

The Reliability Of Facial Soft Tissue Landmarks With Photogrammetry

Michael G. Payne
Marquette University

Recommended Citation

Payne, Michael G., "The Reliability Of Facial Soft Tissue Landmarks With Photogrammetry" (2013). *Master's Theses (2009 -)*. Paper 198.
http://epublications.marquette.edu/theses_open/198

THE RELIABILITY OF FACIAL SOFT TISSUE LANDMARKS WITH
PHOTOGRAMMETRY

By

Michael G. Payne, D.D.S.

A Thesis submitted to the Faculty of the Graduate School,
Marquette University,
In partial fulfillment of the Requirements for
the Degree of Master of Science

Milwaukee, WI

May 2013

ABSTRACT
THE RELIABILITY OF FACIAL SOFT TISSUE LANDMARKS WITH
PHOTOGRAMMETRY

Michael G. Payne, D.D.S.

Marquette University, 2013

Introduction: With attention being given to the deleterious effects of radiation exposure from dental radiographs and inaccuracies in cephalometric soft tissue measurements, an alternative method of facial analysis with sufficiently reliable soft tissue landmarks should be developed. The goals of this study were threefold: (1) to define a new, low-cost method for capturing standardized frontal and sagittal facial images, (2) to determine on which photographic view that landmarks can be more reliably located, and (3) to determine which landmarks are appropriate for quantitative facial analysis.

Materials and Methods: Simultaneous frontal and right sagittal facial images of 10 male and 10 female dental student subjects were captured using high-definition webcams as part of a low-cost set-up. Seventeen identical predefined facial soft tissue landmarks were located by 5 examiners on both types of images and were recorded as coordinate values. These coordinate values were used to calculate the best estimate of the true value for each landmark, mean deviation from this best estimate, and reliability in the X- and Y-axes using the Shrout-Fleiss intraclass correlation coefficient with corresponding 95% confidence intervals. Two examiners repeated the landmark location to evaluate intra-examiner reliability.

Results: With a 95% confidence interval range of >0.950 , nose and mouth landmarks were among the most reliable landmarks on frontal and sagittal facial images. Conversely *right soft tissue gonion* was one of the least reliable landmarks located in this study. In general, landmarks located by a single examiner showed greater reliability than when there were multiple examiners.

Conclusions: This low-cost method yielded frontal and sagittal images sufficient for landmark identification. The magnitude of error varies between landmarks, is largest for poorly demarcated landmarks, and most had a non-circular envelope of error. Certain landmarks were more reliable on sagittal images and others were more reliable on frontal images. All landmarks had greater reliability and less mean deviation when located by a single examiner.

ACKNOWLEDGMENTS

Michael G. Payne, D.D.S.

I would like to thank my wife and family for all the love and support they have given me. They have stood by me throughout my academic career and I would not be where I am today without them. My accomplishments are truly a testament to their support.

I would also like to extend my deepest gratitude and respect to Dr. José Bósio, who spent countless time and energy to help guide me during this project. His drive and dedication to make improvements has helped shape this project into what it is today.

I greatly appreciate the time and effort that were spent on this project by Mr. Tom Wirtz and Dr. Jessica Pruszynski. The knowledge and mastery of their respective fields, computer science and statistics, has been integral in the success of this project.

Lastly, I would like to acknowledge all the faculty and staff of the Marquette University Department of Orthodontics for their support over the past two years. They have helped push me and prepare me for any future obstacles I may face.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	i
INTRODUCTION	1
LITERATURE REVIEW	4
Introduction of Facial Esthetics	4
Methods of Facial Analysis	6
Reliability of Facial Analysis Methods.....	10
Current State of the Problem.....	16
MATERIALS AND METHODS.....	17
Image Capture.....	17
Measurement.....	22
RESULTS	28
Inter-Examiner Reliability	28
Intra-Examiner Reliability	31
Inter-Examiner vs. Intra-Examiner	33
DISCUSSION	35
CONCLUSION.....	46
BIBLIOGRAPHY.....	47
ADDENDUM A - Scattergrams	53

INTRODUCTION

Historically, heavy emphasis has been placed on the evaluation of lateral cephalometric radiographs, using linear and angular analyses of predefined dentoskeletal standards, for orthodontic treatment planning. (Downs 1956, Steiner 1959) Despite being a part of standard orthodontic records, frontal and sagittal photographs are rarely analyzed quantitatively and merely are used as an adjunct to diagnosis and treatment planning. (Proffit 2012) With increased attention being given to radiation exposure from dental radiographs, less reliance on cephalometric analysis and increased utilization of facial photographs in a quantitative manner for diagnosis is justified. (Mupparapu 2005, Fazel 2009, Claus 2012) As patients are not accustomed to interpreting radiographs, facial photographs as diagnostic records may be a more comprehensible tool. Additionally, the variability in the amount of soft tissue covering facial skeletal structures may mask the appearance of dentoskeletal deformities, thus rendering dentoskeletal standards unreliable when attempting to achieve facial balance. (Park 1986)

Improvement of facial esthetics is a main motivating factor for many patients and parents seeking orthodontic treatment, including up to 80% of adult patients. (Pabari 2011, Livas 2012) Traditionally, orthodontists have placed emphasis on the evaluation of the soft tissue profile, with less attention being given to frontal facial analysis. However, when patients judge facial esthetics, they often do so by looking in a mirror from a frontal view and are less concerned with their profile. Consequently, more emphasis should be given to the improvement of total facial esthetics during orthodontic treatment.

Soft tissue facial characteristics have been evaluated in multiple ethnic groups and races using a variety of methods, such as anthropometry (Farkas 1994, Mollov 2012), cephalometry (Legan 1980, Arnett 1999), two dimensional photogrammetry (Neger 1959, Anic-Milosevic 2011, Fernandez-Riveiro 2002), three dimensional imaging(Hajeer 2004, Fourie 2011) and video imaging (Sarver1996). Each of these techniques has certain disadvantages, including, clinical time needed for anthropometry, radiation exposure and limitation to profile outline in cephalometry, cost of 3D imaging systems, and image distortion and limited visibility of landmarks with indirect methods. (Gavan 1952, Baumrind 1971, Farkas 1980) Values for measurements derived from different methods of soft tissue evaluation are not as reliable as those taken with the same method due to different sources of error for each technique. (Phillips 1984, Shaner 1998, Ghoddousi 2007)

The validity of any measurement obtained from cephalometric radiographs is dependent on the reliability of the landmarks identified. (Trpkova 1997) This concept emphasizes the importance of reliable landmarks for cephalometric facial analysis and should be considered for angular and linear soft tissue measurements on facial photographs. The reliability of skeletal landmarks on lateral cephalometric radiographs has been well documented. (Baumrind 1971, Trpkova 1997) However, there is limited evidence about the reliability of facial soft tissue landmarks on photographs, especially inter-examiner reliability. (Phillips 1984, Jorgensen 1991, Muradin 2007) Intra-examiner reliability of facial landmarks is important when attempting to quantify changes during treatment, whereas inter-examiner reliability is important when quantitatively comparing individuals to standards from predefined facial analyses. Further investigation into the

variation of identification error varies between photogrammetric landmarks and the effect of examiner bias is needed. Therefore, better evidence about the reliability of photogrammetric soft tissue landmarks is needed before a reliable facial analysis can be constructed.

LITERATURE REVIEW

Introduction of Facial Esthetics

Facial esthetics is not a new concept and the perception of what characteristics are considered pleasing has changed throughout history. This evolution is represented in the differences in facial characteristics seen in works of art ranging from ancient civilizations to the present. (Peck 1970, Naini 2011) More recently, cross-cultural agreement of facial attractiveness has been shown when assessing features from multiple races. (Edler 2001) This reaffirms that, despite being largely subjective, there is agreement that certain facial characteristics are more pleasing than others.

The facial features most associated with attractiveness appear to be averageness, secondary sex characteristics and symmetry. (Peck 1970, Edler 2001) It is generally viewed that averageness is an important factor in facial attractiveness when highlighted with secondary sexual characteristics, such as a prominent brow ridge, especially in males. General facial symmetry is an attractive feature, yet it has also been shown that a certain amount of facial asymmetry is pleasing. (Peck 1970)

Various methods of facial analysis have yielded many linear and angular measurements. The absolute values of these measurements are not as important as proportionality. (Reynecke 2012) Proportionality is a major component of facial esthetics and is represented in classical sculptures and Renaissance drawings. Facial proportions have been shown to remain relatively stable during growth, with only minor changes with maturation. More recently, the golden proportion, also known as the Fibonacci

proportion, (1:1.618), which has been used to describe beauty in inanimate objects, has been used to describe attractive faces. (Mizumoto 2009, Ferring 2008)

Over the past century, the objective of orthodontic treatment has transitioned from solely focusing on occlusion to also include improvement of facial esthetics. (Proffit 2012) This increased focus has led to multiple qualitative and quantitative analyses of facial esthetics. (Peck 1970, Neger 1959, Reynecke 2012, Arnett 1993a and 1993b, Mizumoto 2009) Studies have shown, however, that the public prefers a more protrusive profile than conventional cephalometric standards and these analyses may be limited in their usefulness. (Peck 1970, Sutter 1998)

Traditionally, orthodontists have relied on cephalometric radiographs to judge facial esthetics, yet this limits evaluation to the soft tissue profile outline. (Steiner 1959, Downs 1956, Holdaway 1983) When judging personal esthetics, patients view themselves in a mirror and are often less concerned with their profile. (Edler 2001, Peck 1970) Additionally, treatment plans based on dentoskeletal cephalometric standards may be unreliable when attempting to improve facial esthetics due to variation in soft tissue thickness. (Park 1986) Soft tissue measurements on cephalometric radiographs have also been shown to be unreliable. (Trpkova 1997) Therefore, a comprehensive understanding and evaluation of total facial esthetics, both frontal and sagittal, using proper diagnostic records is essential for orthodontic diagnosis and treatment planning.

Facial esthetics, when described in the literature, has been subdivided into macro-esthetics (total facial evaluation), micro-esthetics (smile structure) and mini-esthetics (tooth and gingiva proportionality and evaluation). (Sarver 1996, 2003a, 2003b, 2004)

Improvement of these three components is a major goal of orthodontic treatment planning. Measurements for these three components of facial esthetics were obtained by anthropometric, photogrammetric and video imaging techniques. This reinforces the usefulness of a variety of techniques for capturing and measuring different components of the face.

Previous studies have shown that orthodontic treatment has the potential to affect facial esthetics. (Bishara 1995, Cummins 1995, Kocadereli 2002, Stephens 2005) The amount of facial change from orthodontic treatment is typically minimal and confined to the lower third of the face. An understanding of facial esthetics is especially important in orthognathic surgery, which has the potential for major changes to the face. (Tsang 2009) These changes may cause major improvement of the patients' self-perception and quality of life. (Murphy 2011, Rustemeyer 2012, Feu 2012)

Prior to planning an improvement, guidelines defining "ideal" facial proportions need to be established. Arnett et al. (Arnett 1993a, 1993b, 1999, 2004) have offered guidelines for measuring and improving facial esthetics with combined orthodontic and orthognathic surgical treatment. Metallic markers were placed on the faces of 46 patients prior to cephalometric radiographs for use in measuring facial soft tissues.

Methods of Facial Analysis

Craniometry, physical measurement of dry skulls, was one of the first scientific methods for obtaining measurements of the head and neck. This method dates back to ancient Greece, but the use of measurements to compare skulls was not developed until the 17th century. (Findlay 1980) The study of craniometry is responsible for the definition

and identification of many dentoskeletal landmarks that are used with other methods of identification. (Proffit 2012) One of the most used measurements obtained in craniometry was Camper's facial angle, which was used to distinguish ethnicity and was thought to be associated with intelligence. (Findlay 1980) Camper's facial angle was the angle created by the intersection of a line connecting glabella and A point, the deepest part of the anterior maxillary concavity, with a line connecting anterior nasal spine with the center of the external auditory meatus. This angle, along with other measurements was used for simple classification that provided limited information. With craniometry, measurements from an individual skull represent a single time-point and longitudinal data evaluating changes during growth were not possible.

The inability to measure longitudinal changes with craniometry led to the development of anthropometry, the standardized measurement of living individuals over time. Anthropometry is a technique that dates back to ancient times but it was not until the early 20th century that it was routinely used in medicine. Hrdlicka (Hrdlicka 1920), considered the "father of medical anthropology", routinely used calipers and rulers to record direct measurements from individuals over an extended time period. Further anthropometric standards were developed by Farkas (Farkas 1994, 2005) who took measurements of individuals of various ethnicities in addition to 2500 Caucasian Canadians.

The use of anthropometric measurements in orthodontics was also adopted in the early 20th century which offered a standardized and comprehensive method to assess facial esthetics and quantify changes to facial structures during growth and treatment. (Hellman 1939) Anthropometric measurements are still being used to quantitatively

measure various aspects of esthetics in orthodontics, such as dimensions of the teeth and characteristics of the smile. (Sarver 2004) The previously described methods are considered to be direct measurements, whereas, indirect techniques have more frequently been used in the orthodontic field.

Once introduced by Broadbent, (Broadbent 1931) cephalometric radiographs have played a major role in quantitative diagnosis and treatment planning in orthodontics. These radiographs, like craniometry, allow for measurement of dentoskeletal structures, without interference from soft tissues of varying thickness. (Park 1986) Other benefits of cephalometry include the ability to measure the soft tissue profile outline and quantify changes over time from growth and treatment. (Burstone 1959) Additional soft tissue measurements are made possible with the use of radiographic markers placed on the skin prior to exposing the radiograph. (Bjork 1955)

However, the use of cephalometric radiographs for longitudinal evaluation has limitations. Such repeated exposure of patients to ionizing radiation has proven to have detrimental effects, especially when taking progressive radiographs. (Mupparapu 2005, Fazel 2009, Claus 2012) Cephalometric radiographs are a two-dimensional representation of three-dimensional craniofacial structures, which causes distortion of dentoskeletal structures based on the plane in which they lie. (Baumrind 1971)

Photogrammetry, the evaluation of an object by means of a photograph, is an inexpensive and non-invasive method of quantifying facial esthetics. Photogrammetry has long been utilized in orthodontics to evaluate facial proportions and assess changes during treatment. (Stoner 1955, Neger 1959) Peck and Peck (Peck 1970) utilized

photographs of 52 young adults, who were judged to have pleasing facial esthetics, to quantify measurements that correspond with facial beauty. Frontal and sagittal facial photographs are standard pre- and post-treatment orthodontic records, but historically have only been used for qualitative evaluation of treatment goals and outcomes. (Proffit 2012)

Advances in digital photography and computer software have increased the usefulness of photographs for quantitative linear and angular facial analysis. Now, digital photographs may be viewed immediately, rather than waiting for film negatives to be developed, as well as modified and measured using specialized computer programs. Photographs, which may easily be taken from multiple angles, allow facial soft tissue dimensions to be fully evaluated, a benefit not possible with cephalometry. However, just as with cephalometry, photographs are two-dimensional representations of a three-dimensional object with variable enlargement based on lens-to-object distance. (Gavan 1952)

The use of three-dimensional imaging for facial evaluation, first described for use in orthodontics by Thalmann-Degan in 1944, offers the three-dimensional benefit of craniometry and anthropometry with the benefit of indirect measurement. (Burke 1967) The first technique described, stereophotogrammetry, involves the use of multiple photogrammetric angles converging on the face, captured simultaneously with multiplex cameras, to construct the three-dimensional soft-tissue outline. (Hajeer 2004) Advances in technology have led to improvements in stereophotogrammetry and development of additional methods of three-dimensional facial analysis.

The recent improvements of computers have led to more accurate reconstruction of stereophotogrammetric images with smaller pixel size for improved facial measurements. Three-dimensional laser-imaging is another non-invasive method used for soft tissue facial analysis in orthodontics. (Hajeer 2004) A scanning laser records the facial soft tissue envelope and generates a computerized reconstruction of the face on which measurements can be made. Cone-beam computer tomography (CBCT) is another three-dimensional technique used for facial measurements in orthodontics. (Chang 2011) Like craniometry, CBCT allows for measurements of the dentoskeletal structures of the head, but with the convenience of indirect measurement and ability to measure the same individual at different time-points. However, CBCTs expose patients to radiation and facial soft tissues are minimally visible which limit their usefulness for evaluation of total facial esthetics.

More recently, video-imaging has been used to make dynamic measurements of facial soft tissues, rather than static measurements from other indirect techniques. The use of video-imaging in orthodontics has most-notably been used for measuring dynamic movements of the soft tissue during smile animation. (Sarver 1996, 2003a, 2003b) This technique, like photogrammetry, yields two-dimensional images and improved visualization of the soft tissue contour is possible by recording the face from different angles.

Reliability of Facial Analysis Methods

Craniometry is a highly accurate method of measuring the true dentoskeletal distances. This technique is not limited by accuracy, rather by the inability for

longitudinal and soft tissue measurements. (Proffit 2012) This is considered the gold-standard for accuracy and reliability of dentoskeletal measurements and values obtained by other methods, such as cephalometry and CBCT have been compared to craniometry to assess accuracy. (Gribel 2011)

The accuracy reliability of craniofacial anthropometry has been extensively investigated. (Hrdlicka 1920, Farkas 1994) There are two main factors that affect the accuracy of the anthropometric measurements, the skill of the examiner and the quality of the instruments used to make these measurements. Extensive training is necessary, especially when there are multiple examiners, to improve accuracy and reliability of measurements. Studies have recommended that landmarks be marked on the skin to improve inter-examiner reliability, however, inaccuracies may also arise from variable pressure placed on soft tissue landmarks during measurement. Additionally, it has been stated that cooperation of the examinee has a significant effect on accuracy.

Reliability is evaluated by repeating anthropometric measurements twice within a short period of time. A short time period is ideal due to possible changes in facial dimensions over an extended period of time due to environmental factors, such as body mass index, temperature and humidity. (Farkas 1994) These environmental factors may cause measureable soft tissue changes that affect reliability of quantifying growth and treatment changes. The reliability of ethnic craniofacial anthropometric norms depends on the ethnic composition, representativeness, environmental factors and socioeconomic status of the individuals measured. (Farkas 1994, 2005) These factors must be accounted for when comparing measurements of an individual to anthropometric norms.

Anthropometric measurements have been shown high reliability. Marking landmarks on the skin prior to taking measurements has been recommended and was shown to have a significant effect on reliability of anthropometric measurements. (Farkas 1994, Shaner 1998) Studies have found high levels of anthropometric reliability of the face, most measurements varying less than 1mm on repeated measurements. (Chu 1989, Shaner 1998, Ghoddousi 2007) Mollov et al. (Mollov 2012) found high intra-examiner reliability, yet inter-examiner reliability was quite variable. Due to the high level of reliability, anthropometry is widely considered the primary method of craniofacial measurement, against which the accuracy of indirect methods is often measured.

