Title
Superimposition accuracy of various three-dimensional craniofacial data sets

Author
Harner, Andrew Thomas

Publication Date
2005

Peer reviewed|Thesis/dissertation

Superimposition Accuracy of Various Three-Dimensional Craniofacial Data Sets

by

Andrew Thomas Harner, D.D.S.

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

ORAL AND CRANIOFACIAL SCIENCES

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

Date

University Librarian

Dedication

I would like to dedicate this thesis to my wife, Marcella. Without her support this would not have been possible. In addition, I would like to thank all of my family, especially my mother Carol, for their support and encouragement.

Acknowledgement

I would like to acknowledge and sincerely thank my thesis committee: Drs. John Huang, Art Miller and Daniel Fried. Without their patience, input, and guidance, this project would not have been possible. I would also like to acknowledge Jeff Leong for his contribution to this work.

ABSTRACT

Advances in digital radiography, computer graphics, and three-dimensional animation have allowed three-dimensional reconstruction of orthodontic records. However, the ability to generate accurate three-dimensional models is critical for these records to be useful diagnostically. In addition, virtual treatment planning and outcome assessment can only be possible with accurate and reliable registration and superimposition techniques. The purpose of this paper was to measure the accuracy of three dimensional models generated from a white light scanning system and a CB MercuRay cone beam computed tomography (CBCT) system. The reliability (precision) of measurements is also assessed. Techniques to accurately and reliably register three-dimensional dental and craniofacial models were developed and tested. Results from this work demonstrated that very repeatable and accurate measurements can be made with the AMIRA software program on both the three-dimensional model generated from the white light scanner, and the CB MercuRay CBCT. There is a clinically insignificant reduction in size of the 3-D models, when compared to the physical gold standard. There was a 0.80% error and a 1.19% error, respectively, when comparing all measurements from the models generated from the white light scanner and the CBCT to their respective physical models. Registration accuracy of white light scanned models increases when increasing the number of tie points utilized. When performing registrations in this manner, the importance of individual tie points to the overall registration process can be assessed by separately eliminating individual tie points. It appears that registration accuracy increases as the tie points used are separated spatially. In other words, a spatial pattern that optimizes the distance between the tie points will produce the most accurate

registration. Operator error is quite low and clinically insignificant with this technique. Registration of 3-D models generated from the CBCT and the white light scanner can be performed. However, conclusive recommendations can not be made in terms of optimizing the registration for accuracy and reliability. Further work needs to be performed in this area. Potential applications of this work include a more accurate assessment of tooth movement in all planes of space, given specific conditions. As a result, computer models to predict force parameters may be generated to optimize specific types of tooth movements. Finally, more accurate virtual treatment outcomes may become available.

Table of Contents

List of Tables

List of Figures

Introduction

Traditional Methods to Assess Orthodontic Progress and Outcomes

Traditional diagnostic records for orthodontic treatment include extra-oral and intra-oral photographs, dental models, and x-rays typically consisting of a full mouth dental series, a panoramic x-ray, and a lateral cephalometric x-ray. Once records are taken, they are assessed, an orthodontic treatment plan is formulated, and treatment begins. After treatment has commenced, it is important for the clinician to evaluate treatment progress. Quite often, additional records are taken to monitor progress, and at the conclusion of treatment, final records are taken to evaluate overall treatment changes. Orthodontic treatment progress and outcomes are assessed by several methods. Comparisons are made between records that were obtained before treatment began and records taken at subsequent time points. In order to compare pre-treatment records to progress or post-treatment records, a stable reference area on the particular record is needed. The stable reference area serves as a location whereby similar records at different time points can be registered and superimposed. In this manner, any skeletal or dental spatial movements can be assessed in relation to this stable area. Specific methods for superimposing pre-treatment and post-treatment records have been developed for traditional orthodontic records and are currently used to quantify dental and skeletal spatial changes during treatment.

Lateral Cephalometric Superimposition

With respect to lateral cephalometry, several methods have been developed for superimposition whereby stable skeletal structures are aligned/registered, and the resulting changes in the maxilla, mandible, and dentition are determined (Broadbent 1937; Steiner 1953; Bjork 1968; Ricketts 1976) The anterior cranial base is commonly utilized to register the cephalometric tracings because its growth is essentially complete by 6 to 7 years of age and therefore is considered a stable structure (Ford 1958). The Bjork method relies upon a best fit superimposition of the anterior wall of sella turcica, the anterior contours of the middle cranial fossa, the contours of the cribiform plate, and the cortical layers of the frontal bone (Bjork 1968). After superimposition on the cranial base of two lateral headfilms taken at different timepoints, differences observed between the dental and skeletal structures can be assumed to be a result of orthodontic treatment or growth.

There are, however, several limitations to this technique. First, only sagittal and vertical changes can be appreciated due to the x-ray/subject orientation. Second, ideal superimpostion on serial lateral radiographs is only possible if the patient is in the exact same position for all lateral radiographs to be superimposed (rotation of the head obscures anatomic features away from the midline). Finally, lateral radiographs may not be taken as readily as wanted due to concerns of patient radiation exposure. As a result, clinicians cannot check treatment progress as often as they might wish with the use of serial headfilms.

Model Analysis and the use of Occlusograms

Dental casts at various stages of treatment are easily obtained. They can be compared, and methods for two-dimensional superimposition in conjunction with lateral cephalometry have been described. Bjork's occlusogram technique utilizes a two-dimensional rendering (typically a photocopy) of the occlusal surface of dental models in conjunction with a lateral cephalometric tracing taken at the same timepoint (Bjork 1968). Using this technique, tooth movements within either the maxilla or mandible are determined. First a superimposition of each jaw (maxilla or mandible) is made. The pre- and post- or progress maxilla is registered on the anterior wall of the zygomatic arch. Growth of the maxillary sutures and remodeling of the nasal floor is accounted for by positioning the two overlying headfilms such that the floor of the orbit raises 1.5 times as much as the palatal plane drops during growth. The mandibular superimpositions are made by registering the different timepoints on a best fit of the symphysis and the mental nerve. Then, the two-dimensional occlusograms are registered to the corresponding jaw superimposition. This is performed by aligning the labial surface of the anterior teeth from the 2-dimensional occlusogram to the labial surface of the anterior teeth from the lateral headfilm. Since this technique requires the use of a lateral headfilm, x-ray exposure is necessary, and the inherent errors and limitations of lateral cephalometric superimposition apply to this technique as well.

To avoid the necessity of x-ray exposure as a way to register the teeth within the jaws, other investigators have used palatal rugae as a means to discern what tooth movements have occurred during treatment (Almeida, Phillips et al. 1995; Bailey,

Esmailnejad et al. 1996; Abdel-Aziz and Sabet 2001; Hoggan and Sadowsky 2001; Ashmore, Kurland et al. 2002). Palatal rugae are transverse thickenings that occur on the surface of the hard palate. They are comprised of thick dense collagenous tissue covered with a thick orthoketatinized and often parakeratinized stratified squamous epithelium (TenCate 1998). Although each human has a very unique palatal rugae pattern, all humans have palatal rugae consisting of 3 distinct transverse thickenings on each side of the midpalatal raphe (English, Robison et al. 1988). The end points of these thickenings are convenient areas to make measurements from (Figure 1).

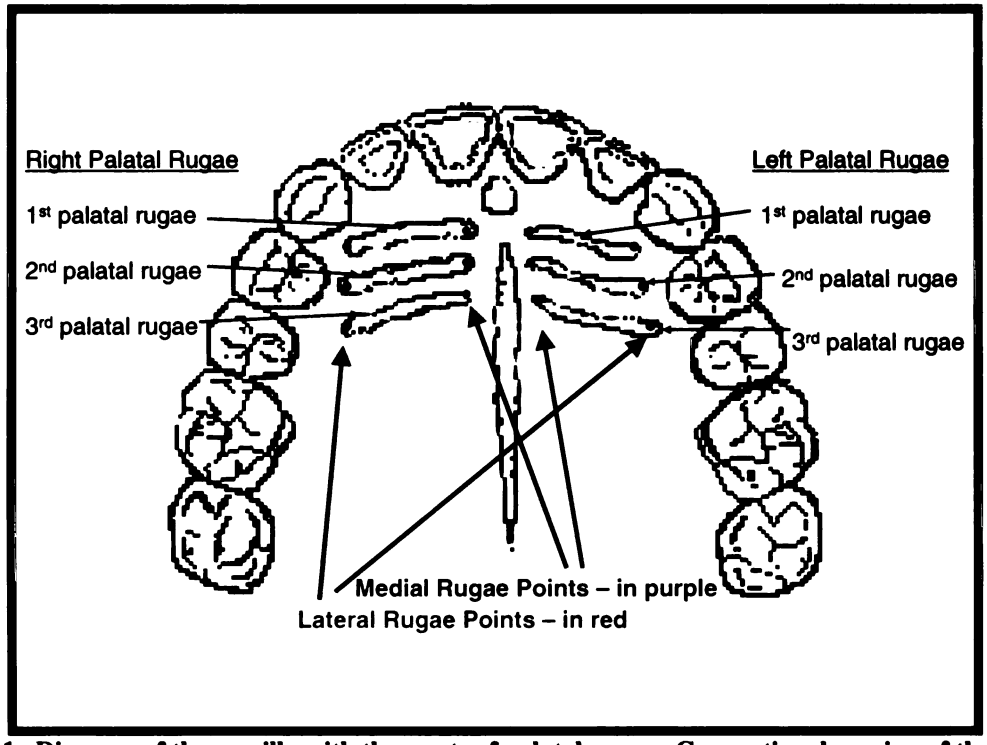


Figure 1. Diagram of the maxilla with three sets of palatal rugae. Conventional naming of the palatal rugae is shown. The medial and lateral end points of all the palatal rugae are convenient points from which measurements can be made.

During growth, palatal rugae are very stable. Results from the University of Michigan Growth Study demonstrate this fact (Moyers 1976). Subjects in this study had impressions taken yearly during their childhood until growth had ceased or orthodontic

treatment began. Ages of the subjects ranged from 3 years to 18 years old. From the annual study models that were obtained, palatal width, palatal depth, and palatal height measurements were recorded. Essentially these three measurements assessed the transverse change, the antero-posterior change, and the vertical change, respectively, over the study period. All transverse and antero-posterior measurements were linear measurements from various palatal rugae points to palatal rugae points. Vertical measurements were linear measurements from a single palatal rugae point to the functional occlusal plane. Transversely (palatal width), and antero-posteriorly (palatal depth) the palatal rugae points are remarkable stable, with virtually no change. As a result of the anatomy of the palate and the location of the palatal rugae points, one could argue that the palatal rugae are stable in the x, y and z dimension because both the transverse and the antero-posterior measurements invariably have a vertical component to them (see Figure 2).

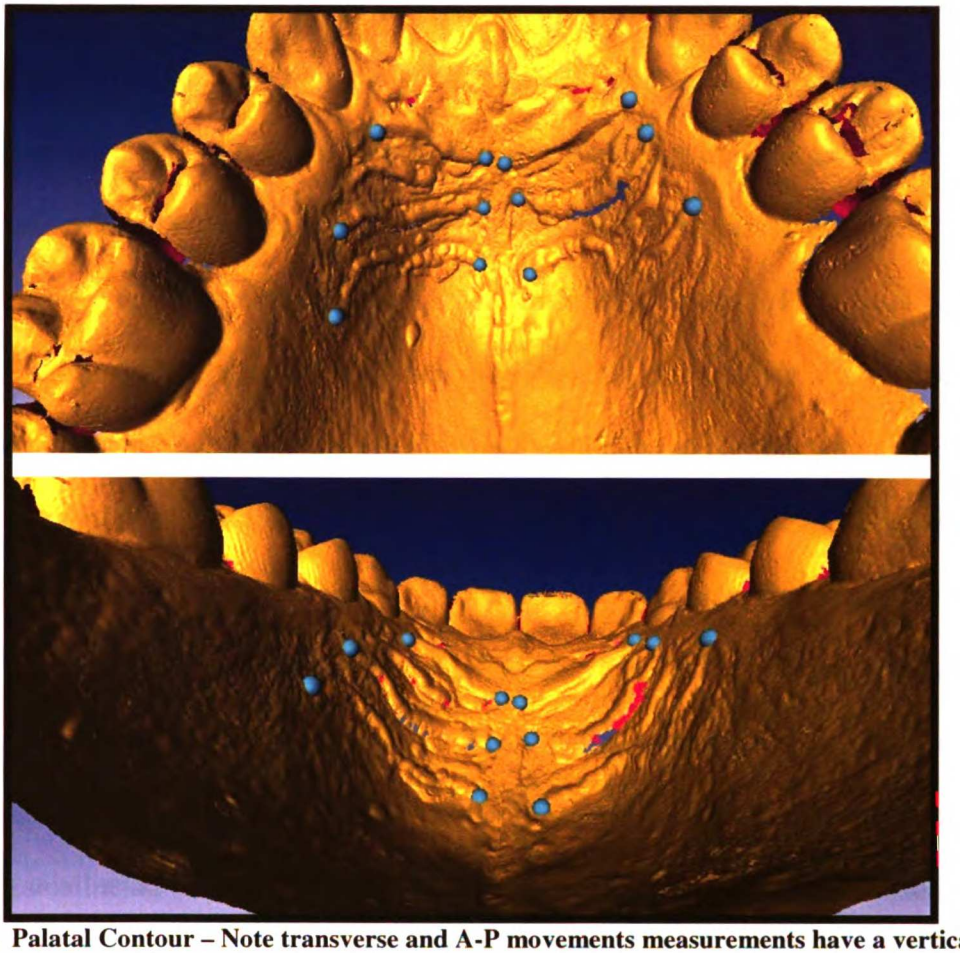


Figure 2. Palatal Contour – Note transverse and A-P movements measurements have a vertical component to them.

In relation to the palatal height, there were slight changes over time. However this could be a result of the method chosen to test this variable. Since vertical measurements were based on the functional occlusal plane, tooth eruption would affect the outcome. With tooth eruption, as one would expect in a growing child, this measurement would invariably increase.

Palatal rugae are relatively stable during orthodontic treatment. (Hoggan and Sadowsky 2001). When comparing molar and incisor movements produced orthodontically, no statistical difference was found when evaluated by lateral headfilms

as compared to dental models using palatal rugae as superimposition landmarks. This study suggests that palatal rugae are at least as reliable as anterior cranial base registrations to evaluate tooth movement. Other investigators have looked at the relative stability of palatal rugae during orthodontic treatment by assessing their relative movement as referenced to their pre-treatment position (Almeida, Phillips et al. 1995; Bailey, Esmailnejad et al. 1996). When investigating the effects of headgear and functional appliance treatment on palatal rugae, Almeida et al. found that the medial rugae are relatively stable and can be utilized as references for longitudinal dental model analysis. The lateral aspects of the rugae, however, did show significant movement especially in the headgear group (Almeida, Phillips et al. 1995). When evaluating the difference in rugae movement between extraction and non-extraction groups, found that the amount of tooth movement affected the overall movement of the palatal rugae (Bailey, Esmailnejad et al. 1996). However, no significant movement was found in either the extraction or non-extraction group on the medial and lateral aspect of the third rugae, and therefore these areas could be used as a stable reference structure.

Emerging Methods to Assess Orthodontic Progress and Outcomes – 3-D Imaging Techniques in Orthodontics

With the emergence of new three-dimensional imaging systems, the utilization of two-dimensional records for orthodontics will surely decrease. This transition, however, will only occur if the new three-dimensional methods are as accurate as or better than the traditional methods. Advances in digital radiography, computer graphics, and three-dimensional animation have paved the way for three-dimensional reconstruction of

orthodontic records (Mah and Hatcher 2003). Cone-beam computed tomography is well-suited for use in maxillofacial imaging, and three-dimensional volumetric reconstruction is possible with appropriate software (Sukovic 2003). Moreover, surface light and laser scanners allow the acquisition of extremely accurate three-dimensional computer models of the dentition, the face, and other three dimensional objects. As a result, surface light and laser scanner may be used instead of a two-dimensional photocopy of the occlusal surfaces of dental models for post-treatment occlusal analysis.

Surface Light Scanners

Computer models of the dentition are generated by scanning the teeth directly or by scanning a dental model with a surface laser scanner or surface light scanner. These scanners capture the surface topography of the object and generate a 3-D image using mathematical triangulation algorithms. Essentially a light is shined on the object, and two video sensors capture the object from two different orientations. The scanning process results in the acquisition of a collection of digitized points with each point represented by an x, y, and z coordinate. The collection of digitized points is often referred to as a point cloud. In order to create a surface model from the point cloud, 3-D modeling software is necessary.

Surface scanners are already being commercially utilized as a replacement for plaster orthodontic models. Geodigm Corporation's e-model service (Chanhassen, MN) utilizes laser scanning to produce high resolution three-dimensional computer models of the patient's dentition. Orametrix (Richardson, TX) utilizes a handheld intraoral surface

scanner to scan the patient's teeth directly. When used as a replacement for plaster models, computer models eliminate the need for extensive physical storage space. In addition to alleviating storage problems, the use of computer models of the dentition affords new and potentially more accurate methods of treatment analysis (Ayoub, Wray et al. 1997; Bell, Ayoub et al. 2003; Santoro, Galkin et al. 2003; Quimby, Vig et al. 2004). In addition, other investigators have already reported on combining data from surface scanners for use in accuracy studies of impression materials (Shah, Sundaram et al. 2004), facial features (Lee, Han et al. 2004; Littlefield, Cherney et al. 2005).

Cone Beam Computed Tomography

Cone Beam Computed Tomography (CBCT) is a relatively new imaging modality that is uniquely suited for craniofacial and dental imaging (Sukovic 2003). Similar to conventional CT, CBCT utilizes an x-ray source that moves in a circular pattern around the subject. The only difference in regard to the x-ray pattern is that CBCT has a cone-shaped beam whereas the conventional CT has a flat fan shaped beam (Sukovic 2003; Araki, Maki et al. 2004). As the x-rays pass through the subject, the rays are differentially attenuated based on the subject's anatomy. Images are acquired when the attenuated x-rays strike the CBCT's detector. 3-D image reconstruction is possible due to the changing orientation of the x-ray source during the CBCT scan and the complex algorithms used to calculate the spatial changes of the x-ray unit in relation to the stationary subject. The CBCT units are smaller, are a fraction of the cost, and have significantly less radiation exposure per scan than conventional CT systems (Mozzo, Procacci et al. 1998; Vannier 2003). In addition, investigators report that the CBCT

system produce accurate representations of the scanned subject (Mozzo, Procacci et al. 1998; Lascala, Panella et al. 2004). All of these features of the CBCT lend themselves well for use in dental offices.

Purpose

The purpose of this work is to develop accurate and reliable methods of three-dimensional superimposition/registration techniques. The accuracy by which the UCSF three-dimensional acquisition systems represent the scanned object will be tested as well as the accuracy and reliability of registering these three-dimensional data sets.

Materials and Methods

Comparison of the Accuracy and Reliability of Measurements made with a Three-Dimensional Surface Scanner to those made with Digital Calipers.

Three-dimensional renderings of three objects (listed below and shown in Figures 3 and 4) were created. Objects measured include:

a. Phantom created with the maxilla of a plastic skull with a 10 mm ± 1 micron diameter tungsten carbide sphere affixed to the palate.

b. Phantom created with the mandible of a plastic skull with two 10 mm ± 1 micron diameter tungsten carbide spheres affixed distal to the 3rd molars on the alveolar ridge.

c. Phantom of a human skull mandible with various fiducials (stainless steel ball bearings) measuring 1.58 mm placed at specific anatomical landmarks (see table for location of the various fiducials).

Each object was mounted on the robotic stage of the UCSF white light scanner system donated by Align Technology (Santa Clara, CA). This particular piece of equipment consists of a Steinbichler Type C75 White Light Scanner (Neubeuern, Germany) oriented to scan objects that are placed on a stage controlled by a Newport Electro-Optics robotic servo system (Irvine, California). In order to acquire detailed surface architecture of the object in question, 38 individual scans were performed on each object at various orientations. Each individual scan generated a specific point cloud which, when combined with all the other scans of the object, created a complete three-dimensional computer model of the scanned object. Registered point clouds were then converted to a polygonal surface using Polyworks software (Innovmetric, Sainte-Foy, Quebec Canada). Once a surface rendering (.stl file) was created, linear measurements were performed utilizing AMIRA software (Mercury Computer Systems, Richardson, Texas).

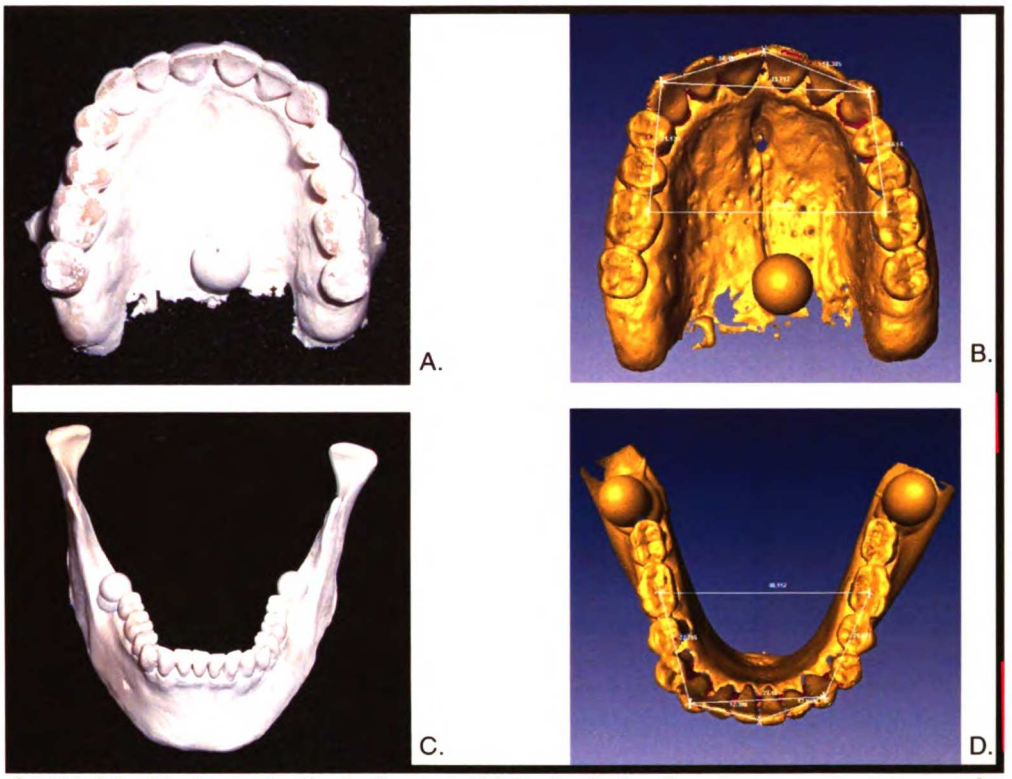


Figure 3. (A). Maxilla from plastic skull upon which physical caliper measurements were taken. (B). 3-D surface model of the maxilla in 1A upon which digital measurements were taken. (C). Mandible from a plastic skull upon which physical caliper measurement were taken. (D). 3-D surface model of the mandible in 1C upon which digital measurements were taken.

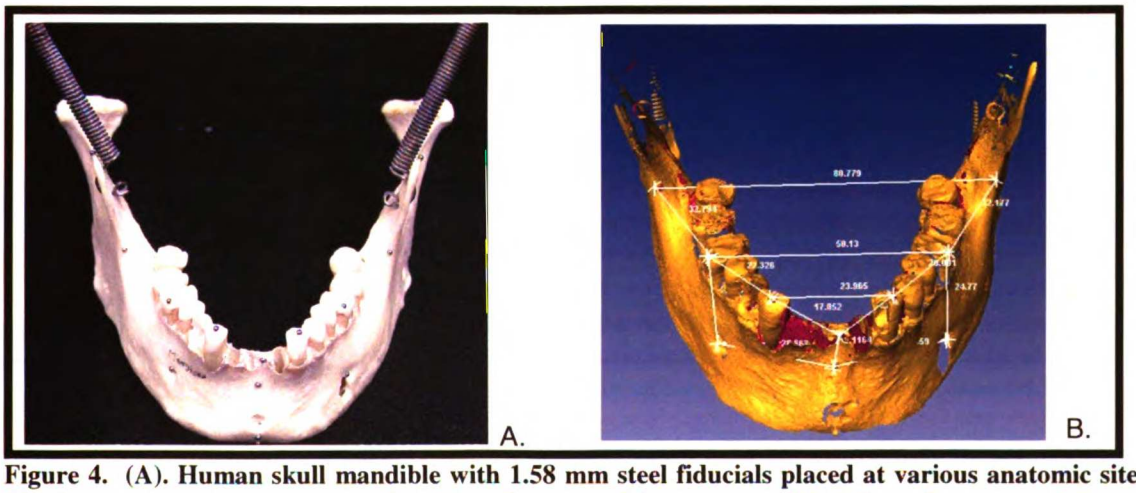


Figure 4. (A). Human skull mandible with 1.58 mm steel fiducials placed at various anatomic sites. Physical measurements were made with calipers on this specimen. (B). 3-D surface generated from the human mandible in 2A. Digital measurements as shown were made from this 3-D surface model.

In addition, physical measurements on the objects themselves were performed with a calibrated digital caliper (Mitutoyo Digimatic Calipers Model No CD-6"CS, Aurora, Illinois).

Measurements made are shown in Figures 3B, 3D, and 4B and listed in Tables 1.

