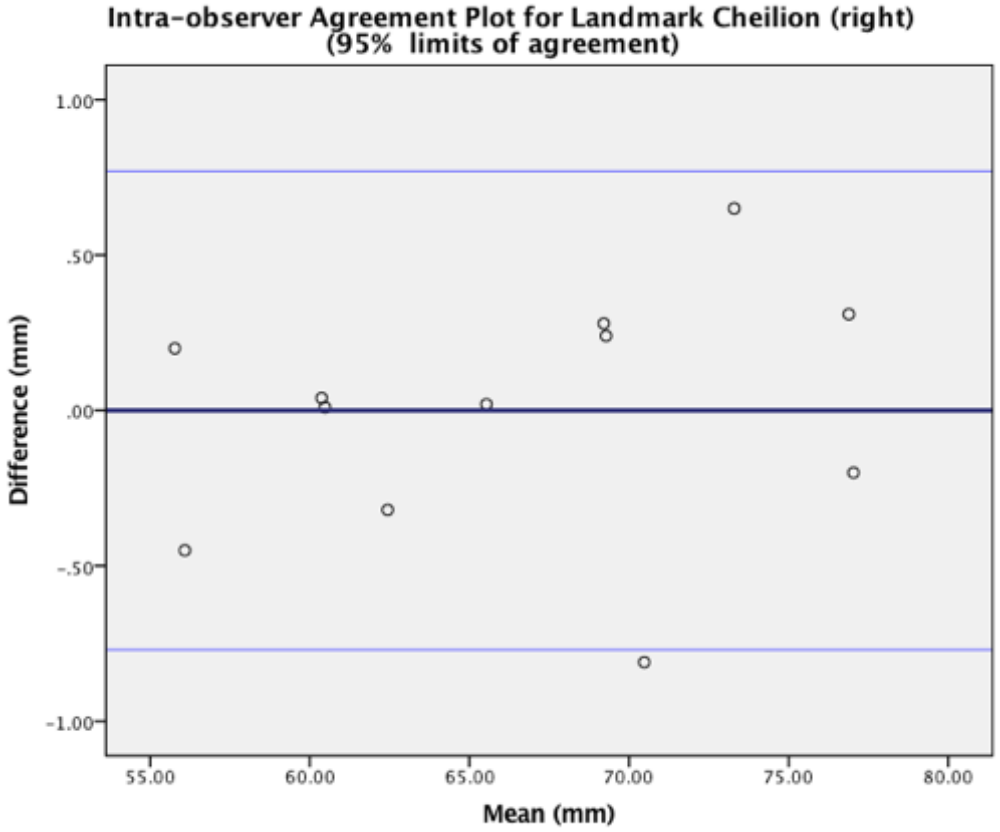

Intra-observer Agreement Plot for Landmark Palpebrale Inferious (left)
(95% limits of agreement)
Difference (mm)
.60
.40
.20
.00
-.20
-.40
-.60
30.00
32.00
34.00
36.00
Mean (mm)

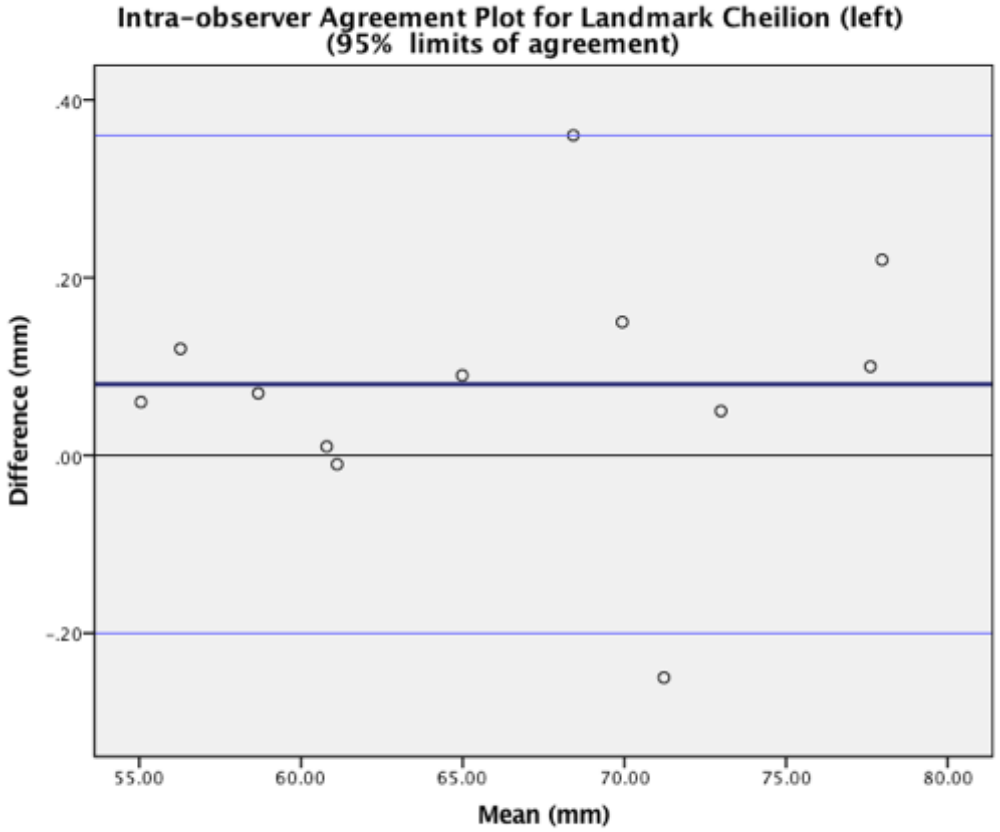

Intra-observer Agreement Plot for Landmark Alare (right)
(95% limits of agreement)
Difference (mm)
.00
-.20
-.40
-.60
-.80
-1.00
30.00
35.00
40.00
45.00
Mean (mm)

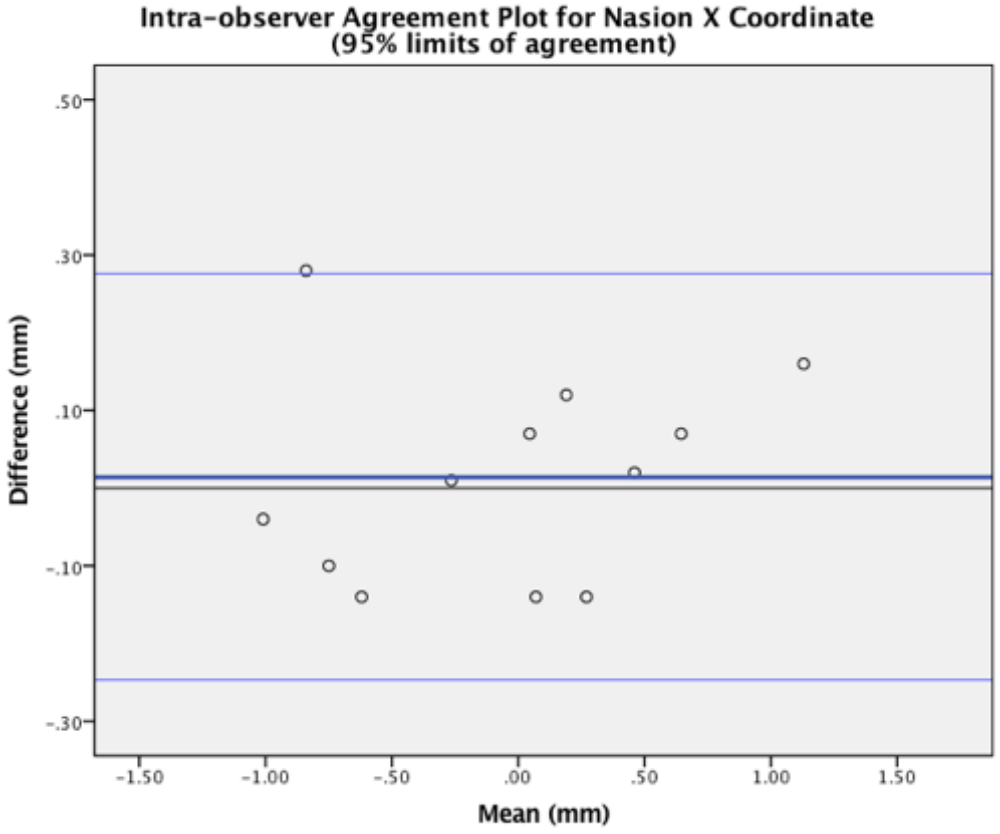

Intra-observer Agreement Plot for Landmark Alare (left)
(95% limits of agreement)
Difference (mm)
.20
.00
-.20
-.40
-.60
-.80
-1.00
30.00
35.00
40.00
45.00
Mean (mm)

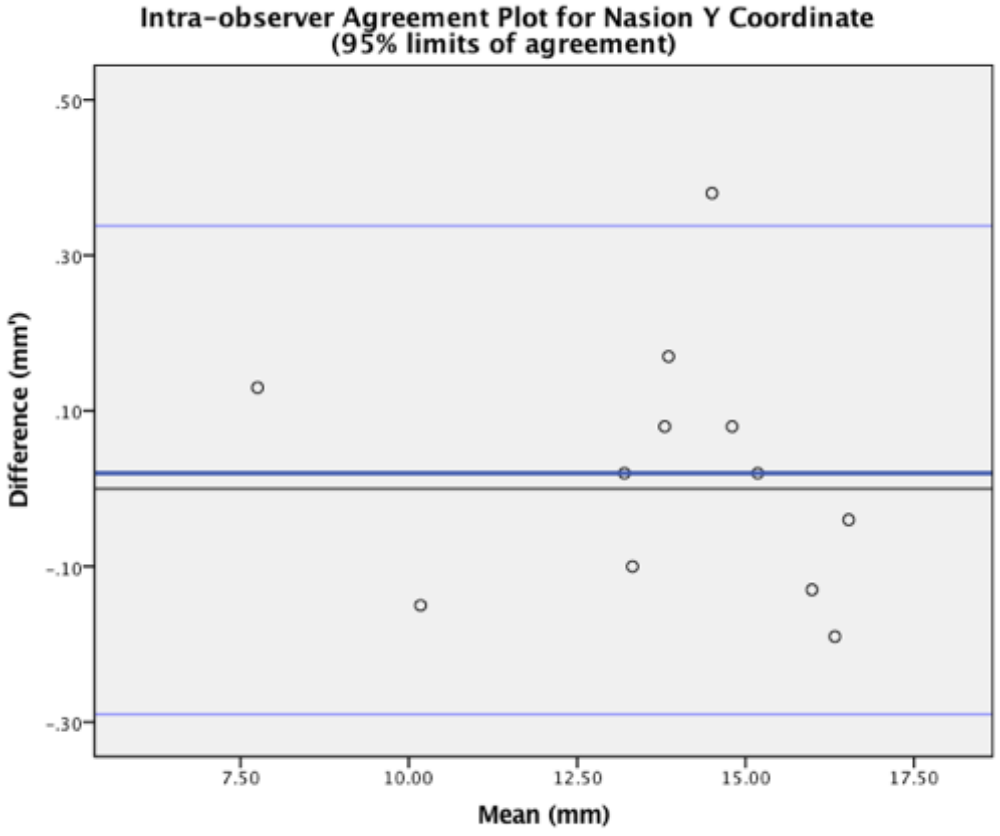

Intra-observer Agreement for Landmark Christa Philtri (right)
(95% limits of agreement)
Difference (mm)
.69
.39
.09
-.21
-.51
-.81
40.00
45.00
50.00
55.00
60.00
65.00
70.00
Mean (mm)