There are certain concerns that must be accounted for when comparing the reliability of indirect methods of craniofacial measurements to the previously mentioned direct techniques. Most notably, cephalometric radiographs and photographs are two-dimensional representations of the three-dimensional structures of the head and neck, with variable distortion of landmarks lying in different planes. (Gavan 1952, Baumrind 1971) Therefore, data gathered by different techniques can be expected to yield different results and absolute values of measurements should not be compared to each other. (Phillips 1984, Shaner 1998)

The reliability of dentoskeletal landmark identification on cephalometric radiographs has been widely investigated. (Baumrind 1971, Trpkova 1997) Two major sources of error occur when locating cephalometric landmarks, errors of projection, a two-dimensional representation of three-dimensional structures, and errors of identification, differences in locating landmarks. Additionally, representativeness of radiographs, representativeness of examiners, machine errors in point identification and

errors in superimposition of tracings have been cited as possible sources of error that may affect cephalometric reliability. (Baumrind 1971)

The reliability of cephalometric measurements depends on the reliability of landmarks possibly affected by these sources of error. This was investigated in a meta-analysis of cephalometric landmark reliability by Trpkova et al. (Trpkova 1997) It was recommended that a total error of less than 0.59 mm in the X-axis and 0.56 mm in the Y-axis be achieved for a landmark to be considered sufficiently reliable. Of the 15 landmarks investigated, only 5 landmarks, *B point*, *A point*, *pterygomaxillary fissure inferior*, *sella* and *gonion*, reached this level of reliability in the X-axis. In the Y-axis, only *pterygomaxillary fissure inferior*, *A point* and *sella* exhibited sufficient reliability. This meta-analysis, however, did not investigate the reliability of soft tissue cephalometric landmarks, which have been found to be fairly unreliable. (Burstone 1958, Park 1986)

Much attention has been given to the reliability of facial photogrammetric measurements. The accuracy and reliability of photogrammetric facial soft tissue measurements was most notably investigated by Farkas et al. (Farkas 1980) Of 104 anthropometric measurements of the head face and ears, 62 could be duplicated on frontal and sagittal photographs. Only 26 of the 62 photogrammetric measurements were considered reliable, differing from the anthropometric measurements by no more than 1 mm or 2 degrees. Measurements corresponding to the lips and mouth yielded the greatest number of reliable measurements. Additionally, measurements were found to be more accurate when the landmarks used for the measurement were the same distance from the camera lens.

The accuracy of photogrammetric measurements is affected by the differential distortion of photographs that are two-dimensional representations of three-dimensional structures. (Phillips 1984, Shaner 1998) Therefore, repeatability of photogrammetric landmark measurements is a more suitable method for evaluating the reliability than comparing absolute values to other methods of facial evaluation. Many studies investigating the reliability of facial photogrammetry physically marked the landmarks to be identified prior to capturing the images. (Farkas 1980, Shaner 1998, Muradin 2007) This process may improve the reliability of facial measurements. (Shaner 1998)

Phillips et al. (Phillips 1984) investigated errors of projection and landmark identification on both frontal and sagittal photographs of 12 female orthodontic patients. Reliability of photogrammetric landmark identification was defined by an absolute mean error less than 2 mm, a level found in 66% of the 47 frontal landmarks and 57% of the 26 sagittal landmarks located. *Right pupil* and *pronasale* were the most reliable, while *left malar point* and *right soft tissue gonion* were the least reliable facial landmarks on frontal and sagittal images, respectively. When identifying facial photogrammetric landmarks, it was found that magnitude of error varied from landmark to landmark, the envelope of error for most landmarks was not circular and the largest errors occurred on surfaces that are a gradual curvature. (Phillips 1984)

When assessing nasolabial soft tissues, Muradin et al. (Muradin 2007) concluded that standardized frontal facial photographs were an acceptable medium for reliable measurements. All landmarks showed a mean square error of less than 0.38 mm and a reliability coefficient greater than 99%. The sources of error for photogrammetry were

found to be 38.64% due to intra-examiner differences of landmark location and 61.36% due to patient variance and posturing.

The reliability of landmarks is of utmost importance because linear and angular measurements are affected by points used for the measurements. Photogrammetry was found to have good repeatability, though measurements obtained from photographs were more variable than anthropometric measurements. (Ghoddousi 2007) The resolution of the images used for photogrammetry, have an effect on the reliability of landmark identification. There has been less attention given to the reliability of measurements from video-images, however, measurements on a single frame of a video is comparable to photogrammetry.

Three-dimensional imaging techniques have been shown to yield reliable facial measurements. The accuracy of measurements obtained from three-dimensional reconstructions has been assessed by comparing them to anthropometric measurements. The accuracy from stereophotogrammetry has been considered acceptable since it was first described and recent advances in camera technology has further improved accuracy to be within 0.5mm of anthropometric measurements. (Burke 1967, Fourie 2011) Images generated by laser scanning and CBCT also produce accurate reproductions of the face on which soft tissue measurements within 1 mm of corresponding anthropometric measurements have been found. (Hajeer 2004, Fourie 2011) The reliability of measurements made on all types of three-dimensional images has been show to be very high, and has even exceeded anthropometric reliability. (Hajeer 2004, Ghoddousi 2007, Fourie 2011) The high level of accuracy and reliability of measurements on three-dimensional images combined with the benefits of indirect measurements are clear

benefits of these techniques for facial measurements. However, the major drawbacks to these methods are the radiation exposure in CBCT and the high cost of all three-dimensional techniques.

Current State of the Problem

The previously described methods of linear and angular facial evaluation are all suitable for quantitative analysis. However, many of the previously described techniques are either invasive or resource intensive, which limit their applicability in orthodontics. Photogrammetry is a non-invasive method that uses standardized facial photographs, which are already standard orthodontic records. Cephalometry, which is the primary method of facial evaluation in orthodontics, has been shown to yield unreliable soft tissue measurements. (Park 1986) A reliable photogrammetric facial analysis may provide an accurate method of facial soft tissue evaluation to complement cephalometric dentoskeletal evaluation. Before a quantitative facial analysis is constructed, more information about inter- and intra-examiner reliability is needed to determine which landmarks should be utilized.

The purpose of this study was threefold:

- 1) to define a new, low-cost method for capturing standardized frontal and sagittal facial images
- 2) to determine on which photographic view that landmarks can be more reliably located
- 3) to determine which landmarks are optimal for quantitative facial analysis.

MATERIALS AND METHODS

Prior to recruiting subjects, this study was approved by the Marquette University Institutional Review Board, protocol number HR-2083. Twenty dental student volunteers, 10 males and 10 females, aged 22-33 years were recruited and randomly selected to participate in the study. This study was designed to analyze normal facial landmarks and there were no exclusions based on race or skeletal jaw pattern. The exclusion criteria included subjects who: (1) had history of any congenital or acquired abnormality, (2) were considered obese, (3) had undergone medical/pharmacological treatment that would affect the facial complex or (4) had facial hair that would mask landmarks to be identified.

Each of the volunteers was seated with Frankfort Horizontal plane parallel to the floor. Their lips were then manipulated into a relaxed-lip posture, which was attained by (1) asking the patients to relax their lips, (2) gently jiggling the mandible, as if trying to achieve centric relation position, slowly closing until the teeth first touch, and visualizing the amount of lip separation at first contact of the teeth, and (3) gently stroking the lips. (Burstone 1967) Relaxed-lip posture was used because it is a reasonably reproducible position independent of teeth and supporting alveolar processes. Once a relaxed-lip posture was attained two simultaneous photographs, frontal and right sagittal, were taken by one investigator in the standardized manner described below.

Image Capture

Each of the volunteers was seated over a mechanical beam scale (Health o meter, Boca Raton, FL, model #402LB) with the height rod extended horizontally over the

volunteers' heads and a plain white background behind and to the left of the subject. Two ABFO No. 2 bitemark scales (Forensics Source, Jacksonville, FL) were fixed perpendicular to each other and positioned at the end of the height rod, one parallel and one perpendicular to the height rod. The subjects were positioned directly under the height rod with the frontal facial and the midsagittal planes coincident with the bitemark scales.

Two high-definition webcams (Logitech HD Pro Webcam C910, Newark, CA), mounted on tripods, were positioned 1 meter in front and 1 meter to the right of the subjects with the optical axis at the same height and parallel to Frankfurt horizontal. The webcams were connected to a computer and the frontal and right sagittal images were captured simultaneously using a specialized software program, OrthoCapture (Tom Wirtz, Milwaukee, WI). The photographic set-up can be seen in Fig. 1. The images were captured on two separate days, for volunteer convenience, with 11 subjects on the first day and the other 9 subjects on the second day. Once the images were obtained, they were evaluated for image quality and whether the subjects were in a relaxed-lip posture.



Fig. 1 Set-up for simultaneous frontal and right sagittal image capture with high-definition webcams

The ABFO No. 2 bitemark scales were used to determine the magnification factor of each image and calibrate the pixel size. The captured images, containing a subject's face and ABFO bitemark scale, were cropped and scaled using Adobe Photoshop CS4. The width of the document size of the new image was transformed to 13 inches with the changes to width and height locked in Adobe Photoshop CS4 and the "scale Styles", "Constrain Proportions", and "Resample Image" boxes checked. For each image,

calibration was performed to determine the magnification factor and pixel size in millimeters. Using the ABFO bitemark scale, the number of pixels in 10 mm, both horizontally and vertically, were measured by two separate examiners. Pixel size was then calculated by dividing 10 mm by mean number of pixels from these 4 measurements. Following this, the images were saved in JPEG format at the maximum quality setting and imported into a software program designed for manual identification of facial landmarks, OrthoMeasure (Tom Wirtz Milwaukee, WI). This process produced 1950 X 5246 pixel facial images with 150 dpi.

OrthoMeasure is a software application developed with Delphi 2010 (Embarcadero, San Francisco, CA). First, the user selected an image for landmark identification and the program displayed two views of that image on the screen. One view showed the full head image, displaying the image with proportional scale “true” in an image object 1442H x 536w. The other view showed a close-up portion of the same image, with proportional scale “false” in an image object 5500H and 2000W. Only a small portion of this image was visible (550H x 535W). The OrthoMeasure program can be seen below, Figs. 2 and 3.

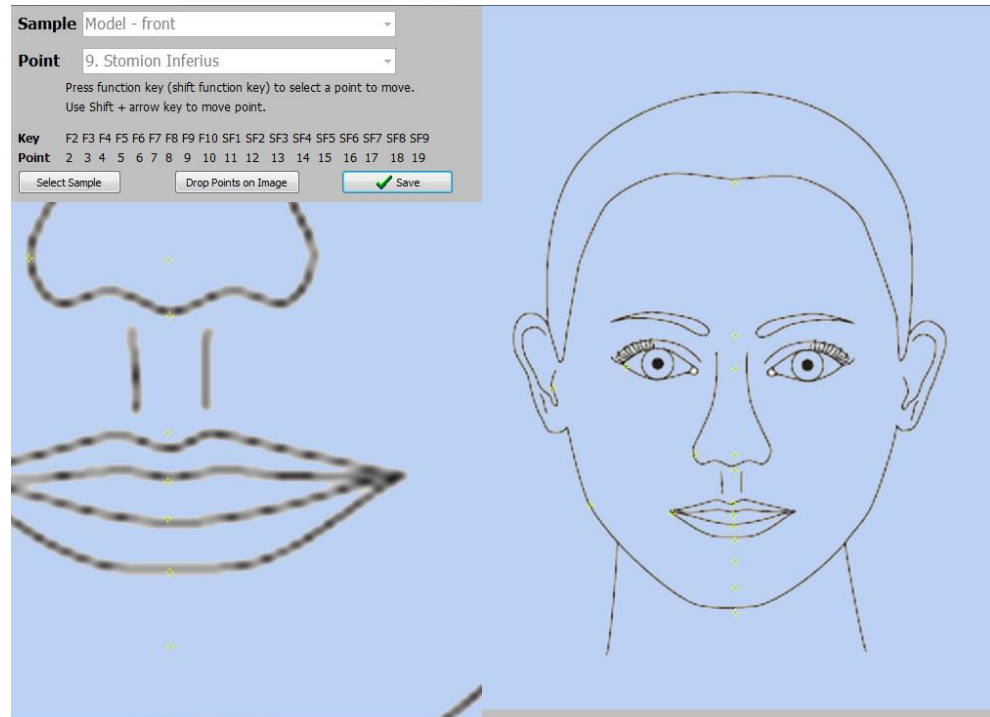


Fig. 2 OrthoMeasure with landmarks on frontal images

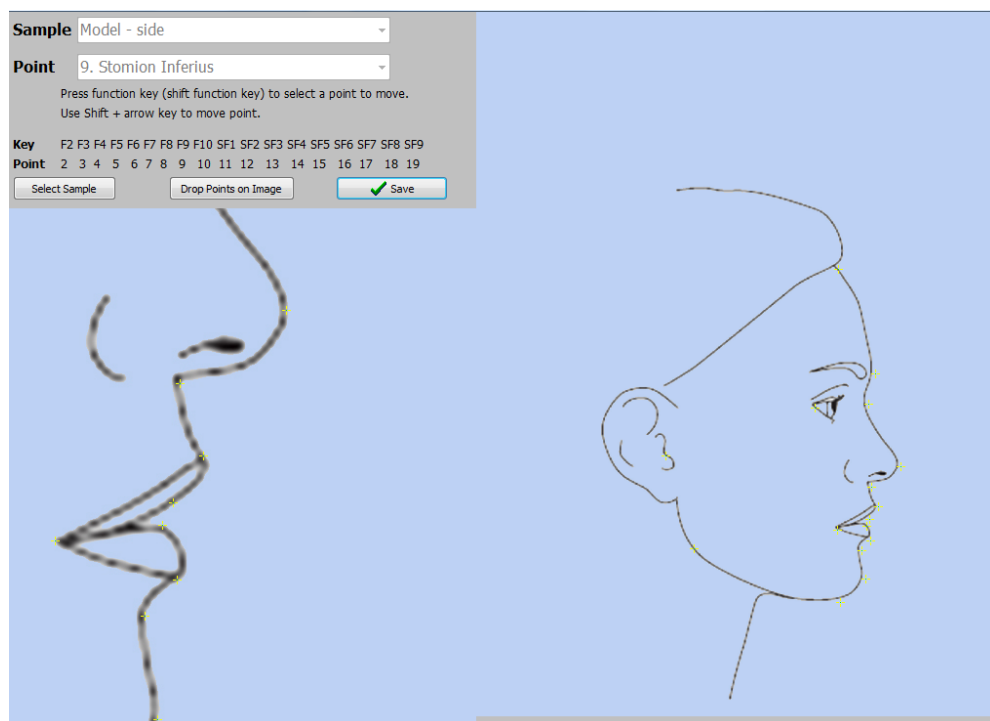


Fig. 3 OrthoMeasure with landmarks on sagittal images

Measurement

The landmark to be identified was selected with either a drop-down control box or a function key. The examiners first identified the approximate location of soft tissue facial landmarks on the full head images with a left mouse click. Once the approximate location of the landmark was determined, the marker appears on both the full head and close-up images. When a landmark was located on the full sized images, the close-up image was shifted to center the current landmark marker in the window. The arrow keys were then used to move the marker one pixel at a time, while viewing the close-up window, to what the examiner believed to be the precise location of the landmark. While making the precise adjustments, the examiners viewed the close-up images. For each point marker, the corresponding X- and Y- pixel coordinates were recorded.

A group of 5 examiners, four orthodontic residents and one orthodontist, located seventeen predetermined, previously defined facial soft tissue landmarks (Farkas 1980, 1994, Phillis 1984, Mollov 2012) on both the frontal and the sagittal images using OrthoMeasure, Figs. 4 and 5. The examiners were given a list of the landmarks with definitions, table 1, and were shown which landmarks to locate. If a landmark could not be visualized on one or both of the images, such as the subject's hair covering trichion, the examiners were instructed to estimate the location of the landmark. The examiners first located the seventeen landmarks on the 20 frontal facial images, and then recorded the same landmarks on the 20 right sagittal images. Two of the 5 examiners (MGP, JAB) repeated the landmark identification, two weeks after the initial session, to determine intra-examiner reliability. After the initial instruction, landmark identification was not influenced by other examiners by any means.

Table 1. Definitions of Soft Tissue Landmarks

Soft Tissue Landmark	Abbreviation	Definition
Trichion	tr	the point on the hairline in the midline of the forehead
Soft Tissue Glabella	g'	the most prominent midline point between the eyebrows
Sellion	se	the deepest midline point of the nasofrontal angle
Pronasale	prn	the most protruded point of the nose
Subnasale	sn	the midpoint where the columella base and the upper lip meet
Philtrum Point	pp	the midline point on the imaginary line connecting the intersection of the philtrum columns with the vermillion border of the upper lip
Stomion Superius	stos	the midpoint of the lowermost point of the upper lip
Stomion Inferius	stoi	the midpoint of the uppermost point of the lower lip
Labrale Inferius	li	the midpoint of the lower border of the lower vermillion line
Supramentale	sm	the deepest point of the labiomenal sulcus at the midline
Soft Tissue Pogonion	pg'	the most anterior midpoint of the chin
Soft Tissue Menton	me'	the most inferior midpoint of the chin
Right Soft Tissue Gonion	go'	the most lateral point on the right mandibular angle close to the bony gonion
Center of Right Tragus	t	the most lateral and posterior point of the right tragus of the ear
Right Exocanthion	ex	the point at the outer commissure of the right eye fissure
Right Cheilion	ch	the lateral most point located at the right labial commissure
Right Alare	al	the lateral most point on the right alar contour

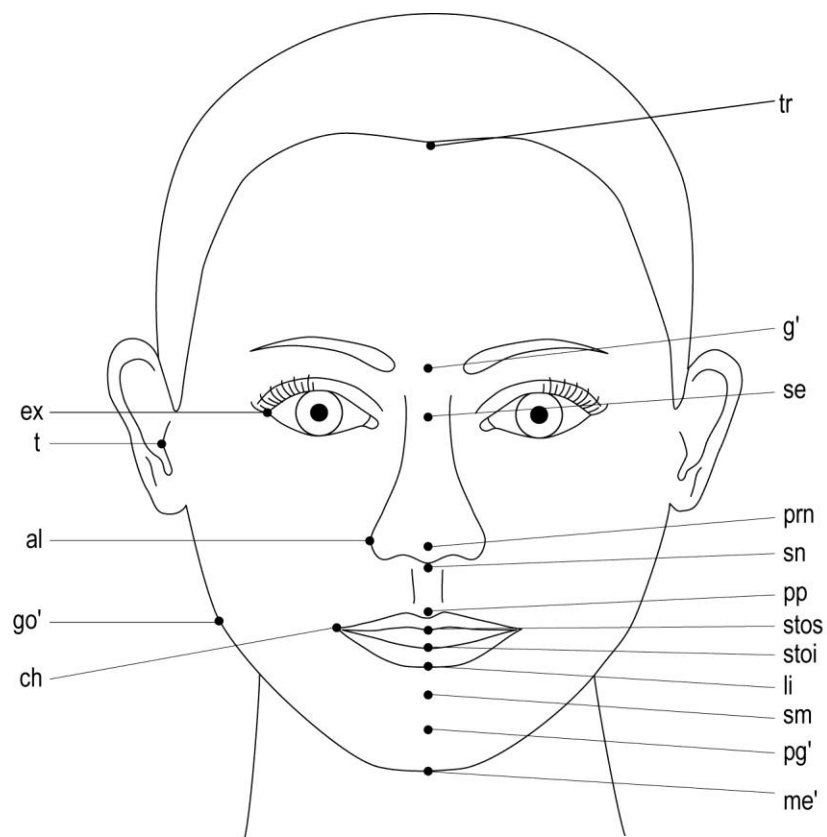


Fig. 4. Diagram of frontal facial soft tissue landmark location

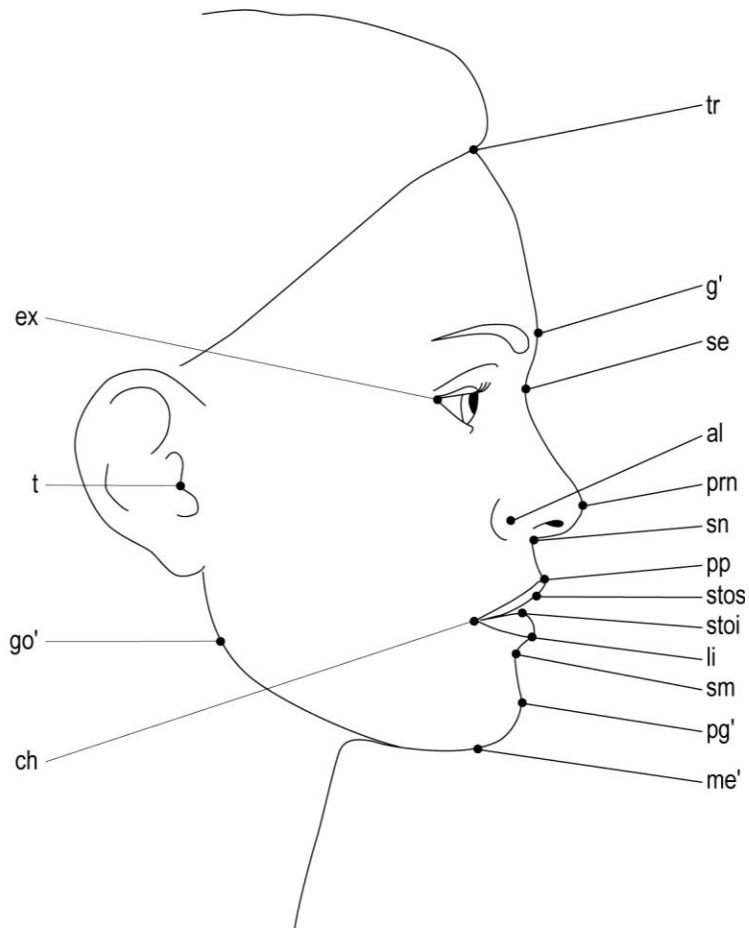


Fig. 5. Diagram of frontal facial soft tissue landmark location

As the examiners located the landmarks using OrthoMeasure, the identified points were recorded as (X,Y) coordinate values based on the location of the single pixel selected on the image. This yielded five sets of 17 landmark coordinates, one set from each examiner, for each of the 40 images to be used for evaluating inter-examiner reliability. Using the five coordinate values that correspond to the same landmark on the same image, the arithmetic mean coordinate value was calculated. This mean value represents the best estimate of the true value for each landmark on that specific image. (Baumrind 1971) Next, all arithmetic mean (X,Y) coordinate values were converted to (0,0). Each of the five individual points was then assigned a new, adjusted (X,Y)

coordinate value based on the distance, in X- and Y-planes, that each point deviated from this mean value. This conversion was repeated for all 17 landmarks on all 40 images.