Table 1. Locations of Linear Measurements (Part 1)

Sample A - Plastic Maxilla Phantom	
Cusp tip of #6 to the cusp tip of #11	U3-U3
Mesiolingual cusp tip of #3 to the mesiolingual cusp tip of #14.	U6-U6
Mesiolingual cusp tip of #3 to the cusp tip of #6.	U3-U6 R
Cusp tip of #6 to the incisive embrasure between #8 and #9.	U1-U3 R
Incisive embrasure between #8 and #9 and the cusp tip of #11	U1-U3 L
Cusp tip of #11 to the mesiolingual cusp tip of #14	U3-U6 L
Sample B - Plastic Mandible Phantom	
Cusp tip of #22 to the cusp tip of #27	L3-L3
Central fossa of #19 to the central fossa of #30	L6-L6
Central fossa of #19 to the cusp tip of #22.	L3-L6 L
Cusp tip of #22 to the incisive embrasure between #24 and #25	L1-L3 L
Incisive embrasure between #24 and #25 to the cusp tip of #27	L1-L3 R
Cusp tip of #27 to the central fossa of #30	L3-L6 R
Sample C – Human Skull Mandible Right	Right
Mandibular Canine Cusp Tip - Mandibular 1st Molar Cusp Tip	L3-L6R
Mandibular 1st Molar Cusp Tip - Mandibular Point R1	L6-R1R
Infradentale - Mandibular Canine Cusp Tip	Id-L3R
Mental Foramen - Mandibular 1st Molar Cusp Tip	Mf-L6R
Mental Foramen - B point	Mf-BR
Left	Left
Mandibular Canine Cusp Tip - Mandibular 1st Molar Cusp Tip	L3-L6L
Mandibular 1st Molar Cusp Tip - Mandibular Point R1	L6-R1L
Infradentale - Mandibular Canine Cusp Tip	Id-L3L
Mental Foramen - Mandibular 1st Molar Cusp Tip	Mf-L6L
Mental Foramen - B point	Mf-BL
Right-to-Left	Right-to-Left
Mandibular Canine Cusp Tip - Mandibular Canine Cusp Tip	L3
Mandibular 1st Molar Cusp Tip - Mandibular 1st Molar Cusp Tip	L6
Mandibular Point R1- Mandibular Point R1	R1
Mental Foramen - Mental Foramen	Mf
Midline	Midline
B point - Infradentale	B-Id

Each of the above measurements was repeated five times by the same investigator (AH). To assess intra-examiner reliability, all measurements were repeated exactly 1 week after the initial series of 5 measurements were taken. Means and standard

deviations were determined for each series of measurements. To compare the digital measurements to the caliper measurements, paired t-tests were performed with a $p < 0.05$. A regression plot was generated to assess the correlation between the physical and digital measurements. Intra-examiner reliability was determined with Lin's concordance and a Pearson-Product Correlation Coefficient.

Comparison of the Accuracy and Reliability of Measurements made with a Cone Beam CT to those made with Digital Calipers.

A human skull with stainless steel ball bearing fiducials placed at various anatomical locations (Figure 5a) was created. Each of the steel fiducials measured 1.58 mm in diameter. A cone beam CT scan was performed on the phantom with a CB MercuRay™ cone beam CT unit (Hitachi Medical Corporation - Tokyo, Japan) using the following parameters: 100 kVp and 15 mA. Once the scan was completed, a surface model was created by CBworks software (CyberMed, Seoul, Korea) and three-dimensional measurements were made with the AMIRA software. Measurements made are shown in Figures 5B and 5D and listed in Table 2.

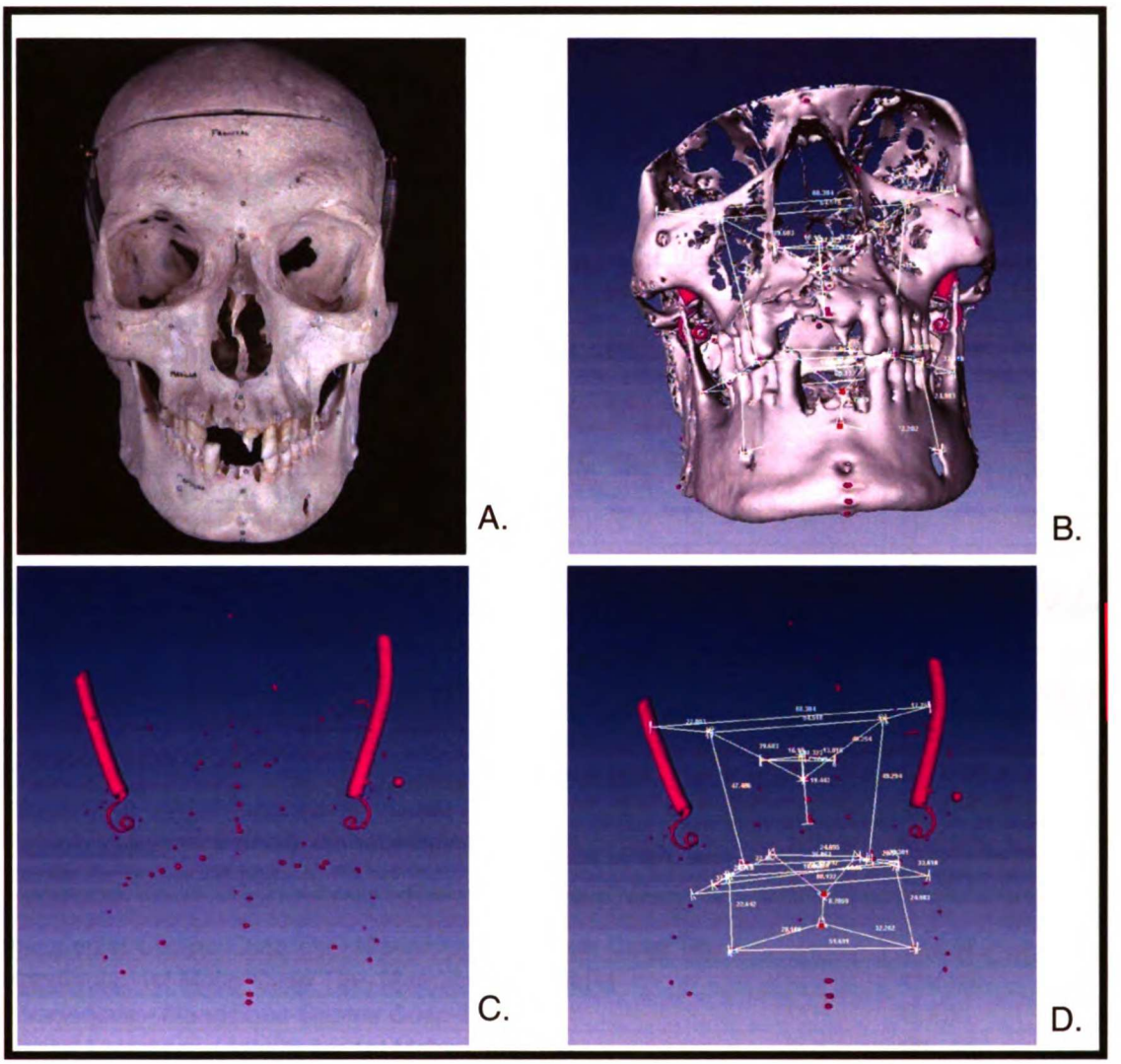


Figure 5. (A). Human skull with 1.58mm steel fiducials placed at various anatomic sites for measurement purposes. Physical measurements were taken with a digital caliper on this specimen. (B). 3-D surface of the skull created after CBCT imaging. (C). In addition to the 3-D surface of the skull, a 3-D surface model of the steel fiducials (and springs attached to the skulls mandible) was created by changing the threshold values of the CBCT scan. (D). From the 3-D model of just the steel fiducials, digital 3-D measurements were taken.

Table 2 Locations of Linear Measurements on Human Skull Phantom (Part II)

Right	Right
Orbitale - Infraorbital Foramen	Or-IOFR
Infraorbital Foramen - Maxillary Canine Cusp Tip	IOF-U3R
Piriform Rim - Anterior Nasal Spine	PR-ANSR
Infraorbital Foramen - A point	IOF-AR
Left	Left
Maxillary Canine Cusp Tip - Maxillary 1st molar ML Cusp Tip	U3-U6L
Orbitale - Infraorbital Foramen	Or-IOFL
Infraorbital Foramen - Maxillary Canine Cusp Tip	IOF-U3L
Piriform Rim - Anterior Nasal Spine	PR-ANSL
Infraorbital Foramen - A point	IOF-A (IS)L
Right-to-Left	Right-to-Left
R Maxillary Canine Cusp Tip - L Maxillary Canine Cusp Tip	U3
Piriform Rim - Piriform Rim	PR
Orbitale - Orbitale	Or
Infraorbital Foramen - Infraorbital Foramen	IOF
Midline	Midline
Anterior Nasal Spine - A point	ANS-A
Anterior Nasal Spine - Prostheon	ANS-Pr
Right	Right
Mandibular Canine Cusp Tip - Mandibular 1st Molar Cusp Tip	L3-L6R
Mandibular 1st Molar Cusp Tip - Mandibular Point R1	L6-R1R
Infradentale - Mandibular Canine Cusp Tip	Id-L3R
Mental Foramen - Mandibular 1st Molar Cusp Tip	Mf-L6R
Mental Foramen - B point	Mf-BR
Left	Left
Mandibular Canine Cusp Tip - Mandibular 1st Molar Cusp Tip	L3-L6L
Mandibular 1st Molar Cusp Tip - Mandibular Point R1	L6-R1L
Infradentale - Mandibular Canine Cusp Tip	Id-L3L
Mental Foramen - Mandibular 1st Molar Cusp Tip	Mf-L6L
Mental Foramen - B point	Mf-BL
Right-to-Left	Right-to-Left
Mandibular Canine Cusp Tip - Mandibular Canine Cusp Tip	L3
Mandibular 1st Molar Cusp Tip - Mandibular 1st Molar Cusp Tip	L6
Mandibular Point R1- Mandibular Point R1	R1
Mental Foramen - Mental Foramen	Mf
Midline	Midline
B point - Infradentale	B-Id

In addition to the three-dimensional measurements, physical measurements were performed on the skull itself with digital calipers (Mitutoyo Digimatic Calipers; CD-6"CS, Aurora, Illinois). Each of above measurements was repeated five times by the same investigator (AH). To assess intra-examiner reliability, all measurements were repeated exactly one week after the initial series of 5 measurements were taken. Means and standard deviations were determined for each series of measurements. To compare the digital measurements to the caliper measurements, paired t-tests were performed with a $p < 0.05$. A regression plot was generated to assess the correlation between the physical and digital measurements. Intra-examiner reliability was determined with Lin's concordance and a Pearson-Product Correlation Coefficient.

Comparison of Various Methods to Register Three-Dimensional Computer Generated Dental Models

In order to develop accurate and reproducible methods to register three-dimensional computer generated dental models, several techniques were employed. A maxillary stone dental model was scanned with the white light scanner and processed as described in part 1 of the study. From this procedure a three-dimensional computer dental model was generated. Two identical models were loaded into the AMIRA software program, and various registration techniques were compared and optimized as described below (Figure 6).

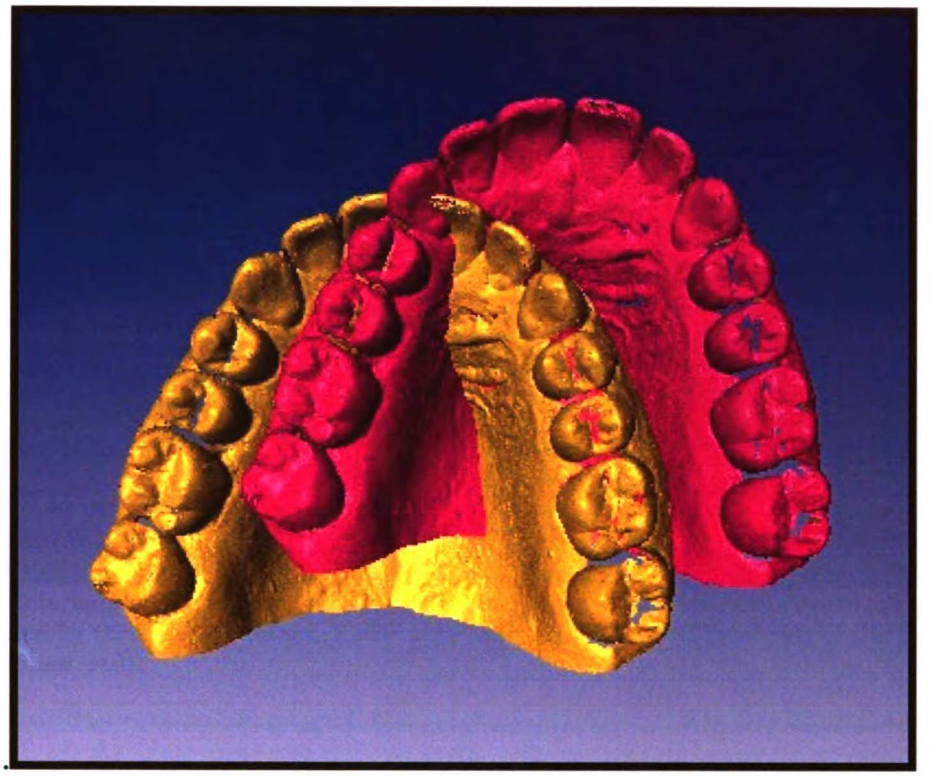


Figure 6. Two identical 3-D surface models before superimposition techniques were employed.

Computer Generated Automatic Alignment

An automatic alignment algorithm was utilized initially. This process allows the AMIRA software to automatically align the two identical model surfaces. The algorithm minimizes the root mean square distance between the vertices of one model and the closest surface on other model. This is also known as the Procrustes method. Results were assessed by calculating the mean surface distance between all the corresponding points of the two models after alignment. In addition, color coded images (colormaps) of the surfaces were generated to illustrate the distance between the two models in various region(s) of the models (Figure 7).

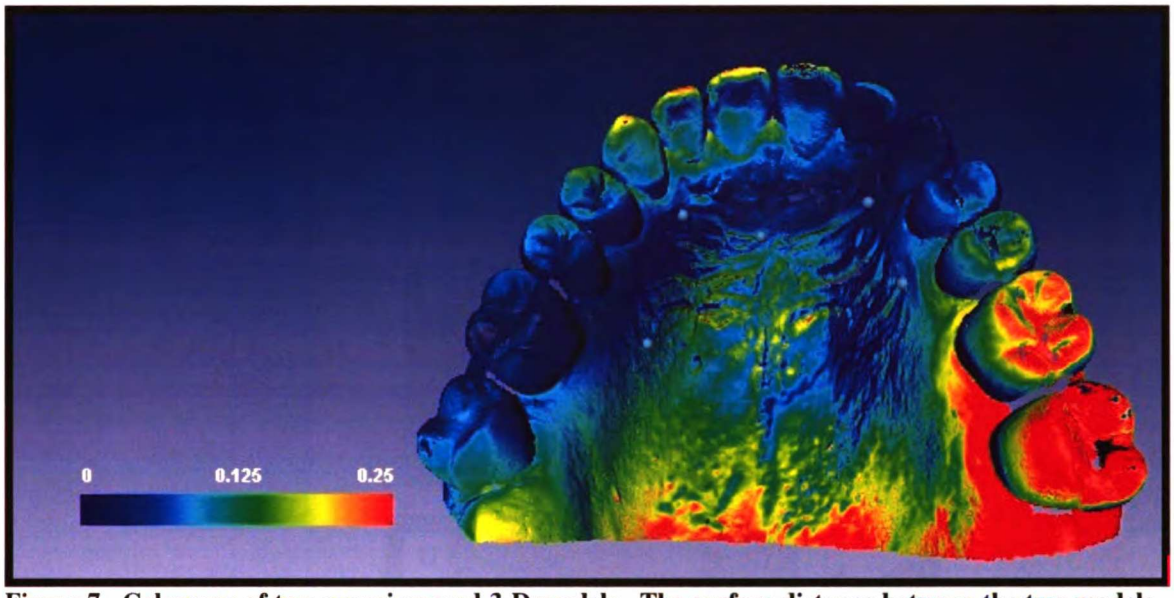


Figure 7. Colormap of two superimposed 3-D models. The surface distance between the two models is represented by the differing colors. Color scale is shown on the left hand side of the image: 0mm (blue) – 0.25mm (red).

Tie Point Registration

Although the automatic alignment algorithm is an extremely powerful tool, it has several limitations. Therefore, a more practical method of registration (utilizing tie points) was assessed. Tie points are operator chosen reference points on a 3-D model/surface. In this particular method, they were placed separately on specific features of the two identical models. After selection of all the tie points necessary, a registration function ("landmark surface warp") aligns the two sets of tie points and concomitantly both identical models (Figure 8)

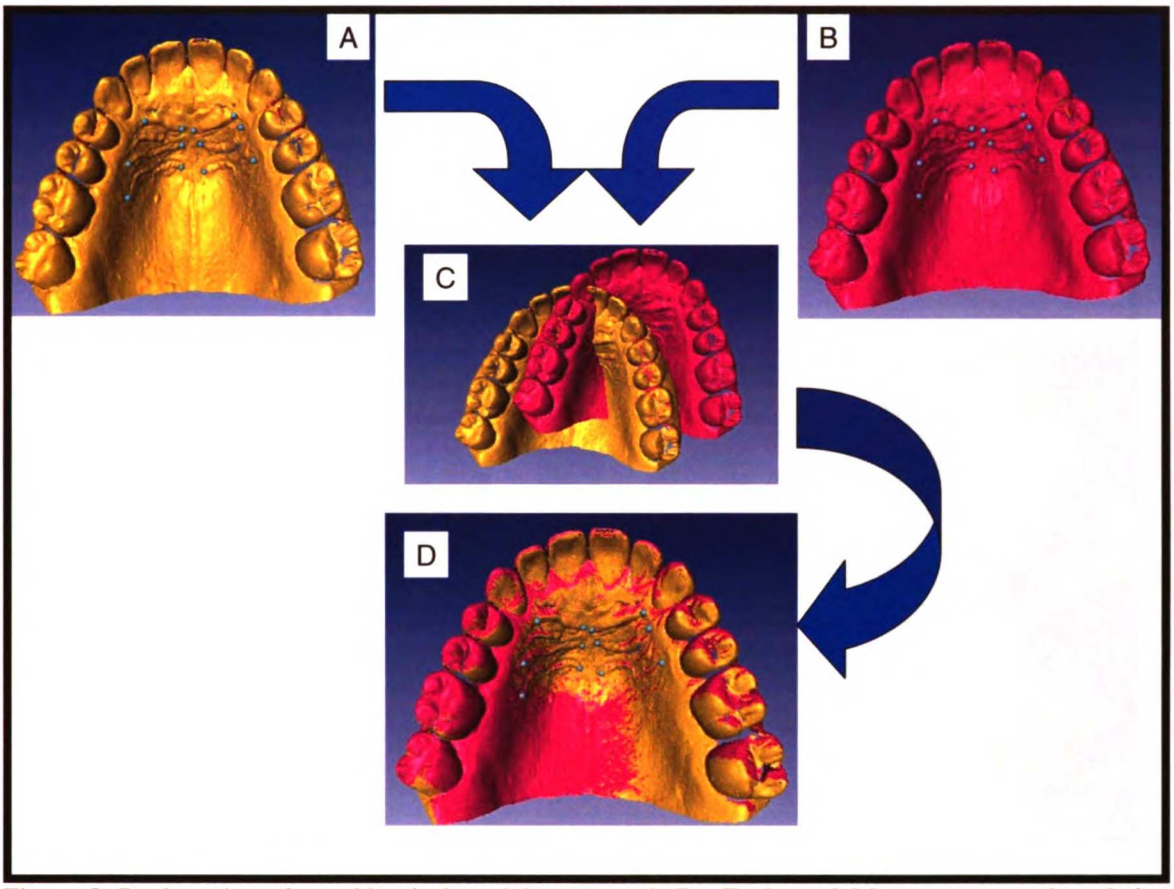


Figure 8. Registration of two identical models. (A) and (B): Each model has operator selected tie points placed at specific sites on each model. (C) and (D). Models are registered with the Amira software program by translating and rotating the models based on the tie point positions on each of the models.

For all tie point registration techniques, a rigid transformation was chosen so that a global translation and rotation occur during the registration process. Since the medial and lateral aspects of the maxillary palatal rugae are easily identifiable, they were used exclusively for tie point registration. To facilitate notation in all experiments using palatal rugae, the conventional naming system was converted to a numerical system as shown (Figure 9).

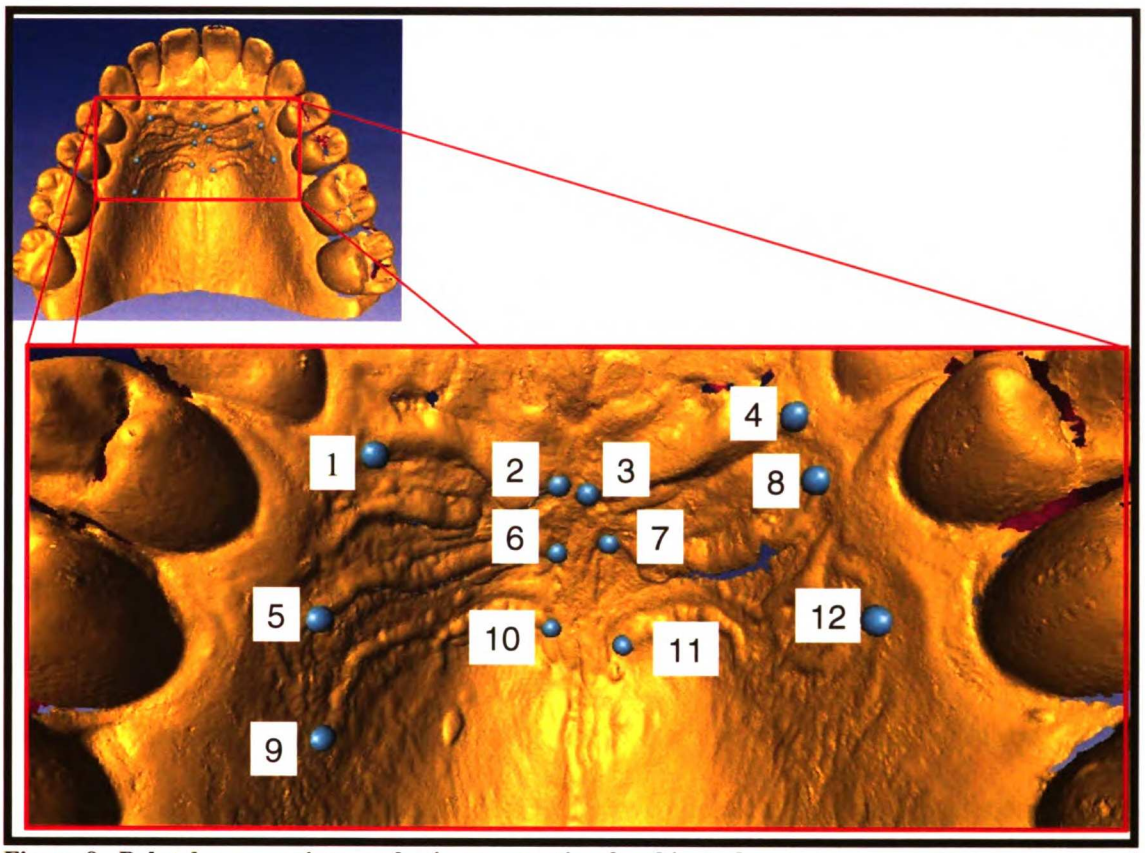


Figure 9. Palatal rugae point numbering convention for this study.

In order to account for and control all the variables that could potentially affect the registration of the identical surfaces, various series of registrations were performed. The following series of registrations were performed:

1. Increasing the number of tie points: It was hypothesized that increasing the number of tie points will result in a more accurate superimposition.
2. Variation of the spatial distribution of the tie points selected: This series tests the hypothesis that choosing tie points spatially distant from each other will result in a more accurate superimposition.
3. Elimination of a single tie point: This "elimination series" was performed to assess how individual tie points affect the registration process either positively or negatively.

4. Variables tested included increasing the number of tie points, spatial distribution of tie points, isolation/elimination of individual tie points, and operator error.

Increasing the Number of Tie Points

The following experiment was conducted in order to determine the relationship between the accuracy of the registered models and the number of tie points chosen. Initially the tie points on palatal rugae numbers 1, 2, and 4 were used to register the identical models. Once registered, the mean surface distance between the registered models was determined. This computation measures the distance from each vertex of one three-dimensional model to the closest surface of the other three-dimensional model. The mean distance between all the vertices of the two identical models was reported. This measurement was used to assess the accuracy of the registration method. In addition, colormaps of the registered models were generated. The spectrum of colors on the colormap represents the various distances between the two registered models according to the color scale in mm. (Figure 10)

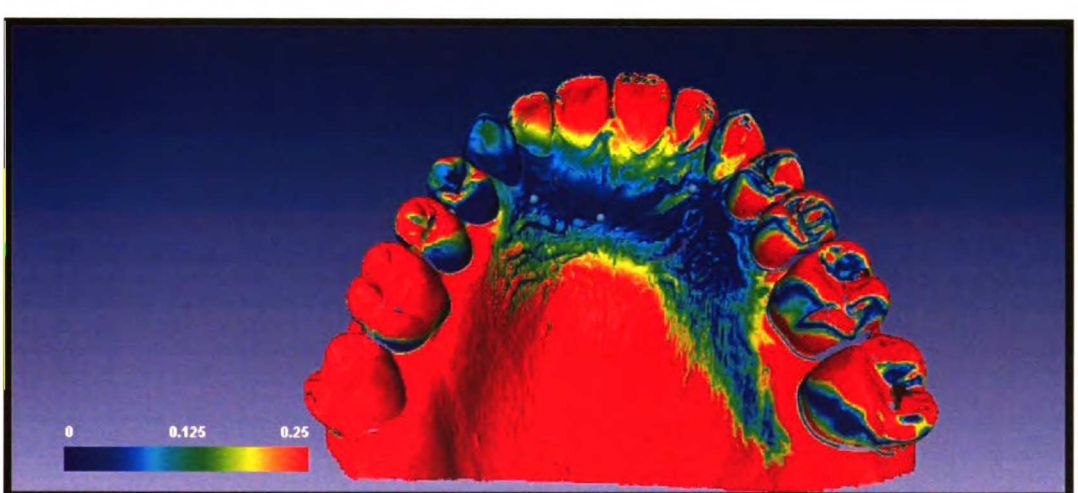


Figure 10. Colormap of two superimposed 3-D models. Colorscale: 0 mm (blue) – 0.25 mm (red).