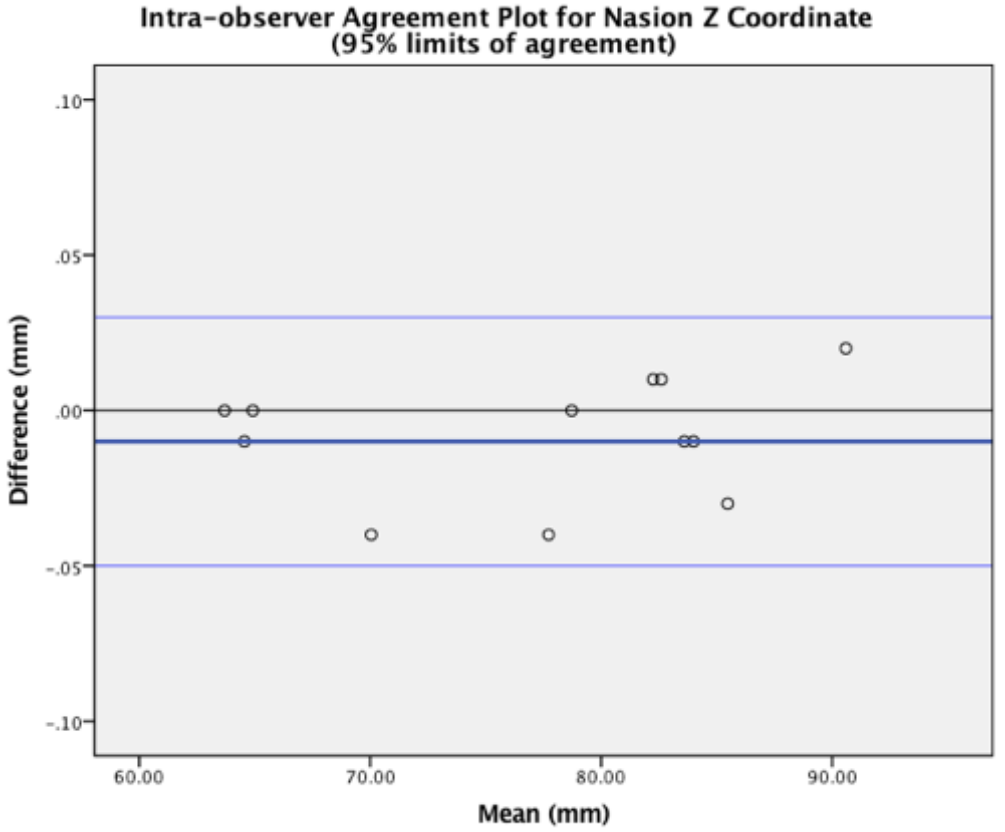

Intra-observer Agreement Plot for Landmark Christa philtri (left)
(95% limits of agreement)
Difference (mm)
.50
.25
.00
-.25
-.50
40.00
45.00
50.00
55.00
60.00
65.00
70.00
Mean (mm)


Intra-observer Agreement Plot for Landmark Cheilion (right)
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
55.00
60.00
65.00
70.00
75.00
80.00
Mean (mm)

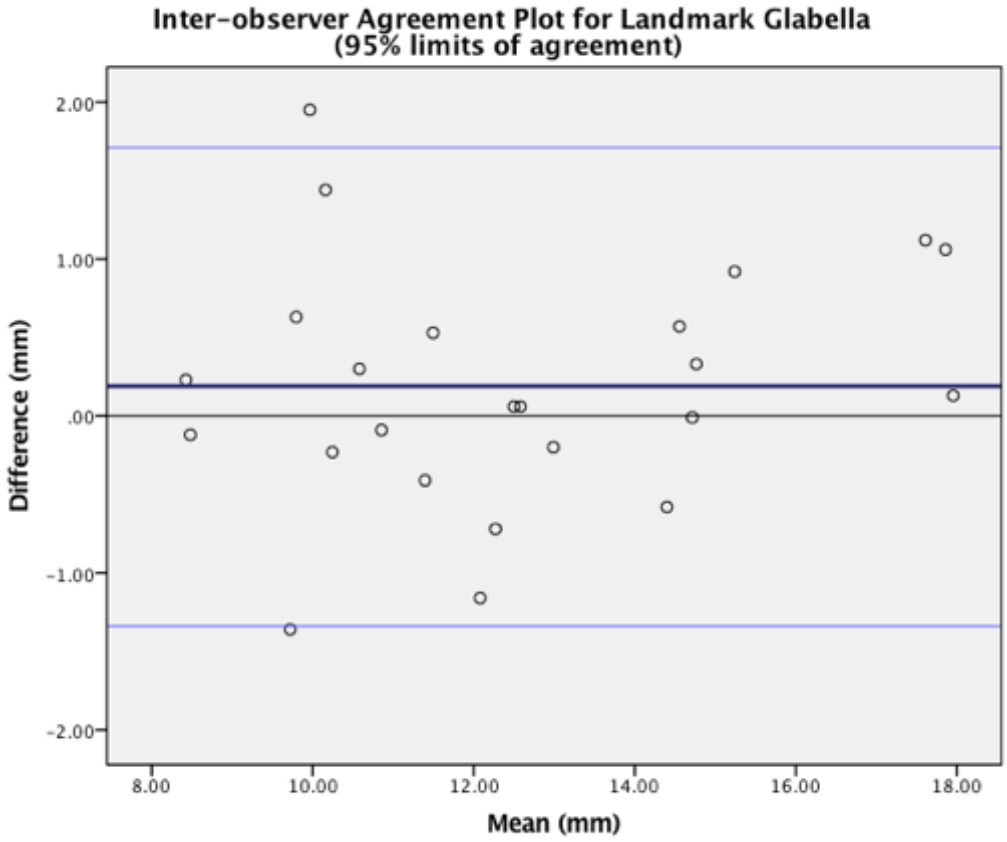

Intra-observer Agreement Plot for Landmark Cheilion (left)
(95% limits of agreement)
Difference (mm)
.40
.20
.00
-.20
55.00
60.00
65.00
70.00
75.00
80.00
Mean (mm)

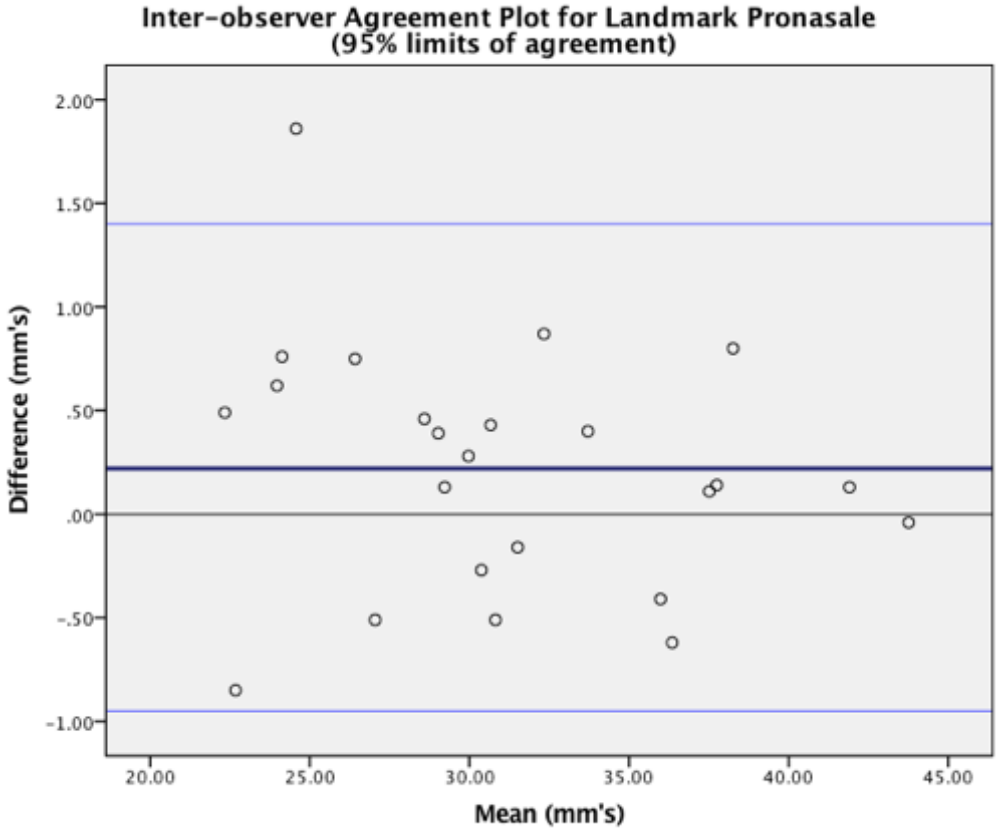

Intra-observer Agreement Plot for Nasion X Coordinate
(95% limits of agreement)
Difference (mm)
.50
.30
.10
-.10
-.30
-1.50
-1.00
-.50
.00
.50
1.00
1.50
Mean (mm)