These adjusted coordinate values were then converted from number of pixels to millimeters by multiplying the number and size of pixels. Using the adjusted coordinate values in millimeters, the absolute distance from the best estimate of the landmark was calculated using the Pythagorean Theorem.

Next, all points representing the same landmark on the same type of image, either frontal or right sagittal, were compiled. This yielded 100 adjusted coordinate values, five examiners' points on twenty images, corresponding to each specific landmark on each type of image. These 34 sets of values, one for each of the 17 landmarks on both types of images, were then used to calculate inter-examiner reliability of landmark identification.

After the adjusted coordinate values were calculated, 5-point scattergrams were constructed for each landmark on each image with the best estimate of the landmark at the origin. Then, twenty 5-point scattergrams, representing the same landmark on the same type of image, were superimposed at their origins to create a 100-point scattergram for each landmark, see Addendum A. This process yielded 34 scattergrams, one for each of the 17 landmarks on frontal and sagittal images. These scattergrams represent the dispersion of estimating errors for a single landmark around the best estimate of the true value.

Using two sets of points from each of the two examiners who repeated the facial landmark identification, the previously described process was repeated. The best estimates of the true value of each landmark were calculated separately for each

examiner. This produced two separate sets of adjusted coordinate values, one from each examiner. Each of these sets consisted of 40 adjusted coordinate values for each of the 34 landmarks. These were combined to calculate mean distance from the best estimate and intra-examiner reliability. The combined values were used to create 80-point intra-examiner scattergrams, see Addendum A.

The Shrout-Fleiss intraclass correlation coefficient (ICC) was used to determine reliability estimates for landmark identification. (Shrout 1979) ICC was calculated for both adjusted X- and Y-coordinate values for each landmark on both frontal and sagittal images. In addition, 95% confidence interval (CI) was calculated for these reliability estimates. If there was no overlap of 95% CI values corresponding to the reliability of X- or Y- coordinates of landmarks, a significant difference could be claimed. When 95% CI values overlap, there was determined to be no significant difference of landmark identification.

RESULTS

The previously described method of capturing simultaneous frontal and sagittal facial images, yielded images with pixel sizes that ranged from 0.0862 mm to 0.1399 mm. These images were satisfactory for examiners to sufficiently locate the 17 predefined facial landmarks. The examiners approximated the location of the landmark if it was visually obstructed on an image (i.e. covered by hair). This obstruction occurred nearly exclusively for *trichion* and *center of right (rt.) tragus*.

Inter-examiner Reliability

The absolute value of distance from the best estimate of a landmark was used to evaluate the average amount of error when locating a specific landmark. *Supramentale* on the sagittal images, was the only landmark with an average of error of less than 0.5 mm, with a mean distance from the best estimate of 0.43(0.24) mm. The next group of landmarks had an average ranging from 0.5 mm to 1.0 mm and included, *pronasale*, *subnasale*, *philtrum point*, *stomion inferius*, *labrale inferius*, *soft tissue (s.t.) pogonion*, and *rt. cheilion* on the sagittal images, as well as, *stomion superius*, *stomion inferius*, and *rt. alare*, on the frontal images. All landmarks of the lips and nose had an average deviation of less than 1.50 mm on both types of images, except for *rt. alare* on the sagittal images. Other landmarks, *trichion*, *s.t. glabella*, and *rt. s.t. gonion*, were not as consistent with a large mean deviation from the best estimate on both types of images, ranging from 2.80 mm up to 6.60 mm, table 2.

Table 2. Inter-examiner Data

Landmark	Frontal						Sagittal					
	MD (mm)	SD (mm)	ICCx	95% CIx	ICCy	95% CIy	MD (mm)	SD (mm)	ICCx	95% CIx	ICCy	95% CIy
tr	2.80	1.68	0.771	(0.682, 0.905)	0.892	(0.845, 0.940)	3.11	2.89	0.931	(0.900, 0.971)	0.921	(0.886, 0.968)
g'	4.08	2.11	0.948	(0.909, 0.985)	0.783	(0.591, 0.944)	4.01	2.71	0.984	(0.977, 0.992)*	0.799	(0.712, 0.921)
se	2.83	1.77	0.963	(0.939, 0.988)	0.876	(0.801, 0.959)	1.33	1.01	0.995	(0.993, 0.998)*†	0.972	(0.955, 0.991)
prn	1.10	0.61	0.969	(0.955, 0.987)	0.989	(0.984, 0.995)	0.83	0.66	0.995	(0.993, 0.997)†	0.986	(0.976, 0.996)
sn	1.31	0.77	0.980	(0.972, 0.990)	0.977	(0.961, 0.993)	0.73	0.61	0.968	(0.954, 0.985)	0.992	(0.987, 0.997)*
pp	1.15	0.55	0.988	(0.983, 0.994)	0.984	(0.966, 0.996)	0.52	0.42	0.994	(0.991, 0.997)	0.994	(0.991, 0.997)
stos	0.67	0.44	0.982	(0.973, 0.990)†	0.996	(0.995, 0.998)*	1.19	1.00	0.874	(0.799, 0.959)	0.996	(0.994, 0.998)*
stoi	0.72	0.79	0.978	(0.969, 0.989)	0.993	(0.989, 0.996)	0.82	0.62	0.958	(0.939, 0.983)	0.997	(0.996, 0.999)*
li	1.03	1.10	0.969	(0.955, 0.986)	0.986	(0.979, 0.992)	0.94	0.62	0.974	(0.963, 0.989)	0.990	(0.986, 0.996)
sm	1.63	2.57	0.955	(0.933, 0.983)	0.927	(0.887, 0.956)	0.43	0.24	0.998	(0.997, 0.999)†	0.997	(0.995, 0.998)†
pg'	1.96	0.97	0.938	(0.909, 0.975)	0.960	(0.942, 0.983)	0.55	0.41	0.999	(0.998, 0.999)*†	0.991	(0.985, 0.995)†
me'	1.84	1.20	0.896	(0.855, 0.947)	0.959	(0.938, 0.984)	1.66	1.47	0.932	(0.904, 0.966)	0.981	(0.973, 0.991)*
go'	4.99	4.05	0.860	(0.786, 0.950)†	0.574	(0.444, 0.807)	6.60	4.00	0.480	(0.331, 0.765)	0.684	(0.578, 0.858)
t	2.08	1.24	0.942	(0.918, 0.975)	0.938	(0.912, 0.970)	1.95	1.20	0.983	(0.971, 0.995)	0.959	(0.915, 0.990)
ex	1.52	1.05	0.821	(0.727, 0.937)	0.993	(0.990, 0.996)*	1.36	0.98	0.892	(0.844, 0.957)	0.996	(0.994, 0.998)*
ch	1.38	3.15	0.870	(0.818, 0.929)	0.919	(0.876, 0.952)	0.65	0.35	0.985	(0.977, 0.995)†	0.997	(0.995, 0.998)†
al	0.81	0.68	0.992	(0.987, 0.995)†	0.988	(0.982, 0.995)	2.45	1.59	0.658	(0.365, 0.914)	0.963	(0.935, 0.989)*

MD, Mean distance from best estimate; SD, standard deviation; ICC, intraclass correlation coefficient; 95% CI, 95% confidence interval

* denotes significantly greater reliability than 95% CI for opposite plane on same type of image

† denotes significantly greater reliability than 95% CI for same plane on different type of image

With an ICC 95% CI > 0.950, 41.2% (7/ 17) and 47.1% (8/ 17) of the frontal facial landmarks showed high levels of reliability in the X- and Y-planes, respectively. Conversely, with a 95% IC \leq 0.950, 29.4% (5/ 17) of the landmarks in the X-plane and 17.6% (3/ 17) in the Y-plane are significantly less reliable than those previously mentioned. On sagittal images, 58.8% (10/17) of the landmarks in the X-plane and 70.6% (12/ 17) in the Y-plane showed high levels of reliability; whereas, in both X- and Y-planes, 11.8% (2/17) of landmarks were significantly less reliable.

With an ICC 95% CI > 0.950, mouth and lip landmarks: *pronasale*, *subnasale*, *philtrum point*, *stomion superius*, *stomion inferius*, *labrale inferius*, and *rt. alare* were significantly more reliably located on frontal images in both the X- and Y-planes than *rt. cheilion*, *rt. s.t. gonion*, and *trichion*, which have an ICC 95% CI completely less than 0.950. On the sagittal images, *sellion*, *pronasale*, *subnasale*, *philtrum point*, *labrale inferius*, *supramentale*, *s.t. pogonion*, and *rt. cheilion* were significantly more reliable in both planes than *rt. s.t. gonion*. When comparing the reliability in the different planes, on frontal images, *stomion superius* and *rt. exocanthion* were significantly more reliable in the Y-plane. On the sagittal images, *s.t. pogonion*, *s.t. glabella*, and *sellion* were significantly more reliable in the X-plane; whereas, *subnasale*, *stomion superius*, *stomion inferius*, *s.t. menton*, *rt. exocanthion*, and *rt. alare* were significantly more reliable in the Y-plane.

Three landmarks, *s.t. pogonion*, *supramentale* and *rt. cheilion* were significantly more reliable in both the X- and Y-planes when located on sagittal, rather than frontal, images. Additionally, *sellion* and *pronasale* were more reliable in the X-plane when located on sagittal images while *stomion superius*, *rt. s.t. gonion* and *rt. alare* were

significantly more reliable in the X-plane on frontal images. The remaining landmarks showed no significant differences when comparing landmarks located on frontal and sagittal images.

Intra-Examiner Reliability

In the intra-examiner portion of this study, more than half of the landmarks (18/34) had an average deviation from the best estimate less than 0.5 mm and only *rt. s.t. gonion*, on both frontal and sagittal images, had an average deviation greater than 1.5 mm, table 3.

All facial soft tissue landmarks had ICC 95% CIs completely > 0.950 , except for *rt. s.t. gonion*, on both types of images, and *trichion*, on the frontal images, table 3. *Rt. s.t. gonion* on sagittal images was the only landmark with an ICC 95% CI that failed to reach 0.950 in the X-axis. Therefore, all landmarks, except for *rt. s.t. gonion* and *trichion* on frontal images, were significantly more reliable than the location of *rt. s.t. gonion* in the X-axis on sagittal images.

Table 3. Intra-examiner Data

Landmark	Frontal						Sagittal					
	MD (mm)	SD (mm)	ICCx	95% CIx	ICCy	95% CIy	MD (mm)	SD (mm)	ICCx	95% CIx	ICCy	95% CIy
tr	1.45	1.16	0.958	(0.920, 0.978)‡	0.981	(0.964, 0.990)‡	1.04	1.11	0.990	(0.981, 0.995)†‡	0.990	(0.981, 0.995)‡
g'	1.33	0.69	0.993	(0.988, 0.997)‡	0.992	(0.984, 0.996)‡	0.75	0.65	0.999	(0.999, 1)*†‡	0.994	(0.989, 0.997)‡
se	0.81	0.52	0.998	(0.996, 0.999)‡	0.995	(0.990, 0.997)‡	0.38	0.30	0.999	(0.999, 1)‡	0.999	(0.998, 0.999)†‡
prn	0.55	0.39	0.995	(0.991, 0.997)‡	0.999	(0.998, 0.999)*‡	0.31	0.27	0.999	(0.999, 1)†‡	0.999	(0.998, 1)‡
sn	0.48	0.31	0.997	(0.995, 0.998)‡	0.999	(0.998, 0.999)‡	0.30	0.23	0.998	(0.997, 0.999)‡	0.999	(0.999, 1)‡
pp	0.27	0.35	0.998	(0.995, 0.999)‡	0.999	(0.999, 1)‡	0.28	0.20	0.999	(0.998, 0.999)‡	0.999	(0.998, 1)‡
stos	0.33	0.21	0.997	(0.995, 0.999)‡	0.999	(0.999, 1)‡	0.20	0.14	0.999	(0.998, 0.999)‡	0.999	(0.999, 1)‡
stoi	0.38	0.27	0.997	(0.995, 0.999)‡	0.999	(0.999, 1)‡	0.40	0.29	0.995	(0.991, 0.998)‡	0.999	(0.999, 1)*
li	0.42	0.32	0.996	(0.992, 0.998)‡	0.999	(0.999, 1)*‡	0.31	0.17	0.998	(0.996, 0.999)‡	0.999	(0.999, 1)‡
sm	0.65	0.43	0.994	(0.988, 0.997)‡	0.998	(0.997, 0.999)‡	0.19	0.11	0.999	(0.999, 1)†	0.999	(0.999, 1)‡
pg'	1.06	1.67	0.988	(0.978, 0.994)‡	0.986	(0.974, 0.993)	0.22	0.19	0.999	(0.999, 1)†	0.998	(0.997, 0.999)†‡
me'	0.96	0.69	0.985	(0.971, 0.992)‡	0.994	(0.989, 0.997)‡	0.66	0.66	0.994	(0.989, 0.997)‡	0.998	(0.998, 0.999)*†‡
go'	3.11	2.67	0.976	(0.955, 0.987)†‡	0.938	(0.882, 0.967)‡	4.26	2.04	0.835	(0.689, 0.913)	0.952	(0.909, 0.974)‡
t	1.02	1.12	0.992	(0.985, 0.996)‡	0.986	(0.973, 0.992)‡	0.25	0.21	0.999	(0.999, 1)†‡	0.999	(0.999, 1)†‡
ex	0.57	0.46	0.984	(0.970, 0.992)‡	0.999	(0.998, 0.999)*‡	0.68	0.76	0.983	(0.967, 0.991)‡	0.999	(0.999, 1)*‡
ch	0.44	0.36	0.989	(0.980, 0.994)‡	0.999	(0.999, 1)*‡	0.21	0.13	0.999	(0.998, 0.999)†‡	0.999	(0.998, 0.999)
al	0.42	0.33	0.998	(0.997, 0.999)†‡	0.999	(0.998, 0.999)‡	0.59	0.39	0.992	(0.985, 0.996)‡	0.998	(0.997, 0.999)*‡

MD, Mean distance from best estimate; SD, standard deviation; ICC, intraclass correlation coefficient; 95% CI, 95% confidence interval

* denotes significantly greater reliability than 95% CI for opposite plane on same type of image

† denotes significantly greater reliability than 95% CI for same plane on different type of image

‡ denotes significantly greater intra-examiner reliability than inter-examiner 95% CI for same landmark in the same plane on same type of image

Four landmarks on the frontal images, *pronasale*, *labrale inferius*, *rt. exocanthion* and *rt. cheilion*, were significantly more reliable in the Y-axis than the X-axis. On sagittal images, *stomion inferius*, *s.t. menton*, *rt. exocanthion*, and *rt. alare* were significantly more reliable in the Y-axis; whereas, *s.t. glabella* was more reliable in the X-axis.

When comparing the reliability of landmarks on frontal and sagittal images, *pogonion* and *rt. tragus* were significantly more reliable in both X- and Y-planes on sagittal images. In the X-axis, *trichion*, *s.t. glabella*, *pronasale*, *supramentale* and *rt. cheilion* were significantly more reliable on sagittal images, while *rt. s.t. gonion* and *rt. alare* were more reliably located on frontal images. A couple of landmarks, *sellion* and *menton*, were significantly more reliable on sagittal than frontal images only in the Y-axis and the remaining landmarks failed to reach significance when comparing the frontal and sagittal images.

Inter-examiner vs. Intra-examiner

Overall, the reliability of facial soft tissue landmarks was significantly greater for the intra-examiner trials than for inter-examiner trials, tables 2 and 3. When compared to inter-examiner data, all landmarks located in this study had decreased intra-examiner mean distance from the best estimate. On the frontal images, all landmarks were significantly more reliable in both planes when the same examiner located the point compared to when there were multiple examiners, except *s.t. pogonion*, which only reached significance in the X-axis. On the sagittal images, two landmarks, *stomion inferius* and *rt. cheilion*, were significantly more reliable in the X-axis, whereas, *s.t. pogonion*, *supramentale*, and *rt. cheilion* had significantly greater reliability in the Y-

axis, when compared to the inter-examiner data. The remaining sagittal landmarks showed significantly greater reliability in both the X- and Y-planes when the same examiner located the landmarks than when there were multiple examiners.

The 100-point inter-examiner and 80-point intra-examiner scattergrams created for each facial landmark represent the envelope of error and aid in visualization of their reliability. The variation in landmark identification error from landmark to landmark and between frontal and sagittal images can be visualized, Addendum A. Additionally, these scattergrams show the decreased amount of error when a landmark is identified by a single examiner.

DISCUSSION

Many previous studies that examined the reliability of facial measurements marked landmarks on the subjects' face prior to the directly or indirectly obtaining measurements by various techniques. (Farkas 1980, 1994, Muradin 2007) Physically marking landmarks on the skin prior to facial photographs is time consuming and may be objectionable to some orthodontic patients. For these reasons, facial images were captured without marking the landmarks on the subjects' skin to simulate a procedure that is more appropriate for an orthodontic practice.