After the accuracy assessment of the three-tie point registration technique was complete, an additional two tie points were included before the registration was performed, so that a total of five tie points was assessed. Tie points were selected on rugae numbers: 1, 2, 4, 9, and 12. The exact same tie point position for tie points 1, 2, and 4 were maintained from the first registration, so that the only change in the outcome of the second registration was the result of the addition of the two extra tie points (9 and 12). Increasing the number of tie points continued until all 12 tie points were selected. The various tie points selected for each of the registrations are shown in Figure 11.

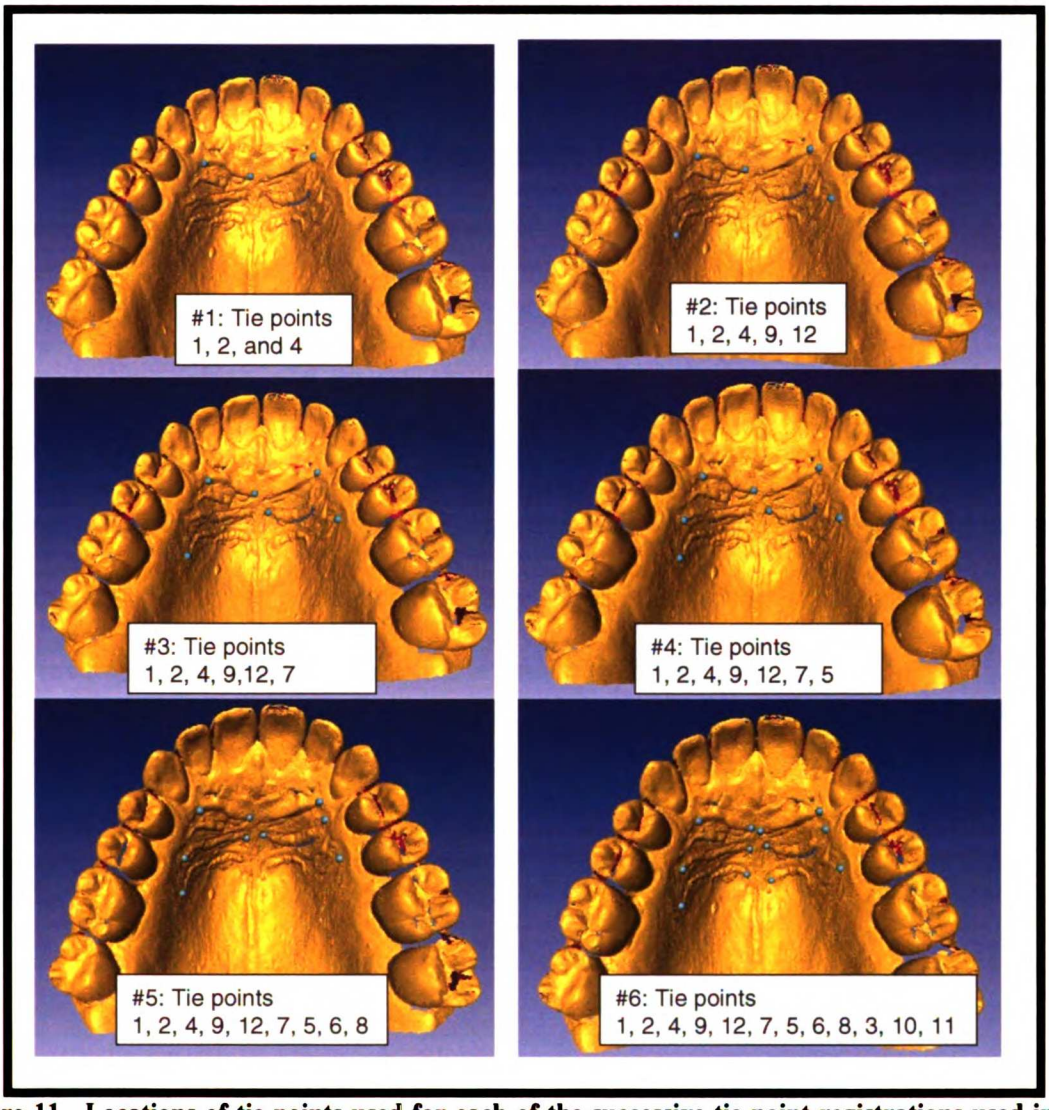


Figure 11. Locations of tie points used for each of the successive tie point registrations used in the increasing number of tie point series

Mean surface distances and colormaps were generated for all six registrations in this tie point series. The results of this initial series of registrations are obviously dependent upon the number of tie points utilized in each registration since this is the variable being tested. However, the results could also be affected by the position and order of the tie points chosen. For example, the first three tie points happened to be numbers 1, 2, and 4. If three different tie points were utilized, would the results differ? In other words, the spatial arrangement of the tie points may also affect the results in this

series of registrations. To control for this potential effect, a second series of registrations were performed whereby the tie point order used for the first registration series was reversed. Therefore, the first three tie points used in the second registration series were the last three tie points used in the first registration series. Then, the fourth and fifth tie points selected in the second series were the eighth and ninth tie points selected in the first registration series, and so on, until all 12 tie points were utilized. Again, by reversing the order of the first tie point series, the effect of the spatial distribution of the tie points was hoped to be eliminated, and the only variable tested was the increasing number of tie points with each successive registration in the series. The tie points utilized in the "reverse-order" registration series is shown in Figure 12.

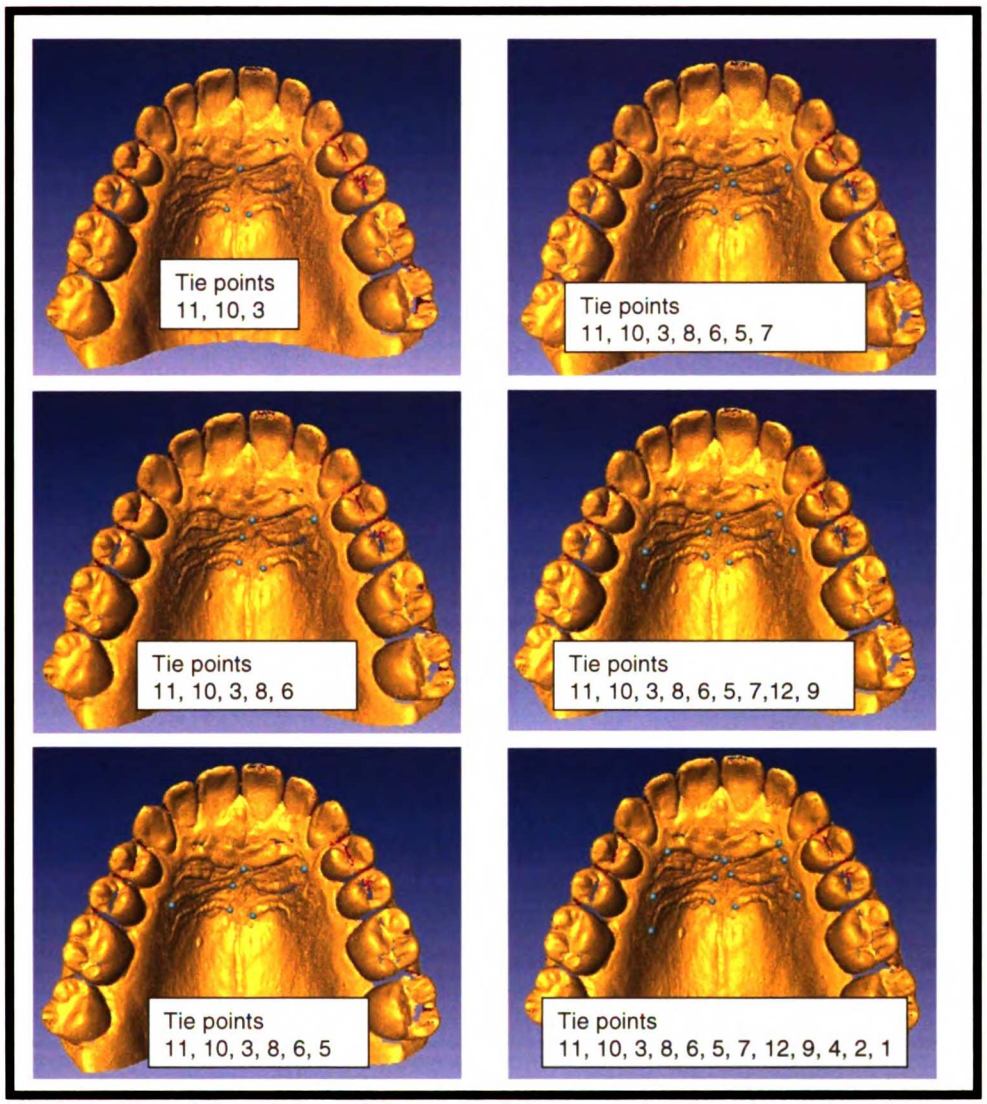


Figure 12. Locations of tie points used for each of the successive tie point registrations used in the Reversed Tie Point Series

Elimination Series

In order to assess if a particular tie point position affected the overall registration process positively or negatively, a registration series was performed whereby one of the 12 possible positions of tie points was eliminated from the registration process. Twelve registrations were performed in total. In the first registration, tie point number 1 was eliminated and only tie points 2 through 12 were used to register the two identical

models. In the second registration of this series, tie point number 2 was eliminated and only tie points 1 and 3 through 12 were used. In the third registration, tie point number 3 was eliminated. In the fourth registration, tie point 4 was eliminated; and so on until each of the 12 tie points were eliminated. Following the registrations, mean surface distances between the registered models were determined and colormaps were generated.

Spatial Patterning

In order to determine what effect tie point spatial patterning has on the registration outcome, a series of registrations were performed whereby six tie points were utilized in order to determine which pattern might produce the most accurate model registration. Of the many different potential patterns of 6 different tie points in 12 potential locations, the patterns utilized were based on several factors described for each of the registrations. The patterns of tie points are shown in Figures 13 and 14.

Figure 13. This series depicts which tie points were used for the spatial assessment series. Tie points used are listed above the 3-D models.

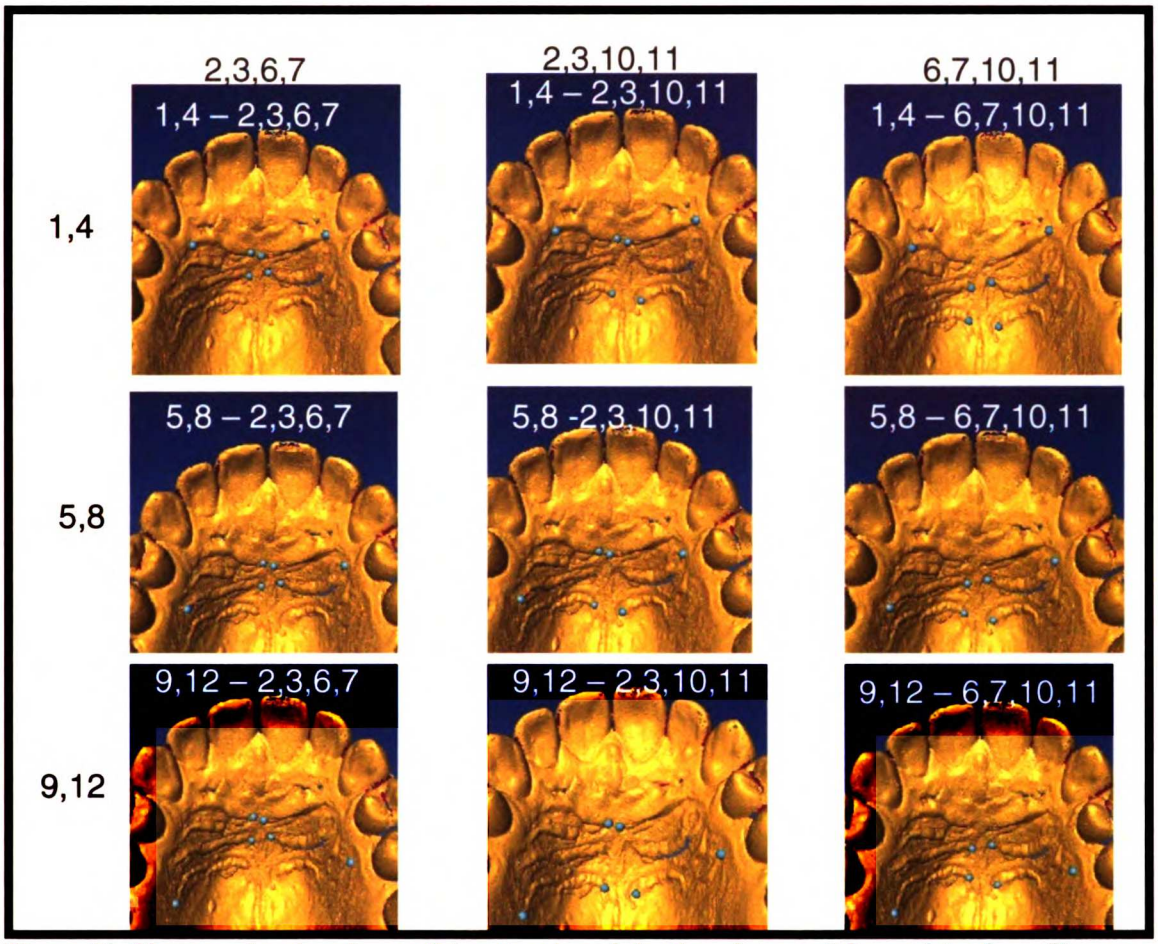


Figure 14. Diagram showing the spatial patterning of the tie points used for registration. The same two lateral palatal rugae tie points are used in each of the three rows. The same four medial palatal rugae tie points are used in each of the three columns.

Tie points 2, 3, 6, 7, 10, 11 represent all the medial palatal rugae. It is thought that these are very stable during orthodontic treatment, and therefore, would be desirable tie point positions. Tie points 1, 4, 5, 8, 9, and 12 are all the lateral tie points. As a result of the medial tie points being very stable, the rest of the spatial pattern series was based on registrations where 4 of the 6 tie points were placed on medial palatal rugae points (potentially numbers: 2, 3, 6, 7, 9, and 10), and 2 of the tie points were placed on the lateral aspects of the palatal rugae (potentially numbers 1, 4, 5, 8, 9, and 12)

Tie Point Selection Operator Error

In order to determine the variation in registration outcomes based on operator error, a series of registrations were performed. In this series, twelve tie points were utilized for each of the five registrations in this series. However, as opposed to all previous registrations where the position of the each tie point was maintained with successive registrations, in this series, all 12 tie points were generated from scratch for each of the five registrations. In this manner, operator error could be assessed by determining the amount of variance in the mean surface distance in each of the five registrations.

Development of Methods to Register Three-Dimensional Data Acquired from the White Light Scanner and the CB MercuRay CBCT

Initial work on the development of accurate and reproducible methods to register three-dimensional data acquired via the white light scanner and the CB MercuRay CBCT is presented below. The mandible from the human skull phantom utilized in Part 1 and 2 of this study was scanned by the white light scanner and the CBCT. Three-dimensional models were generated as shown in Figure 15.

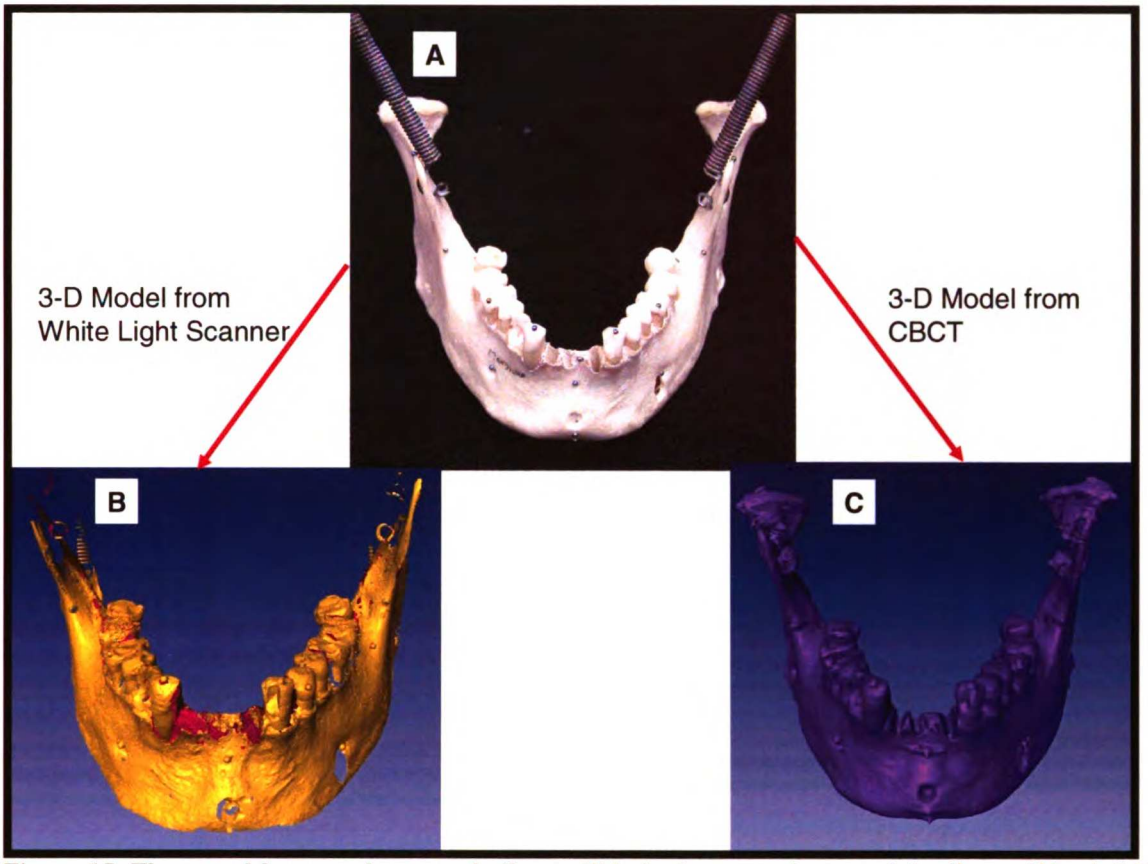


Figure 15. The actual human phantom skull mandible is shown at the top of the figure (A). A 3-D model was generated from the white light scanner (B) and from the CBCT (C)

Tie Point Registration

Both three-dimensional models that were generated with the white light scanner and the CBCT were loaded into the AMIRA software program. Tie points were utilized to register/superimpose the two models. The selected tie points were based on the position of steel fiducials placed on the original human skull mandible (Figures 15A and 16).

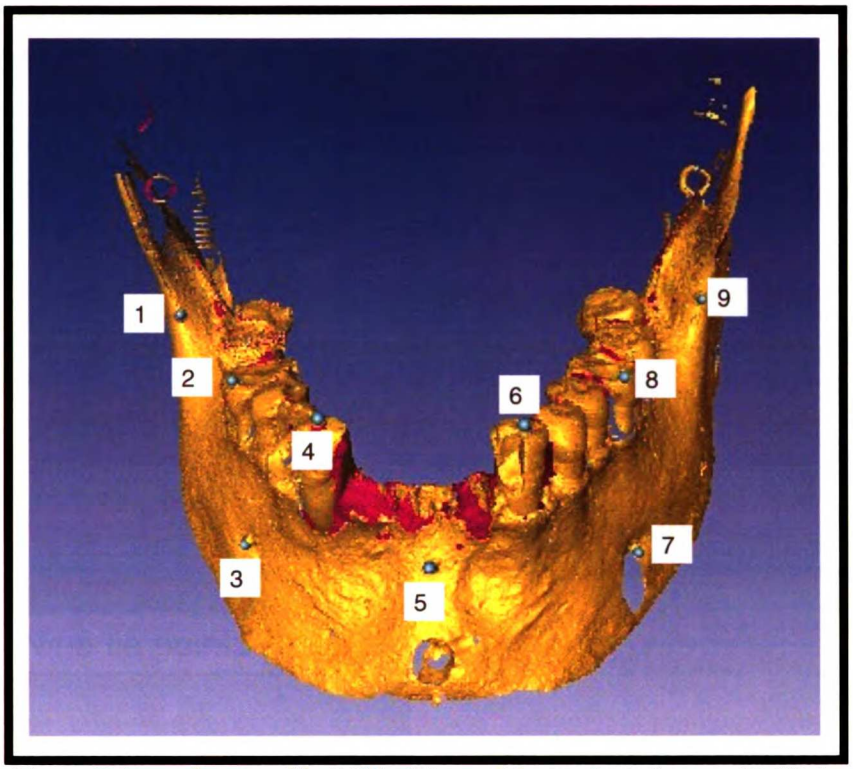


Figure 16. Positions of the various steel ball fiducials used for registration are highlighted in blue and numbered.

In total, nine fiducials could be easily identified in both of the models, and therefore, these were used solely as the tie point positions for all registrations tests. As was performed in part 3 of the study, increasing the number of tie points used for each superimposition was performed in this part of the study to ascertain the minimum number of tie points that would result in an accurate registration. Three series of registrations were performed. The first registration utilized 3 tie points, the second registration utilized 6 tie points and the third registration utilized 9 tie points (see Figure 17).

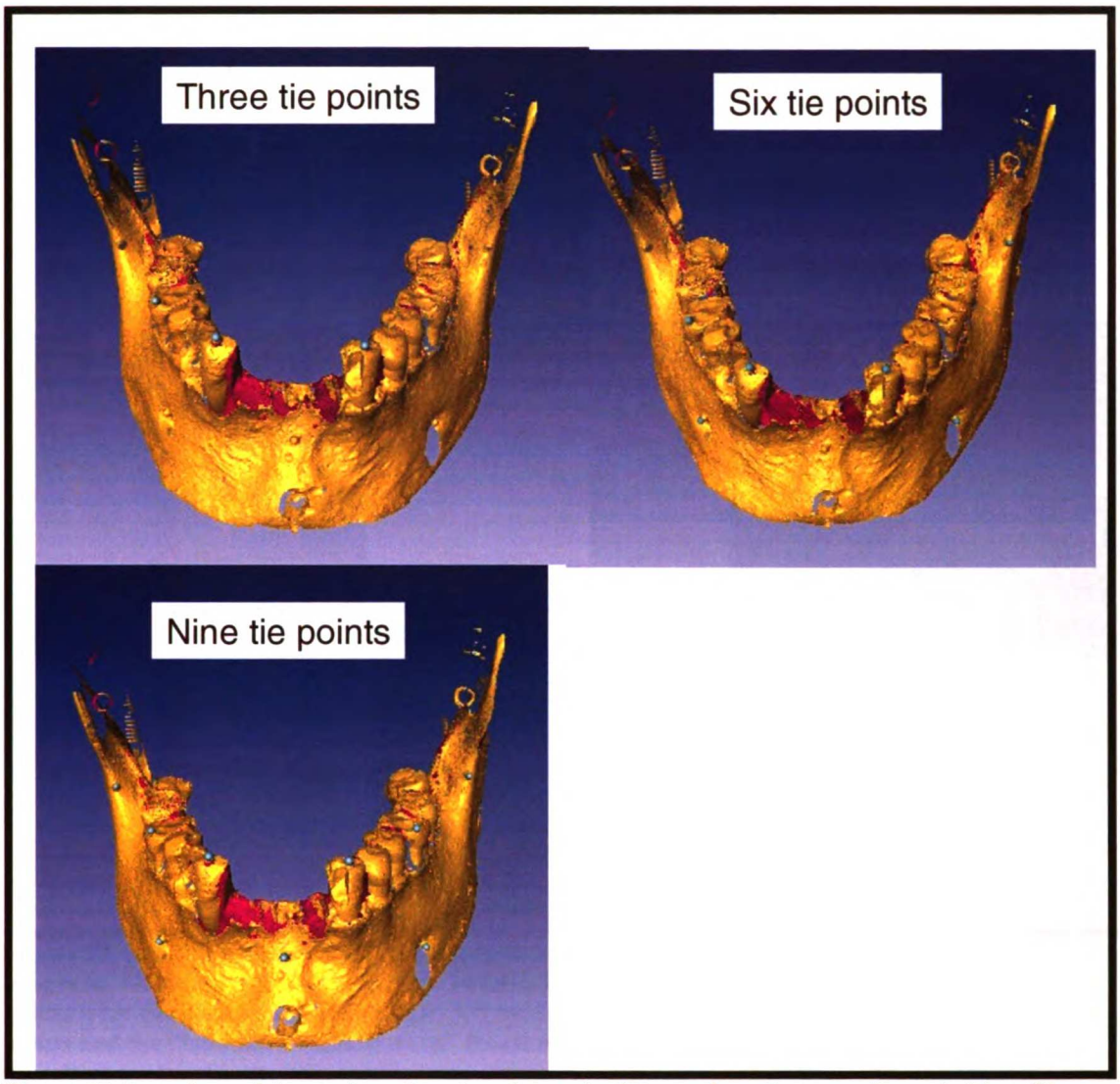


Figure 17. Locations of the various tie points for each of the registrations are shown in blue. The number of tie points is labeled on the image.

With this particular method, tie points were placed separately on each of the corresponding steel fiducials on both models. After selecting all the tie points necessary, the "landmark surface warp" registration function was performed to align the two sets of tie points and concomitantly the two models as shown in Figure 18.