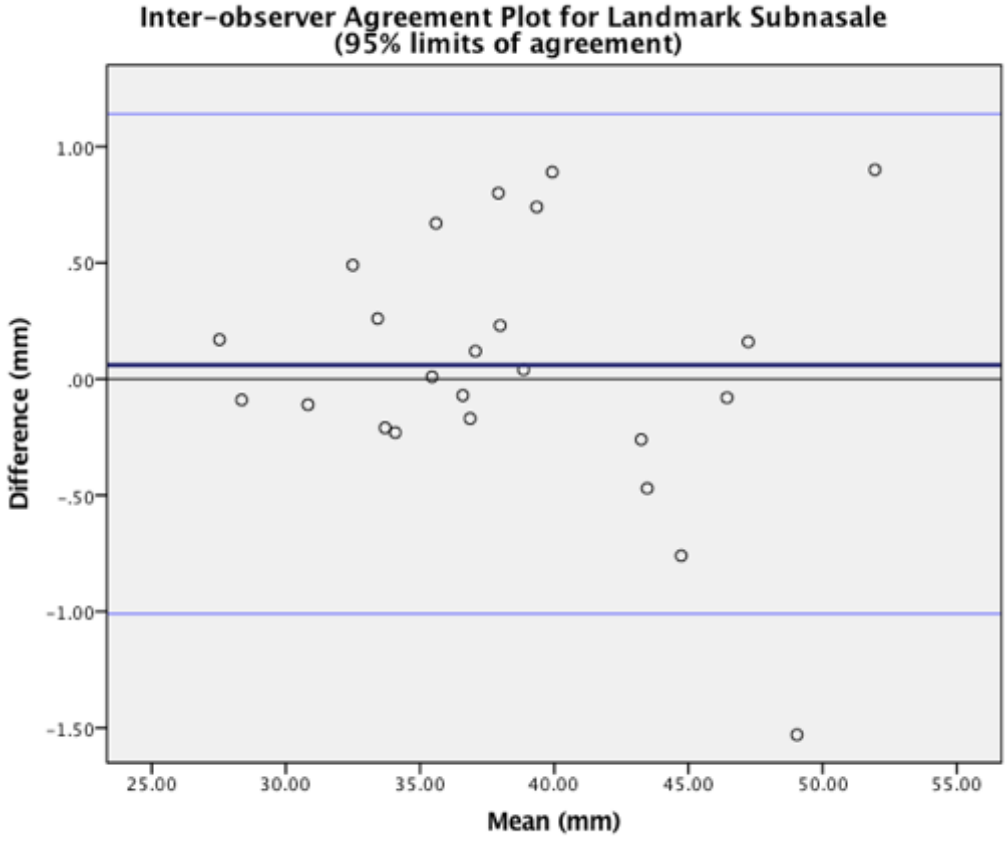

Intra-observer Agreement Plot for Nasion Y Coordinate
(95% limits of agreement)
Difference (mm')
.50
.30
.10
-.10
-.30
7.50
10.00
12.50
15.00
17.50
Mean (mm)

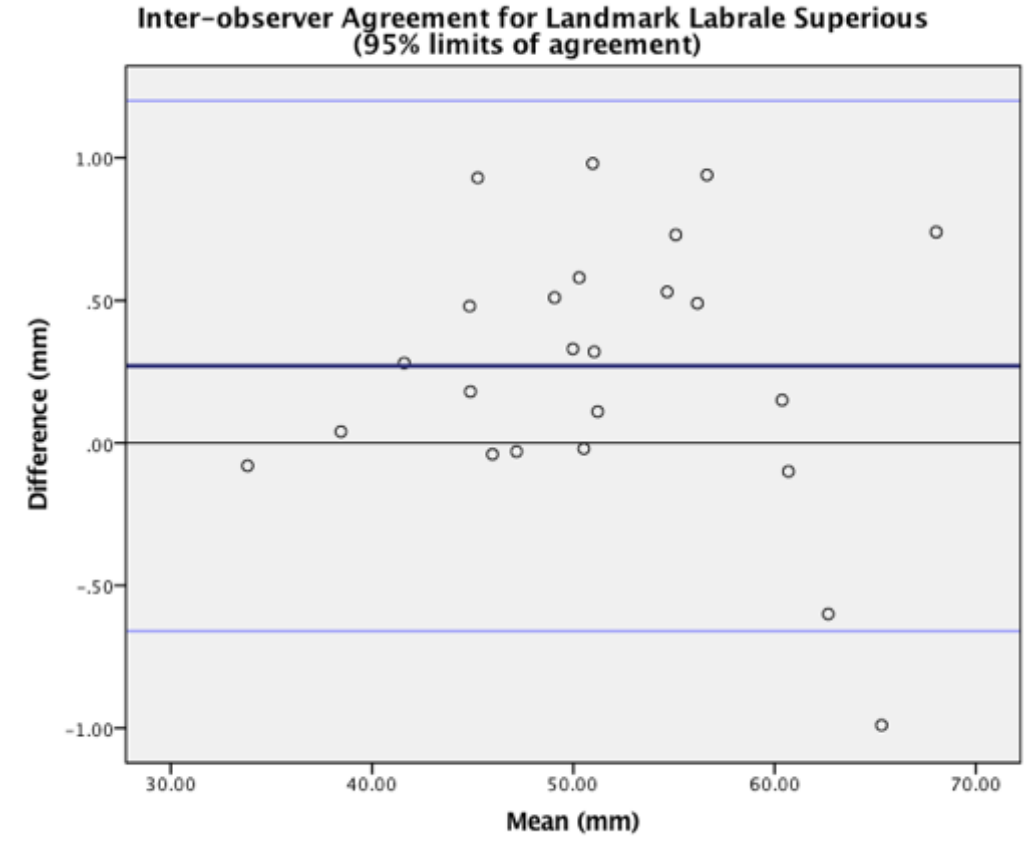

Intra-observer Agreement Plot for Nasion Z Coordinate
(95% limits of agreement)
Difference (mm)
.10
.05
.00
-.05
-.10
60.00
70.00
80.00
90.00
Mean (mm)

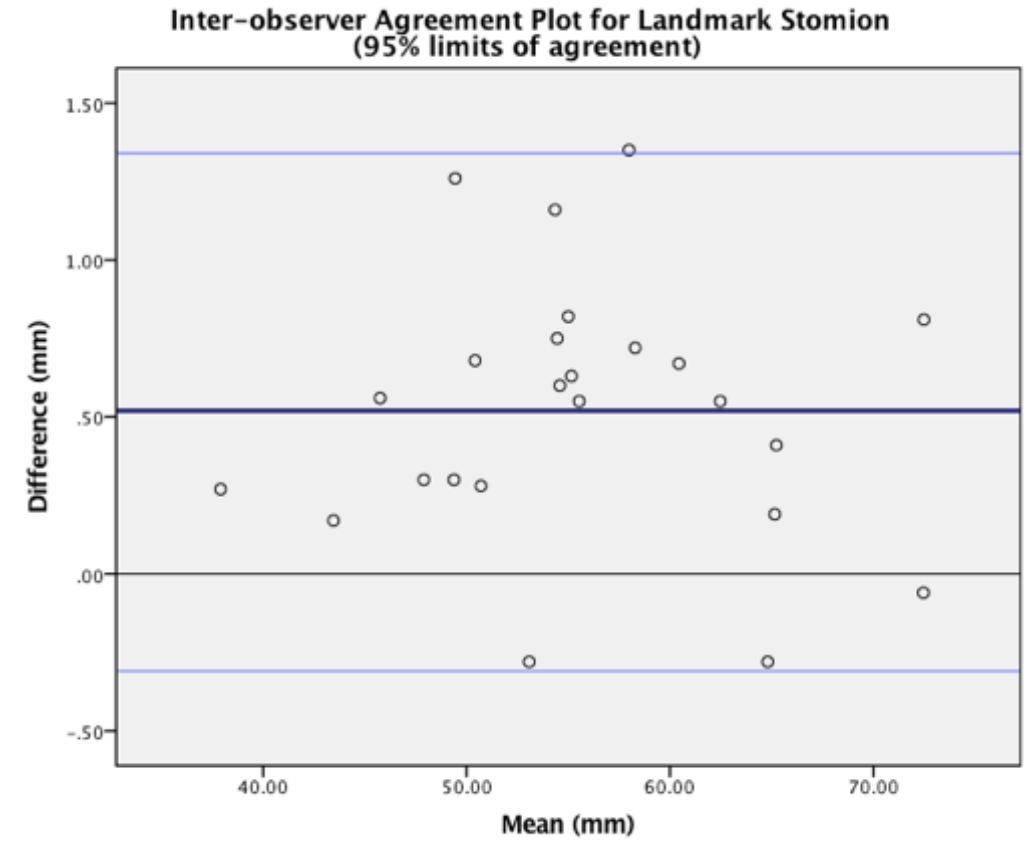
Appendix 3: Bland-Altman Plots for Inter-reliability