When analyzing the reliability of facial landmarks, it is important that location of landmarks by examiners is not influenced by others or previous attempts at landmark location on the same image. Therefore, the examiners were commissioned to locate facial landmarks, rather than simply locating ink marks on the subjects' skin. However, without marking the landmarks on the subjects' skin, mean deviations from the actual landmarks could not be measured. Rather, the best estimates of the true location of the landmarks on each image were derived from the examiners' attempts to locate each landmark and used to calculate mean deviation. (Baumrind 1971)

To minimize the effect of previous attempts on landmark location, two weeks separated the first and second session used for intra-examiner reliability. This two-week interval was within the time range for repeated measures seen in previous studies, which ranged from immediately following the initial measurement to a two month period between measurements. (Fernandez-Riveiro 2002, 2003, Anic-Milosevic 2008, 2011, Schimmel 2010, Lee 2010, Mollov 2012)

Two major types of errors have been noted when evaluating cephalometric radiographs: errors of projection and errors of identification. (Baumrind 1971) The former occurs because cephalometric radiographs are two-dimensional distorted enlargements of the head with the magnification varying based on the plane in which the landmark lies. The latter originates when specific landmarks are identified on the radiographs. These two sources of error may also occur when locating landmarks on facial images due variable distortion of the two-dimensional image based on the distance of landmarks from the webcam lens and the errors of identification. (Gavan 1952) Furthermore, the magnitude of landmark identification error is influenced by many factors, including: the quality of the image, the care and skill of the examiner, and the method of identification used. (Baumrind 1971, Phillips 1984, Trpkova 1997)

The previously described, low-cost method of capturing simultaneous frontal and sagittal facial images with two high-definition webcams produced satisfactory images for locating facial landmarks. The images were captured simultaneously to minimize bias that may have arisen from differences in the tonicity of facial musculature between the frontal and sagittal images. Subjects were seated in ambient light and the images were captured without an additional light source. The use of a supplementary light source could have improved the definition of landmarks and therefore improved the reliability of identification. Additionally, certain landmarks, such as *trichion* and *rt. tragus*, were visually obstructed on some images and may have affected reliability if fully visible on all images.

In order to minimize distortion, the webcams were positioned at a distance of 1 m from the subjects, which decreased the resolution of the resultant facial images once

cropped and scaled. (Gavan 1952) The lens-to-subject distance used in this study is similar to those used in previous studies for photographs. (Phillips 1984, Fernandez-Riveiro 2002, 2003, Anic-Milosevic 2008) ABFO No. 2 bitemark scales were used as an inexpensive and simple instrument for measuring the magnification factor of facial images, and when placed in a plane common to the majority of landmarks, worked well for the purpose of this study.

One possible solution to improve the quality of the images would be the use of high-resolution professional digital single lens reflex cameras to capture the facial images, instead of webcams. Additionally, previous studies have demonstrated reliability of facial measurements with various methods of three-dimensional imaging. (Schimmel 2010, Chang 2011, Fourie 2011) However, both of these would have defeated one of the purposes of this study, to design a low-cost system that may be easily utilized by orthodontic practitioners.

The examiners had a similar level of skill and understanding of facial soft tissue landmarks. In order to minimize the variability between examiners due to vague definitions or differing interpretation of landmark location, an extensive calibration period has been recommended for studies with multiple examiners. (Phillips 1984, Mollov 2012) In this study, each examiner received identical verbal, visual, and written instruction on the exact facial landmarks and how to use OrthoMeasure. However, likely variation in the care taken by the examiners was found due to differing motivation for accurately locating the landmarks into the specialized software program designed for this study.

Possible sources of error may have arisen if care was not taken by an examiner when locating landmarks in OrthoMeasure. Increased error may have occurred if examiners only located the approximate location of landmarks on the full head images, and did not refine landmarks to what they believed to be the precise location on the close-up images. Additionally, if care was not taken as to the order of landmark identification, a landmark different from the one indicated in OrthoMeasure may have been located. This potential increase in error from a single examiner, would have affected the best estimate from the true value, which was an average of the points from all attempts. The distance an examiner's attempt deviates from the best estimate is affected by all other identification attempts for the same landmark on the same image. Therefore, these sources of error from a single examiner could have affected the distance from the best estimate for all other corresponding points.

The statistical analysis in this study examined two separate measures, mean distance from the best estimate and reliability. The adjusted coordinate values were used to calculate the mean distance that the points, located by examiners, deviated from the best estimate of the true value for the corresponding landmark. This measure was used to quantify distance of deviation, whereas ICC measured the reliability of this deviation in the X- and Y-axes. (Shrout 1979) An acceptable landmark for facial analysis should have minimal mean deviation, as well as an ICC close to 1, indicating great reliability. Therefore, data from both of these statistical analyses are needed to claim whether landmarks are sufficient for use in soft tissue facial analysis.

The results of this study support the findings of previous cephalometric and photogrammetric studies that examined the reliability of skeletal and soft tissue facial

landmarks. The magnitude of error varies between landmarks, the largest amount of error was seen on landmarks that were not well demarcated, and most landmarks have an envelope of error that is non-circular. (Baumrind 1971, Phillips 1984, Trpkova 1997)

With increased focus on cephalometric evaluation, several numeric analyses based on the soft tissue profile of these radiographs have been developed. (Legan 1980, Scheideman 1980, Holdaway 1983, Arnett 1999) Soft tissue changes that occur during growth and during treatment, especially in the lower third of the face, may influence treatment decisions. (Burstone 1959, Stephens 2005, Kachiwala 2009, Tsang 2009) Despite, the importance of soft tissue measurements in orthodontic diagnosis, the bulk of the research has focused solely on the reliability of skeletal landmarks and evidence has shown soft tissue measurements from cephalometric radiographs to be unreliable. (Burstone 1959, Park 1986)

The reliability of cephalometric identification varies between landmarks and previous studies recommended less than 0.59 mm in the X- axis and 0.56 mm in the Y- axis, or roughly 0.81mm of total error, to be considered acceptable for evaluation of cephalometric radiographs. (Baumrind 1971, Trpkova 1997) Reliability of photogrammetric facial soft tissue landmarks was previously defined by an absolute mean error less than or equal to 2.00 mm and found 57% of frontal landmarks and 66% of sagittal landmarks reached this level of consistency. (Phillips 1984)

With a mean deviation from the best estimate of the landmark between those defined by previous studies, (Phillips 1984, Trpkova 1997) less than 1.50 mm, 55.9% (19/34) of the landmarks had sufficient inter-examiner consistency and 94.1% (32/34)

had sufficient intra-examiner consistency, tables 2 and 3. The facial soft tissue covering generally has a softer contour than the underlying skeletal and dental structures, resulting in less distinct facial landmarks and generally decreased reliability. (Park 1986, Arnett 1999) Another possible reason for greater total error of landmark identification is the resolution of the facial images and lighting used in this study.

The error of landmark identification varied depending on the delineation of the landmarks on both the frontal and sagittal images. On frontal images, poorly-delineated or visually obstructed landmarks, *trichion*, *s.t. glabella*, *rt. s.t. gonion* and *rt. tragus*, had greater mean distances from the best estimate; conversely, well-demarcated landmarks, *stomion superius*, *stomion inferius* and *right alare*, had less mean deviation., Poorly-distinguished or visually obstructed landmarks on sagittal images, *trichion*, *s.t. glabella*, *rt. s.t. gonion*, had a large mean deviation, while well-delineated landmarks, *pronasale*, *subnasale*, *philtrum point*, *stomion inferius*, *labrale inferius*, *supramentale*, *s.t. pogonion* and *rt. cheilion*, yielded decreased mean deviations. These differences in reliability are represented by the variable dispersion of points seen in the scattergrams, Addendum A. This difference in mean distance from the best estimate between well- and poorly-delineated landmarks supports the findings of previous studies investigating reliability of facial landmarks and measurements. (Phillips 1984, Lagravere 2010, Mollov 2012)

This current study revealed certain landmarks to be significantly more reliable than others for both inter- and intra-examiner tests. On the frontal images, the majority of mouth and nose landmarks: *pronasale*, *subnasale*, *philtrum point*, *stomion superius*, *stomion inferius*, *labrale inferius* and *rt. alare* reached a 95% CI > 0.950, tables 2 and 3. These landmarks were significantly more reliable and more suitable for use in frontal soft

tissue analysis than *rt. s.t. gonion* and *trichion*, which failed to reach a 95% CI > 0.950 in at least one plane.

For the inter-examiner trials, *sellion*, *pronasale*, *subnasale*, *philtrum point*, *labrale inferius*, *supramentale*, *s.t. pogonion* and *rt. cheilion* were the most reliable landmarks and are satisfactory for use in soft tissue profile analysis, whereas, *rt. s.t. gonion* was not sufficiently reliable and should not be used for profile analysis. These results support the findings of other studies that reliability of soft tissue landmarks vary between points and greater reliability was found when a landmark was delineated well. (Phillis 1984, Lagravere 2010, Mollov 2012) When located by the same examiner, all sagittal landmarks, except *rt. s.t. gonion*, had sufficient reliability.

Each facial landmark is utilized quite differently by numeric and angular measurements, some relying on greater precision in the X-plane, such as *s.t. pogonion* and others in the Y-plane, *s.t. menton*. (Stoner 1955, Peck 1970, Phillips 1984, Baumrind 1971, Arnett 1993a, 1993b, Trpkova 1997, Fernandez-Riveiro 2002, Anic-Milosevic 2011) Therefore, with an envelope of error with a noncircular shape, it was more important for a landmark to have greater reliability in the plane that was more utilized for angular or linear analysis. Certain landmarks that are often used for vertical measurements in profile analysis, *stomion superius*, *stomion inferius* and *s.t. menton* reached a 95% CI greater than 0.950 in the Y-axis and are considered sufficiently reliable when used for vertical measurements. This reinforces the suitability of these landmarks for evaluation of lower facial third vertical proportions. (Anic-Milosevic 2010)

The only landmark that was more reliable in the Y-axis than the X-axis on both the frontal and sagittal images was *rt. exocanthion*, indicating that this landmark may be more useful for vertical rather than horizontal measurements. On sagittal images only, *stomion inferius*, *s.t. menton* and *rt. alare* were also more reliable in the Y- than the X-axis, and should be used for measurements that rely more on the vertical position of these landmarks. Conversely, on sagittal images, *s.t. glabella*, was significantly more reliable in the X- than the Y-plane, indicating greater reliability when using this landmark for measurements in the horizontal plane. These findings coincide with the use of these landmarks and the reliability of facial measurements reported by previous studies. (Fernandez-Riveiro 2002, Dimaggio 2007, Muradin 2007, Anic-Milosevic 2010, 2011, Lee 2010, Mollov 2012)

Many soft tissue facial landmarks are used for both frontal and sagittal facial analysis; however, the reliability of landmarks differs depending on the type of image. (Farkas 1980, 1994, Arnett 1993a, 1993b) The inter-examiner tests showed *supramentale*, *s.t. pogonion* and *rt. cheilion* to have significantly greater reliability on sagittal than on frontal images, in both planes. When located by a single examiner, *s.t. pogonion* and *rt. tragus*, are significantly more reliable on sagittal images than frontal images. Therefore, these landmarks may be more suitable for use in soft tissue profile analyses. Additionally, for measurements relying on the horizontal position of the landmark, *pronasale* may be more suitable for profile measurements, whereas *stomion superius*, *rt. s.t. gonion* and *rt. alare* may be more fitting for frontal analysis. These differences in reliability are reflected in the utilization of different landmarks for frontal

and profile facial analysis in previous studies. (Arnett 1993a, 1993b, Ferrario 1993, Reynecke 2012)

When comparing inter- and intra-examiner reliability, all landmarks located have significantly greater intra-examiner reliability, in one or both planes. Additionally, all intra-examiner landmarks have less mean deviation from the best estimate of the true value than inter-examiner landmarks. Many previous studies investigating the reliability of facial landmarks or measurements compared the repeated measures from a single examiner and did not examine the variation of landmark identification between examiners. (Phillips 1984, Fernandez-Riveiro 2002, Ghoddousi 2007, Anic-Milosevic 2008, Lee 2010, Ferring 2008) The bias of a single examiner results in greater consistency of landmark identification than the variable interpretation from multiple examiners. This opposes the claim that the number of examiners and occasions of identification do not play a significant role in the reliability of landmarks. (Trpkova 1997)

Scattergrams visually represent the dispersion of points from the examiners' attempts at landmark identification, superimposed on the best estimate for each landmark. (Baumrind 1971, Phillips 1984) The great variation in mean deviation and reliability between landmarks can be appreciated when comparing the sagittal scattergrams of one of the most reliable, *supramentale*, and one of the least reliable points, *rt. s.t. gonion*, Addendum A. These scattergrams reinforce that well-demarcated facial soft tissue landmarks are significantly more reliable than poorly-demarcated landmarks. Additionally, the scattergrams representing *s.t. pogonion* exhibit the variation in reliability based on the type of image on which a landmark is located. This reinforces that *s.t. pogonion* is more useful for profile analysis than frontal analysis.

Considering the results from the inter- and intra-examiner portion of this study, only nose and mouth landmarks, *pronasale*, *subnasale*, *philtrum point*, *stomion superius*, *stomion inferius*, *labrale inferius* and *rt. alare*, were consistently located and would be sufficient for frontal facial analysis. Additionally, all landmarks located in this study, except *rt. s.t. gonion*, are sufficiently reliable for frontal analysis, when located by a single examiner.

The findings of this study support previous claims that standardized frontal photographs are acceptable for measuring changes in the nasolabial area. (Muradin 2007) For repeated facial measurements, a mean coefficient of repeatability of 0.68 mm was found, indicating that 95% of intra-examiner measurement error was less than 0.68 mm. One major difference in the present study was that landmarks were not marked on the subjects' skin prior to capturing the images, which may account for the difference in total error.

On sagittal images, *sellion*, *pronasale*, *subnasale*, *philtrum point*, *labrale inferius*, *supramentale*, *s.t. pogonion* and *rt. cheilion* are reliable landmarks suitable for profile analysis. For intra-examiner reliability of landmark identification on sagittal images, all landmarks located, except *rt. s.t. gonion*, are acceptable for soft tissue profile analysis. Many of these landmarks have been utilized when developing ethnic standards for linear and angular soft tissue profile measurements. (Fernandez-Riveiro 2002, Dimaggio 2007, Anic-Milosevic 2011, Wamalwa 2011) Most of the angular measurements relying on these landmarks had a low amount of random error, within 2.5 degrees. The linear profile measurements fell within a range of 2-4 mm of total error. When these reliable facial landmarks are used for profile analysis, they yield reproducible measurements.

The landmarks located in this study have greater reliability when located by a single examiner than by multiple examiners. All landmarks located by a single examiner yielded smaller mean distances from the best estimate and significantly greater reliability in one or both X- and Y-planes. This greater reliability for intra-examiner identification indicates that quantifying facial changes may be more accurate if soft tissue landmarks are identified by a single examiner.

Pre-defined standards for linear and angular facial soft tissue analyses used in orthodontic diagnosis and treatment planning are not as reliable when there are multiple examiners, due to differences of interpretation of landmark location. (Arnett 1993a, 1993b, Fernandez-Riveiro 2002, 2003, Anic-Milosevic 2011) Therefore, standards from quantitative facial soft tissue analyses should be used with caution. Each practitioner should determine how linear and angular measurements are affected by personal differences in landmark identification, prior to diagnosing and treatment planning based on pre-defined standards.

CONCLUSION

1. The low-cost method for capturing simultaneous frontal and sagittal facial images described in this study yielded images of sufficient quality for soft tissue landmark identification.
2. The magnitude of error varies between landmarks, is largest for poorly-demarcated landmarks, and most landmarks have a non-circular envelope of error.
3. Several landmarks are more reliable on sagittal images and should be considered for profile analysis, whereas others are more suitable for frontal analysis.
4. When located by a single examiner, all facial soft tissue landmarks from this study showed greater reliability and less mean deviation from the best estimate than when located by multiple examiners.
5. Mouth and nose landmarks are most suitable for photogrammetric facial analysis.
6. *Right soft tissue gonion* is not sufficiently reliable for use in photogrammetric facial analysis.

BIBLIOGRAPHY

- Abed Y, H.-Z. G. (2009). Lip posture following debonding of labial appliances based on conventional profile photographs. *Angle Orthodontist*, 79(2), 235-9.
- Anic-Milosevic S, L.-V. M. (2008). Analysis of the soft tissue facial profile by means of angular measurements. *European Journal of Orthodontics*, 30(2), 135-40.
- Anic-Milosevic S, M. S. (2010). Proportions in the upper lip-lower lip-chin area of the lower face as determined by photogrammetric method. *Journal of Cranio-Maxillo-Facial Surgery*, 38(2), 90-5.
- Anic-Milosevic S, M. S.-V. (2011). Analysis of the soft tissue profile in Croatians with normal occlusions and well-balanced faces. *European Journal of Orthodontics*, 33(3), 305-10.
- Arnett GW, B. R. (1993). Facial keys to orthodontic diagnosis and treatment planning. Part I. *American Journal of Orthodontics and Dentofacial Orthopedics*, 103(4), 299-312.
- Arnett GW, B. R. (1993). Facial keys to orthodontic diagnosis and treatment planning. Part II. *American Journal of Orthodontics and Dentofacial Orthopedics*, 103(5), 395-411.
- Arnett GW, G. M. (2004). Facial Planing For Orthodontist ans Oral Surgeons. *American Journal of Orthodontics and Dentofacial Orthopedics*, 126(3), 290-5.
- Arnett GW, J. J. (1999). Soft tissue cephalometric analysis: Diagnosis and treatment planning of dentofacial deformity. *American Journal of Orthodontics and Dentofacial Orthopedics*, 116(3), 239-53.
- Arnett GW, W. C. (1999). The treatment motivation survey: Defining patient motivation for treatment. *American Journal of Orthodontics and Dentofacial Orthopedics*, 115(3), 233-8.
- Baldwin, D. (1980). Appearance and aesthetic in oral health. *Community Dental Oral Epidemiology*, 9, 244-56.
- Baumrind S, F. R. (1971). The reliability of head film measurements 1. Landmark identification. *American Journal of Orthodontics*, 60(2), 111-27.
- Benson PE, R. S. (1997). A critical appraisal of measurements of the soft tissue outline using photographs and video. *European Journal of Orthodontics*, 19(4), 397-409.
- Bishara SE, C. D. (1995). A computer assisted photogrammetric analysis of soft tissue changes after orthodontic treatment. Part I: Methodology and reliability. *American Journal of Orthodontics and Dentofacial Orthopedics*, 107(6), 633-9.