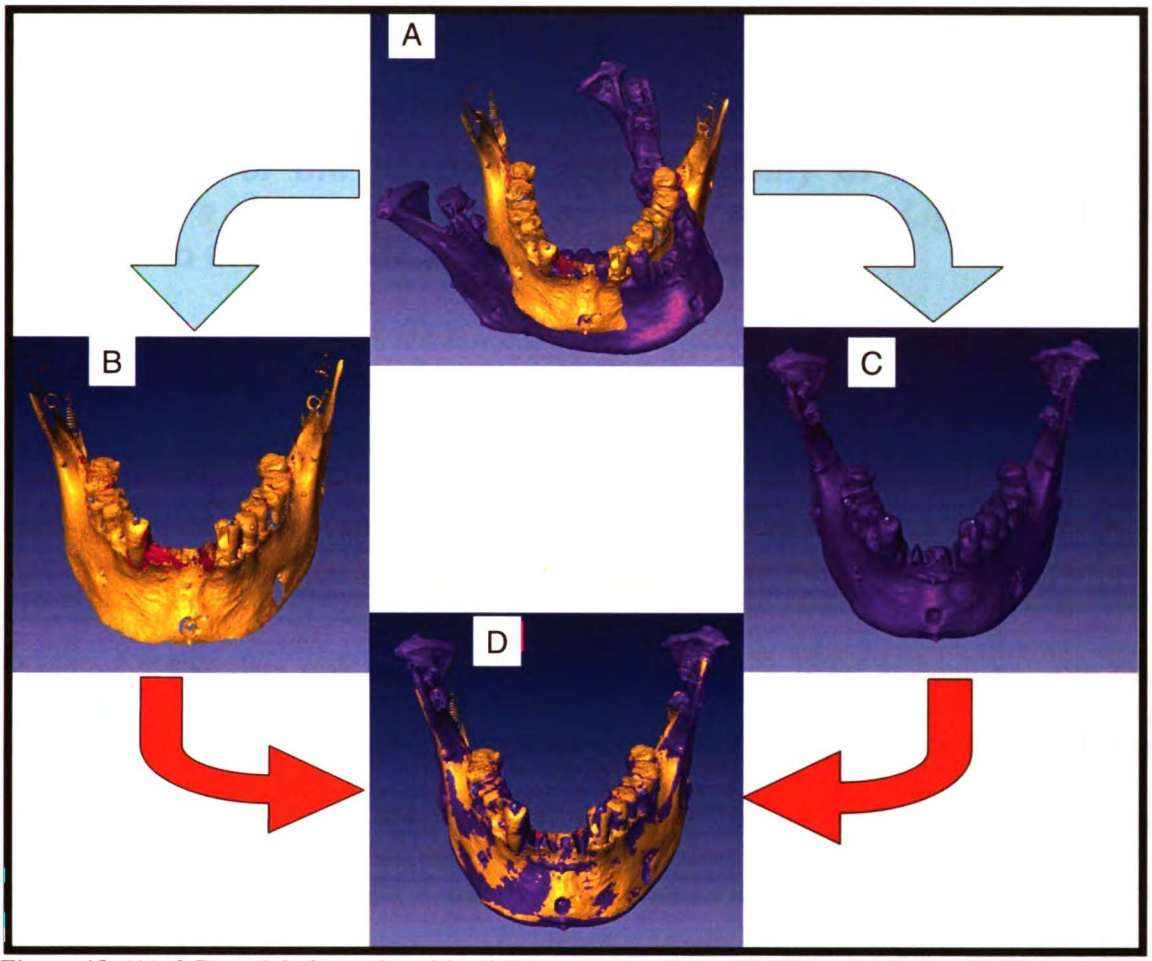


Figure 18. (A). 3-D models from the white light scanner and the CBCT loaded in the AMIRA program. Note they are not registered. (B&C). 3-D models were isolated and corresponding tie points were chosen on each model. (D). 3-D models were registered utilizing the corresponding tie points and the "landmark surface warp" function.

For all tie point registrations, a rigid transformation was utilized so that global translation and rotation of the surfaces occurred. Once registered, colormaps representing the various distances between the two registered models were generated.

Results

Comparison of the Accuracy and Reliability of Measurements made on a Surface Generated from a White Light Surface Scanner to those made with Digital Calipers.

In order to determine the accuracy of the measurements made on the three-dimensional computer models generated from the white light scanner, linear measurements were made on both the 3-D models and the phantoms themselves. Table 3 lists all five successive physical measurements obtained with digital calipers on the three phantoms. This data was treated as the gold standard in order to compare with the measurements made on the three-dimensional computer models. Table 4 lists all five successive measurements obtained on corresponding three-dimensional surfaces generated from the white light scanner. In addition, each table lists the measurements obtained one week after the initial five measurements to assess intra-rater reliability. Each physical measurement was compared with its corresponding three-dimensional measurement taken in the Amira software program. Means and standard deviations are presented in Table 5. In addition, the mean difference between the two measurement types and paired t-tests are presented to show which groups were significantly different. Univariate scattergrams are also presented in Appendix 1 for each measurement pair.

Table 3 Physical Measurements of Plastic Maxilla, Plastic Mandible, and Human Mandible Phantoms (all units are in mm)

Measurement	1	2	3	4	5	1 wk measures
PLASTIC MAXILLA	-	-	-	-	-	-
Right						
U3-U6	21.38	21.68	21.41	21.58	21.53	21.30
U1-U3	17.57	17.29	17.46	17.58	17.55	17.45
Left						
U3-U6	19.6	19.94	19.84	19.96	20.29	20.26
U1-U3	17.57	17.88	17.67	17.69	17.69	17.34
Right-Left						
U3-U3	33	33.4	33.37	33.23	33.24	32.58
U6-U6	38.09	38.35	38.17	38.28	38.35	38.04
PLASTIC MANDIBLE						
Right						
L3-L6	22.06	22.19	22.1	22.23	22.28	22.51
L1-L3	12.42	12.37	12.19	12.47	12.32	12.31
Left						
L3-L6	21.84	21.83	21.83	21.27	21.61	21.63
L1-L3	12.39	12.25	12.05	12.32	12.21	11.84
Right-Left						
L3-L3	24.12	23.61	23.85	23.7	23.93	23.92
L6-L6	39.81	39.73	39.67	39.65	39.8	39.94
HUMAN MANDIBLE						
Right						
L3-L6	22.55	22.46	22.51	22.44	22.45	22.46
L6-R1	33.04	32.81	32.77	32.79	32.85	32.90
Id-L3	18.59	18.50	18.44	18.50	18.39	18.50
Mf-L6	23.15	23.21	23.41	23.18	23.17	23.23
Mf-B	29.54	29.58	29.55	29.56	29.58	29.54
Left						
L3-L6	20.56	20.56	20.57	20.65	20.53	20.57
L6-R1	33.55	33.47	33.49	33.48	33.58	33.76
Id-L3	15.23	15.27	15.31	15.28	15.30	15.24
Mf-L6	25.66	25.60	25.75	25.57	25.79	25.64
Mf-B	32.46	32.50	32.48	32.47	32.48	32.48
Right-Left						
L3	24.05	24.03	24.07	24.07	24.01	24.16
L6	50.01	50.00	50.01	49.98	50.03	50.16
R1	81.03	80.99	81.03	81.03	81.03	81.04
Mf	52.24	52.24	52.23	52.24	52.23	52.26
Midline						
B-Id	9.24	9.19	9.21	9.24	9.25	9.27

Table 4. AMIRA Measurements of Plastic Maxilla, Plastic Mandible, and Human Mandible Phantoms (all units are in mm)

Measurement	1	2	3	4	5	1 wk measures
PLASTIC MAXILLA	-	-	-	-	-	-
Right						
U3-U7	20.99	20.92	20.88	20.98	21.06	20.97
U7-Tub	17.36	17.20	17.67	17.47	17.59	17.46
Left						
U3-U6	19.26	19.48	19.74	19.38	19.62	19.61
U6-Tub	17.84	17.96	17.83	18.10	17.90	18.04
Right-Left						
U3-U3	33.69	33.69	33.66	33.97	33.73	33.72
U6-U6	38.48	38.60	38.71	38.74	38.29	38.51
PLASTIC MANDIBLE						
Right						
L3-L6	22.07	22.21	22.10	21.93	22.18	21.89
L6-R1	12.36	12.49	12.47	12.40	12.44	12.31
Left						
L3-L6	22.08	22.15	21.76	21.95	21.85	21.89
L6-R1	12.22	12.10	11.94	12.14	12.18	12.08
Right-Left						
L3-L3	23.49	23.41	23.70	23.56	23.76	23.42
L6-L6	40.06	40.24	40.13	40.07	40.12	40.11
HUMAN MANDIBLE						
Right						
L3-L6	22.33	22.28	22.33	22.36	22.27	22.32
L6-R1	32.79	33.08	32.79	32.71	32.96	32.67
Id-L3	18.28	18.20	18.18	18.18	18.31	18.47
Mf-L6	23.05	22.80	22.82	22.94	23.04	22.93
Mf-B	29.35	29.31	29.20	29.35	29.31	29.31
Left						
L3-L6	20.49	20.62	20.34	20.59	20.49	20.48
L6-R1	33.51	33.60	32.93	33.14	32.66	33.13
Id-L3	14.87	15.02	14.88	15.07	15.04	14.94
Mf-L6	25.63	25.67	25.58	25.68	25.49	25.38
Mf-B	32.22	32.34	32.21	32.33	32.28	32.24
Right-Left						
L3	23.97	23.97	23.96	24.00	23.98	23.82
L6	50.13	50.19	50.13	50.15	50.19	49.99
R1	80.78	80.81	80.78	80.79	80.70	80.56
Mf	51.88	52.06	52.02	52.04	52.09	51.92
Midline						
B-Id	9.12	9.16	9.12	9.13	9.14	9.11

Table 5. Descriptive Statistics for Part I Amira and physical measurements. Paired t-test comparing the two measures. * - statistically significant with $p < 0.05$.

	Physical Measurements		Amira Measurements			Paired t-test
PLASTIC MAXILLA	Mean of 5 Measurements	Standard Deviation	Mean of 5 Measurements	Standard Deviation	Mean Difference	p-value
Right						
U3-U6	21.52	0.12	20.966	0.068	0.550	0.0009*
U1-U3	17.49	0.12	17.460	0.187	0.030	0.7021
Left						
U3-U6	19.93	0.25	19.496	0.191	0.430	0.0124*
U1-U3	17.70	0.11	17.924	0.114	-0.224	0.0164*
Right-Left						
U3-U3	33.25	0.16	33.749	0.129	-0.501	0.0065*
U6-U6	38.25	0.12	38.563	0.183	-0.315	0.0401*
PLASTIC MANDIBLE						
Right						
L3-L6	22.17	0.09	22.097	0.109	0.075	0.2768
L1-L3	12.35	0.11	12.431	0.054	-0.077	0.3041
Left						
L3-L6	21.68	0.25	21.959	0.157	-0.283	0.0773
L1-L3	12.24	0.13	12.116	0.109	0.128	0.0094*
Right-Left						
L3-L3	23.84	0.20	23.583	0.146	0.259	0.0489*
L6-L6	39.73	0.07	40.123	0.071	-0.391	0.0012*
HUMAN MANDIBLE						
Right						
L3-L6	22.48	0.05	22.313	0.037	0.169	0.0022*
L6-R1	32.85	0.11	32.867	0.148	-0.015	0.8709
Id-L3	18.48	0.08	18.231	0.061	0.253	0.0050*
Mf-L6	23.22	0.11	22.927	0.119	0.297	0.0323*
Mf-B	29.56	0.02	29.306	0.061	0.256	0.0008*
Left						
L3-L6	20.57	0.05	20.505	0.111	0.069	0.2197
L6-R1	33.51	0.05	33.168	0.392	0.346	0.1356
Id-L3	15.28	0.03	14.976	0.093	0.302	0.0016*
Mf-L6	25.67	0.09	25.608	0.079	0.066	0.4385
Mf-B	32.48	0.01	32.276	0.061	0.202	0.0012*
Right-Left						
L3	24.05	0.03	23.974	0.017	0.072	0.0074*
L6	50.01	0.02	50.156	0.028	-0.150	0.0003*
R1	81.02	0.02	80.774	0.042	0.248	0.0005*
Mf	52.24	0.01	52.018	0.079	0.218	0.0039*
Midline						
B-Id	9.23	0.03	9.133	0.0193	0.0931	0.0052*

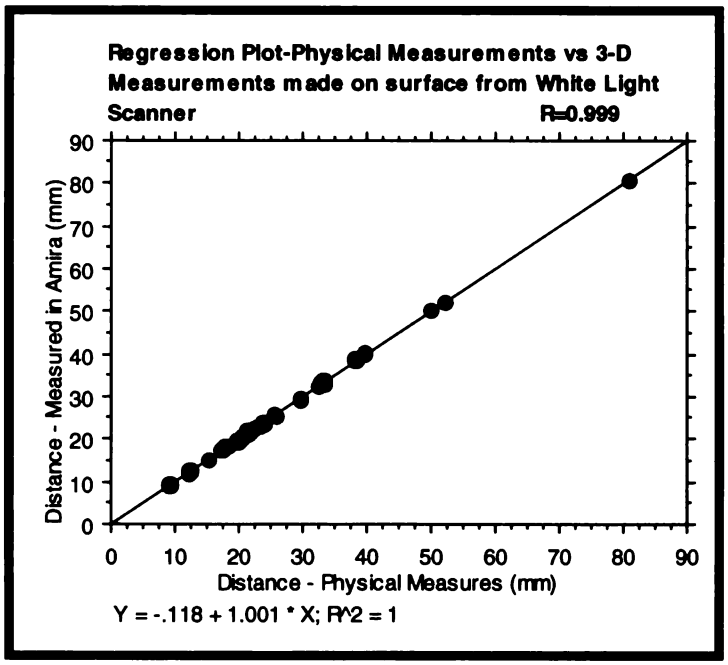


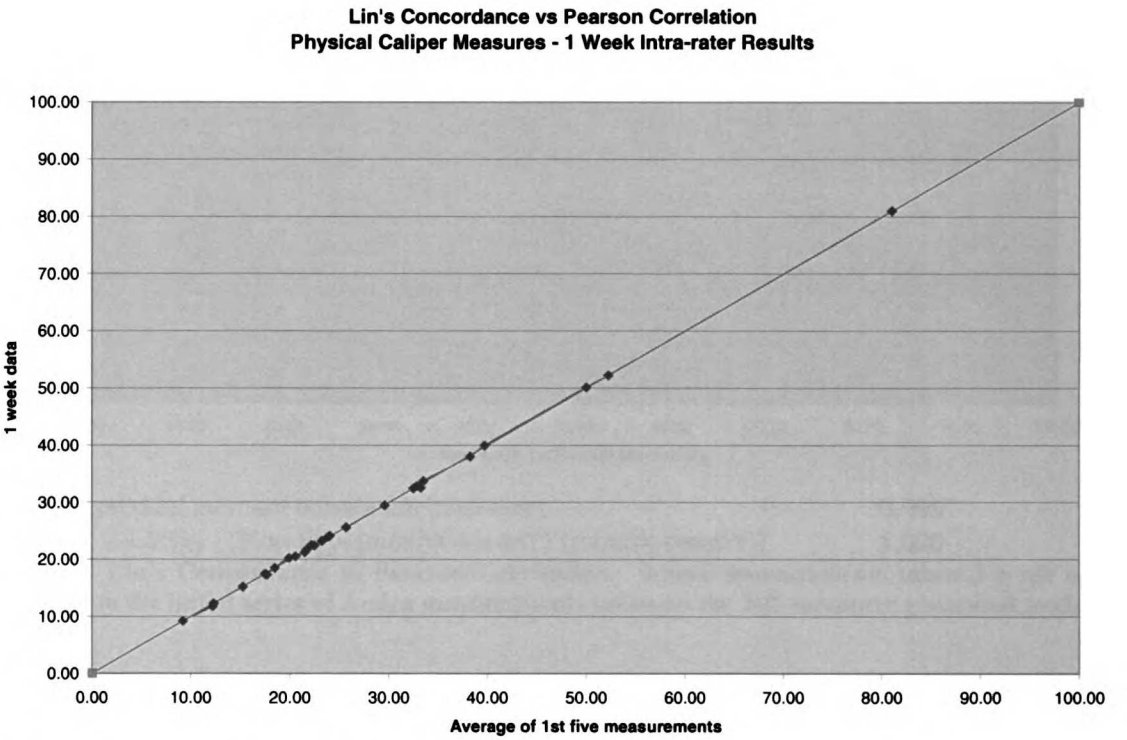
Figure 19. Linear regression analysis comparing the 3-D measures of the computer based model to the gold standard digital caliper measures of the physical model.

The absolute difference between the average of all the measurements made with calipers and with the Amira program are show below in Table 6. The average distance measured for this series was also calculated, and the overall percent error of the Amira measurement method compared to the caliper measurement method as the gold standard was determined.

Table 6: Average descriptive statistics and percent error (Part 1)

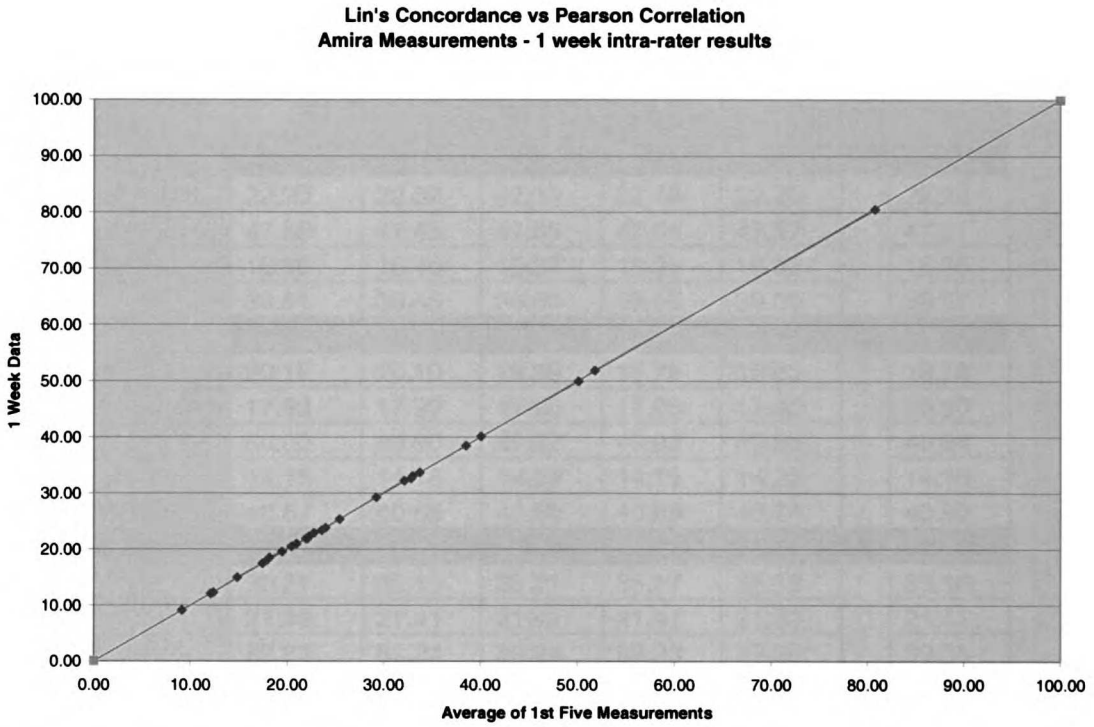
Table X	
Average absolute mean difference between Amira and physical measurements	0.22 mm
Average distance measured	27.77 mm
Percent error	0.80%

To assess intra-rater reliability, Lins concordance was performed, and the results are show below for both the physical measurements and the Amira measurements (Figure 20 and Figure 21, respectively).



Pearson product moment correlation coefficient	0.999
rho(_c_) = 2Sxy / [Sxx+Syy+(meanX-meanY)*(meanX-meanY)]	1.000

Figure 20. Lin's Concordance vs Pearson Correlation. Physical measurements taken 1 week apart compared to the initial series of physical measurements.



Pearson product moment correlation coefficient	0.999
rho(_c_) = 2Sxy / [Sxx+Syy+(meanX-meanY)*(meanX-meanY)]	1.000

Figure 21. Lin's Concordance vs Pearson Correlation. Amira measurements taken 1 week apart compared to the initial series of Amira measurements taken on the 3-D computer generated model.

Comparison of the Accuracy and Reliability of Measurements Obtained from a Reconstructed Three-Dimensional Surface Generated from a Cone Beam CT to those Made with Digital Calipers.

Measurements made to compare the accuracy and reliability of the 3-D model generated from the CBCT to that of the physical gold standard are presented. Table 7 lists all five successive physical measurements obtained with digital calipers. Table 8 lists all five successive measurements obtained via the three-dimensional reconstructed surface in the Amira software program. In addition, each table lists the measurements obtained 1 week after the initial five measurements for intra-rater reliability.

Table 7. Physical measures taken on the human skull phantom using digital calipers (all measurements in mm)

Measurement	1	2	3	4	5		1 week repeat
Right							
Or-IOF	22.20	22.32	22.19	22.18	22.20		22.28
IOF-U3	47.59	47.43	47.59	47.64	47.57		47.61
PR-ANS	16.50	16.46	16.37	16.35	16.33		16.35
IOF-A	39.51	39.45	39.64	39.65	39.56		39.57
Left							
U3-U6	20.18	20.10	19.99	19.75	19.83		19.75
Or-IOF	17.93	17.92	17.95	17.95	17.90		18.00
IOF-U3	50.00	49.90	49.87	49.92	49.96		49.94
PR-ANS	14.15	14.16	14.22	14.15	14.22		14.16
IOF-A (IS)	40.67	40.88	40.68	40.69	40.74		40.90
Right-Left							
U3	35.21	35.19	35.21	35.17	35.19		35.16
PR	21.42	21.41	21.41	21.41	21.39		21.41
Or	89.24	89.24	89.24	89.23	89.36		89.35
IOF	54.99	55.25	54.98	55.02	55.23		55.05
Midline							
ANS-A	7.55	7.54	7.54	7.54	7.55		7.54
ANS-Pr	20.03	20.01	20.04	20.03	20.05		20.03
Right							
L3-L6	22.55	22.46	22.51	22.44	22.45		22.46
L6-R1	33.04	32.81	32.77	32.79	32.85		32.90
Id-L3	18.59	18.50	18.44	18.50	18.39		18.50
Mf-L6	23.15	23.21	23.41	23.18	23.17		23.23
Mf-B	29.54	29.58	29.55	29.56	29.58		29.54
Left							
L3-L6	20.56	20.56	20.57	20.65	20.53		20.57
L6-R1	33.55	33.47	33.49	33.48	33.58		33.76
Id-L3	15.23	15.27	15.31	15.28	15.30		15.24
Mf-L6	25.66	25.60	25.75	25.57	25.79		25.64
Mf-B	32.46	32.50	32.48	32.47	32.48		32.48
Right-Left							
L3	24.05	24.03	24.07	24.07	24.01		24.16
L6	50.01	50.00	50.01	49.98	50.03		50.16
R1	81.03	80.99	81.03	81.03	81.03		81.04
Mf	52.24	52.24	52.23	52.24	52.23		52.26
Midline							
B-Id	9.24	9.19	9.21	9.24	9.25		9.27

Table 8. 3-D Measurements on a human skull obtained in Amira software program (all measurements in mm)

Measurement	1	2	3	4	5	1 week repeat
Right						
Or-IOF	21.98	21.96	21.92	21.92	22.02	22.00
IOF-U3	47.37	47.35	47.04	47.31	47.34	47.41
PR-ANS	16.16	16.17	16.09	16.14	16.16	16.19
IOF-A	39.64	40.69	39.55	39.56	39.63	39.60
Left						
U3-U6	20.13	20.14	20.12	20.05	20.19	20.30
Or-IOF	17.73	17.71	17.65	17.70	17.74	17.76
IOF-U3	49.32	49.29	49.04	49.12	49.30	49.29
PR-ANS	13.78	13.74	13.74	13.83	13.80	13.82
IOF-A	40.12	40.26	40.17	40.19	40.27	40.25
Right-Left						
U3	35.06	35.04	34.81	35.06	35.07	35.05
PR	21.41	21.42	21.28	21.40	21.42	21.32
Or	88.17	88.27	87.76	88.31	88.28	88.30
IOF	54.52	54.54	54.31	54.53	54.48	54.52
Midline						
ANS-A	7.17	7.18	7.06	7.19	7.17	7.17
ANS-Pr	19.45	19.44	19.39	19.50	19.49	19.44
Right						
L3-L6	22.34	22.40	22.22	22.33	22.37	22.34
L6-R1	33.01	33.00	32.93	32.95	33.08	33.03
Id-L3	18.53	18.58	18.82	18.80	18.68	18.67
Mf-L6	22.63	22.69	22.56	22.74	22.68	22.64
Mf-B	29.19	29.20	28.96	29.17	29.17	29.19
Left						
L3-L6	20.41	20.33	20.39	20.32	20.36	20.55
L6-R1	33.63	33.62	33.52	33.60	33.58	33.62
Id-L3	14.90	14.87	14.85	14.94	14.98	14.86
Mf-L6	24.98	24.99	24.77	24.96	24.99	24.98
Mf-B	32.19	32.19	32.03	32.21	32.17	32.20
Right-Left						
L3	24.05	24.08	23.85	24.04	24.12	24.10
L6	49.82	49.83	49.63	49.86	49.85	49.84
R1	80.14	80.13	79.87	80.14	80.15	80.13
Mf	51.72	51.65	51.71	51.68	51.72	51.69
Midline						
B-Id	8.64	8.67	8.72	8.70	8.69	8.71

Each physical measurement was compared with its corresponding three-dimensional measurement taken in the Amira software program. The mean difference between the physical and Amira measurements is presented as well as paired t-tests to show which

groups were significantly different (Table 9). Univariate scattergrams are also presented for each measurement pair in Appendix 2.

Table 9. Descriptive Statistics for Part II Amira and physical measurements. Paired t-test comparing the two measures. * statistically significant with $p<0.05$.