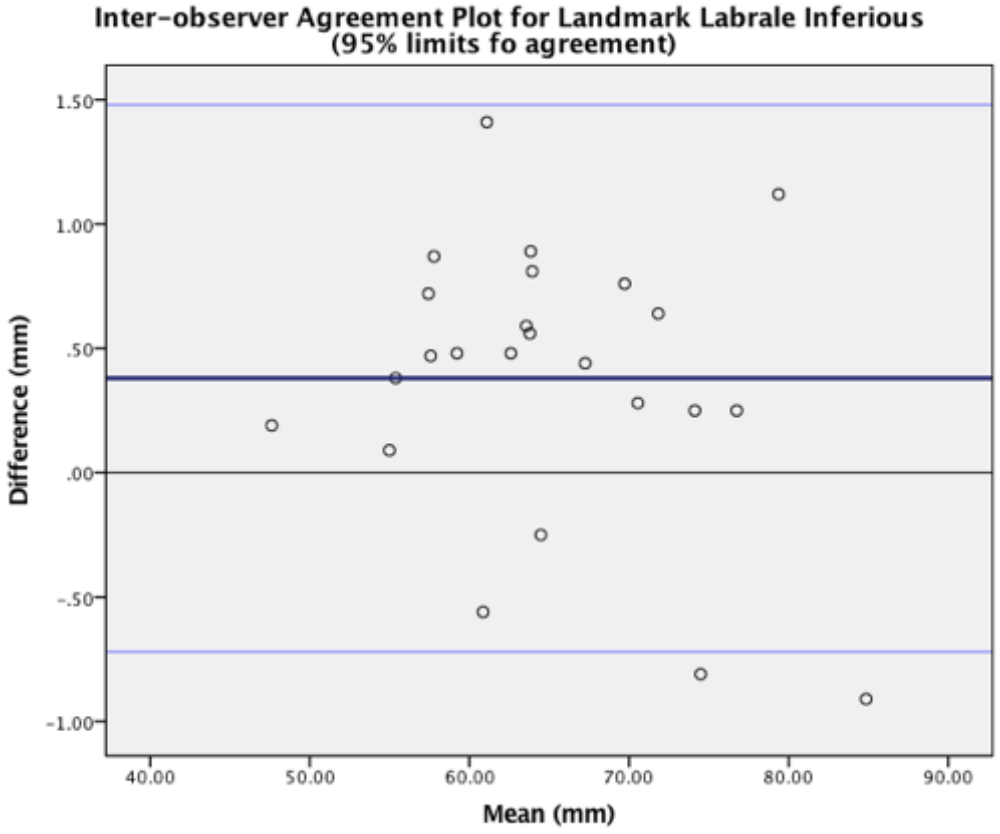


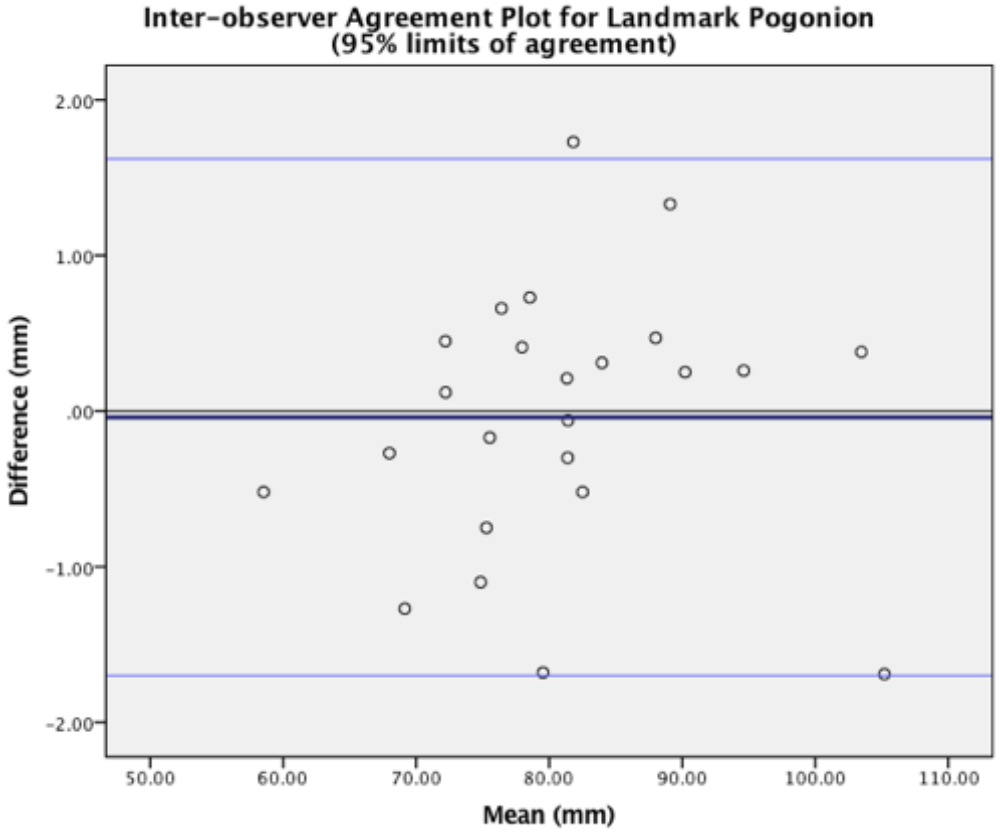

Inter-observer Agreement Plot for Landmark Pronasale
(95% limits of agreement)
Difference (mm's)
2.00
1.50
1.00
.50
.00
-.50
-1.00
20.00
25.00
30.00
35.00
40.00
45.00
Mean (mm's)

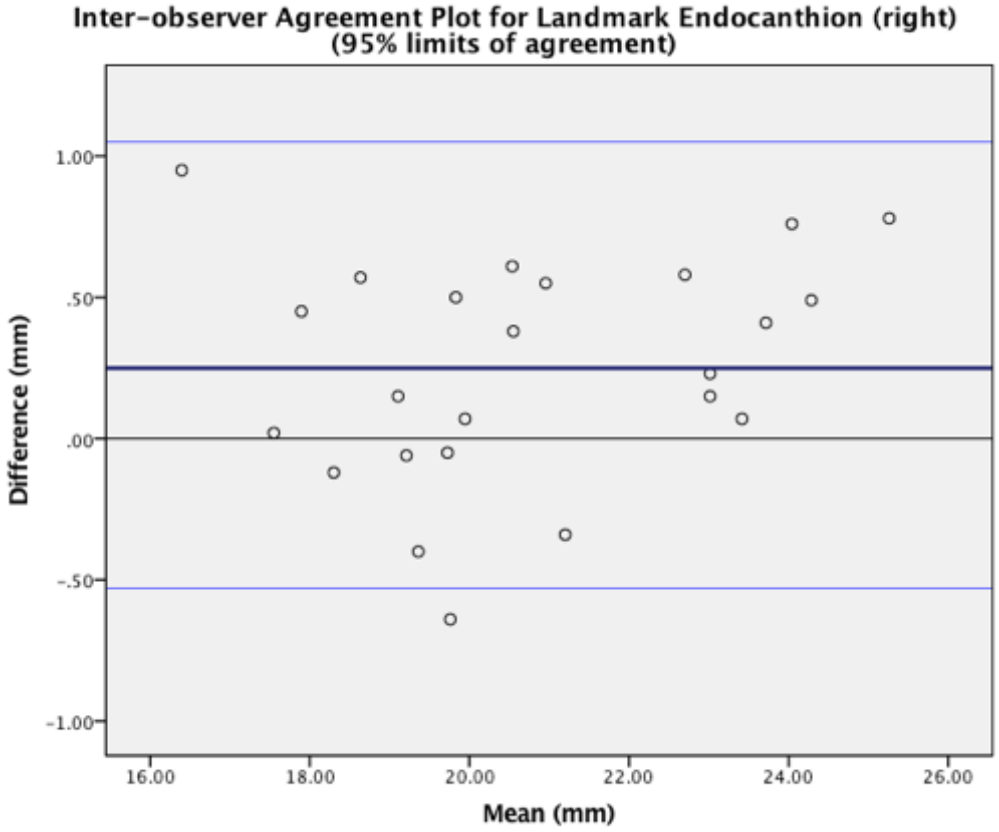

Inter-observer Agreement Plot for Landmark Subnasale
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
-1.50
25.00
30.00
35.00
40.00
45.00
50.00
55.00
Mean (mm)

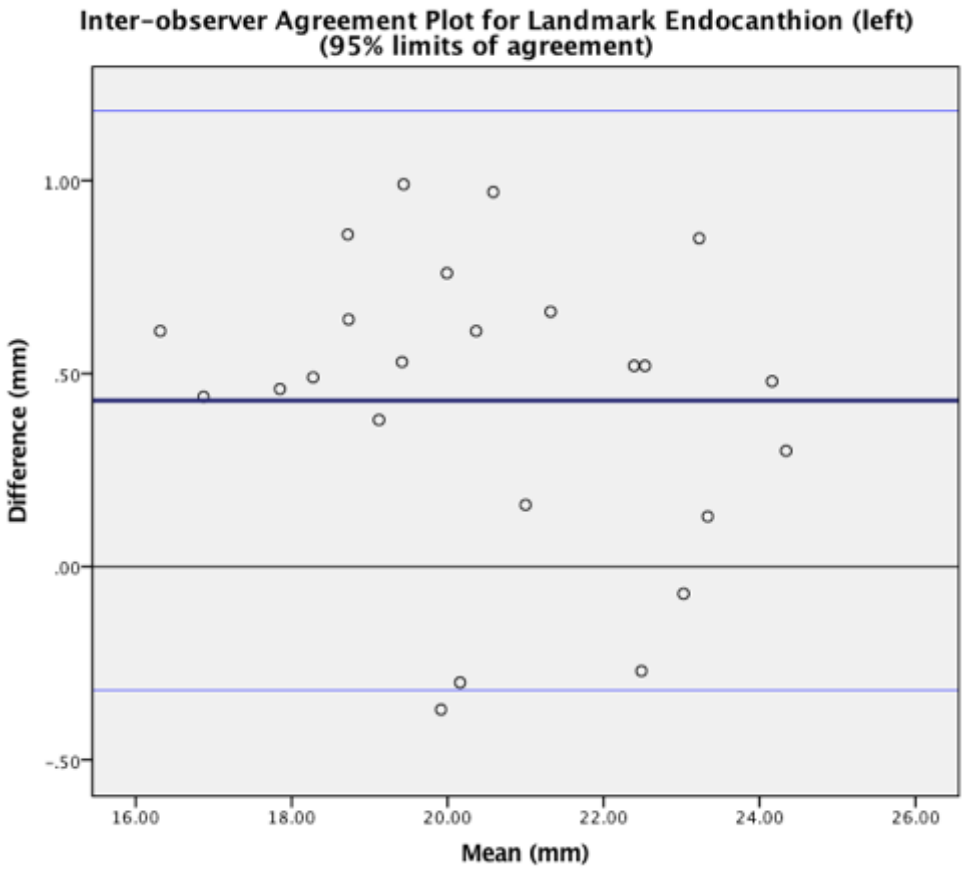

Inter-observer Agreement for Landmark Labrale Superious
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
30.00
40.00
50.00
60.00
70.00
Mean (mm)

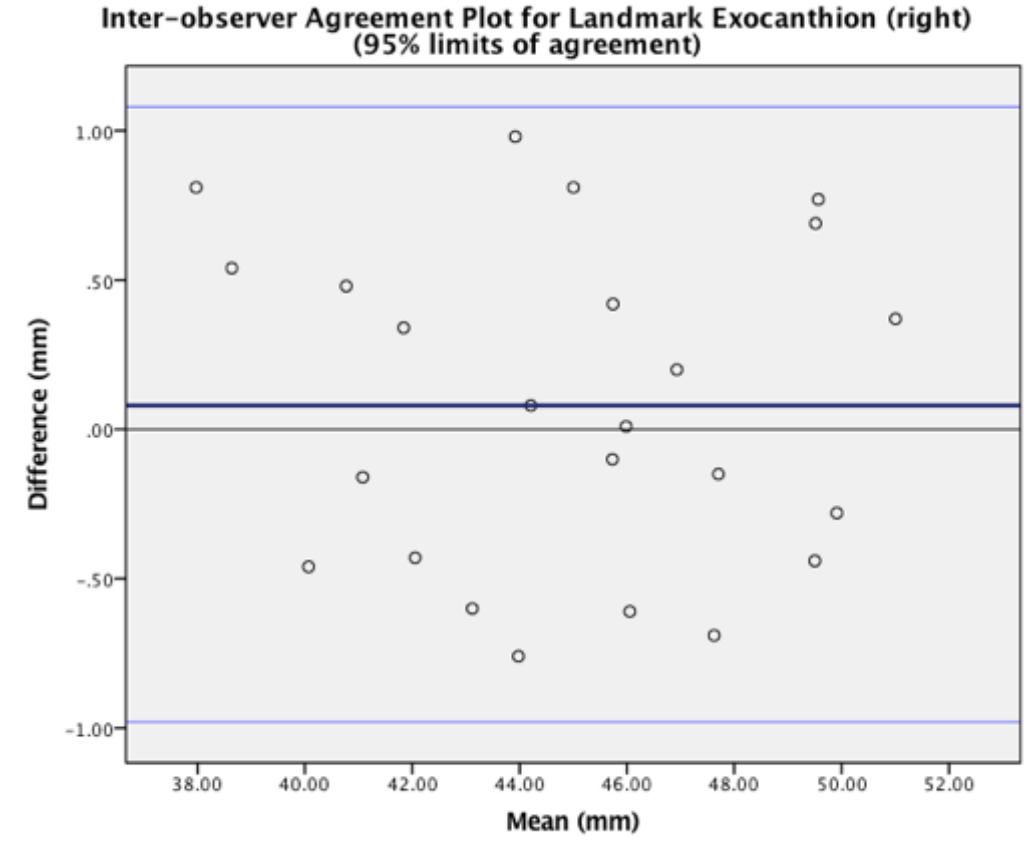

Inter-observer Agreement Plot for Landmark Stomion
(95% limits of agreement)
Difference (mm)
1.50
1.00
.50
.00
-.50
40.00
50.00
60.00
70.00
Mean (mm)