- Bjork, A. (1955). Facial growth in man, studied with the aid of metallic implants. *Acta Odontol Scand*, 13(1), 9-34.
- Broadbent, B. (1931). A new x ray technique and its application to orthodontics. *Angle Orthod*, 145-66.
- Burke PH, B. F. (1967). Stereophotogrammetry of the face. A preliminary investigation into the accuracy of a simplified system evolved for contour mapping by photography. *Am J Orthod*, 53(10), 769-82.
- Burstone, C. (1958). The integumental profile. *American Journal of Orthodontics*, 44(1), 1-25.
- Burstone, C. (1959). Integumental Contour and Extension Patterns. *Angle Orthodontist*, 29(2), 93-104.
- Burstone, C. (1967). Lip posture and its significance in treatment planning. *American Journal of Orthodontics*, 53(4), 262-84.
- Chang ZC, H. F. (2011). Landmark identification errors on cone-beam computed tomography-derived cephalograms and conventional digital cephalograms. *American Journal of Orthodontics and Dentofacial Orthopedics*, 140(6), e289-e297.
- Chu, M. (1989). The accuracy of photogrammetric measurements employed in the analysis of facial morphology. *Masters Thesis*.
- Claus EB, C. L. (2012). Dental X-Rays and Risk of Meningioma. *Cancer*, 118, 4530-7.
- Cummins DM, B. S. (1995). A computer assisted photogrammetric analysis of soft tissue changes after orthodontic treatment. Part II: Results. *American Journal of Orthodontics and Dentofacial Orthopedics*, 108(1), 38-47.
- Dimaggio FR, C. V. (2007). Photographic soft-tissue profile analysis in children at 6 years of age. *American Journal of Orthodontics and Dentofacial Orthopedics*, 132(4), 475-80.
- Downs, W. (1956). Analysis of the dentofacial profile. *Angle Orthodontist*, 26(4), 191-212.
- Edler, R. (2001). Background considerations to facial aesthetics. *Journal of Orthodontics*, 28(2), 159-68.
- Farkas LG, B. W. (1980). Is photogrammetry of the face reliable? *Plastic and Reconstructive Surgery*, 66(3), 346-55.
- Farkas LG, K. M. (2005). International anthropometric study of facial morphology in various ethnic groups/races. *Journal of Craniofacial Surgery*, 16(4), 615-46.
- Farkas, L. (1994). *Anthropometry of the Head and Face*. New York: Raven Press.

- Fazel R, K. H. (2009). Exposure to low-dose ionizing radiation from medical imaging procedures. *New England Journal of Medicine*, 361(9), 849-57.
- Fernandez-Riveiro P, S.-C. E.-Q.-C. (2003). Angular photogrammetric analysis of the soft tissue facial profile. *European Journal of Orthodontics*, 25(4), 393-9.
- Fernandez-Riveiro P, S.-Q. D.-C.-C. (2002). Linear photogrammetric analysis of the soft tissue facial profile. *American Journal of Orthodontics and Dentofacial Orthopedics*, 122(1), 59-66.
- Ferrario VF, S. C. (1993). Craniofacial morphometry by photographic evaluations. *American Journal of Orthodontics and Dentofacial Orthopedics*, 103(4), 327-37.
- Ferring V, P. H. (2008). Divine Proportions in the Growing Face. *American Journal of Orthodontics and Dentofacial Orthopedics*, 134(4), 472-9.
- Feu D, O. B. (2012). Influence of orthodontic treatment on adolescents' self-perceptions of esthetics. *American Journal of Orthodontics and Dentofacial Orthopedics*, 141(6), 743-50.
- Finlay, F. (1980). Craniometry and Cephalometry: A History Prior to the Advent of Radiography. *Angle Ortho*, 50(4), 312-21.
- Fourie Z, D. J. (2011). Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. *Forensic Science International*, 207(1), 127-34.
- Gavan JA, W. S. (1952). Photography: An anthropometric tool. *American Journal of Physical Anthropology*, 10, 331-51.
- Ghoddousi H, E. R. (2007). Comparison of three methods of facial measurement. *International Journal of Oral Maxillofacial Surgery*, 36(3), 250-8.
- Gribel BF, G. M. (2011). Accuracy and reliability of craniometric measurements on lateral cephalometry and 3D measurements on CBCT scans. *Angle Orthod*, 81(1), 26-35.
- Hajeer MY, D. D. (2004). Applications of 3D imaging in orthodontics: Part I. *Journal of Orthodontics*, 31(1), 62-70.
- Halazonetis, D. (2007). Morphometric evaluation of soft-tissue profile shape. *American Journal of Orthodontics and Dentofacial Orthopedics*, 131(4), 481-9.
- Hellman, M. (1939). Some facial features and their orthodontic implication. *Am J Orthod Oral Surgery*, 25(10), 927-51.
- Holdaway, R. (1983). A soft-tissue cephalometric analysis and its use in orthodontic treatment planning. Part I. *American Journal of Orthodontics*, 84(1), 1-28.
- Hrdlicka, A. (1920). *Practical Anthropometry*. Philadelphia: Wistar Institute of Anatomy and Biology.

- Islam R, K. T. (2009). Lip morphological changes in orthodontic treatment: Class II division 1 malocclusion and normal occlusion at rest and on smiling. *Angle Orthodontist*, 79(2), 256-64.
- Jorgensen, G. (1991). Changes in the facial dimensions of children 5 to 12 years of age: a longitudinal study of photographs. [Master thesis]. University of Iowa.
- Kachiwala VA, K. A. (2009). Soft tissue changes associated with first premolar extractions in adult females. *Australian Orthodontic Journal*, 25(1), 24-9.
- Kocadereli, I. (2002). Changes in soft tissue profile after orthodontic treatment with and without extractions. *American Journal of Orthodontics and Dentofacial Orthopedics*, 122(1), 67-72.
- Lagravère MO, L. C.-M. (2010). Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. *137*, 598-604.
- Lee MS, C. D. (2010). Assessing soft-tissue characteristics of facial asymmetry with photographs. *American Journal of Orthodontics and Dentofacial Orthopedics*, 138(1), 23-31.
- Legan HL, B. C. (1980). Soft tissue cephalometric analysis for orthognathic surgery. *Journal of Oral Surgery*, 38, 744-51.
- Livas C, D. K. (2012, January). Subjective and objective perception of orthodontic treatment need: A systematic review. *European Journal of Orthodontics*, epub.
- Ludlow JB, D.-L. L. (2008). Patient risk related to common dental radiographic examinations: The impact of 2007 International Commission on Radiological Protections recommendations regarding dose calculations. *Journal of the American Dental Association*, 139(9), 1237-43.
- Mizumoto Y, D. T. (2009). Assessment of facial golden proportions among young Japanese women. *Am J Orthod Dentofacial Orthop*, 136(2), 168-74.
- Mollov N, B. J. (2012). Intra and inter-examiner reliability of direct facial soft tissue measurements using digital calipers. *1*, e157-61.
- Mupparapu M. (2005). Radiation protection guidelines for the practicing orthodontist. *American Journal of Orthodontics and Dentofacial Orthopedics*, 128(2), 168-72.
- Muradin MS, R. A. (2007). The reliability of frontal facial photographs to assess changes in nasolabial soft tissues. *International Journal of Oral Maxillofacial Surgery*, 36(8), 728-34.
- Murphy C, K. G. (2011). The clinical relevance of orthognathic surgery on quality of life. *Internatinoal Journal of Oral Maxillofacial Surgery*, 40(9), 926-30.

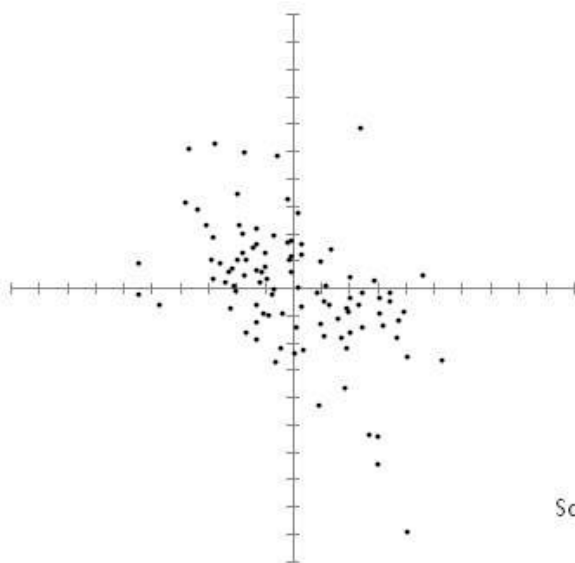
- Naini, F. (2011). *Facial Aesthetics: Concepts & Clinical Diagnosis*. Singapore: Wiley-Blackwell.
- Neger, M. (1959). A quantitative method for the evaluation of the soft-tissue facial profile. *American Journal of Orthodontics*, 45(10), 738-51.
- Pabari S, M. D. (2011). Assessment of motivation and psychological characteristics of adult orthodontic patients. *American Journal of Orthodontics and Dentofacial Orthopedics*, 140(6), e263-72.
- Park YC, B. C. (1986). Soft tissue profile: Fallacies of hard tissue standards in treatment planning. *American Journal of Orthodontics and Dentofacial Orthopedics*, 90(1), 52-62.
- Peck H, P. S. (1970). A concept of facial esthetics. *Angle Orthodontist*, 40(4), 284-318.
- Phillips C, G. J. (1984). Photocephalometry: Errors of projection and landmark location. *American Journal of Orthodontics*, 86(3), 233-43.
- Powell N, H. B. (1984). *Proportions of the esthetic face*. New York: Thieme-Stratton.
- Prabakaran R, S. S. (2012). Motivation for orthodontic treatment investigated with Q-methodology: Patients' and parents' perspectives. *American Journal of Orthodontics and Dentofacial Orthopedics*, 142(2), 213-20.
- Proffit W, F. H. (2012). *Contemporary Orthodontics*. St. Louis: Mosby Elsevier.
- Reyneke JP, F. C. (2012). Clinical Assessment of the Face. *Seminars in Orthodontics*, 18(3), 172-86.
- Rustemeyer J, G. J. (2012). Quality of life in orthognathic surgery patients: Post-surgical improvements in aesthetics and self-confidence. *Journal of Cranio-Maxillo-Facial Surgery*, 40(5), 400-4.
- Sarver DM, A. M. (2003). Dynamic smile visualiation and quantification: Part 2. Smile analysis and treatment strategies. *American Journal of Orthodontics and Dentofacial Orthopedics*, 124(2), 116-27.
- Sarver DM, A. M. (2003). Dynamic smile visualization and quantification: Part 1. Evolution of the concept and dynamic records for smile capture. *American Journal of Orthodontics and Dentofacial Orthopedics*, 124(1), 4-12.
- Sarver, D. (1996). Video cephalometric diagnosis (VCD): A new concept in treatment planning? *American Journal of Orthodontics and Dentofacial Orthopedics*, 110(2), 128-36.
- Sarver, D. (2004). Principles of cosmetic dentistry in orthodontics: Part 1. Shape and proportionality of anterior teeth. *Am J Orthod Dentofacial Orthop*, 126(6), 749-53.

- Sarver, D. (2004). Principles of cosmetic dentistry in orthodontics: Part 1. Shape and proportionality of anterior teeth. *Am J Orthod Dentofacial Orthop*, 126(6), 749-53.
- Scheideman GB, B. W. (1980). Cephalometric analysis of dentofacial normals. *American Journal of Orthodontics*, 78(4), 404-20.
- Schimmel M, C. P. (2010). Distances between facial landmarks can be measured accurately with a new digital 3-dimensional video system. *American Journal of Orthodontics and Dentofacial Orthopedics*, 137(5), 580.e1-580.e10.
- Shaner DJ, B. J. (1998). Technical note: Different techniques, different results--A comparison of photogrammetric and caliper-derived measurements. *American Journal of Physical Anthropology*, 106(4), 547-52.
- Shindoi JM, M. Y. (2013). Soft Tissue Cephalometric Norms for Orthognathic and Cosmetic Surgery. *J Oral Maxillofac Surg*, 71(1), e24-30.
- Shrout PE, F. J. (1979). Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin*, 86(2), 420-8.
- Steiner, C. (1959). Cephalometrics in clinical practice. *Angle Orthodontist*, 29(1), 8-29.
- Stephens CK, B. J. (2005). Long-term profile changes in extraction and nonextraction patients. *American Journal of Orthodontics and Dentofacial Orthopedics*, 128(4), 450-7.
- Stoner, M. (1955). A photogrammetric analysis of the facial profile. *American Journal of Orthodontics*, 41(6), 453-69.
- Subtelny, J. (1959). A longitudinal study of soft tissue facial structures and their profile characteristics, defined in relation to underlying skeletal structures. *American Journal of Orthodontics*, 45(7), 481-507.
- Sutter RE Jr, T. P. (1998). Soft tissue evaluation of contemporary Caucasian and African American female facial profiles. *Angle Orthod*, 68(6), 487-96.
- Trpkova B, M. P. (1997). Cephalometric landmarks identification and reproducibility: A Meta analysis. *American Journal of Orthodontics and Dentofacial Orthopedics*, 112, 165-70.
- Tsang ST, M. L. (2009). Profile changes in orthodontic patients treated with mandibular advancement surgery. *American Journal of Orthodontics and Dentofacial Orthopedics*, 135(1), 66-72.
- Wamalwa P, A. S. (2011). Angular Photogrammetric Comparison of the Soft-Tissue Facial Profile of Kenyans and Chinese. *J Craniofac Surg*, 22(3), 1064-72.

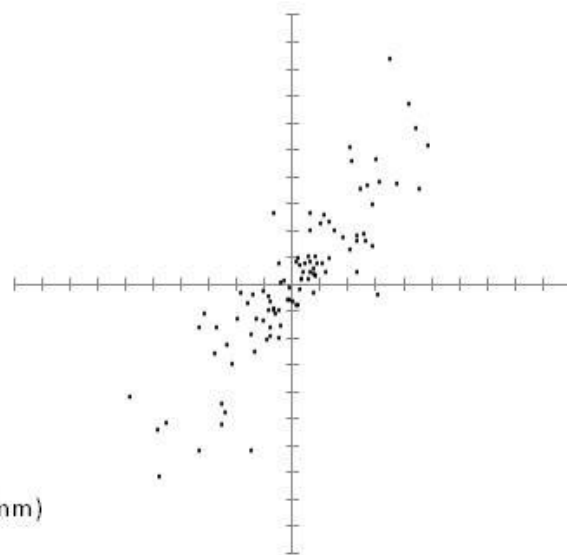
ADDENDUM A

Trichion

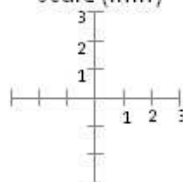
Inter-examiner Frontal



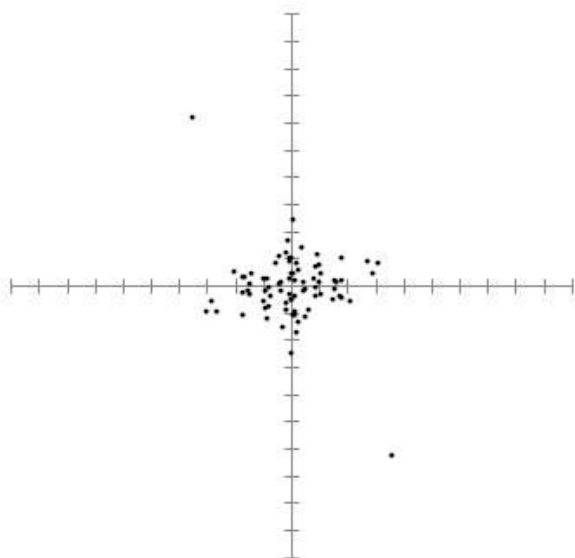
Inter-examiner Sagittal



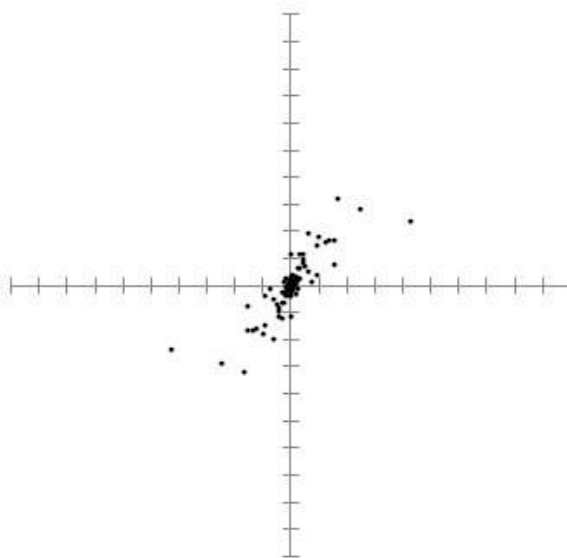
Scale (mm)



Intra-examiner Frontal

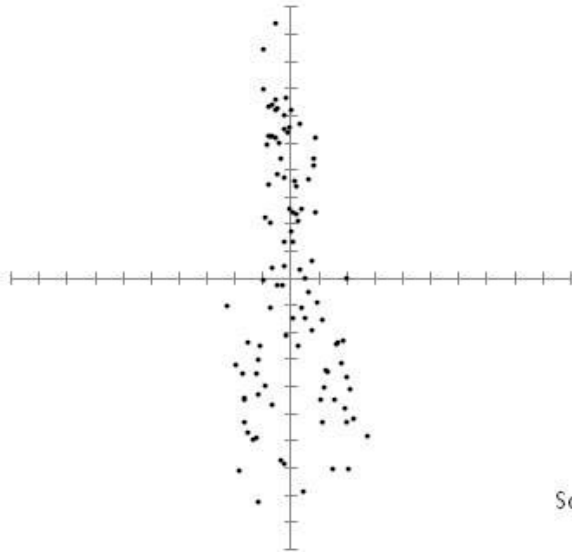


Intra-examiner Sagittal

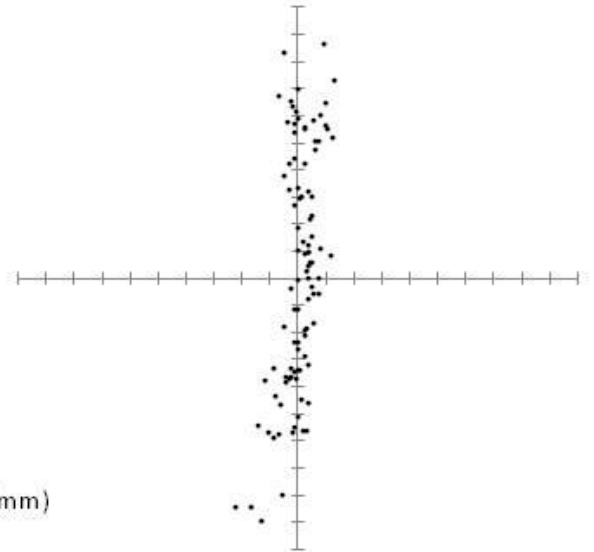


Soft tissue Glabella

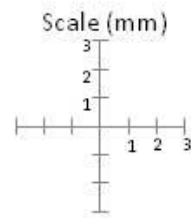
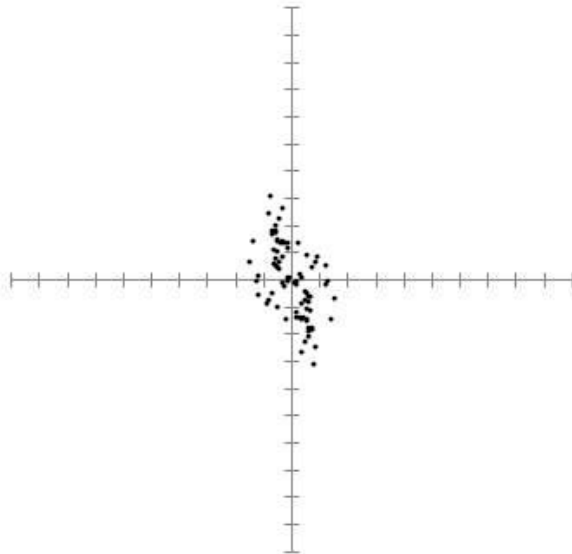
Inter-examiner Frontal