Part II Measures	Physical Measurements		Amira Measurements			paired t-test
Measurement	Mean of 5 measurements	Standard Deviation	Mean of 5 Measurements	Standard Devation	mean difference	p value
Right						
Or-IOF	22.22	0.058	21.96	0.044	0.26	0.000971611*
IOF-U3	47.56	0.079	47.28	0.137	0.28	0.022784392
PR-ANS	16.40	0.074	16.15	0.033	0.26	0.000946625*
IOF-A	39.56	0.085	39.81	0.493	-0.25	0.373388985
Left						
U3-U6	19.97	0.180	20.12	0.051	-0.15	0.118879916
Or-IOF	17.93	0.021	17.71	0.032	0.22	0.000592205*
IOF-U3	49.93	0.051	49.21	0.125	0.72	0.000070280*
PR-ANS	14.18	0.037	13.78	0.037	0.40	0.000104105*
IOF-A	40.73	0.087	40.20	0.064	0.53	0.000027512*
Right-Left						
U3	35.19	0.017	35.01	0.109	0.19	0.024237723*
PR	21.41	0.011	21.39	0.059	0.02	0.507959300
Or	89.26	0.055	88.16	0.229	1.10	0.000367136*
IOF	55.09	0.134	54.48	0.095	0.62	0.000423382*
Midline						
ANS-A	7.54	0.005	7.16	0.054	0.39	0.000082805*
ANS-Pr	20.03	0.015	19.45	0.042	0.58	0.000006873*
Right						
L3-L6	22.48	0.047	22.33	0.068	0.15	0.023997677*
L6-R1	32.85	0.109	32.99	0.058	-0.14	0.035314101*
Id-L3	18.48	0.075	18.68	0.128	-0.20	0.071435350
Mf-L6	23.22	0.106	22.66	0.068	0.56	0.001461768*
Mf-B	29.56	0.018	29.14	0.099	0.42	0.000556343*
Left						
L3-L6	20.57	0.045	20.36	0.038	0.21	0.002799080*
L6-R1	33.51	0.048	33.59	0.046	-0.08	0.053977843
Id-L3	15.28	0.031	14.91	0.053	0.37	0.000155762*
Mf-L6	25.67	0.094	24.94	0.092	0.74	0.000455058*
Mf-B	32.48	0.015	32.16	0.073	0.32	0.000718534*
Right-Left						
L3	24.05	0.026	24.03	0.104	0.02	0.741300271
L6	50.01	0.018	49.80	0.095	0.21	0.009555259*
R1	81.02	0.018	80.09	0.118	0.94	0.000071635*
Mf	52.24	0.005	51.69	0.034	0.54	0.000005304*
Midline						
B-Id	9.23	0.025	8.68	0.030	0.54	0.000007636*

The absolute difference between the average of all the measurements made with calipers and with the Amira program are show below in Table 10. The average distance measured for the series was also calculated, and the overall percent error of the Amira measurement method compared to the caliper measurement method as the gold standard was determined. While many of the measures were significantly different between the two methods, the mean difference was less than 0.5 mm, which can be considered clinically insignificant. The linear regression analysis indicates near perfect correlation (Figure 22).

Table 10. Average descriptive statistics and percent error (Part II)

Average absolute mean difference between Amira and physical measurements	0.38 mm
Average distance measured	32.09 mm
Percent error	1.19%

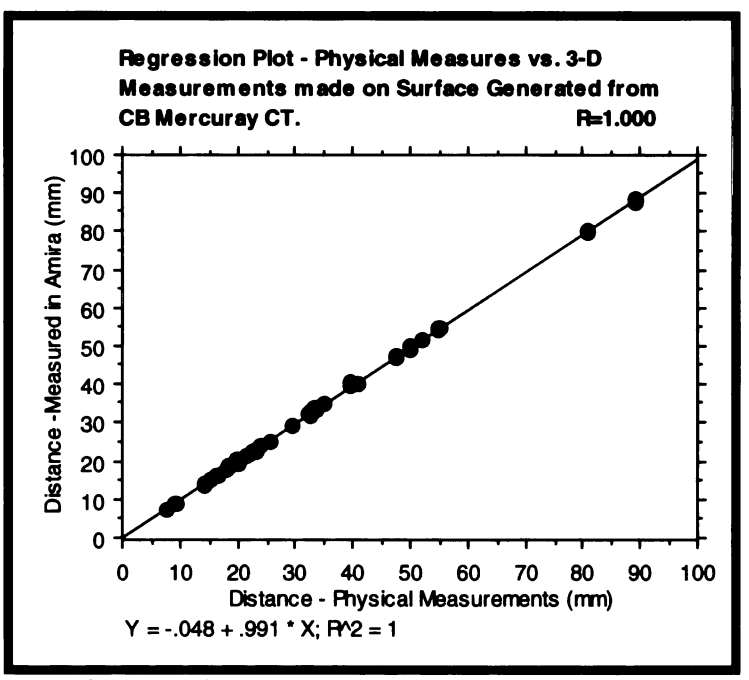
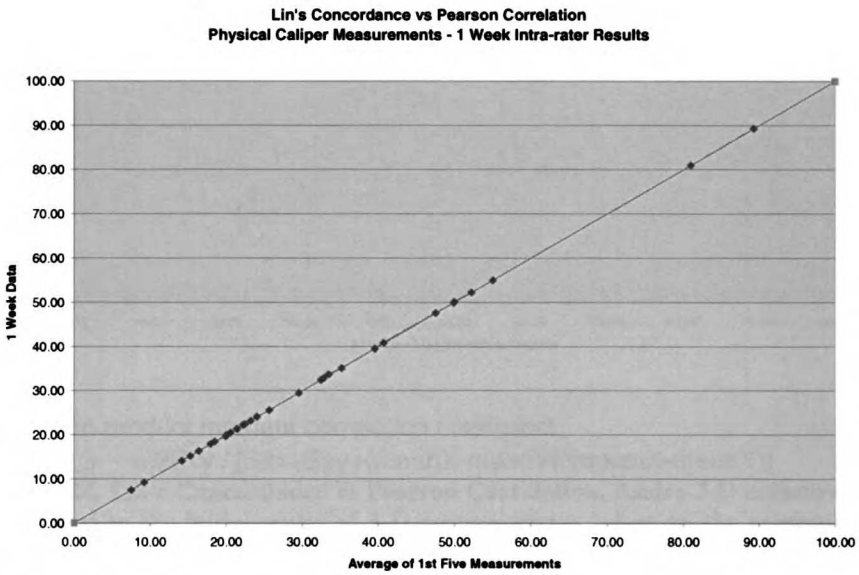


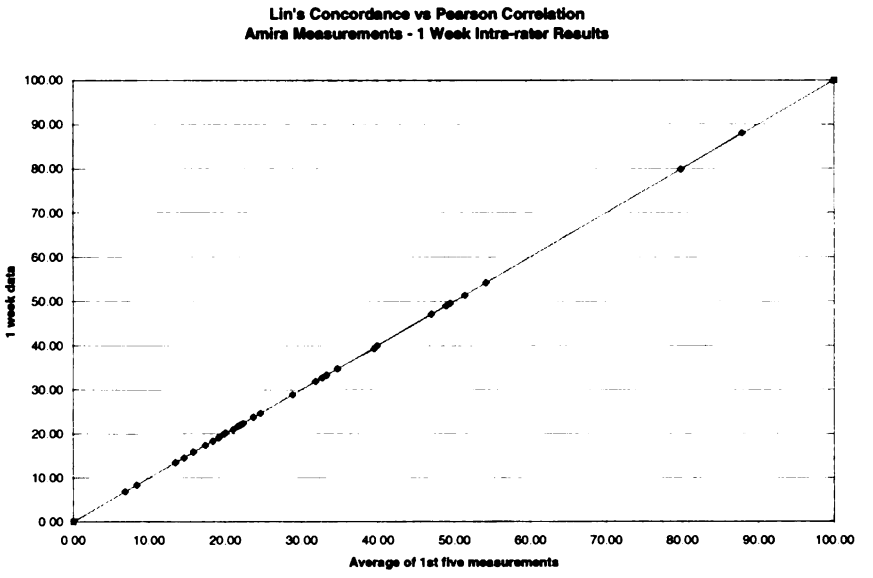
Figure 22. Linear regression analysis comparing the same measures by two different methods where the physical caliper measures are the gold standard.

To assess intra-rater reliability Lin's concordance was performed. Results are shown below for both the physical 1 week measurements and the Amira 1 week measurements (Figures 23 and 24, respectively).



Pearson product moment correlation coefficient	0.999
rho(_c_) = 2Sxy / [Sxx+Syy+(meanX-meanY)*(meanX-meanY)]	1.000

Figure 23. Lin's Concordance vs Pearson Correlation. Physical measurements taken 1 week apart compared to the initial series of physical measurements on the human phantom skull.



Pearson product moment correlation coefficient	0.999
rho(_c_) = 2Sxy / [Sxx+Syy+(meanX-meanY)*(meanX-meanY)]	1.000

Figure 24. Lin's Concordance vs Pearson Correlation. Amira 3-D measurements taken 1 week apart compared to the initial series of 3-D measurements taken on the computer generated model of the human phantom skull.

Comparison of Various Methods to Register Three-Dimensional Dental Models

Computer Generated Automatic Alignment

The results of the automatic alignment of the two identical three-dimensional models are shown below. The alignment algorithm using the Procrustes method produced a mean surface distance between the two surfaces of 0.0003mm+/- 0.00016mm. The colormap (Figure 25) indicates near perfect alignment of the two surfaces.

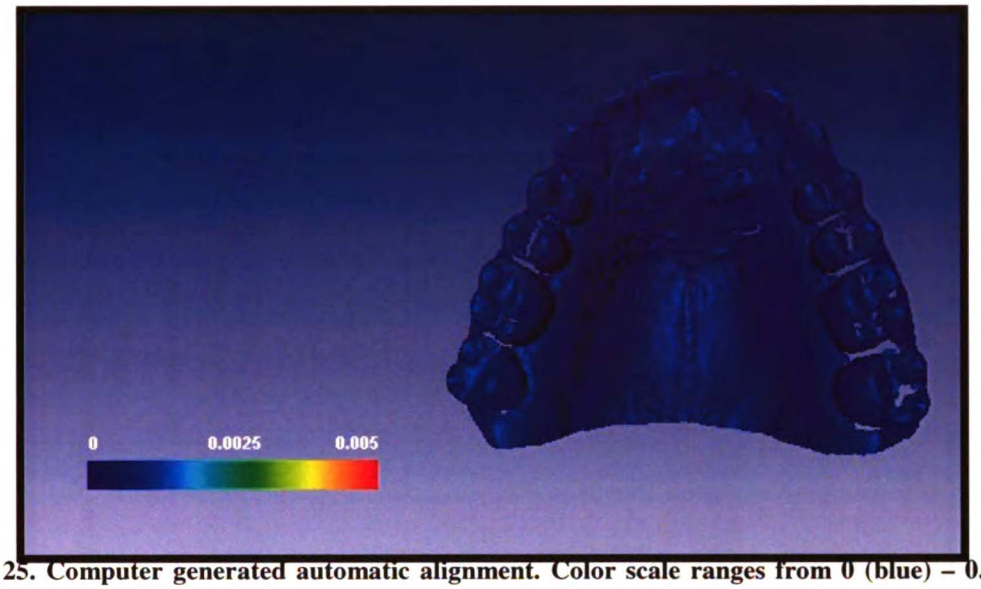


Figure 25. Computer generated automatic alignment. Color scale ranges from 0 (blue) – 0.005mm (red).

Increasing the Number of Tie Points

The effect of increasing the number of tie points used to register the two models is shown in Table 11 and Figure 26. As the number of tie points increases, the mean surface distance between the registered models decreases.

Table 11. Descriptive Statistics for Increasing Number of Tie Points Series

er of tie points	mean surface distance (u	
3	0.268	0.219
5	0.094	0.071
6	0.072	0.053
7	0.072	0.061
9	0.033	0.022
12	0.024	0.017

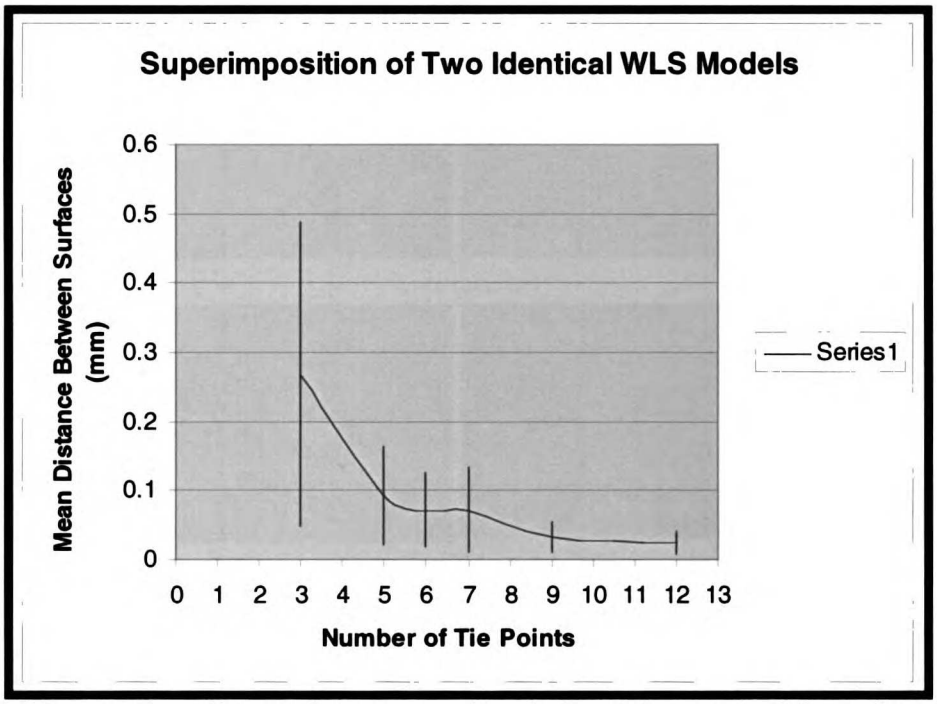


Figure 26. Mean surface distance between two identical models generated from the white light scanner where the number of tie points on the palatal rugae were increased.

The colormaps (Figure 27) demonstrate the same trend as can be seen graphically in Figure 26; namely that the surface distances between the two models decreases with increasing numbers of tie points. As can be seen in the lower left hand corner of the image, the scale of the colormap is from 0mm to 0.25 mm on each of the registered images.

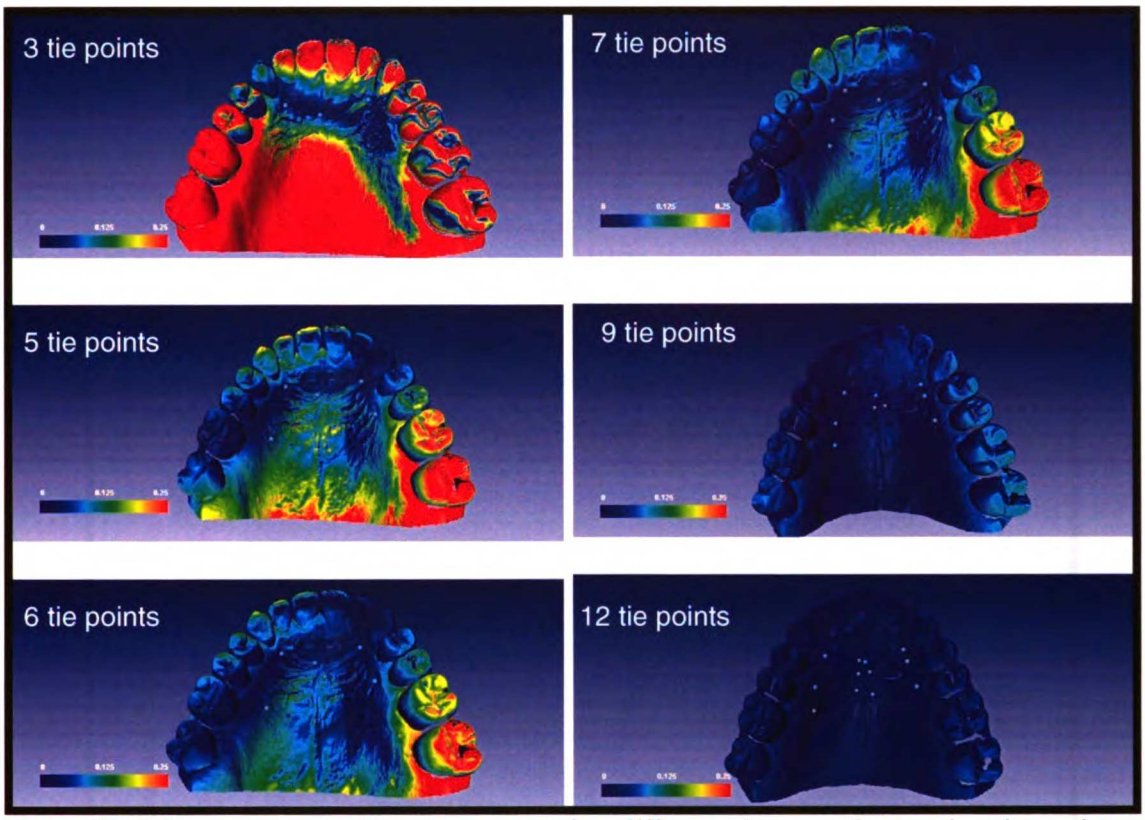


Figure 27. Colormaps representing the mean surface difference between the two tie point registered models. The number of tie points used is listed next to the colormap image. Colorscale: 0 mm (blue) – 0.25 mm (red).

"Reverse-Order" Series

As mentioned in the methods, the results of the "increasing number of tie points" series of registrations, could be affected by the position and order of the tie points used. To control for this effect, the exact same registration procedures were used in this series, while successively increasing the number of tie points. However, the order of tie points was reversed. The results are shown in Table 12, and Figure 28 and 29. Again, as the number as tie points is increased, the mean surface distance between the surfaces decreases.

Table 12. Descriptive Statistics for **Reverse Order Tie Point Series**

Number of Tie Points	Mean Surface Distance	Standard Deviation
3	0.073	0.061
5	0.083	0.065
6	0.070	0.056
7	0.070	0.056
9	0.022	0.016
12	0.024	0.017

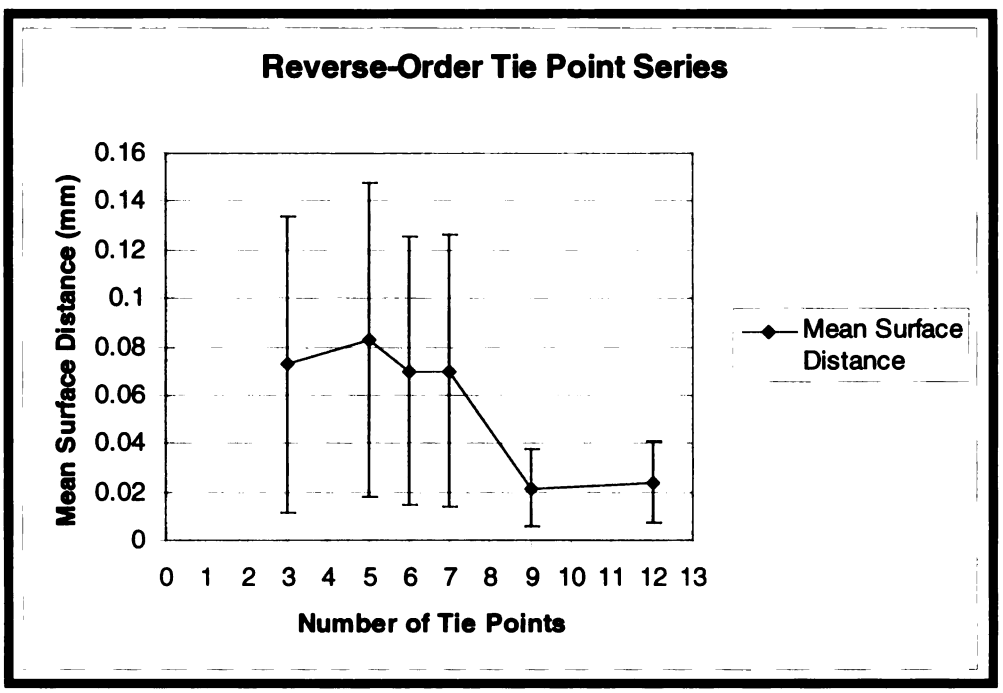


Figure 28. Graph depicts the mean surface distance between two identical models generated from the white light scanner when the tie points chosen to register them were the exact opposite as the one chosen the the "increasing number" of tie points series.

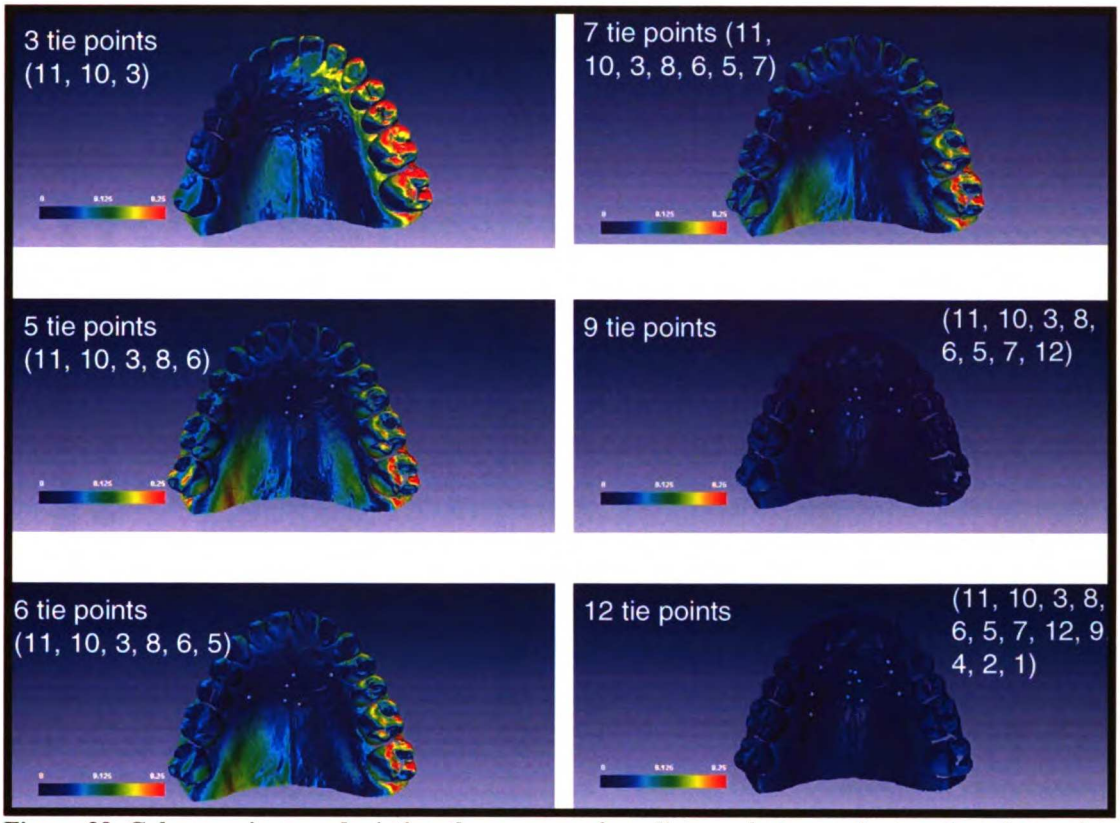


Figure 29. Colormap images depicting the mean surface distance between the two tie point registered models. The number of tie points and the specific tie points utilized is listed. Color scale: 0 mm (blue) – 0.25mm (red)

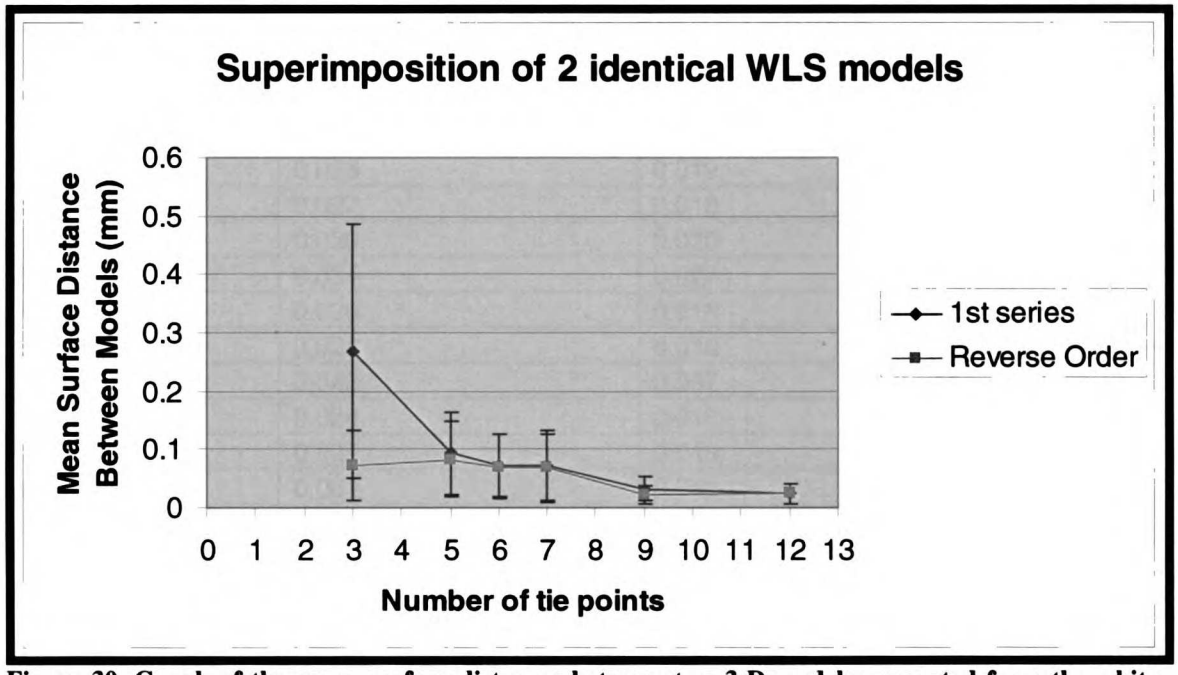


Figure 30. Graph of the mean surface distances between two 3-D models generated from the white light scanner when using a particular order of adding tie points in the first series, and reversing the order for the second series.

Overlaying the graphs of the Increasing Tie Points Series and the Reverse-Order Series demonstrates the overall trend in mean surface distance as the number of tie points increases (Figure 30). In addition, the data show that when using less than five tie points, the selection of the tie points can significantly affect the registration of the 3-D models. It also demonstrates the range that can occur when too few tie points are used.

Tie Point Elimination Series

The "tie point elimination series" was performed in order to ascertain whether a single tie point affected the overall superimposition process, and in what manner it affected the overall superimposition process. The results of the elimination series is shown below in Table 13 and Figures 31 and 32.