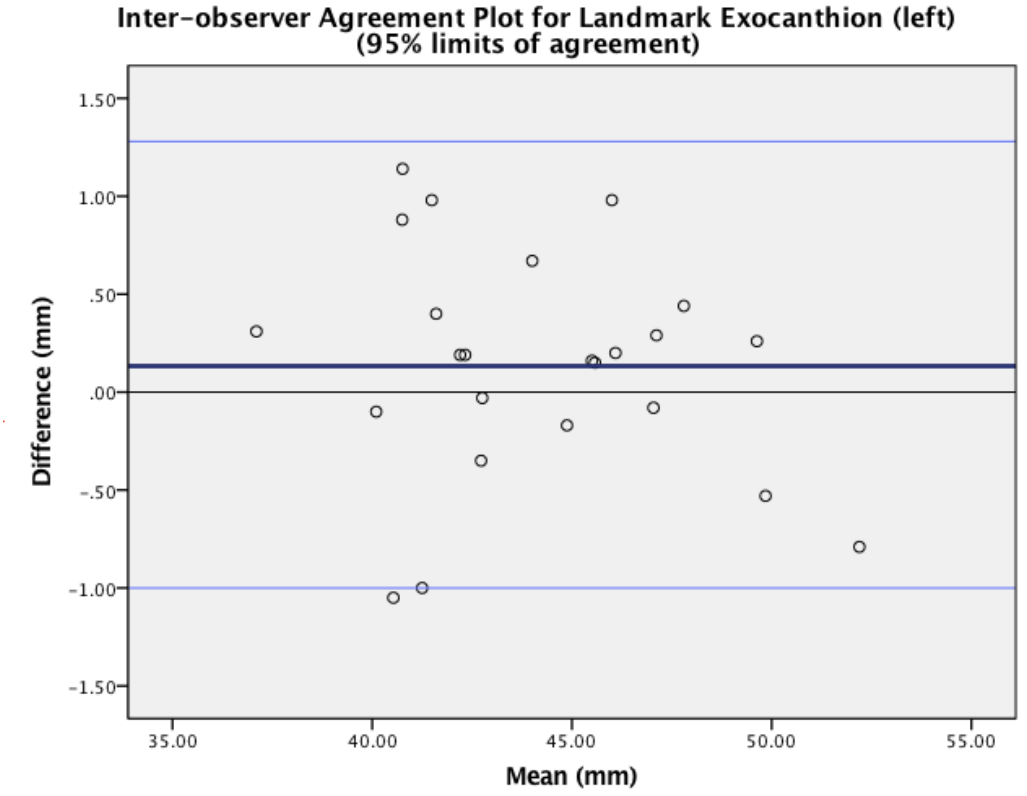

Inter-observer Agreement Plot for Landmark Labrale Inferious
(95% limits fo agreement)
Difference (mm)
1.50
1.00
.50
.00
-.50
-1.00
40.00
50.00
60.00
70.00
80.00
90.00
Mean (mm)

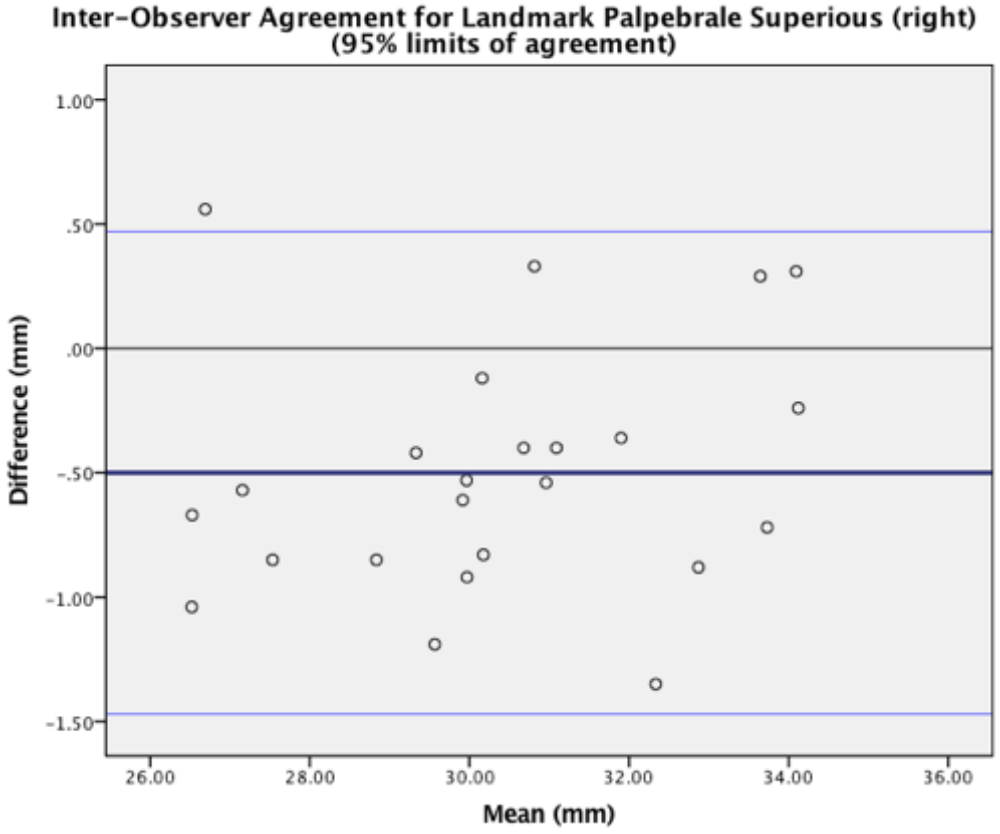

Inter-observer Agreement Plot for Landmark Pogonion
(95% limits of agreement)
Difference (mm)
2.00
1.00
.00
-1.00
-2.00
50.00
60.00
70.00
80.00
90.00
100.00
110.00
Mean (mm)

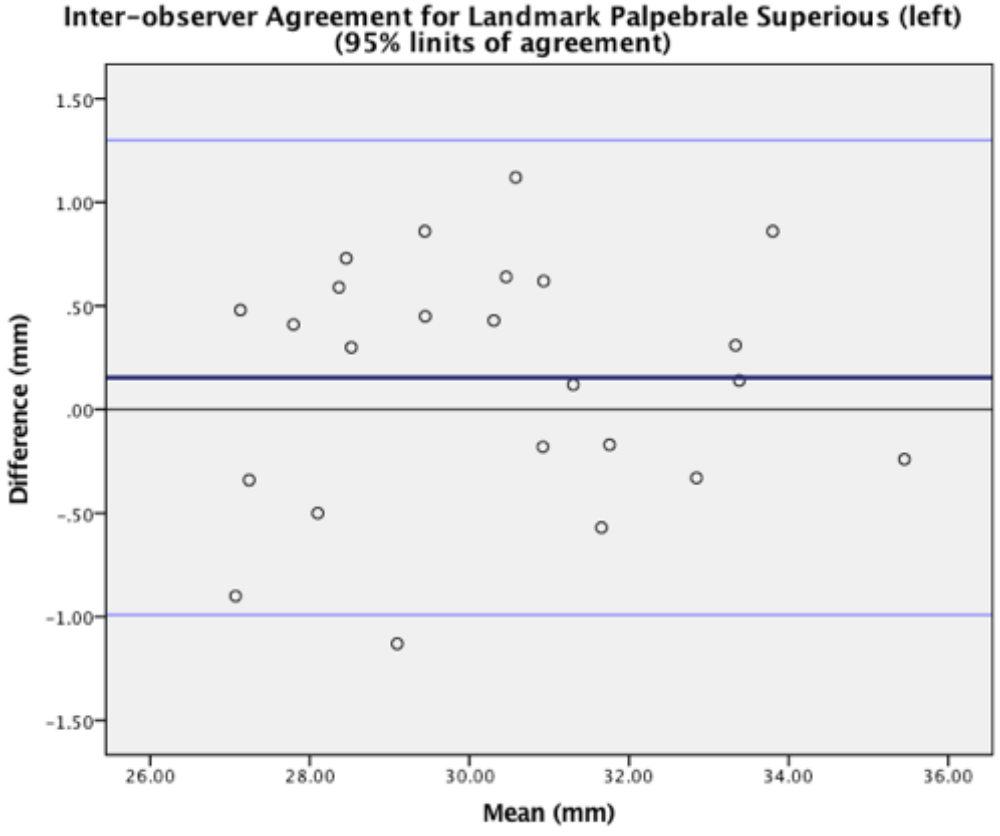

Inter-observer Agreement Plot for Landmark Endocanthion (right)
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
16.00
18.00
20.00
22.00
24.00
26.00
Mean (mm)

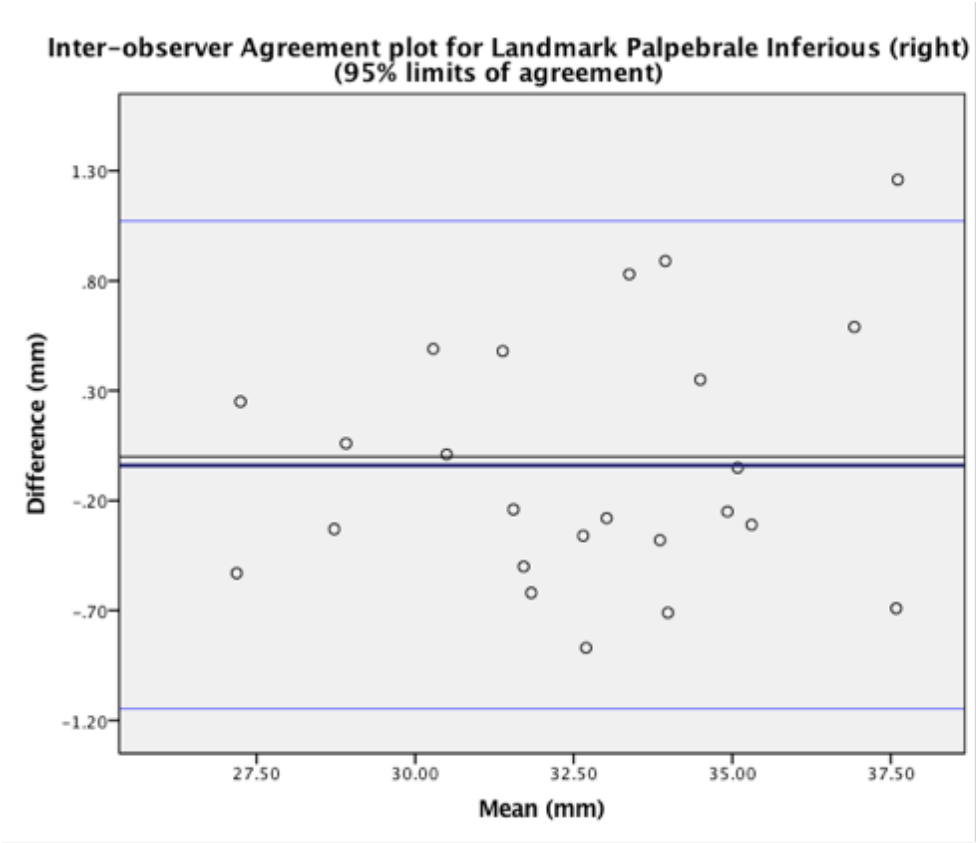

Inter-observer Agreement Plot for Landmark Endocanthion (left)
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
16.00
18.00
20.00
22.00
24.00
26.00
Mean (mm)