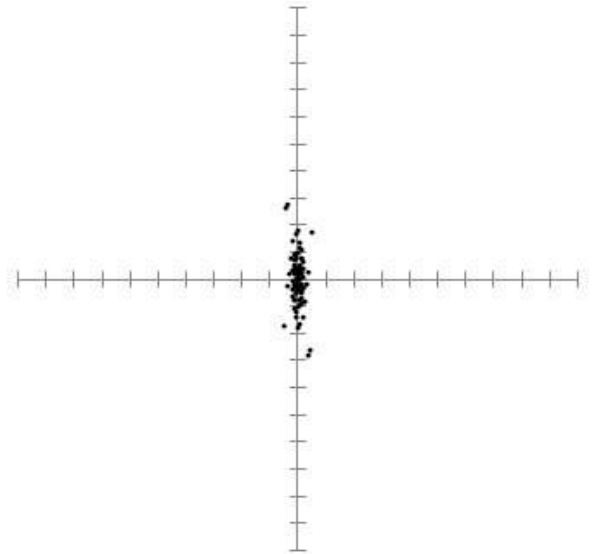
Inter-examiner Sagittal



Intra-examiner Frontal

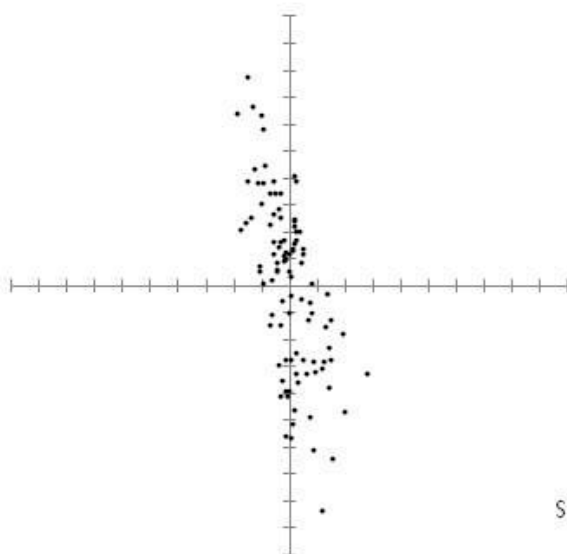


Intra-examiner Sagittal

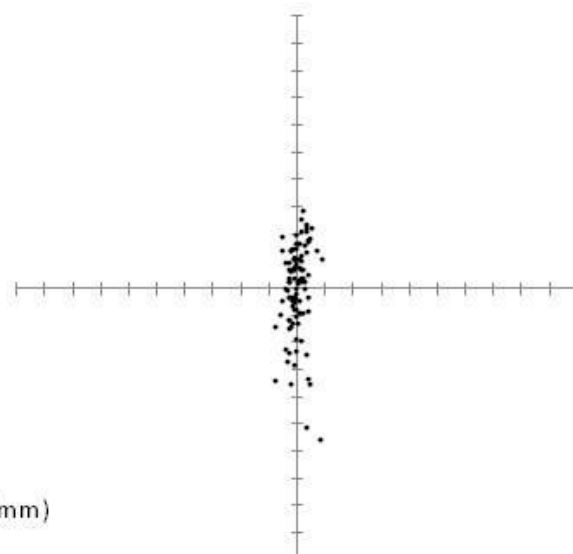


Sellion

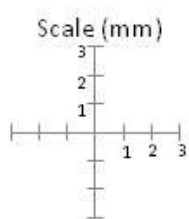
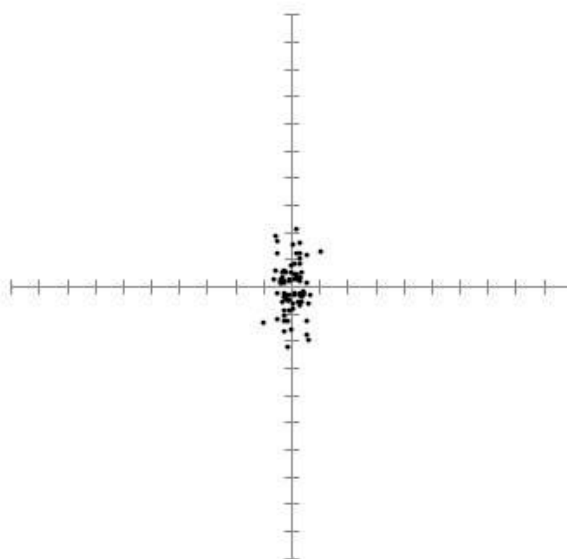
Inter-examiner Frontal



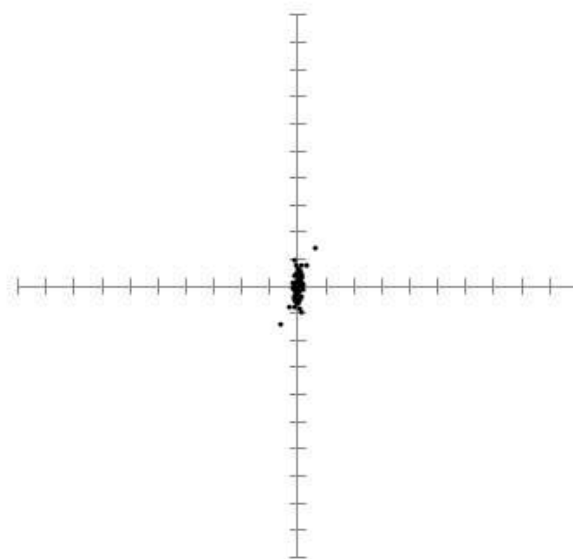
Inter-examiner Sagittal



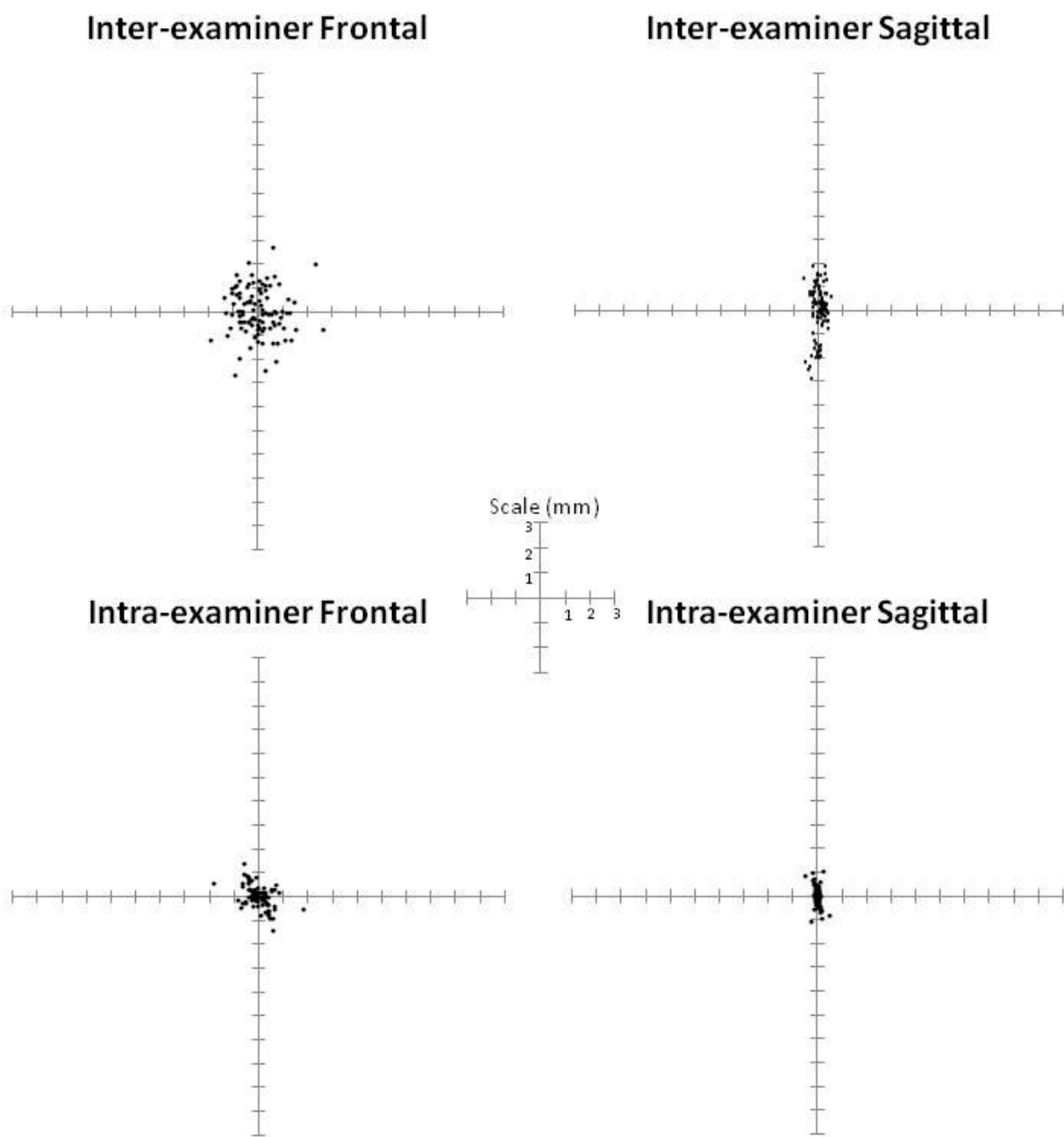
Intra-examiner Frontal



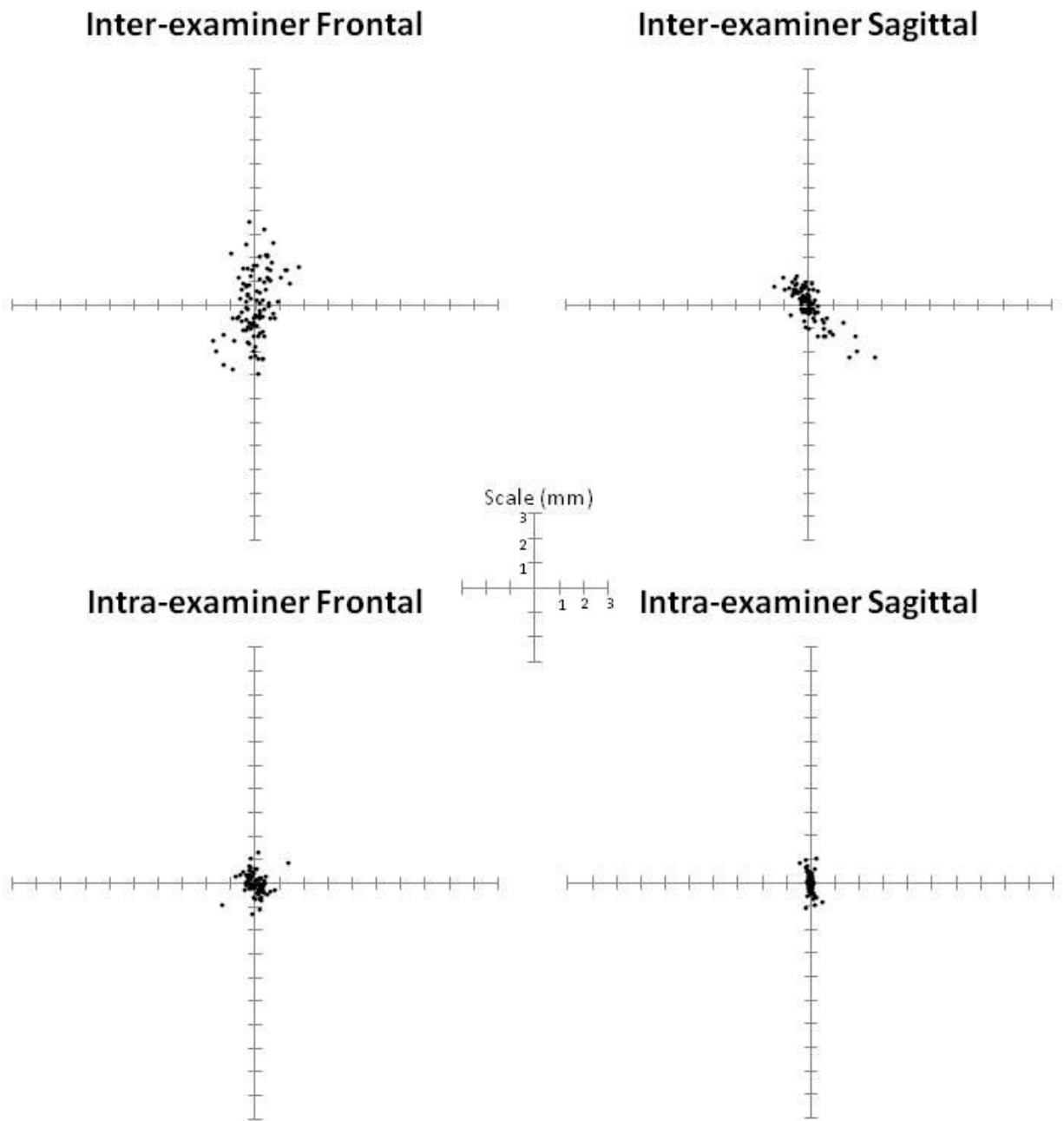
Intra-examiner Sagittal



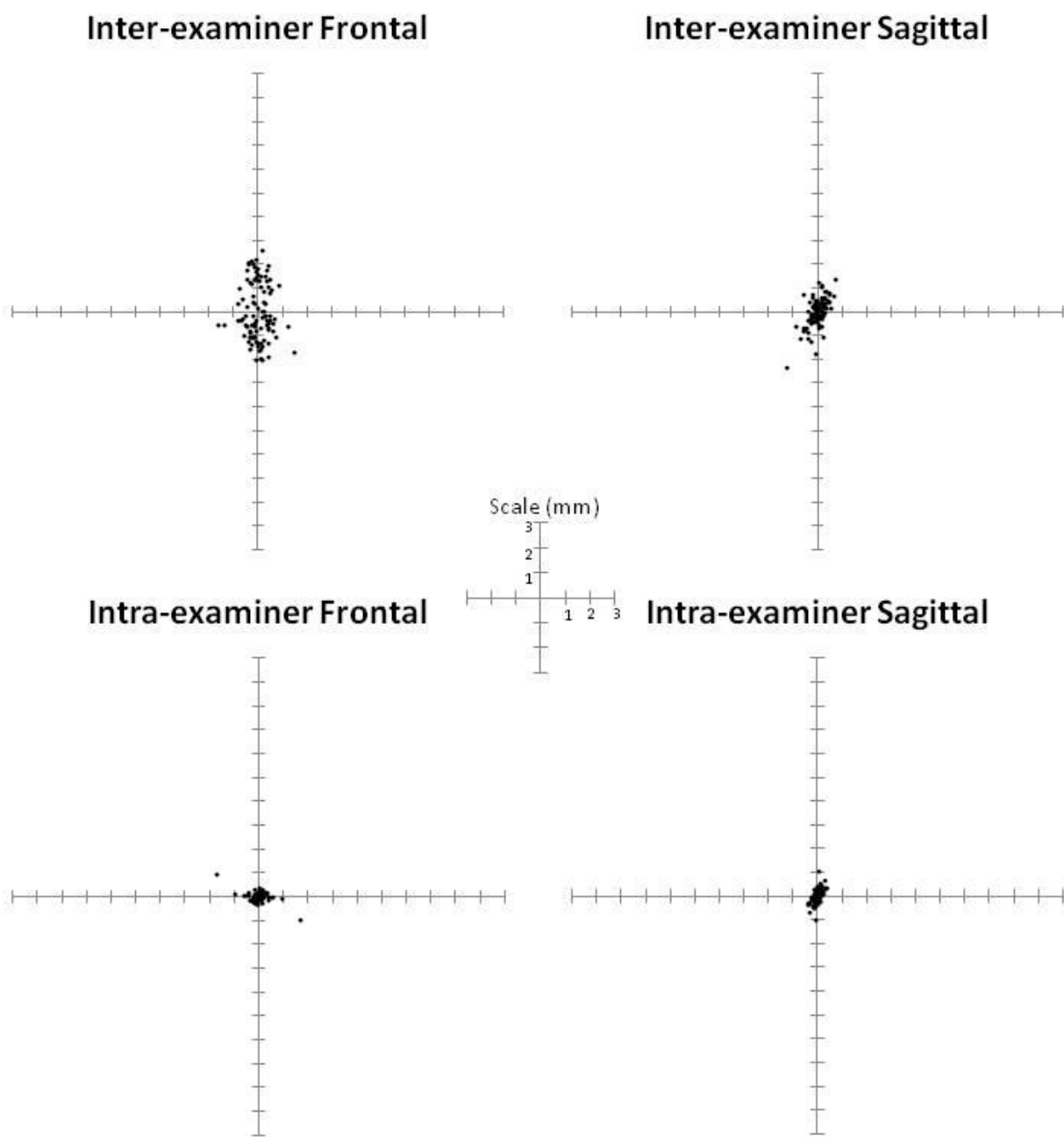
Pronasale



Subnasale

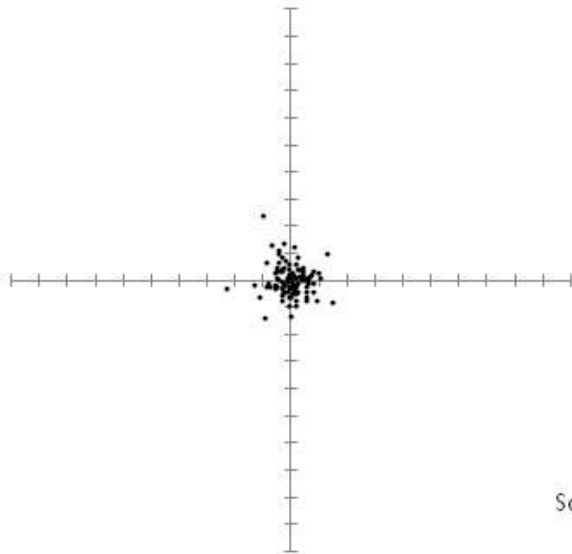


Philtrum Point

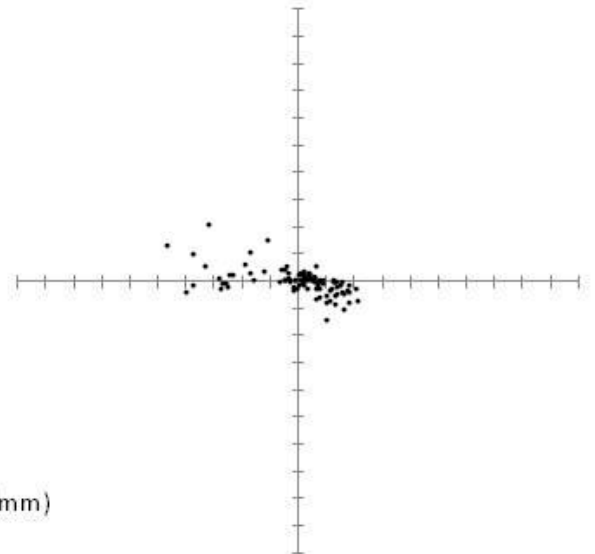


Stomion Superius

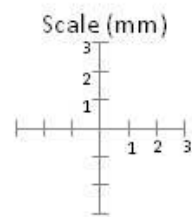
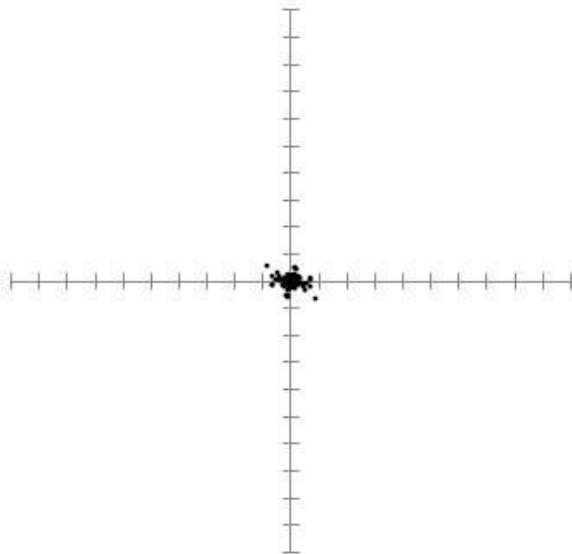
Inter-examiner Frontal