Table 13. Descriptive Statistics for Elimination Series

Tie Point Elimination #	Mean Surface Distance (mm)	Standard Deviation
1	0.032	0.024
2	0.025	0.018
3	0.028	0.019
4	0.027	0.018
5	0.030	0.020
6	0.034	0.027
7	0.026	0.018
8	0.021	0.015
9	0.048	0.037
10	0.024	0.018
11	0.027	0.019
12	0.061	0.045

It should be noted that a tie point that has a positive affect on the superimposition process would create a large mean surface distance when it is eliminated. For example when tie point number 12 was eliminated, the resulting mean surface distance of the two registered models was greater than the mean surface distance of all the other registrations. This suggests that tie point number 12 is relatively more important in the registration process than the other tie points.

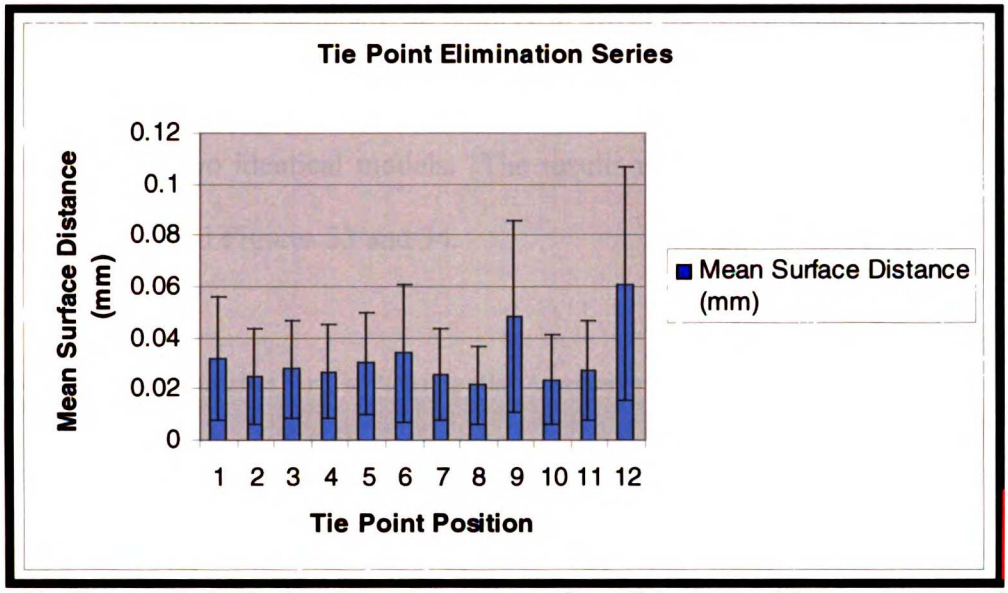


Figure 31. Bar graph indicating how the mean surface distance would vary between the two registered models after removal of a single tie point.

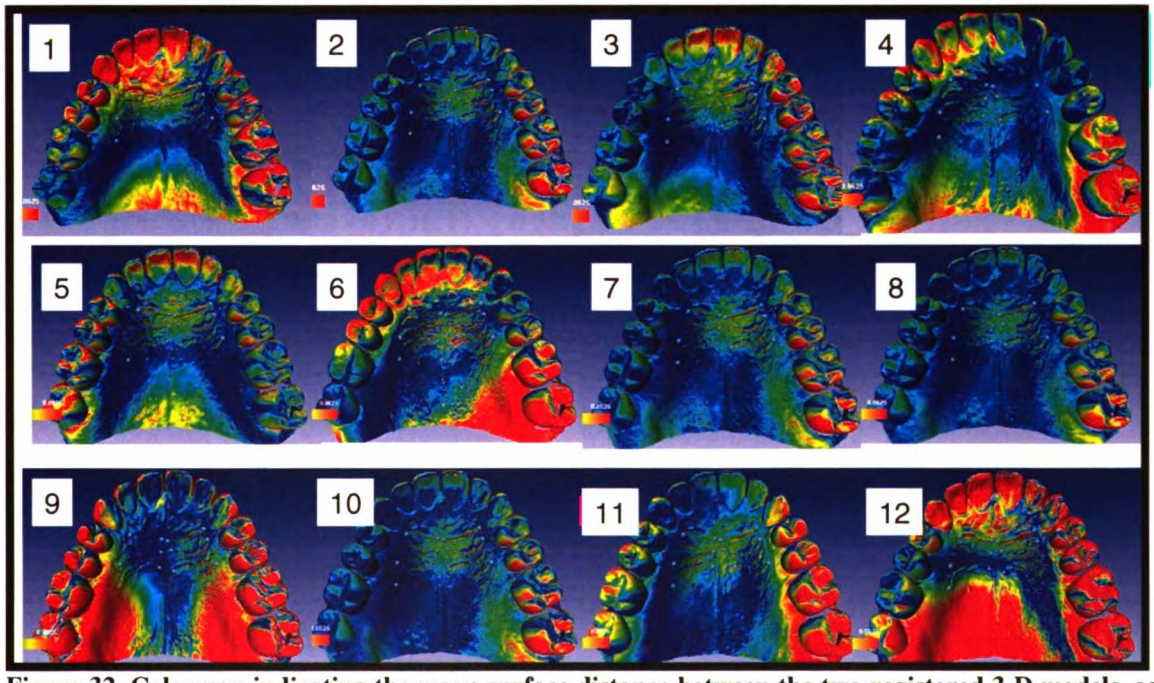


Figure 32. Colormap indicating the mean surface distance between the two registered 3-D models, as one tie point was eliminated. Colormap scale: 0 mm (blue) and 0.0625 mm (red).

Spatial Patterning

After determining that single tie points can affect the registration, it was important to test if particular patterns or orientations of the tie points would affect the overall registration of the two identical models. The results of the spatial patterning is shown below in Table 14 and Figures 33 and 34.

Table 14. Descriptive Statistics for Tie Point Spatial Assessment.

	Tie Points Used	Mean Surface Distance (mm)	Standard Deviation
INSIDE POINTS	2,3,6,7,10,11	0.096	0.081
OUTSIDE POINTS	1,4,5,8,9,12	0.138	0.116
1,4 SERIES	1,4 - 2,3,6,7	0.186	0.154
	1,4 - 2,3,10,11	0.154	0.115
	1,4 - 6,7,10,11	0.156	0.119
5,8 SERIES	5,8 - 2,3,6,7	0.074	0.060
	5,8 - 2,3,10,11	0.065	0.046
	5,8 - 6,7,10,11	0.082	0.064
9,12 SERIES	9,12 - 2,3,6,7	0.076	0.055
	9,12 - 2,3,10,11	0.063	0.051
	9,12 - 6,7,10,11	0.082	0.057

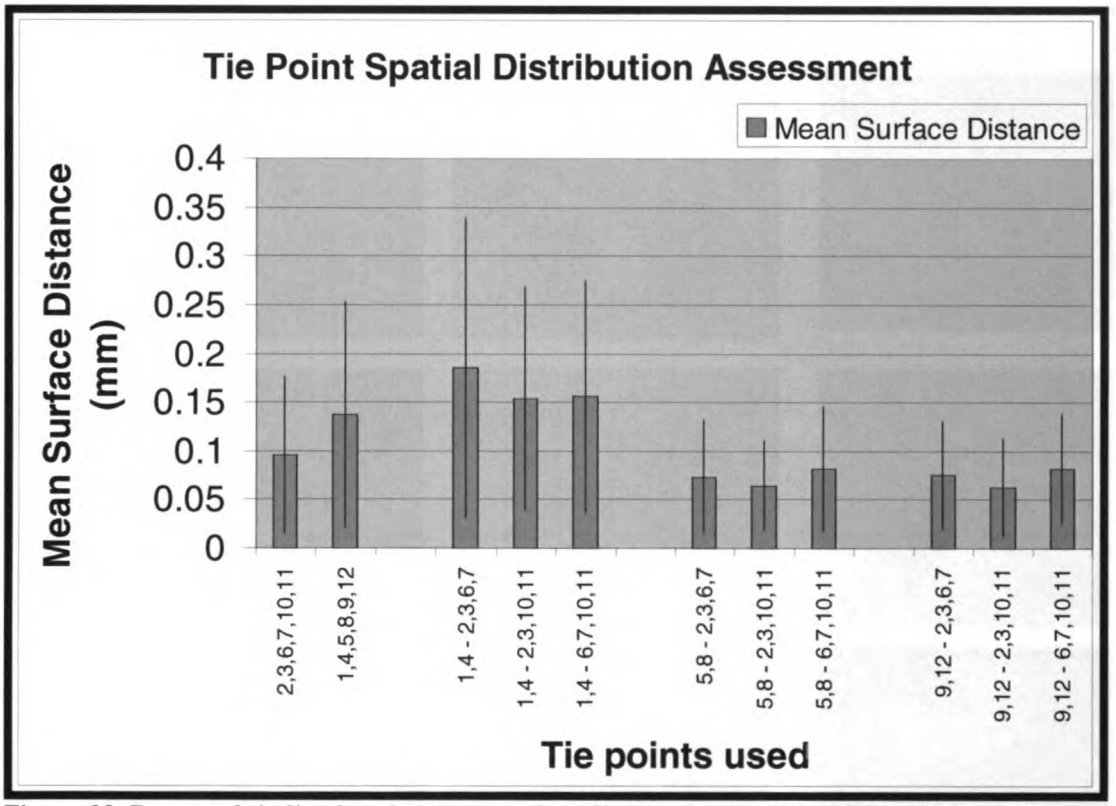


Figure 33. Bar graph indicating the mean surface distance between two 3-D models generated from the white light scanner when different combination of 6 tie points are used.

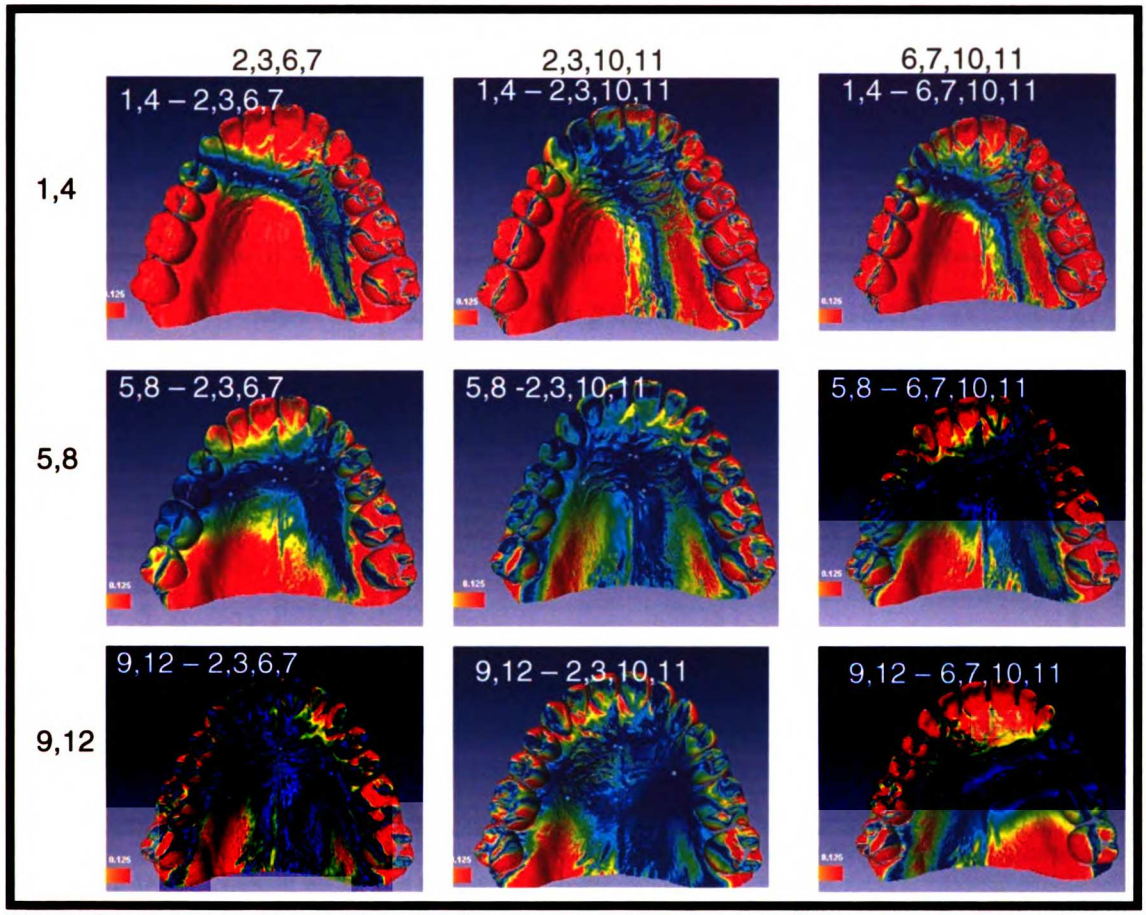


Figure 34. Colormap of the mean surface distances between two 3-D models generated from the white light scanner when different combination of 6 tie points are used. Colormap scale: 0 mm (blue) and 0.125 mm (red).

Operator Error

Operator error was based on the amount of variance when 5 separate 12 tie point registrations were completed. The results are shown below in Table 15 and Figures 35 and 36. As can be seen, the variance (as determined by standard deviation) is quite low. The average of the five mean surface distances between the identical models is 0.034905 mm with a standard deviation of 0.014873.

Table 15 Descriptive Statistics for Operator Error Experiment

	0.024	0.027	0.032	0.030	0.061
	0.017	0.022	0.025	0.024	0.052

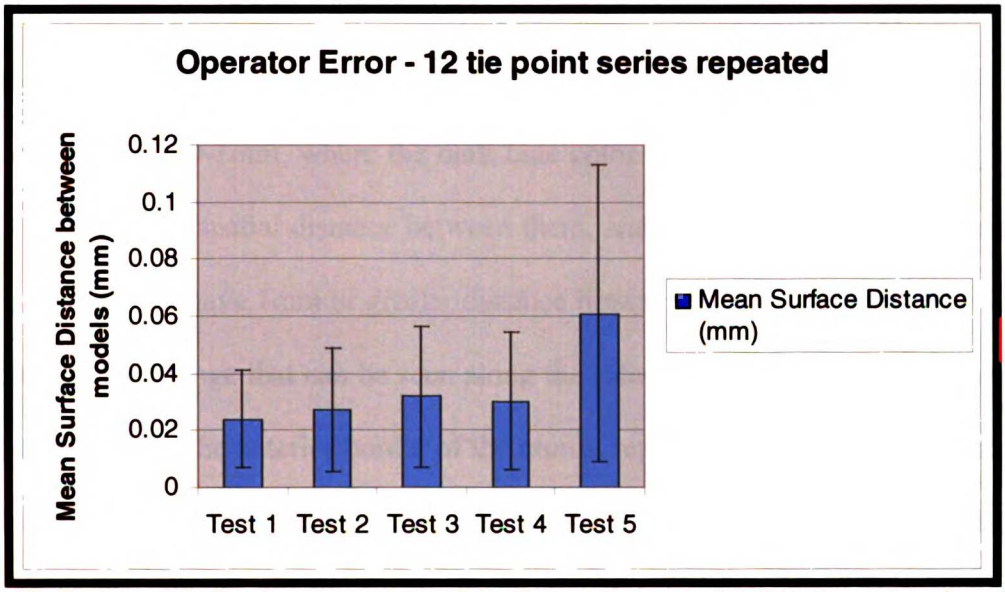


Figure 35. Bar graph indicating the mean surface distance between the five repeated registrations of the same 3-D model using all 12 ties points.

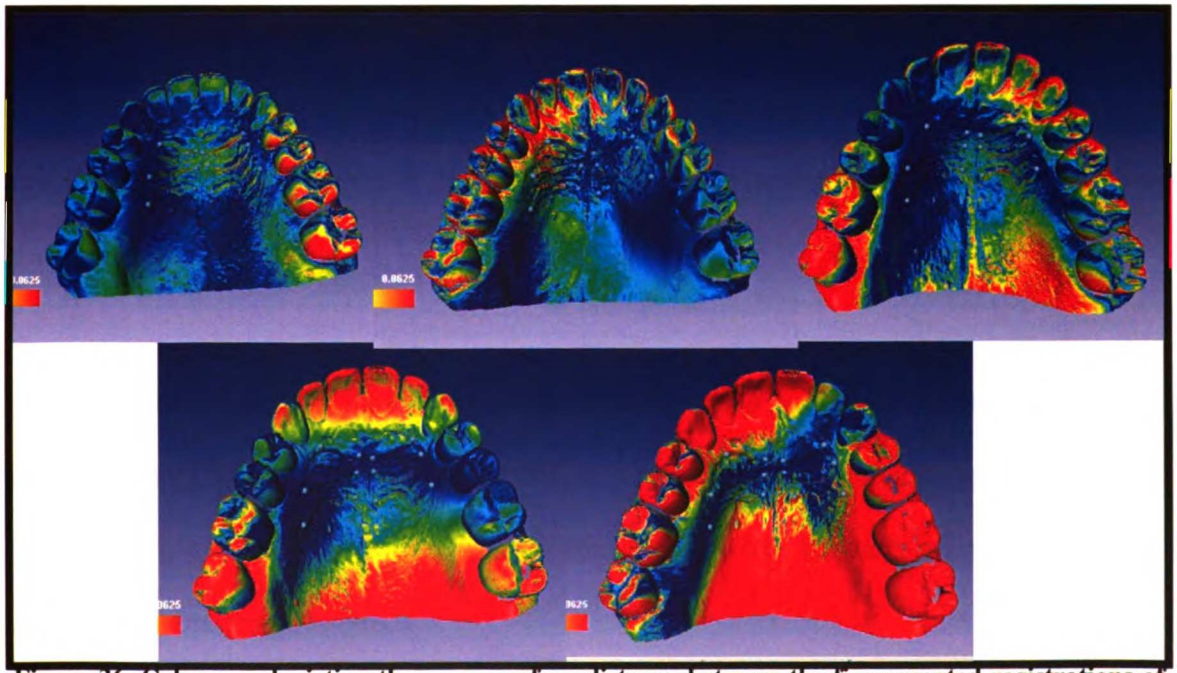


Figure 36. Colormap depicting the mean surface distance between the five repeated registrations of the same 3-D model using all 12 ties points. Scale is 0mm (blue) – 0.0625 mm (red).

Development of Methods to Register Three-Dimensional Data Acquired from the White Light Scanner and the CB MercuRay CBCT

Results from the registration of models generated from the white light scanner and the CBCT portion of the study are shown in Figure 37. Qualitatively, it appears as if the two models register quite well regardless of the number of tie points utilized. The scale of the colormaps is 0-1mm, where the dark blue color represents areas of the registered models that have no spatial distance between them, and the red color represents areas of the two models that have 1mm or greater distance between the two surfaces. The distinct and abrupt color change that can be seen along the inferior border of the entire mandible and vertically along the anterior border of the ramus, represents the edge of the surface of the white light scanned mandible. As a result of the different sized 3-D models, after registering both of them, the portion of the 3-D model that is larger than the other 3-D model will always show colormaps with extremely large surface distances, because the computer program searches for the correct correspondence and never finds one. In addition, mean surface distances between the surfaces were not calculated due to the fact that there were large portions of the models that did not correspond, and therefore skewed the mean surface distance.

Spatially the two three-dimensional models appear to register quite well. This is quite encouraging considering the fact that the models were generated using two different three-dimensional modalities. This is a testament to the accuracy of the two three-dimensional imaging systems.

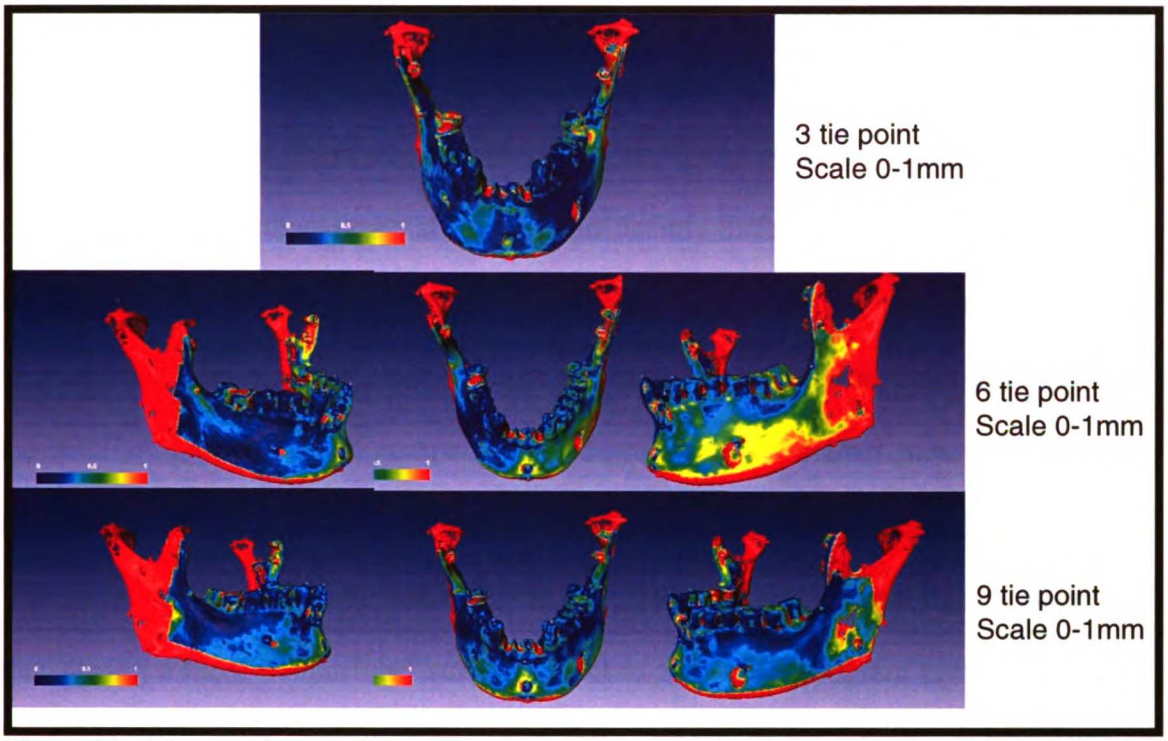


Figure 37. Colormap images depicting the difference in mean surface distance between two 3-D models – one was generated from the white light scanner and the other was generated from the CBCT. The scale is 0 mm (blue) – 0.0625 mm (red)

Discussion

Accuracy of the White Light Scanner

The results from the white light scanner accuracy portion of the study demonstrate that three-dimensional models generated via the white light scanner are fairly representative of the physical model from which it was created. The average physical caliper measurement made in this part of the study was 27.81 mm with a standard deviation of 0.09 mm. With the Amira software to measure the 3-D generated computer models, the average measurement made was 27.73 mm with a standard deviation of 0.11 mm. The absolute mean difference between the physical and computer 3-D measures was 0.22 mm. This discrepancy represents an overall error of less than 1% which

demonstrates the accuracy of this white light scanner. When comparing individual measures, one might expect not to see statistically significant differences between the two measurement groups. However, only 8 of the 27 different measures fit this criterion. As for the 19 measures that were statistically significantly different, this might be explained by the relatively low variance for the 5 measures in each of the groups. Therefore, it can be stated that very repeatable and accurate measures can be made on the 3-D model with the AMIRA program when compared to the gold standard caliper measures; however there seems to be a very slight clinically insignificant reduction in size of the 3-D model. The value of this finding is that scanning a dental cast with a white light scanner and transferring a three-dimensional surface model to the Amira program provides a highly accurate computer based method to store, and do research on orthodontic study models.

Accuracy of the CB MercuRay Cone Beam CT

The results from the CBCT accuracy portion of the study demonstrate that the three-dimensional model generated from the CBCT is fairly representative of the physical model from which it was created. The average physical caliper measurement made was 32.26 mm with a standard deviation of 0.04 mm. With the Amira software, the average measurement made was 31.93 mm with a standard deviation of 0.09 mm. The absolute mean difference between the physical and computer 3-D measures was 0.38 mm. This discrepancy represents a 1.19% overall error, which is extremely low and suggests that our CBCT scanned data is highly accurate.

It is interesting to note that the overall standard deviations of the measurement made in this portion of the study are less than in the white light scanner accuracy portion of the study. This might be due to the fact that all of the measurements made in this portion of the study were performed exclusively on steel fiducials, and therefore could be easily recognized and referenced in both the physical caliper measurements and the three-dimensional computer measurements. Only a subset of the measurements in the white light scanner accuracy portion of the study was performed with steel fiducials, while the rest of the measurements were made on cusp tips, fossae, and embrasures.

Comparison of Various Methods to Register Three-Dimensional Computer Generated Dental Models

In this portion of the study, all series of registrations utilized identical models in order to assess the error associated with the various registration techniques. This conscious decision was made in an effort to control for possible errors due to impression technique, impression materials, distortion from impression material, stone set differences, etc.

When utilizing Amira's automatic alignment function, the mean surface distance between the identical models was extremely small (0.000311 mm ± 0.000158 mm). Since identical models were utilized, this value should theoretically approach zero, and therefore the value obtained can be considered the error associated with the software superimposition technique.

Operator error was assessed by repeating the 12 tie point registration process five times. The average of the five mean surface distances between the identical models was 0.035 mm with a standard deviation of 0.015 mm. The standard deviation represents the operator error, and is well below any clinically relevant measure. In fact, the true operator error may be considered the difference between the standard deviation and the error associated with the automatic alignment, since the automatic alignment standard deviation can be considered the error associated with the software superimposition technique. From a clinical perspective, this amount of error is negligible.

It should be noted that the operator error calculated is a function of the superimposition technique by which 12 tie points were chosen repeatedly, and the models were registered. It is not the error of picking a single point on the model or two points as would be the error of making a linear measurement (as in part 1 and part 2 of the study). Although the calculation of two point operator error (linear distance) was not performed, this error is directly correlated with the overall standard deviations of the measurements in both part 1 and part 2 of this project.