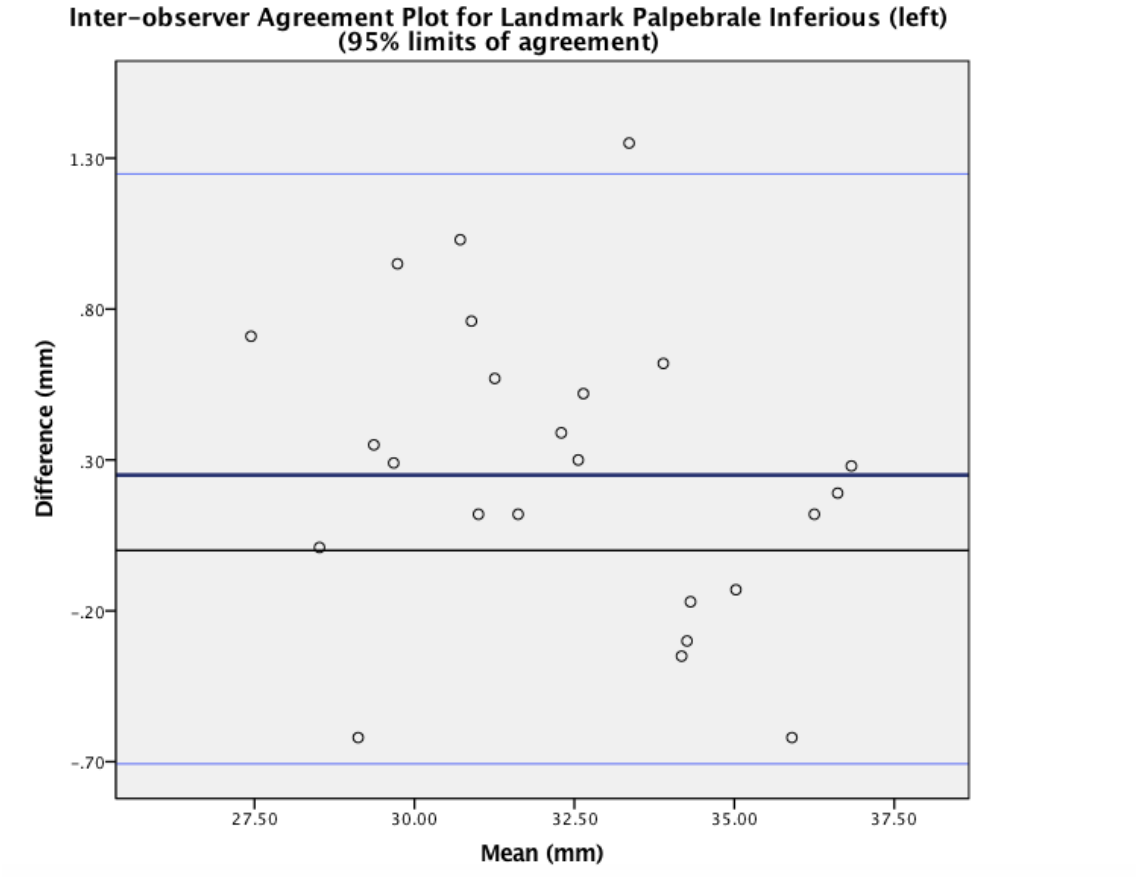

Inter-observer Agreement Plot for Landmark Exocanthion (right)
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
38.00
40.00
42.00
44.00
46.00
48.00
50.00
52.00
Mean (mm)

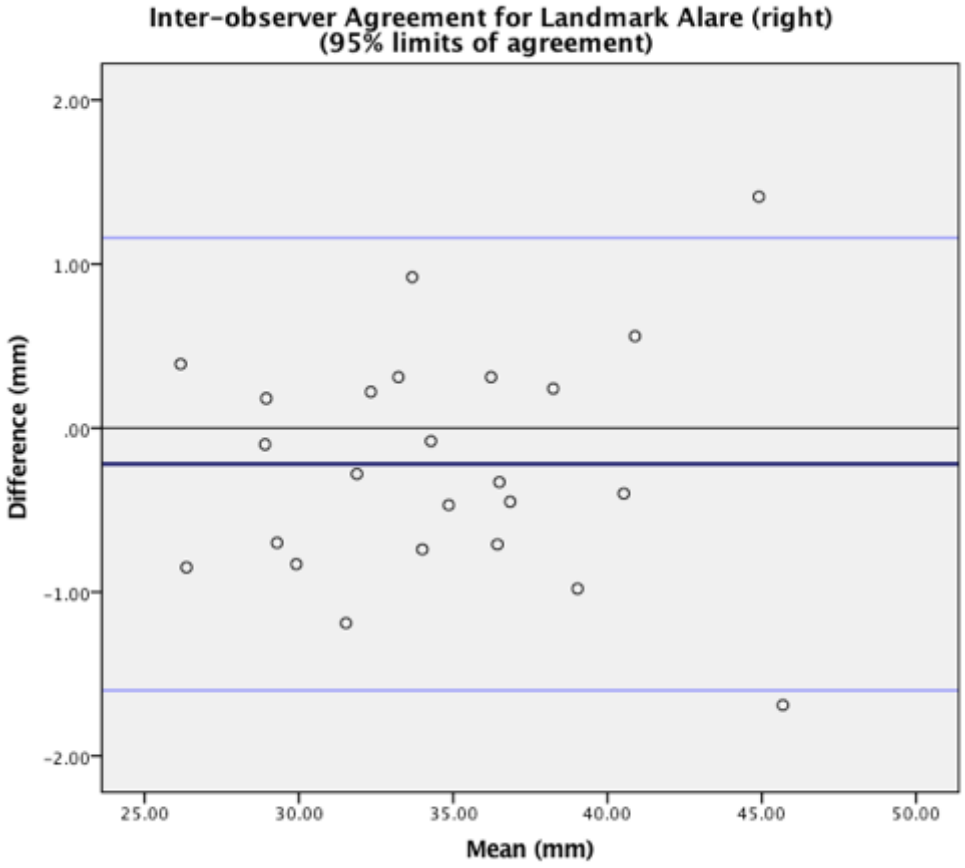

Inter-observer Agreement Plot for Landmark Exocanthion (left)
(95% limits of agreement)
Difference (mm)
1.50
1.00
.50
.00
-.50
-1.00
-1.50
35.00
40.00
45.00
50.00
55.00
Mean (mm)

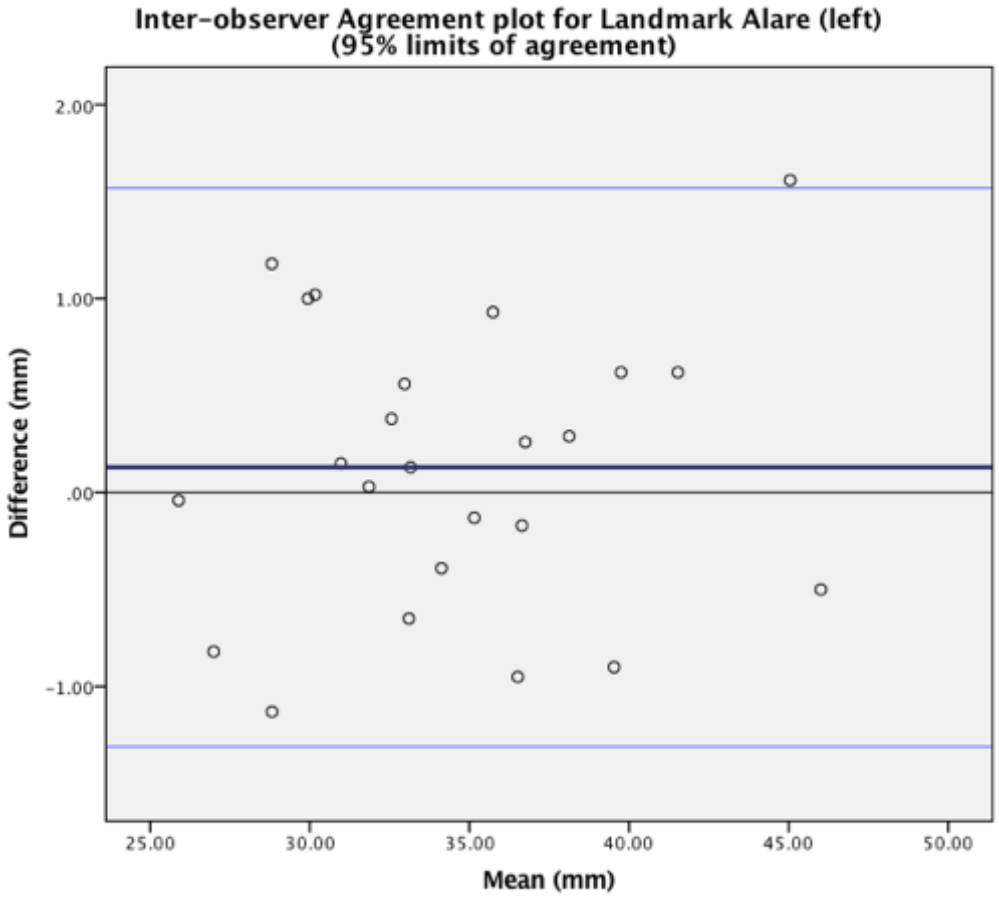

Inter-Observer Agreement for Landmark Palpebrale Superious (right)
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
-1.50
26.00
28.00
30.00
32.00
34.00
36.00
Mean (mm)