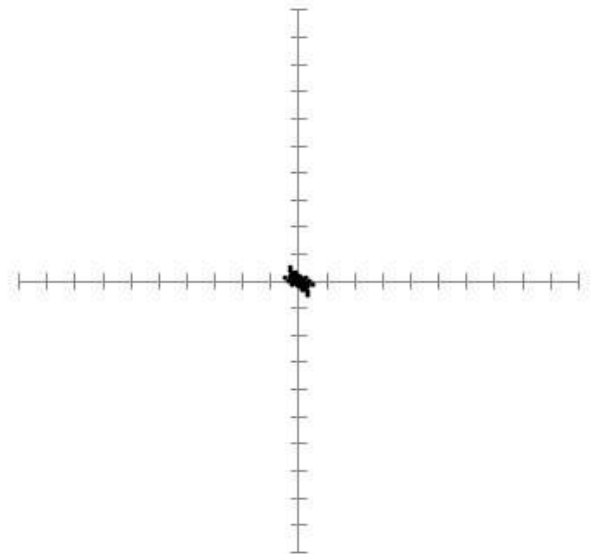
Inter-examiner Sagittal



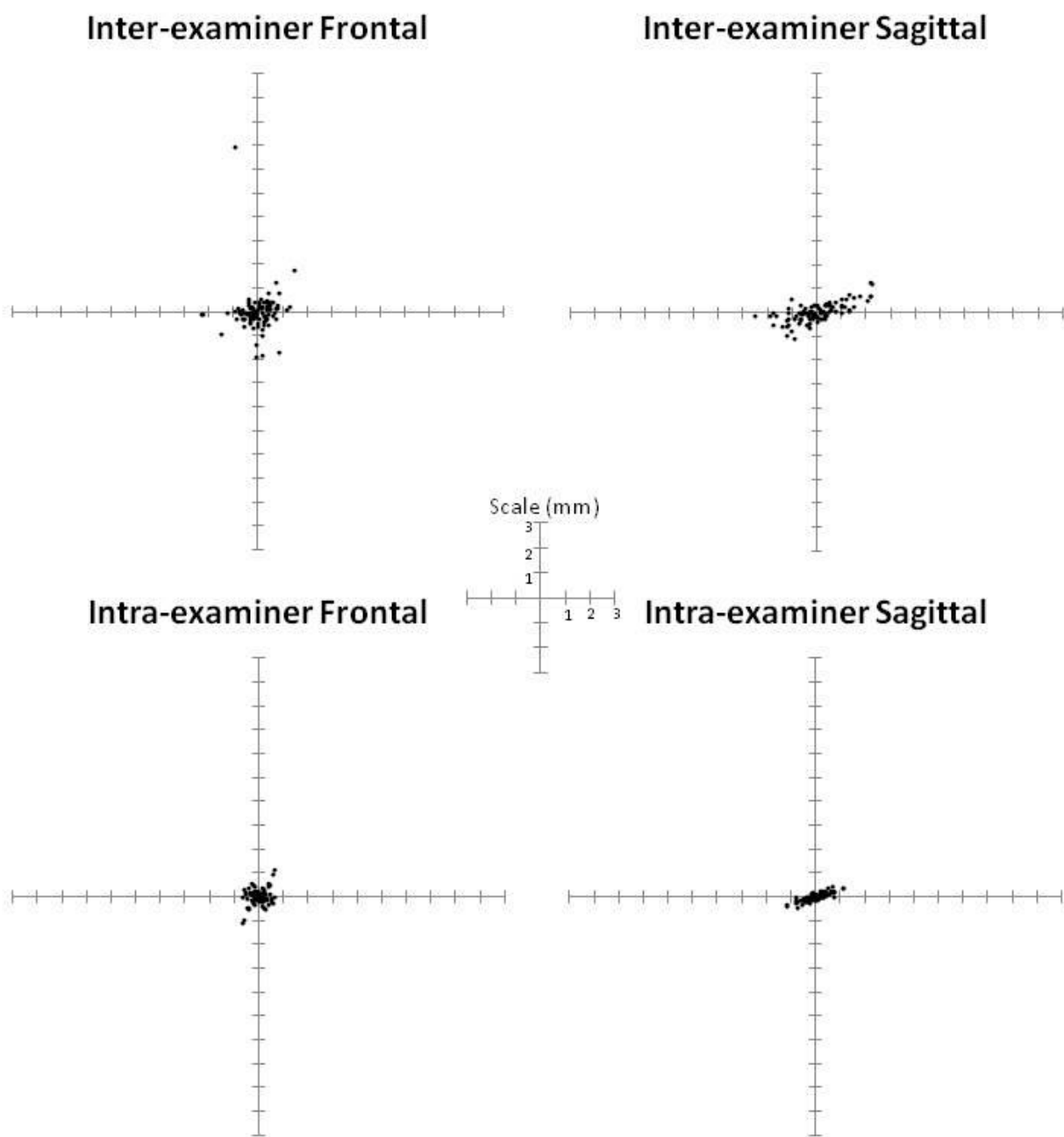
Intra-examiner Frontal



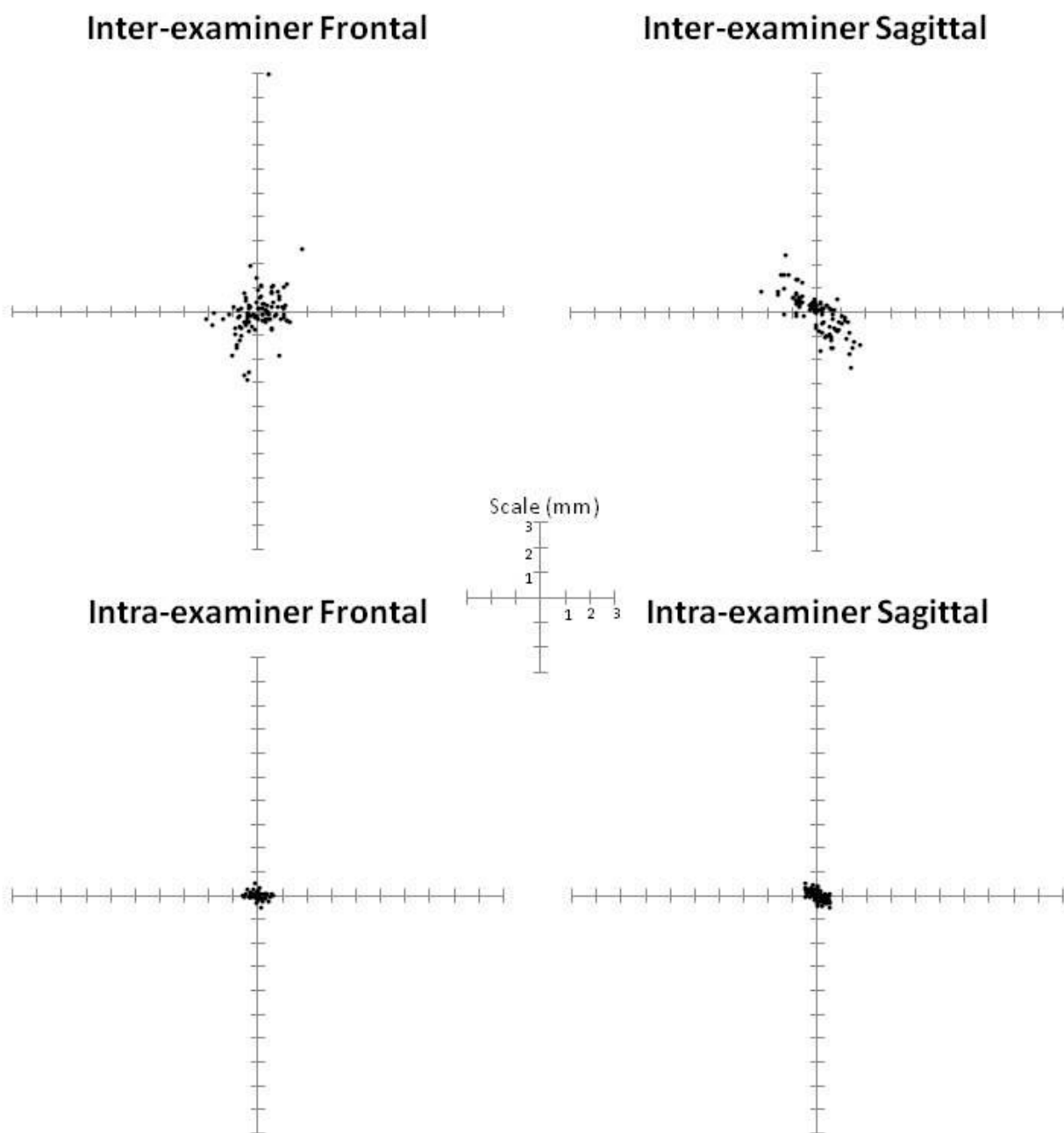
Intra-examiner Sagittal



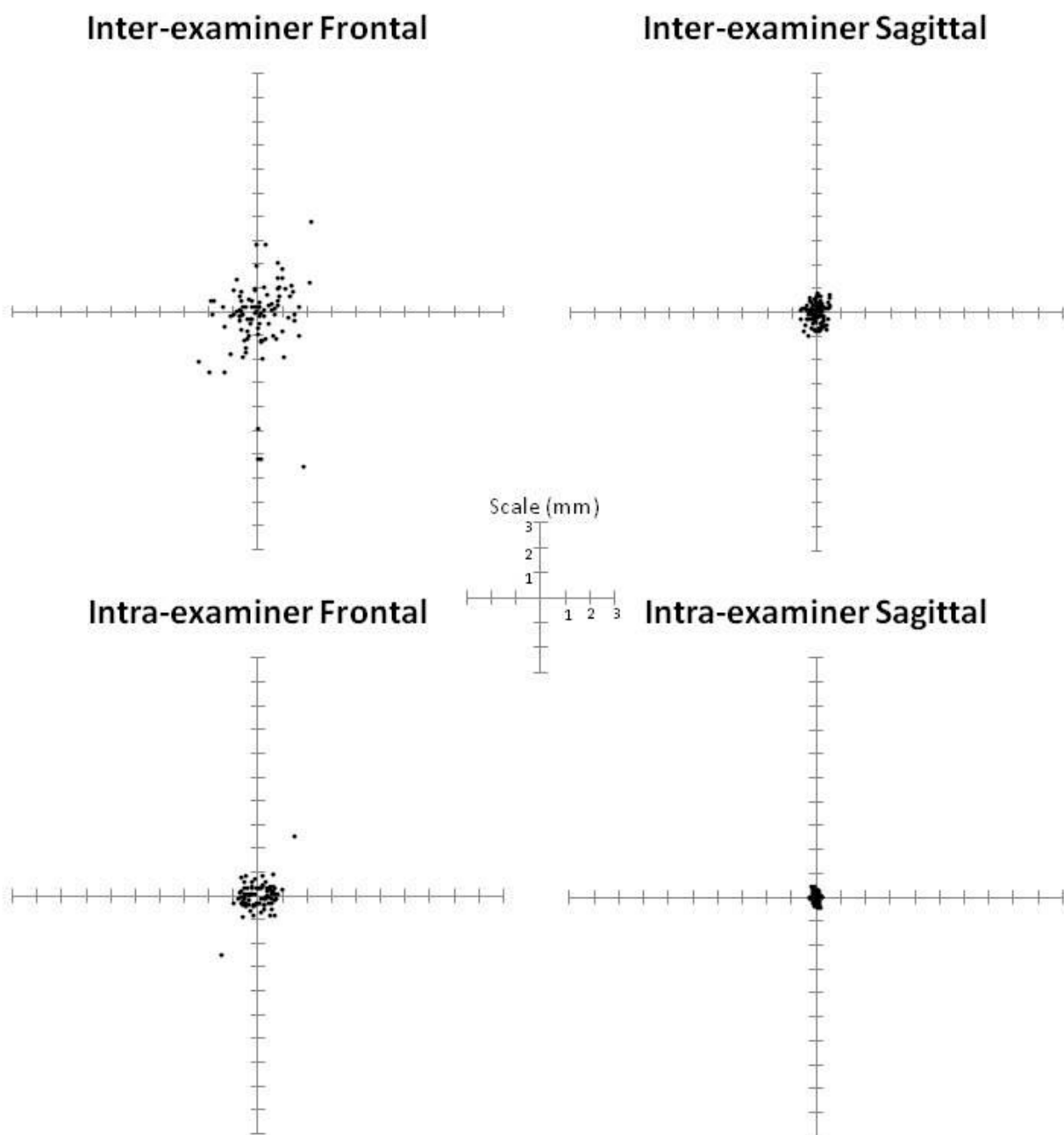
Stomion Inferius



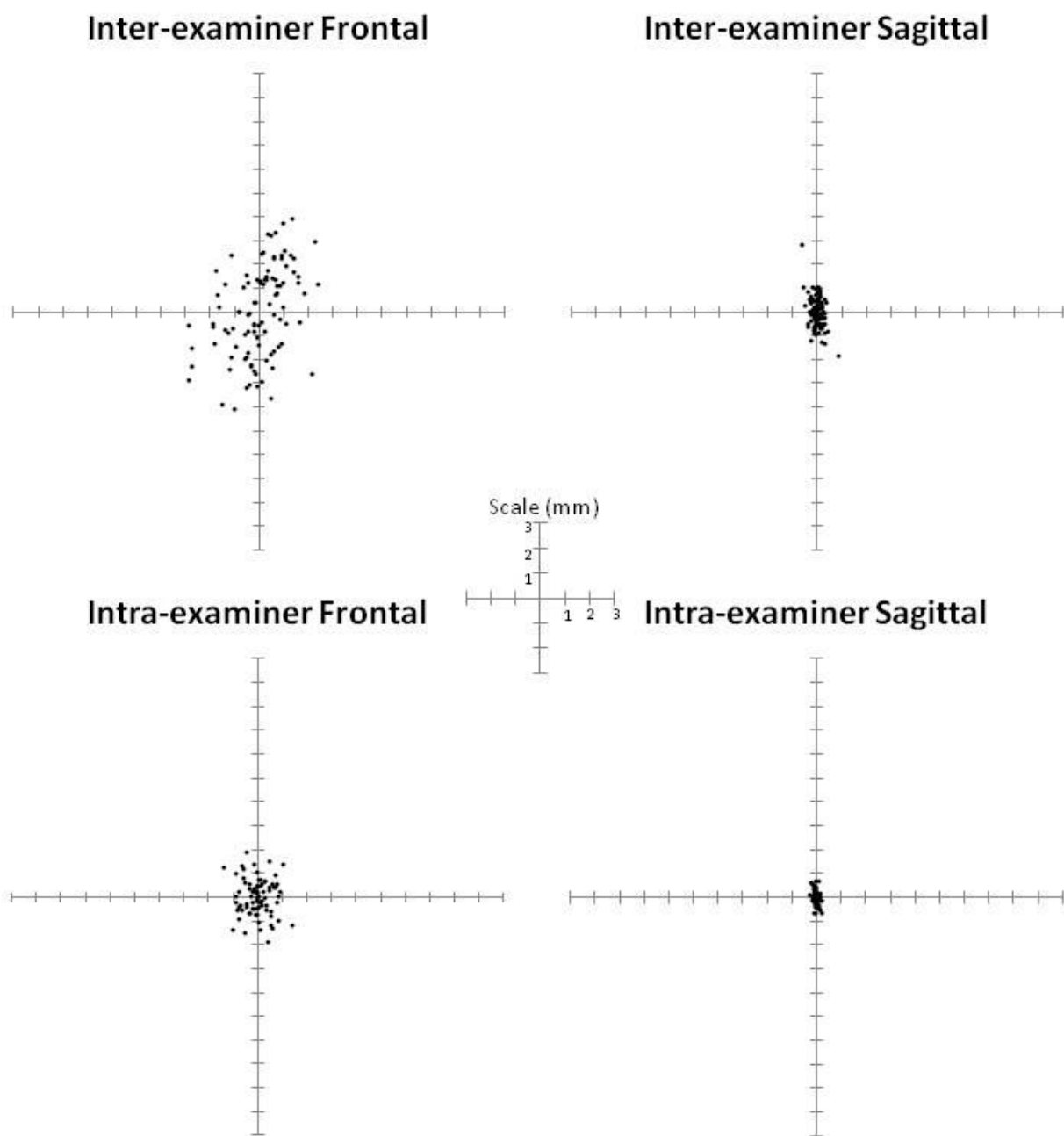
Labrale Inferius



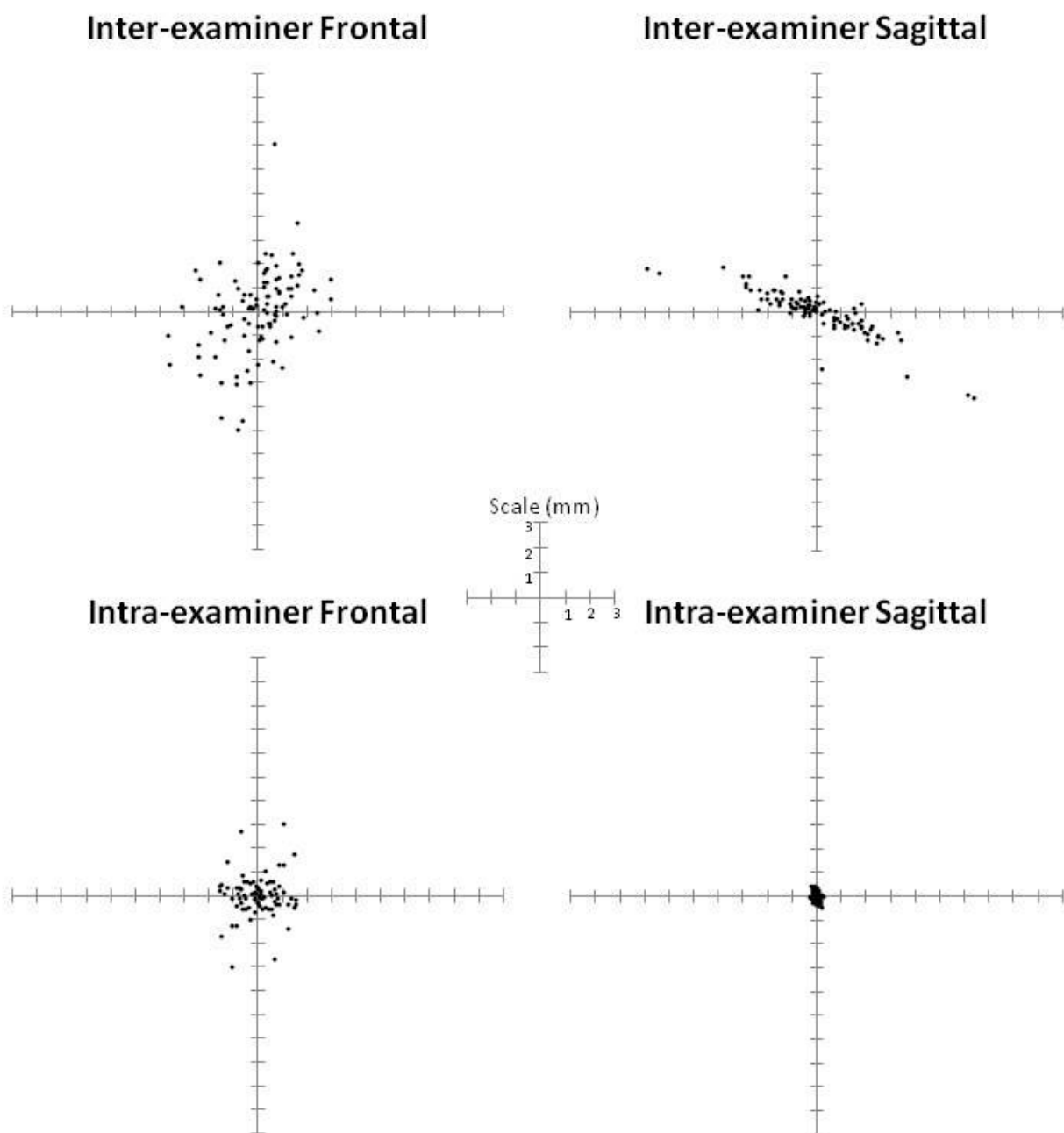
Supramentale



Soft Tissue Pogonion

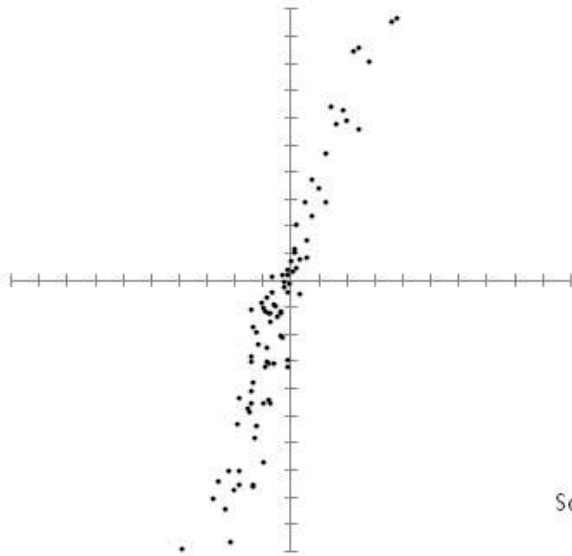


Soft Tissue Menton

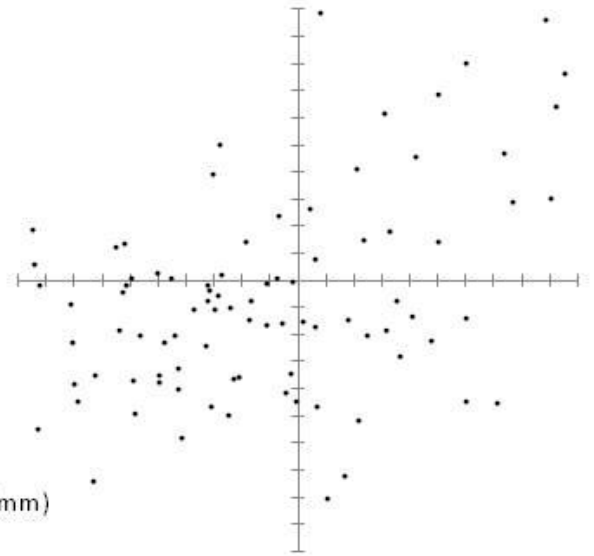


Right Soft Tissue Gonion

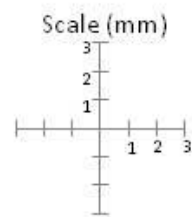
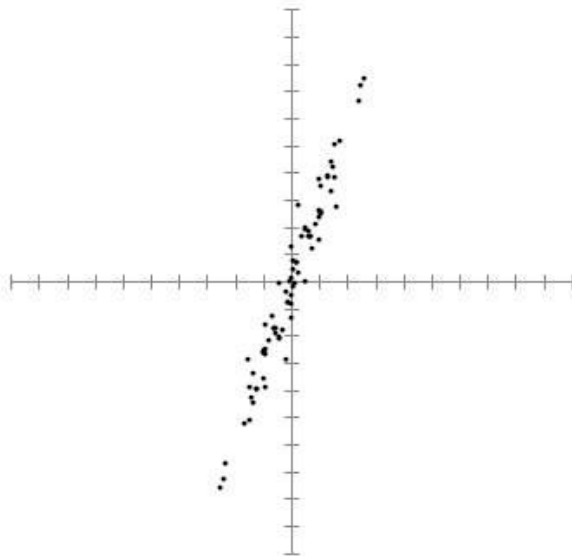
Inter-examiner Frontal



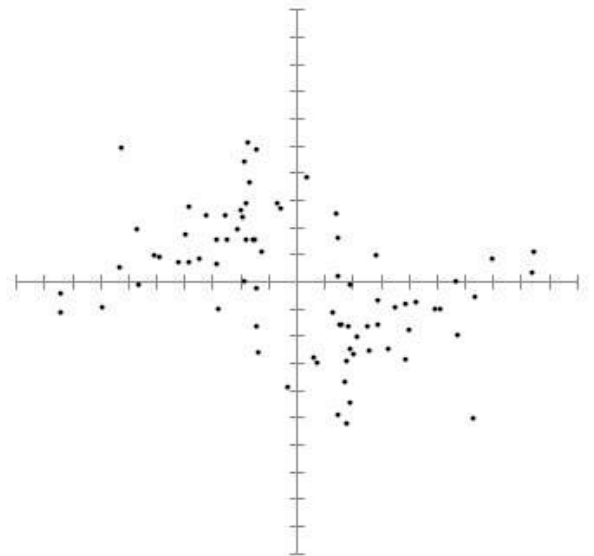
Inter-examiner Sagittal



Intra-examiner Frontal

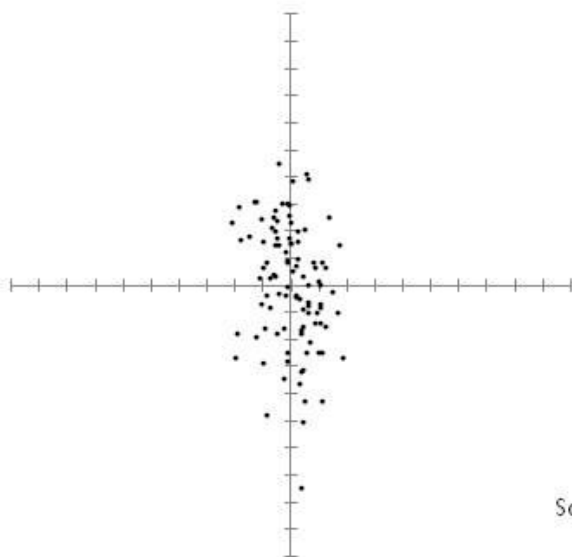


Intra-examiner Sagittal

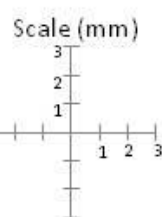
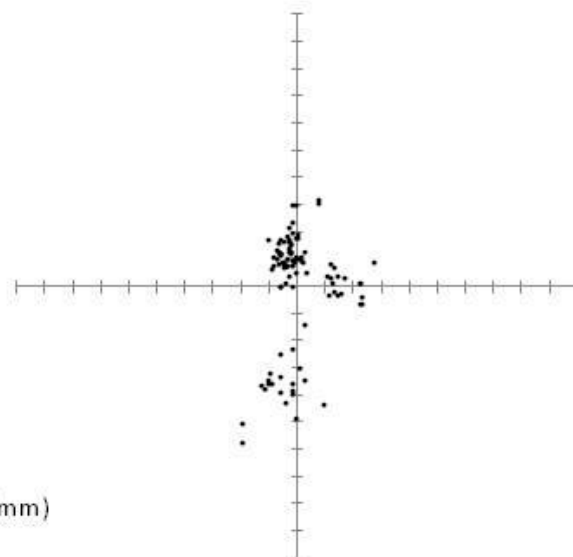