By increasing the number of tie points during the registration process, the accuracy of the superimposed models was increased. As shown graphically, the trend tends toward what was calculated as the error of the method. In fact, when utilizing 9 and 12 tie points, the mean surface distance between the superimposed models was very near the average operator error. This is not surprising, since the operator error was based on registrations with 12 tie points. However, it does demonstrate that as one utilized

more tie points when registering two models, the registration is more accurate. A minimum number of tie points that gives a highly accurate registration are 5 or 6 tie points.

As was mentioned in the methods section, performing a series of tie point registrations whereby the order of tie points chosen in the increasing number series was reversed was important to eliminate any coincidental trends caused by the original increasing order. As was demonstrated by the "reverse order" series, the overall trend was similar to the increasing number series. However, some of the differences, especially in the 3 tie point registrations, demonstrate the variability that can occur if too few tie points are utilized.

The "elimination series" of registrations was performed to determine if a particular tie point position affected the overall registration process positively or negatively. In other words, by eliminating a single tie point and assessing the resulting superimposition, one could ascertain how important the eliminated tie point was to the registration. The results showed that the most important tie point to the registration process was tie point number 12. This was determined by eliminating tie point number 12 and finding that the resulting registration had the largest mean surface distance between the two identical models. Quite simply, it had the least accurate registration. Following tie point number 12 in relative importance are tie points 9, 6, 1, 5, 3, 11, 4, 7, 2, 10 and finally 8. It is interesting to note when assessing the locations of the 4 most important tie points, that 3 out of the four are lateral tie points and one is a medial tie

point. This is the exact opposite for the 4 least important tie points, where 3 out of the 4 tie points are medial and 1 is a lateral palatal rugae tie point. This suggests that the spatial orientation of the individual tie points in relation to the others does have an impact on the registration. Although it would be very tempting to state that this exact trend might apply to all maxillary 3-D model registrations, the results invariably depend on the unique three dimensional pattern of the palatal region on this particular 3-D model. The relationship of at least one lateral tie point to an accurate registration argues to the location of tie points and may account for potential tipping or canting of the registered models in relation to each other.

In an attempt to determine what role various patterns of tie points may have on the overall registration technique, registrations were performed on identical models with various patterns and orientations of the tie points. Six tie points were selected for this series based on two factors. First, utilizing six tie points would create many different patterns because half of the 12 potential tie points would not be used. In addition, as a result of the "increasing number of tie point series" the results from the 6 tie point registrations resulted in a very accurate registration. Next the decision to utilize 4 medial tie points and 2 lateral tie points was based on several clinical studies that found that medial palatal rugae points were very stable even during orthodontic treatment (Almeida, Phillips et al. 1995), and that the third or posterior lateral palatal rugae points were also stable during orthodontic treatment (Bailey, Esmailnejad et al. 1996). In addition, it was decided to have pairs of tie points. Although the results of these studies directed the decision to utilize 4 medial and 2 lateral tie points, all combinations of patterns were

tested regardless of the fact that Bailey et, al. found that only the posterior/third lateral palatal rugae point was stable during orthodontic treatment. Finally, two more registrations were performed on all medial palatal rugae points and all lateral palatal rugae points.

The results of this series of tests demonstrate that when four medial tie points (2 pairs of 2) are utilized, the anterior and posterior pairs produce the most accurate registrations. In regards to the lateral tie point pair, the second (middle) and third (posterior) palatal rugae points produce the most accurate registrations. It is hypothesized that these tie points are spatially the furthest from one another and therefore generate fairly accurate registrations. Future testing will need to be performed to determine this fact.

The registration of three-dimensional models generated from the white light scanner and the CBCT was fairly simple using the AMIRA software. However, developing a quantitative means to determine the accuracy of the registrations was more difficult. Due to the varying size of the models generated with the two different modalities, mean surface measurements between the superimposed models was meaningless. Colormaps were created and are useful to qualitatively assess how well the registration was performed, however, it is very difficult to compare them to each other. As mentioned previously, one encouraging byproduct of the registrations is the 3-D models generated from two different modalities superimposed fairly well, suggesting they are accurate representations of each other and the physical model from which they

were generated. Further work needs to be completed in this area to accurately and reliably register three-dimensional models generated from the white light scanner and the CBCT.

Conclusions

Very repeatable and accurate measurements can be made with the AMIRA software program on both the three-dimensional model generated from the white light scanner, and the CB MercuRay CBCT. There is a clinically insignificant reduction in size of the 3-D models, when compared to the physical gold standard. There was a 0.80% error and a 1.19% error respectively when comparing all measurements from both the 3-D model and the physical model.

Registration accuracy increases with increasing the number of tie points utilized. When performing registrations in this manner, the importance of individual tie points to the overall registration process can be assessed by separately eliminating individual tie points. It appears that registration accuracy increases as the tie points used are separated spatially. In other words, a spatial pattern that optimizes the distance between the tie points in all dimensions (x, y, and z axis) will produce the most accurate registration. Operator error is quite low and clinically insignificant with the tie point registration technique.

Finally, registration of 3-D models generated from the CBCT and the white light scanner can be performed. However, no conclusive recommendations can be made in

References

Abdel-Aziz, H. M. and N. E. Sabet (2001). "Palatal rugae area: a landmark for analysis of pre- and post-orthodontically treated adult Egyptian patients." East Mediterr Health J **7**(1-2): 60-6.

Almeida, M. A., C. Phillips, et al. (1995). "Stability of the palatal rugae as landmarks for analysis of dental casts." Angle Orthod **65**(1): 43-8.

Araki, K., K. Maki, et al. (2004). "Characteristics of a newly developed dentomaxillofacial X-ray cone beam CT scanner (CB MercuRay): system configuration and physical properties." Dentomaxillofac Radiol **33**(1): 51-9.

Ashmore, J. L., B. F. Kurland, et al. (2002). "A 3-dimensional analysis of molar movement during headgear treatment." Am J Orthod Dentofacial Orthop **121**(1): 18-29; discussion 29-30.

Ayoub, A. F., D. Wray, et al. (1997). "A three-dimensional imaging system for archiving dental study casts: a preliminary report." Int J Adult Orthodon Orthognath Surg **12**(1): 79-84.

Bailey, L. T., A. Esmailnejad, et al. (1996). "Stability of the palatal rugae as landmarks for analysis of dental casts in extraction and nonextraction cases." Angle Orthod **66**(1): 73-8.

Bell, A., A. F. Ayoub, et al. (2003). "Assessment of the accuracy of a three-dimensional imaging system for archiving dental study models." J Orthod **30**(3): 219-23.

Bjork, A. (1968). "The use of metallic implants in the study of facial growth in children: method and application." Am J Phys Anthropol **29**(2): 243-54.

Broadbent, B. (1937). "The Face of the Normal Child." Am J Orthod(7): 183-208.

English, W. R., S. F. Robison, et al. (1988). "Individuality of human palatal rugae." J Forensic Sci **33**(3): 718-26.

Ford, E. (1958). "Growth of the Human Cranial Base." Am J Orthod(44): 498-506.

Hoggan, B. R. and C. Sadowsky (2001). "The use of palatal rugae for the assessment of anteroposterior tooth movements." Am J Orthod Dentofacial Orthop **119**(5): 482-8.

Lascala, C. A., J. Panella, et al. (2004). "Analysis of the accuracy of linear measurements obtained by cone beam computed tomography (CBCT-NewTom)." Dentomaxillofac Radiol **33**(5): 291-4.

Lee, J. Y., Q. Han, et al. (2004). "Three-dimensional facial imaging: accuracy and considerations for clinical applications in orthodontics." Angle Orthod **74**(5): 587-93.

Littlefield, T. R., J. C. Cherney, et al. (2005). "Comparison of plaster casting with three-dimensional cranial imaging." Cleft Palate Craniofac J **42**(2): 157-64.

Mah, J. and D. Hatcher (2003). "Current status and future needs in craniofacial imaging." Orthod Craniofac Res **6**(Suppl 1): 10-6; discussion 179-82.

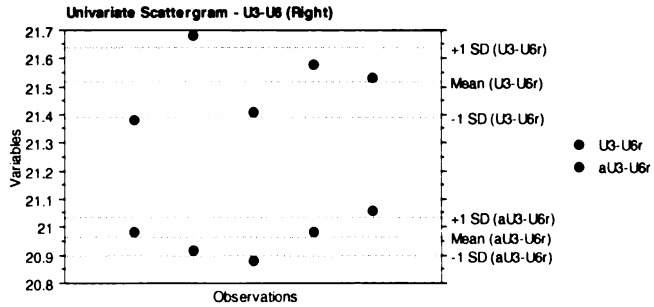
Moyers, R. E. (1976). Standards of Human Occlusal Development. Ann Arbor, Mich., Center for Human Growth and Development.

Mozzo, P., C. Procacci, et al. (1998). "A new volumetric CT machine for dental imaging based on the cone-beam technique: preliminary results." Eur Radiol **8**(9): 1558-64.

Quimby, M. L., K. W. Vig, et al. (2004). "The accuracy and reliability of measurements made on computer-based digital models." Angle Orthod **74**(3): 298-303.
Ricketts RM, S. R., Bagha L (1976). "Orientation Sella-Nasion or Frankfort Horizontal." Am J Orthod(69): 648-54.
Santoro, M., S. Galkin, et al. (2003). "Comparison of measurements made on digital and plaster models." Am J Orthod Dentofacial Orthop **124**(1): 101-5.
Shah, S., G. Sundaram, et al. (2004). "The use of a 3D laser scanner using superimpositional software to assess the accuracy of impression techniques." J Dent **32**(8): 653-8.
Steiner, C. (1953). "Cephalometrics for You and Me." Am J Orthod(39): 729-55.
Sukovic, P. (2003). "Cone beam computed tomography in craniofacial imaging." Orthod Craniofac Res **6**(Suppl 1): 31-6; discussion 179-82.
TenCate, A. R. (1998). Oral Histology; Development, Structure, and Function. St. Louis, Mosby.
Vannier, M. W. (2003). "Craniofacial computed tomography scanning: technology, applications and future trends." Orthod Craniofac Res **6**(Suppl 1): 23-30; discussion 179-82.

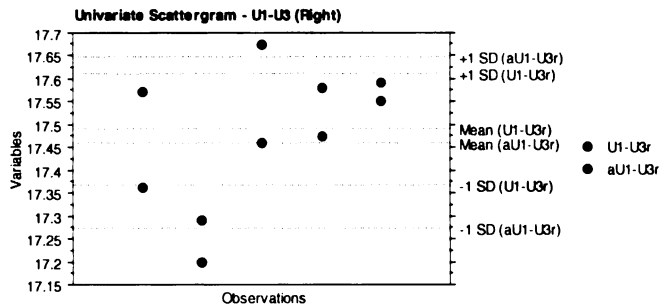
Appendix 1

Univariate plots of each pair of measurements taken to assess the accuracy of measurements made on a 3-D surface generated from the white light scanner as compared to the gold standard physical measurements made with digital calipers.



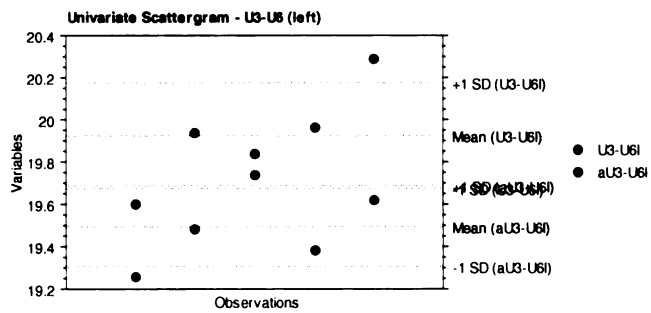
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U3-U6r, aU3-U6r	.550	4	8.843	.0009	.378	.723



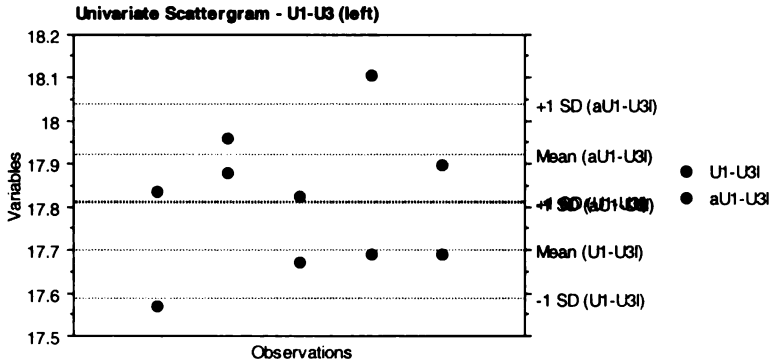
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U1-U3r, aU1-U3r	.030	4	.411	.7021	-.171	.231



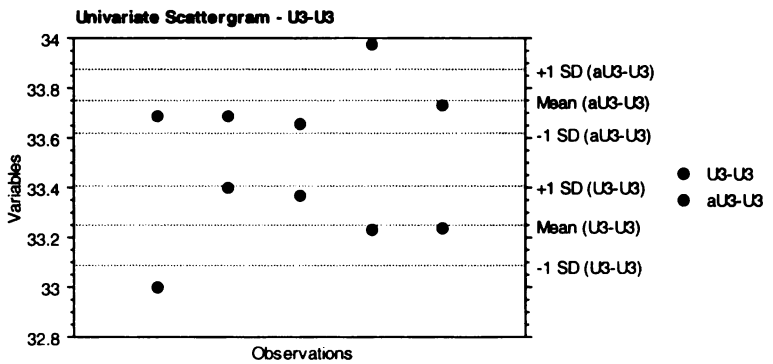
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U3-U6l, aU3-U6l	.430	4	4.327	.0124	.154	.706



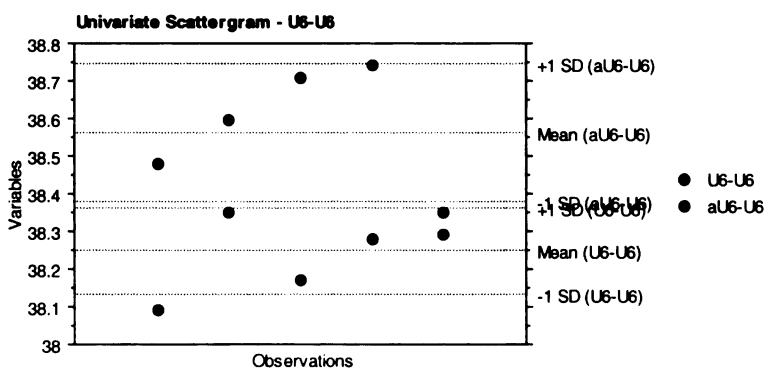
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U1-U3l, aU1-U3l	-.224	4	-3.981	.0164	-.381	-.068



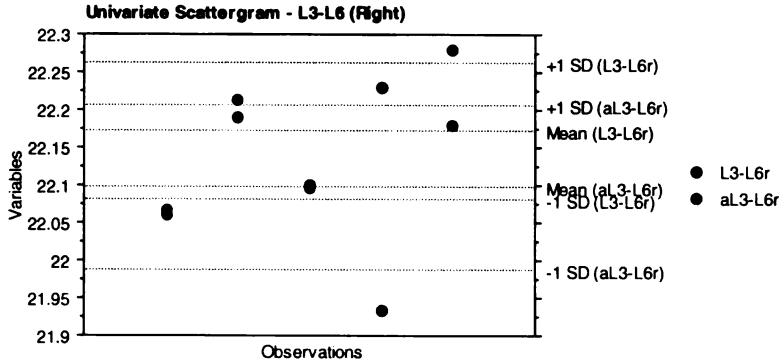
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U3-U3, aU3-U3	-.501	4	-5.204	.0065	-.768	-.234



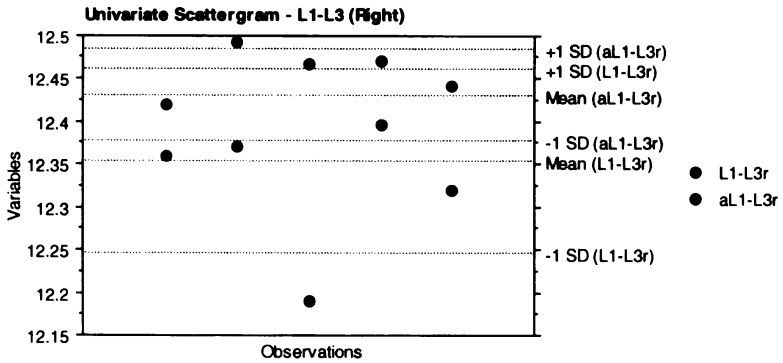
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U6-U6, aU6-U6	-.315	4	-2.996	.0401	-.607	-.023



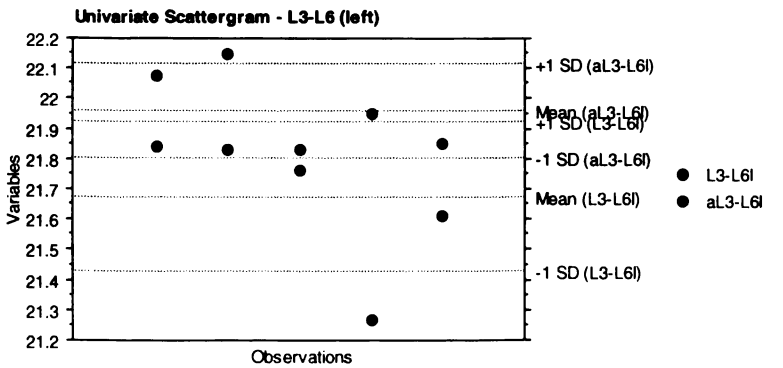
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L3-L6r, aL3-L6r	.075	4	1.258	.2768	-.090	.239



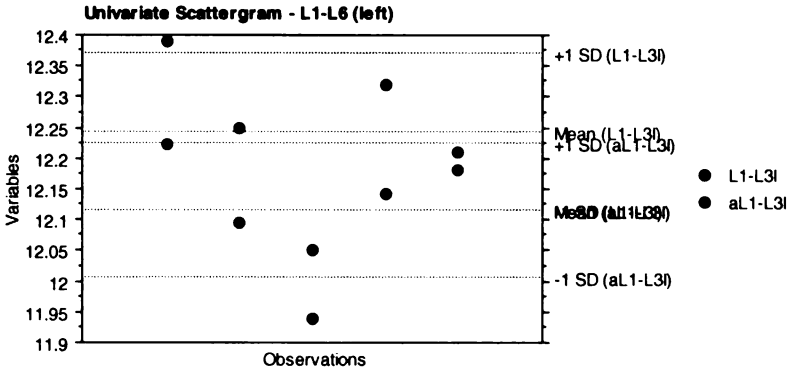
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L1-L3r, aL1-L3r	-.077	4	-1.178	.3041	-.259	.105



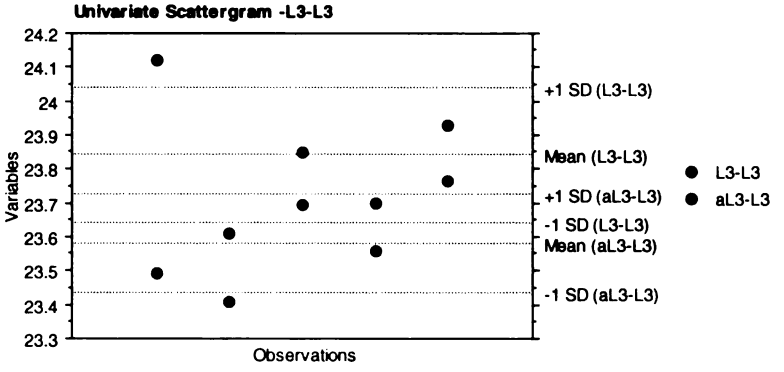
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L3-L6l, aL3-L6l	-.283	4	-2.364	.0773	-.615	.049



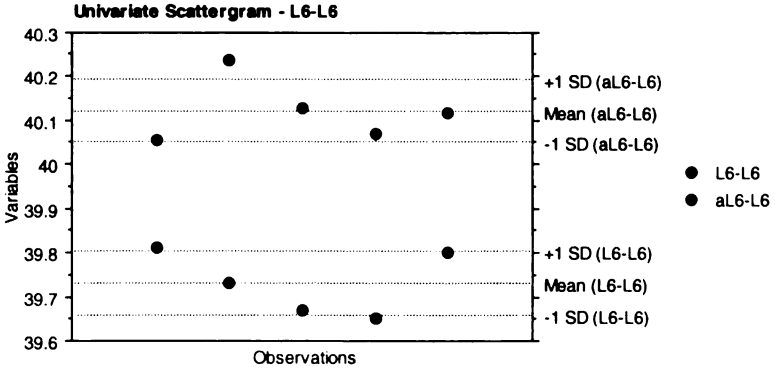
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L1-L3l, aL1-L3l	.128	4	4.686	.0094	.052	.204



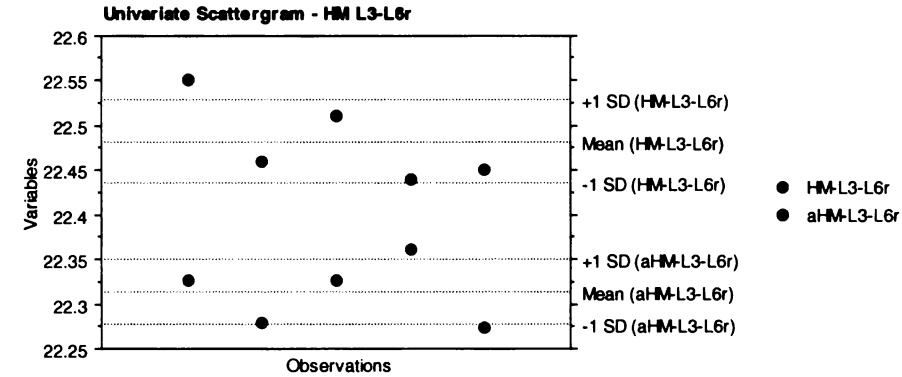
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L3-L3, aL3-L3	.259	4	2.798	.0489	.002	.517



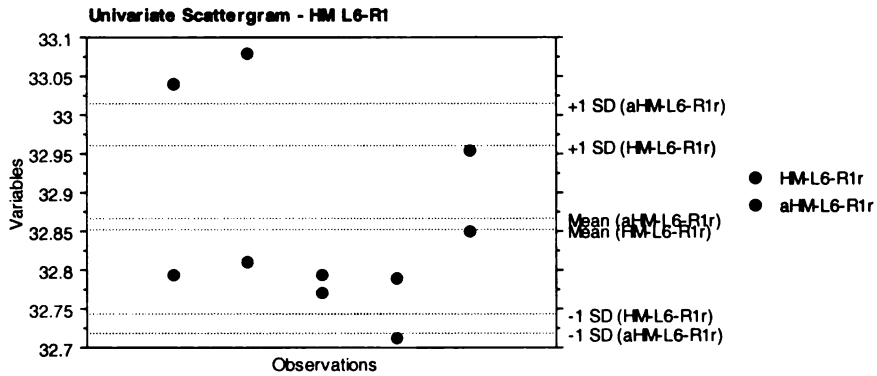
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L6-L6, aL6-L6	-.391	4	-8.218	.0012	-.523	-.259



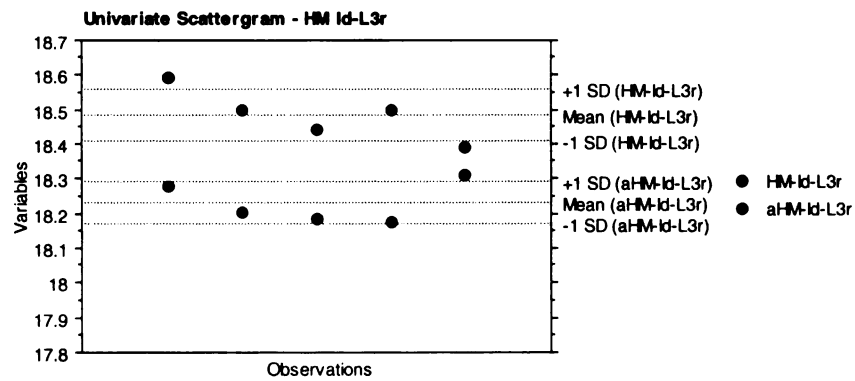
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-L3-L6r, aHM-L3-L6r	.169	4	7.026	.0022	.102	.236



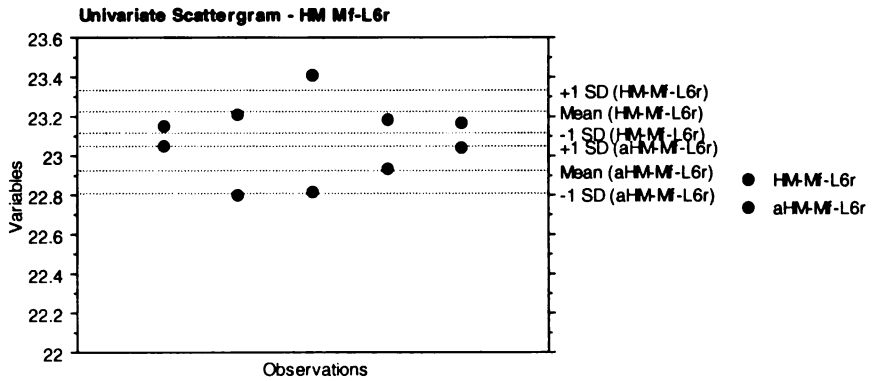
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-L6-R1r, aHM-L6-R1r	-.015	4	-.173	.8709	-.255	.225



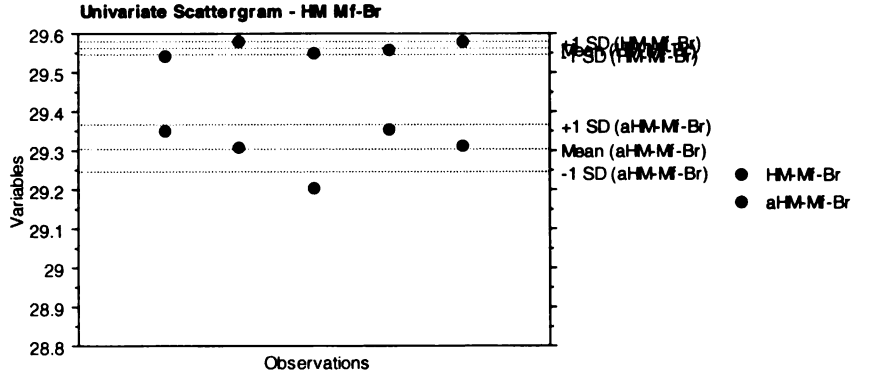
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-Id-L3r, aHM-Id-L3r	.253	4	5.596	.0050	.128	.379