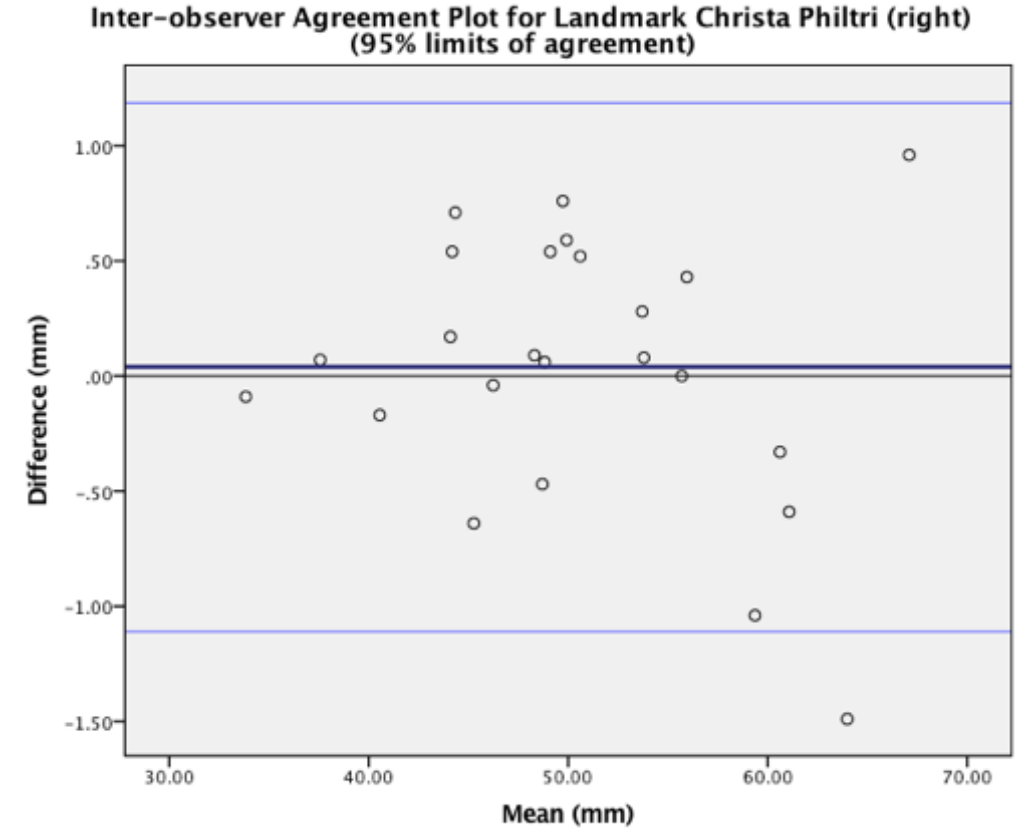

Inter-observer Agreement for Landmark Palpebrale Superious (left)
(95% linits of agreement)
Difference (mm)
1.50
1.00
.50
.00
-.50
-1.00
-1.50
26.00
28.00
30.00
32.00
34.00
36.00
Mean (mm)

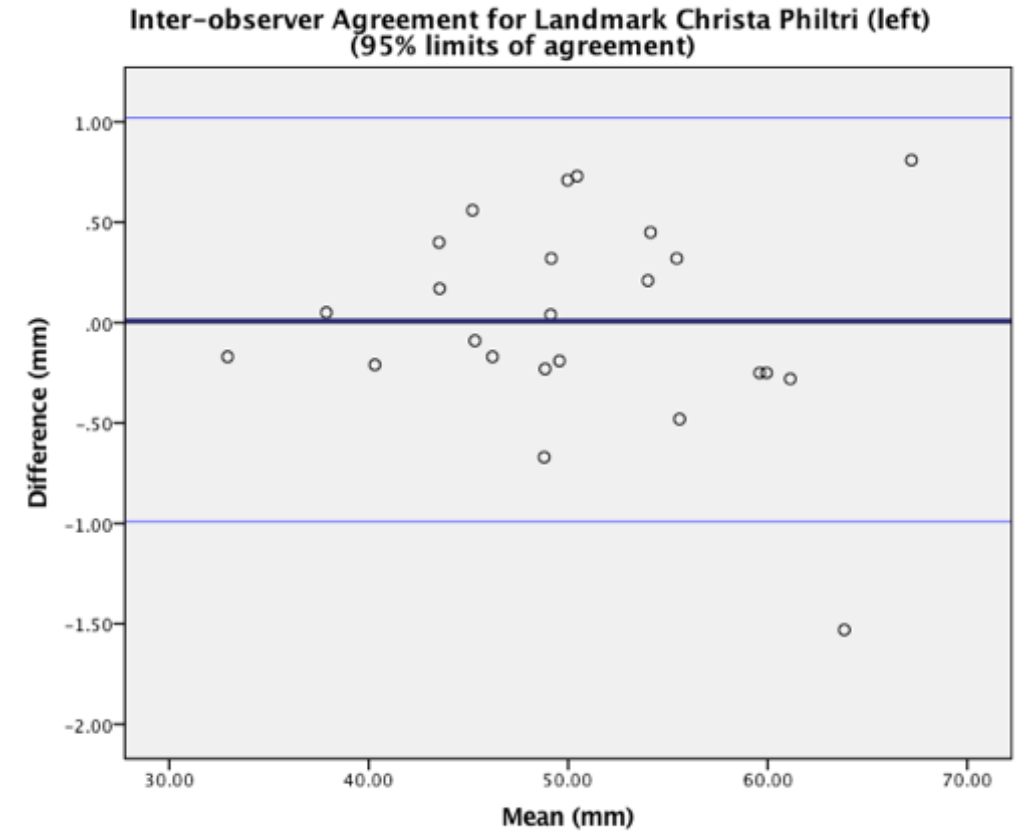

Inter-observer Agreement plot for Landmark Palpebrale Inferious (right)
(95% limits of agreement)
Difference (mm)
1.30
.80
.30
-.20
-.70
-1.20
27.50
30.00
32.50
35.00
37.50
Mean (mm)

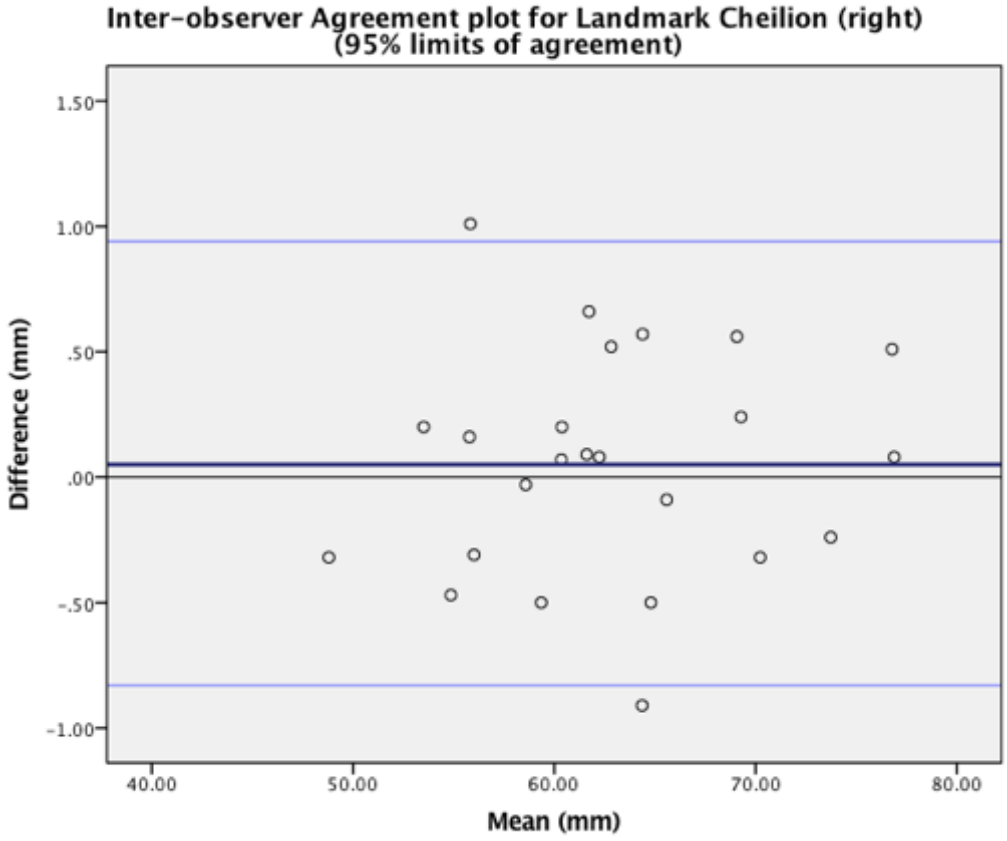

Inter-observer Agreement Plot for Landmark Palpebrale Inferious (left)
(95% limits of agreement)
Difference (mm)
1.30
.80
.30
-.20
-.70
27.50
30.00
32.50
35.00
37.50
Mean (mm)

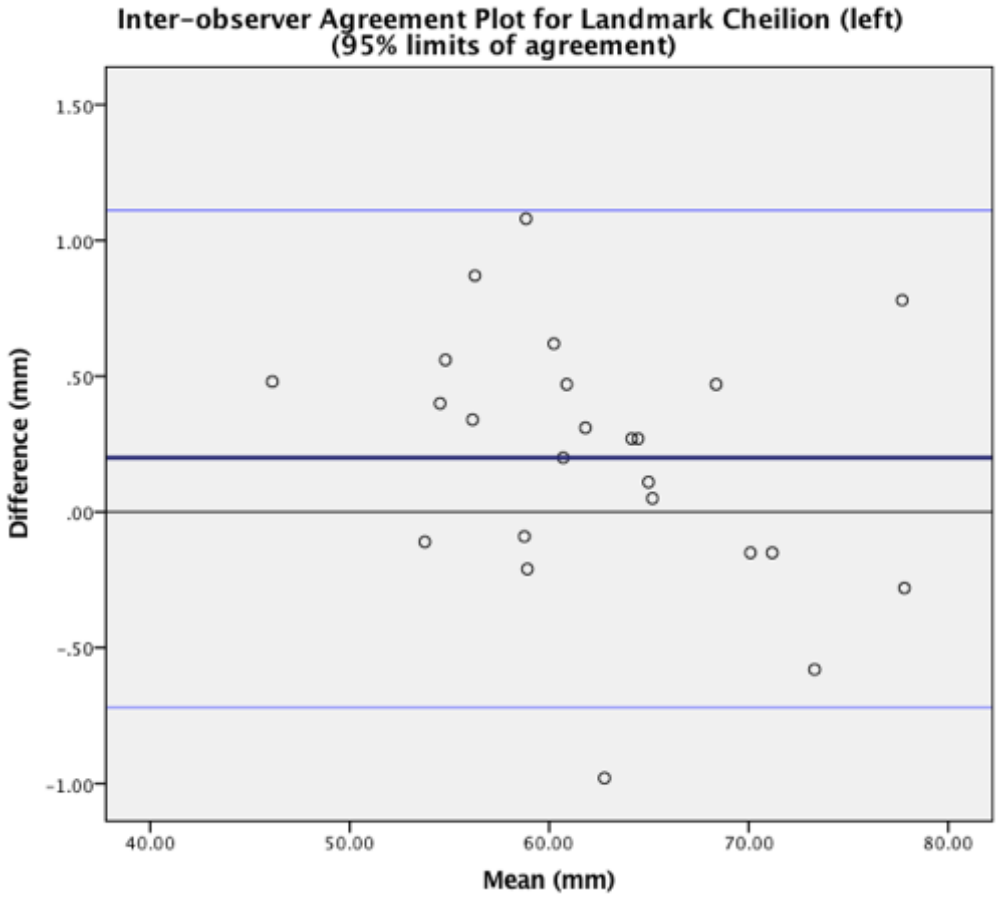

Inter-observer Agreement for Landmark Alare (right)
(95% limits of agreement)
Difference (mm)
2.00
1.00
.00
-1.00
-2.00
25.00
30.00
35.00
40.00
45.00
50.00
Mean (mm)