Center of Right Tragus

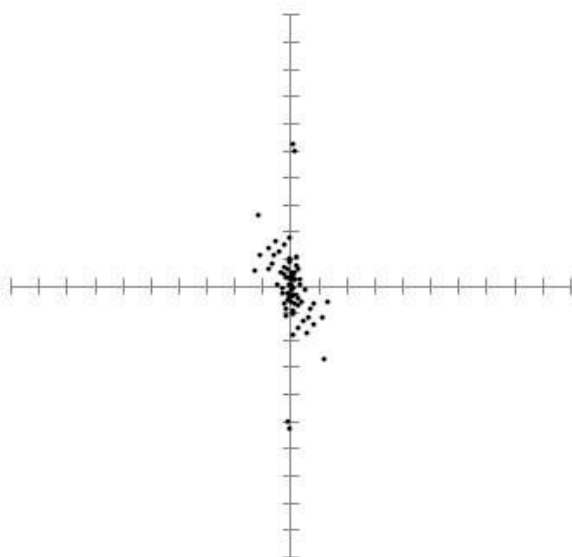
Inter-examiner Frontal



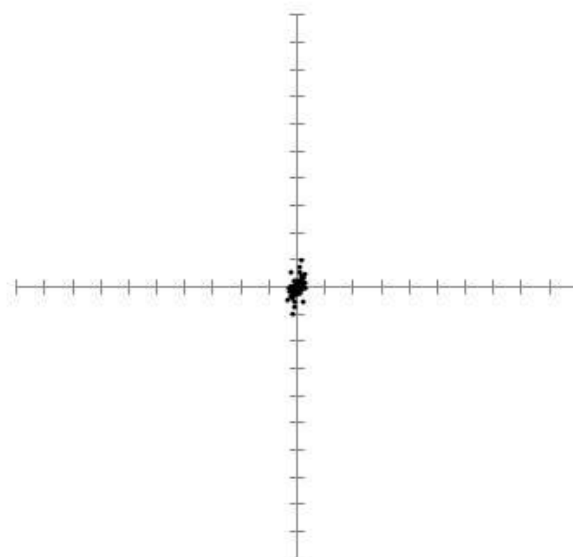
Inter-examiner Sagittal



Intra-examiner Frontal

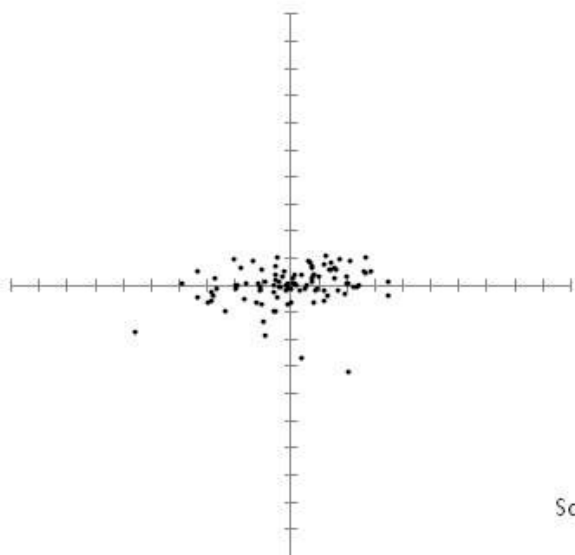


Intra-examiner Sagittal

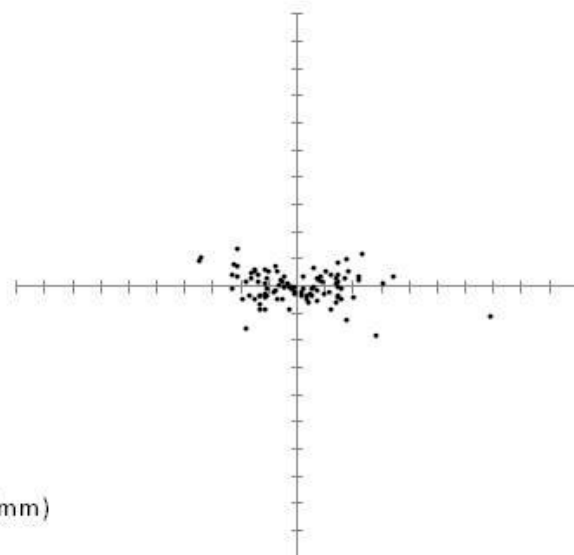


Right Exocanthion

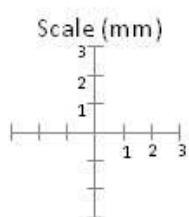
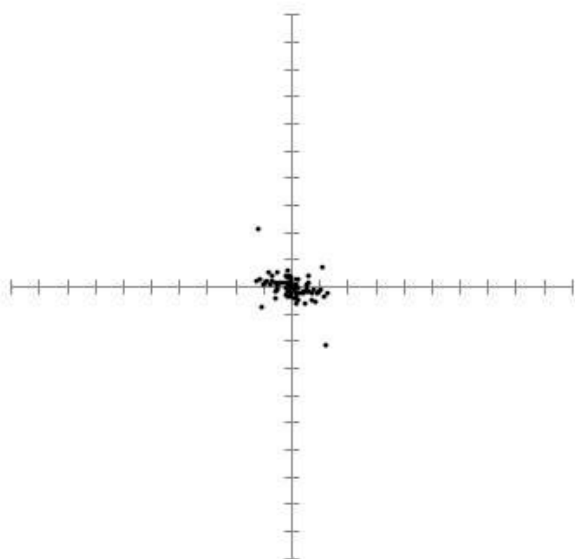
Inter-examiner Frontal



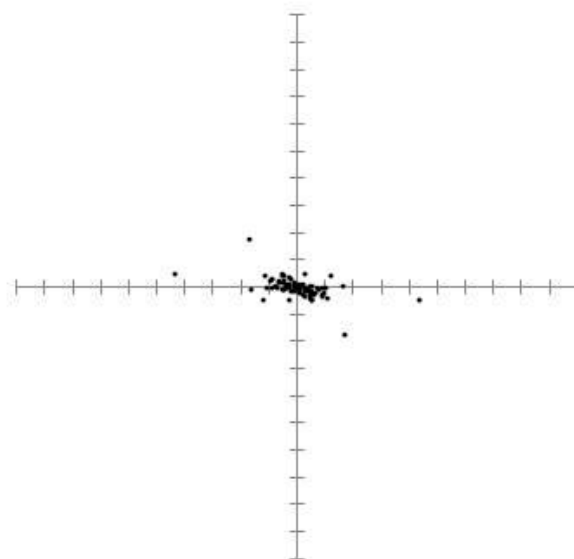
Inter-examiner Sagittal



Intra-examiner Frontal

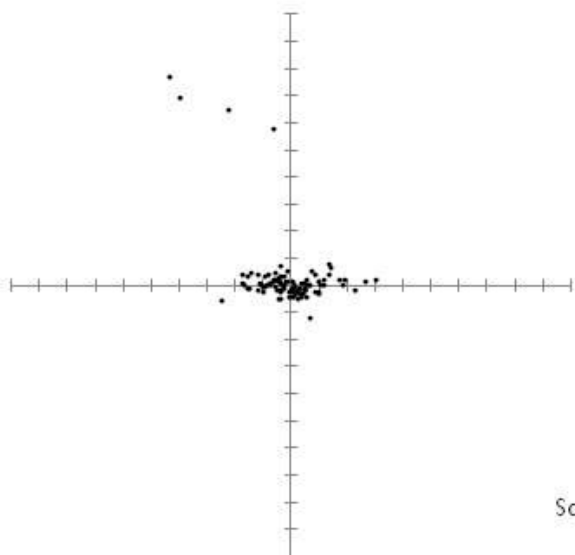


Intra-examiner Sagittal

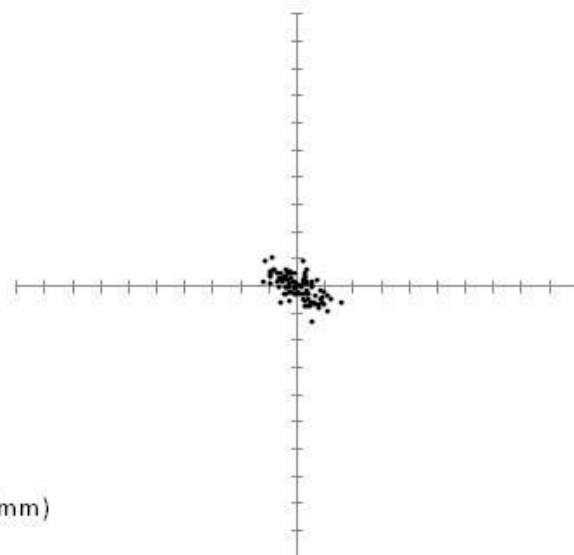


Right Cheilion

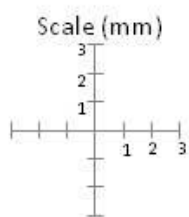
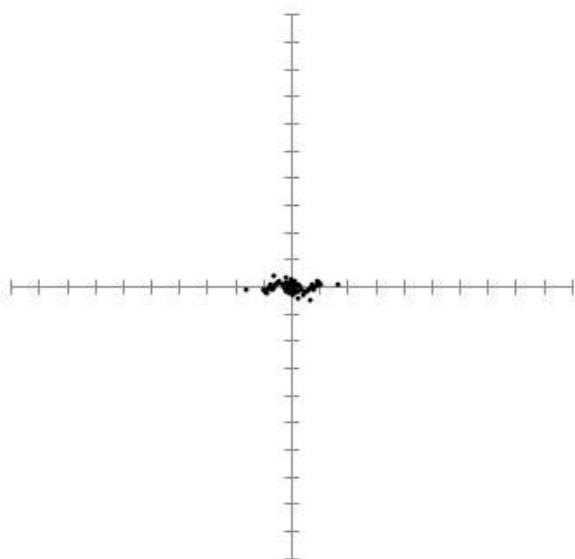
Inter-examiner Frontal



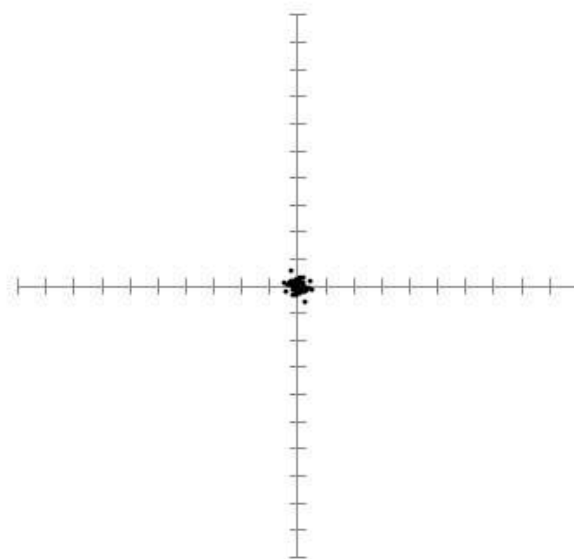
Inter-examiner Sagittal



Intra-examiner Frontal



Intra-examiner Sagittal



Right Alare