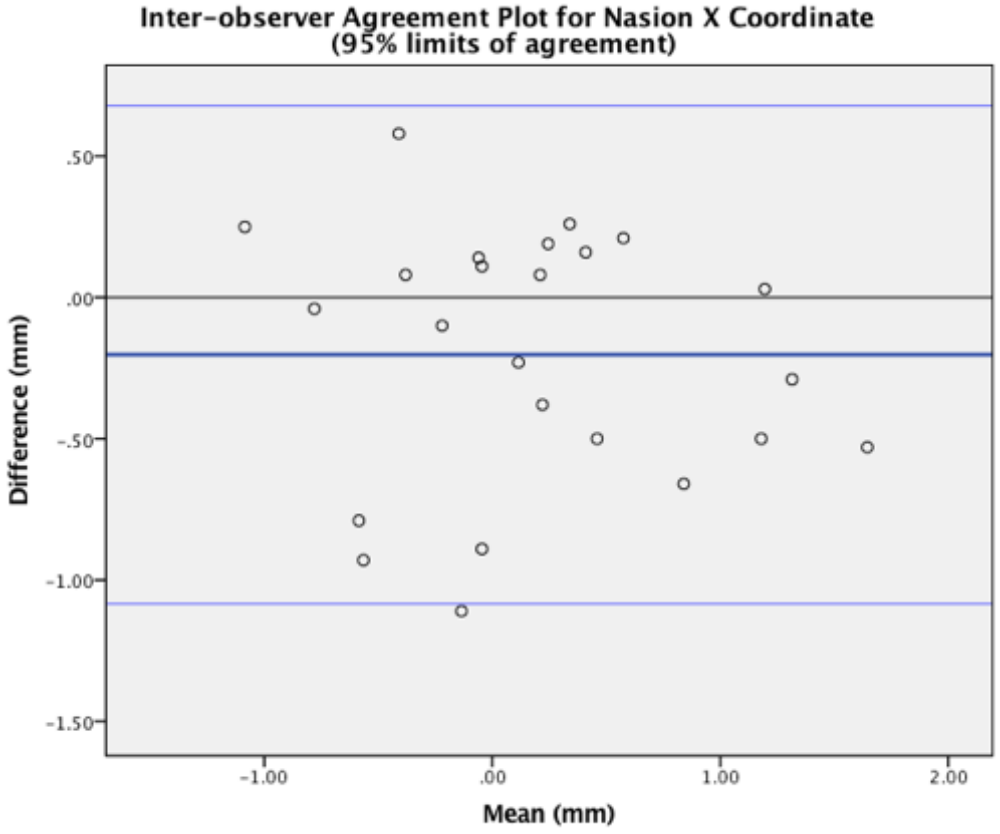

Inter-observer Agreement plot for Landmark Alare (left)
(95% limits of agreement)
Difference (mm)
2.00
1.00
.00
-1.00
25.00
30.00
35.00
40.00
45.00
50.00
Mean (mm)

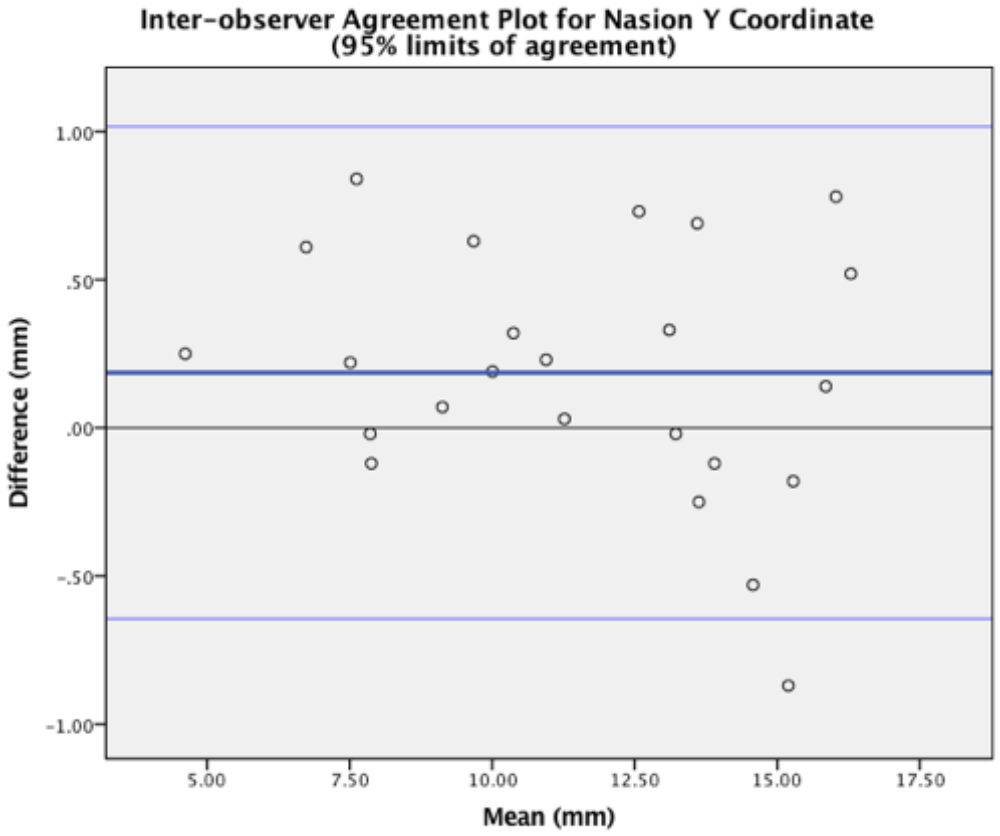

Inter-observer Agreement Plot for Landmark Christa Philtri (right)
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
-1.50
30.00
40.00
50.00
60.00
70.00
Mean (mm)

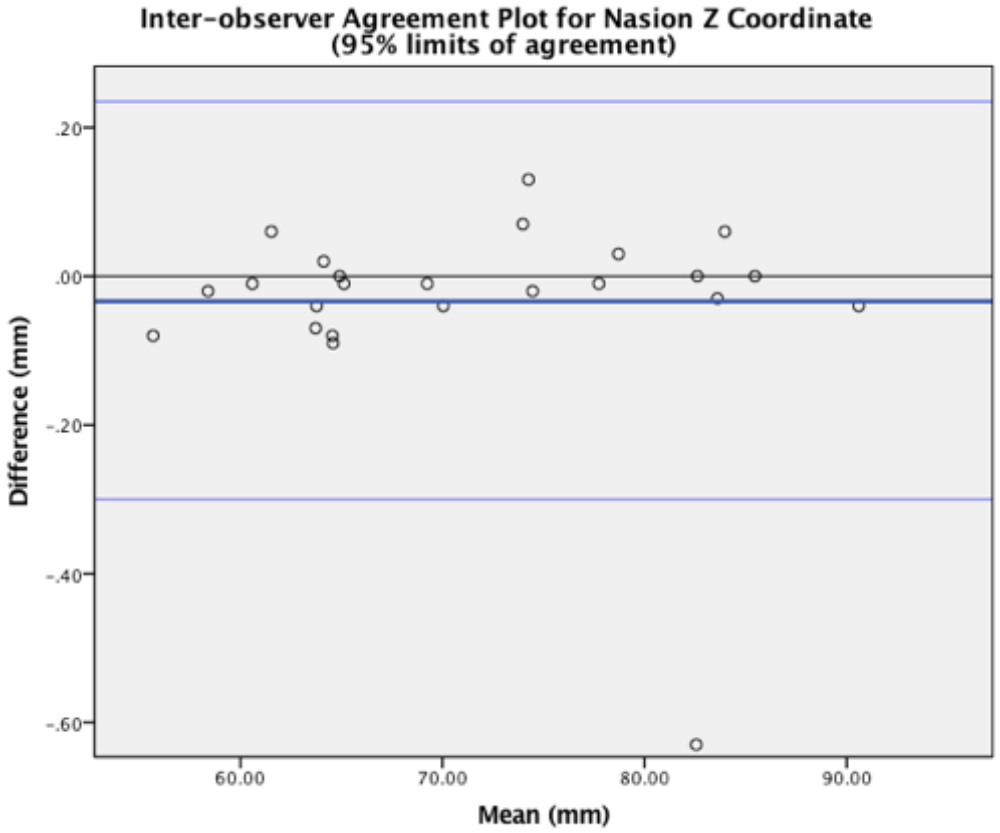

Inter-observer Agreement for Landmark Christa Philtri (left)
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
-1.50
-2.00
30.00
40.00
50.00
60.00
70.00
Mean (mm)


Inter-observer Agreement plot for Landmark Cheilion (right)
(95% limits of agreement)
Difference (mm)
1.50
1.00
.50
.00
-.50
-1.00
40.00
50.00
60.00
70.00
80.00
Mean (mm)


Inter-observer Agreement Plot for Landmark Cheilion (left)
(95% limits of agreement)
Difference (mm)
1.50
1.00
.50
.00
-.50
-1.00
40.00
50.00
60.00
70.00
80.00
Mean (mm)


Inter-observer Agreement Plot for Nasion X Coordinate
(95% limits of agreement)
Difference (mm)
.50
.00
-.50
-1.00
-1.50
-1.00
.00
1.00
2.00
Mean (mm)


Inter-observer Agreement Plot for Nasion Y Coordinate
(95% limits of agreement)
Difference (mm)
1.00
.50
.00
-.50
-1.00
5.00
7.50
10.00
12.50
15.00
17.50
Mean (mm)


Inter-observer Agreement Plot for Nasion Z Coordinate
(95% limits of agreement)
Difference (mm)
.20
.00
-.20
-.40
-.60
60.00
70.00
80.00
90.00
Mean (mm)

Appendix 4: Letter from Director of Research at Alder Hey Children's Hospital.

Alder Hey Children's NHS Foundation Trust
Clinical Research Business Unit
2nd Floor
INSTITUTE IN THE PARK
Eaton Road
Liverpool
L12 2AP

Tel: 0151 252 5570

Dr Susana Dominguez-Gonzalez
Consultant Orthodontist
Alder Hey Children's NHS Foundation Trust

Date: 5th December 2017

Dear Dr Dominguez-Gonzalez,

Re: Headspace and facial symmetry projects

Thank you for discussing with me your Headspace and facial symmetry projects.

I write to confirm in my capacity as Director of Research that Alder Hey that both your projects do not Require REC or HRA approval.

I wish you every success with your studies.

Best wishes.
Yours sincerely

Professor Matthew Peak
Director of Research for Alder Hey