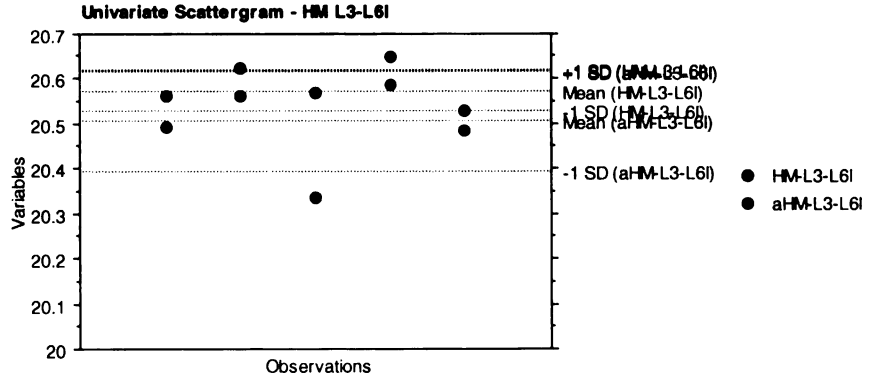
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-Mf-L6r, aHM-Mf-L6r	.297	4	3.218	.0323	.041	.552



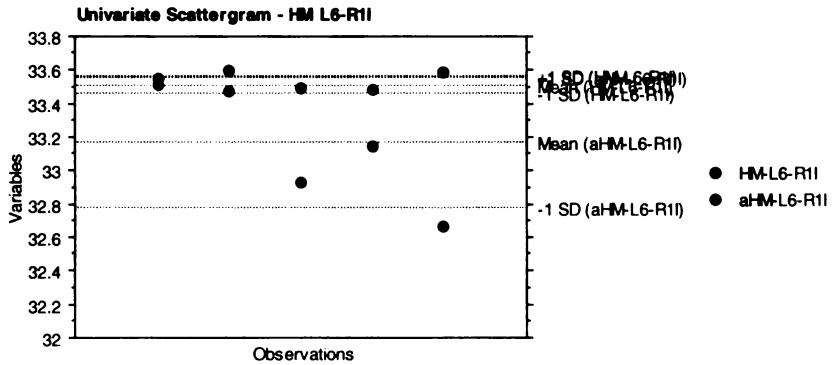
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-Mf-Br, aHM-Mf-Br	.256	4	9.255	.0008	.179	.333



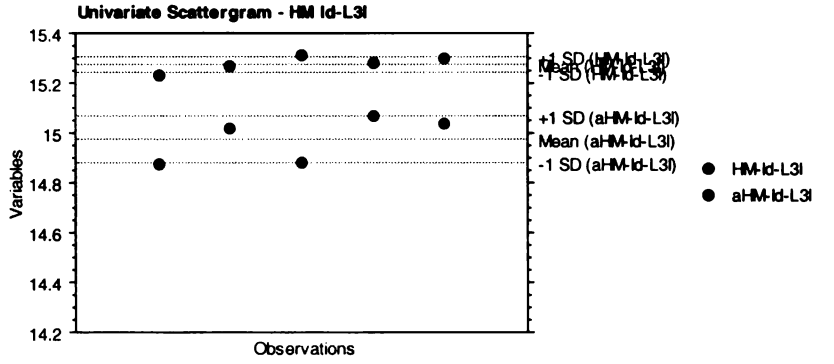
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-L3-L6l, aHM-L3-L6l	.069	4	1.454	.2197	-.063	.201



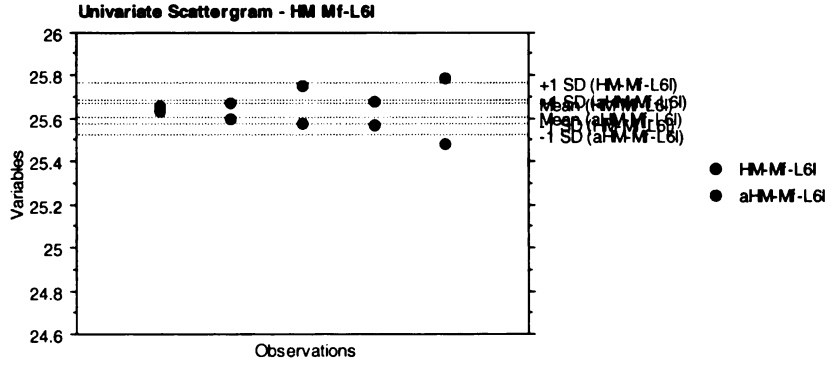
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-L6-R1I, aHM-L6-R1I	.346	4	1.865	.1356	-.169	.862



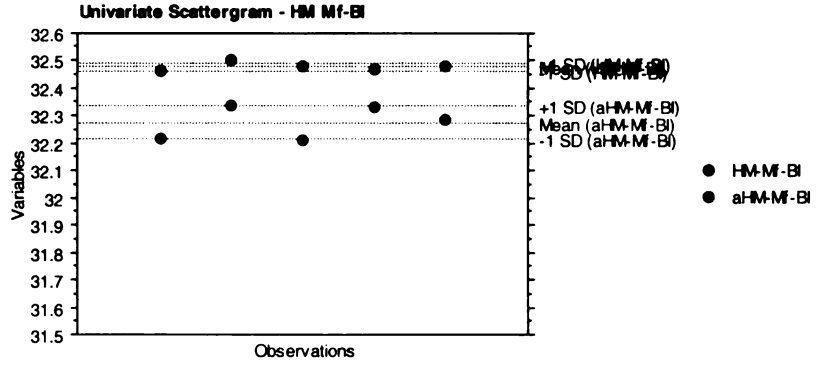
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-Id-L3I, aHM-Id-L3I	.302	4	7.566	.0016	.191	.413



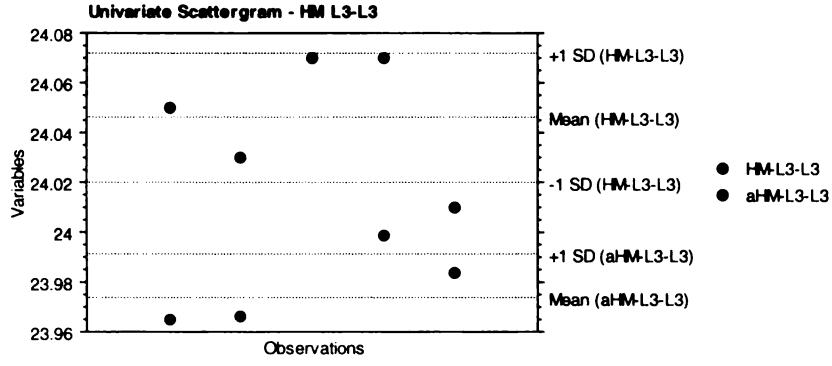
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-Mf-L6I, aHM-Mf-L6I	.066	4	.860	.4385	-.147	.279



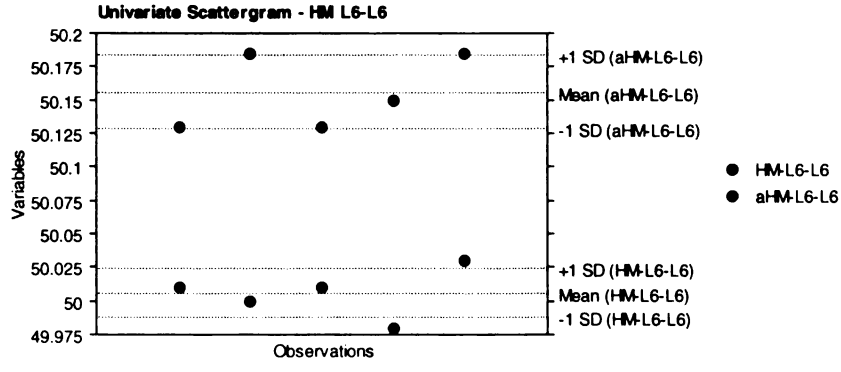
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-Mf-Bl, aHM-Mf-Bl	.202	4	8.300	.0012	.134	.269



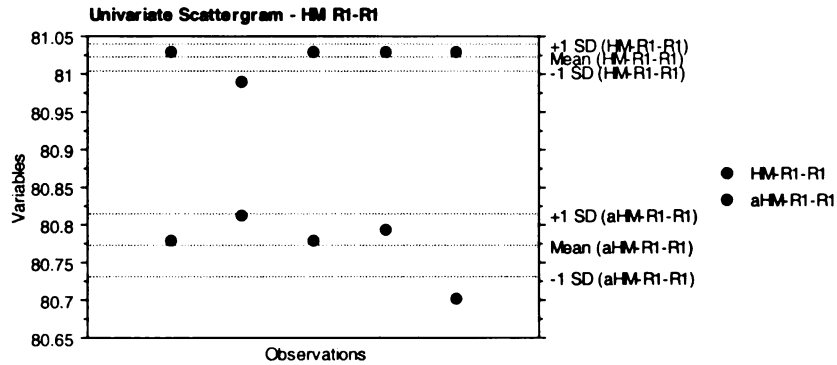
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-L3-L3, aHM-L3-L3	.072	4	5.020	.0074	.032	.112



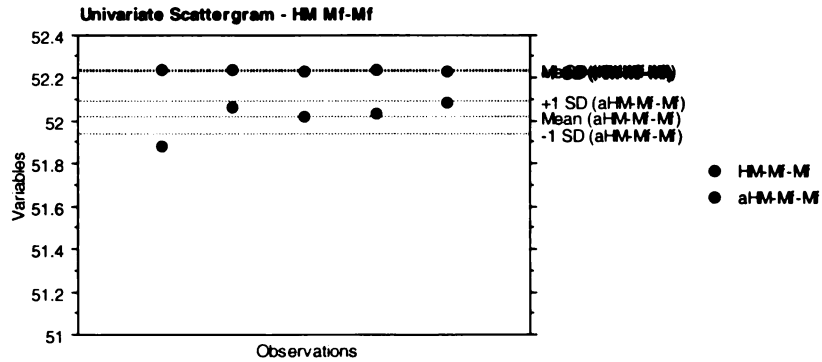
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-L6-L6, aHM-L6-L6	-.150	4	-11.421	.0003	-.186	-.114



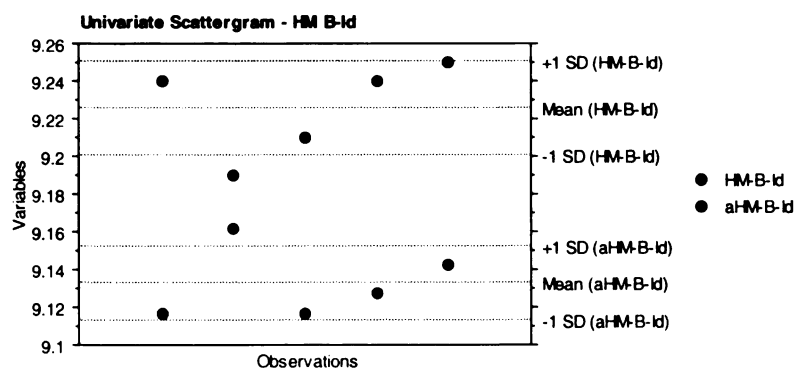
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-R1-R1, aHM-R1-R1	.248	4	10.385	.0005	.182	.315



Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-Mf-Mf, aHM-Mf-Mf	.218	4	6.007	.0039	.117	.319

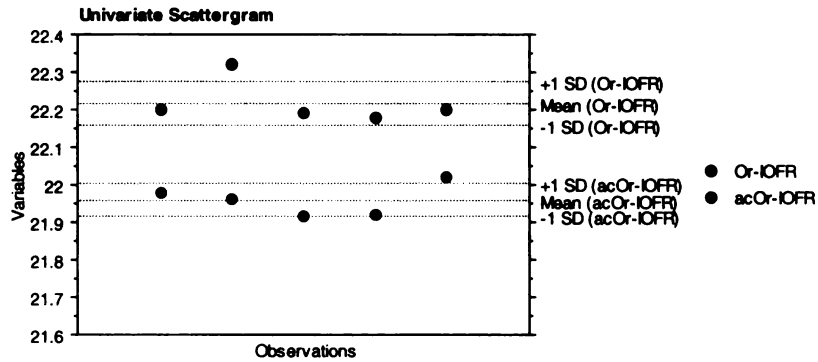


Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
HM-B-Id, aHM-B-Id	.093	4	5.525	.0052	.046	.140

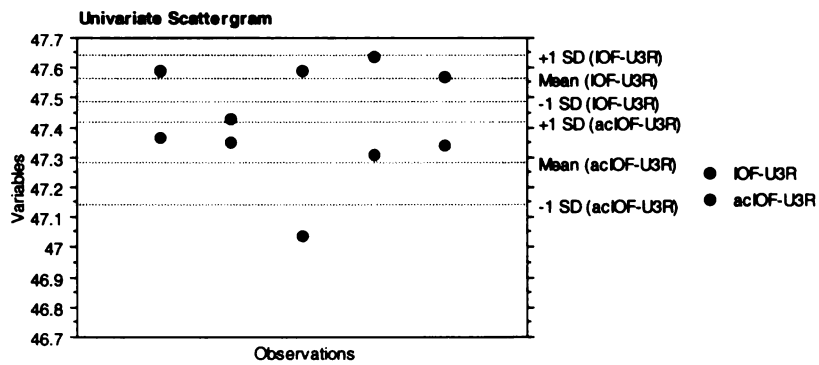
Appendix 2

Univariate plots of each pair of measurements taken to assess the accuracy of measurements made on a 3-D surface generated from the white light scanner as compared to the gold standard physical measurements made with digital calipers



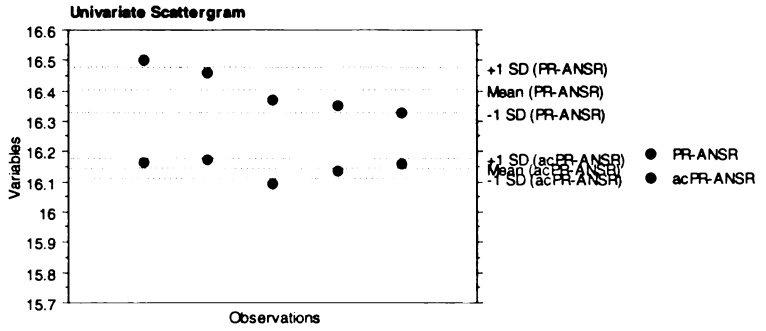
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Or-IOFR, acOr-IOFR	.259	4	8.675	.0010	.176	.342



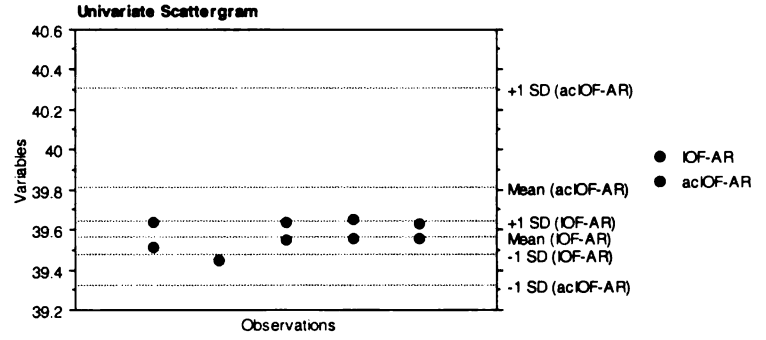
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
IOF-U3R, actOF-U3R	.282	4	3.599	.0228	.065	.500



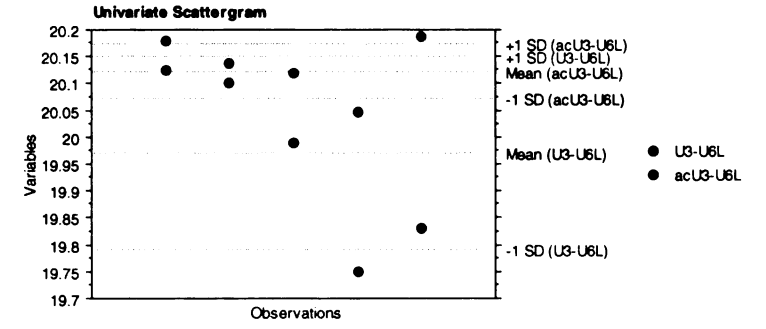
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
PR-ANSR, acPR-ANSR	.257	4	8.734	.0009	.175	.338



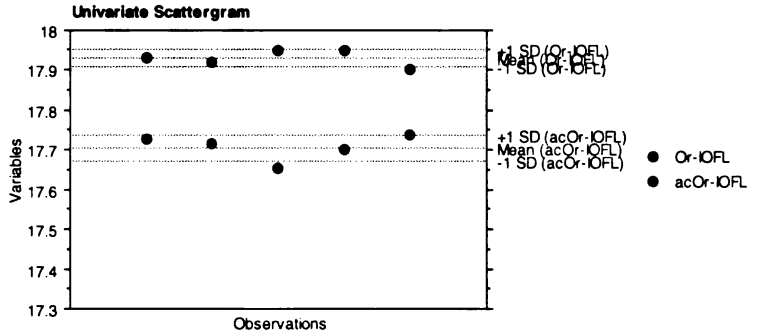
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
IOF-AR, acIOF-AR	-.252	4	-1.001	.3734	-.949	.446



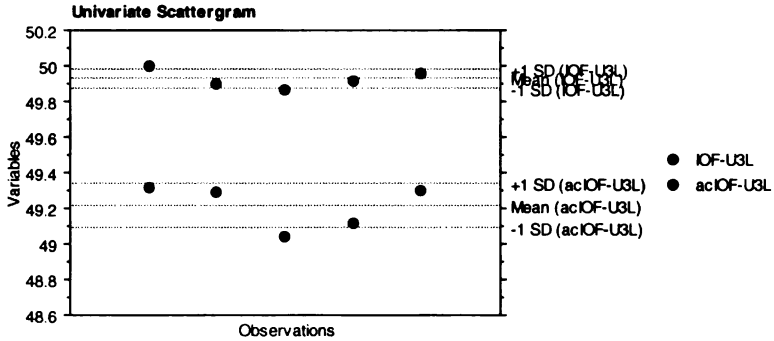
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U3-U6L, acU3-U6L	-.153	4	-1.979	.1189	-.367	.062



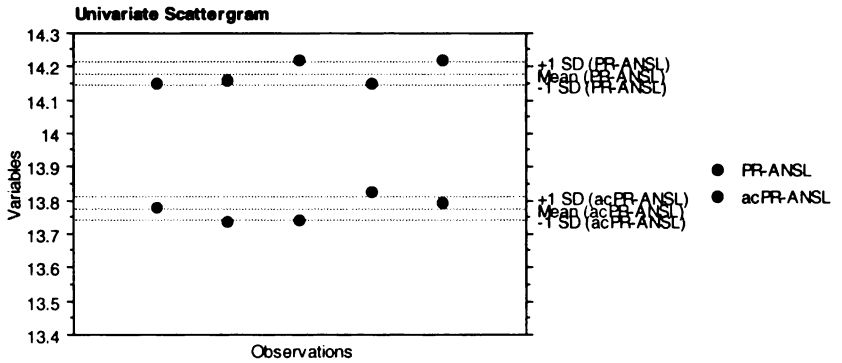
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Or-IOFL, acOr-IOFL	.224	4	9.866	.0006	.161	.287



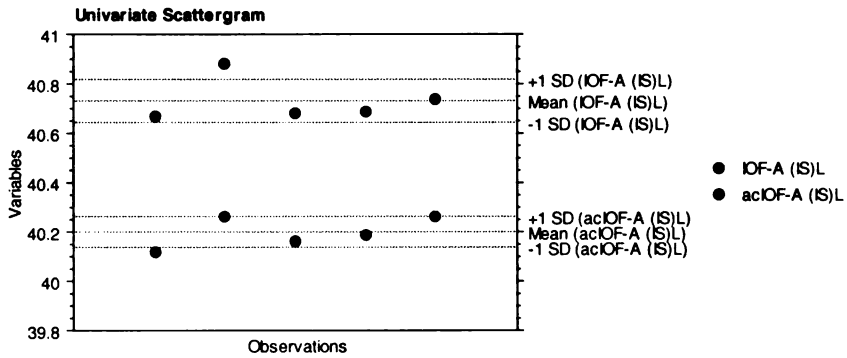
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
IOF-U3L, acIOF-U3L	.717	4	16.996	<.0001	.600	.834



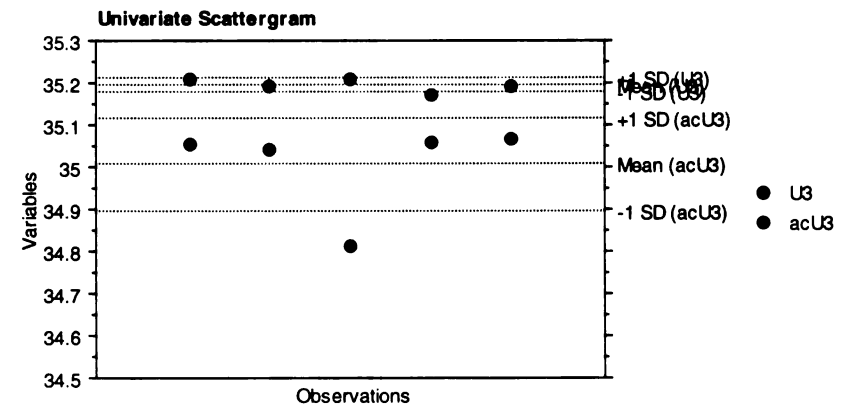
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
PR-ANSL, acPR-ANSL	.403	4	15.386	.0001	.330	.475



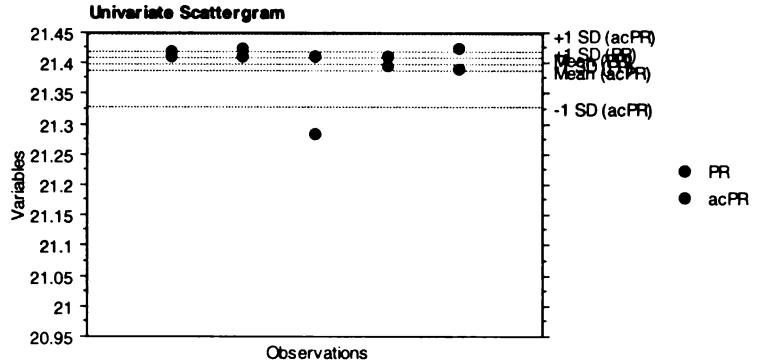
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
IOF-A (IS)L, acIOF-A (IS)L	.533	4	21.533	<.0001	.464	.602



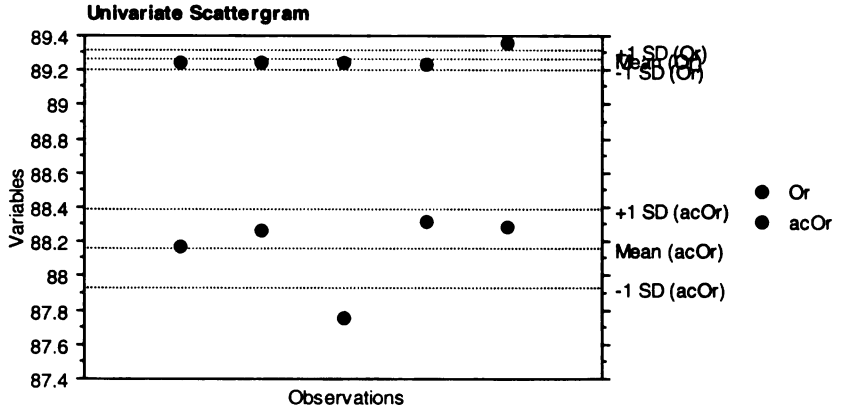
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
U3, acU3	.188	4	3.530	.0242	.040	.335



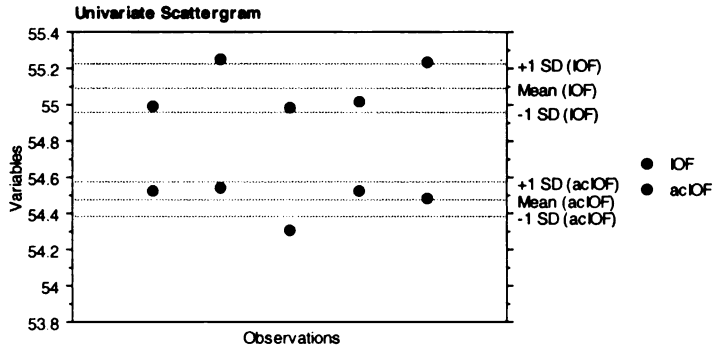
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
PR, acPR	.020	4	.726	.5080	-.057	.097



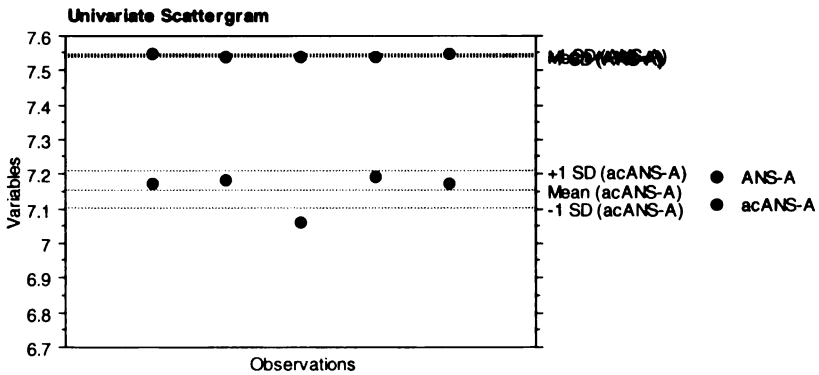
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Or, acOr	1.104	4	11.158	.0004	.829	1.378



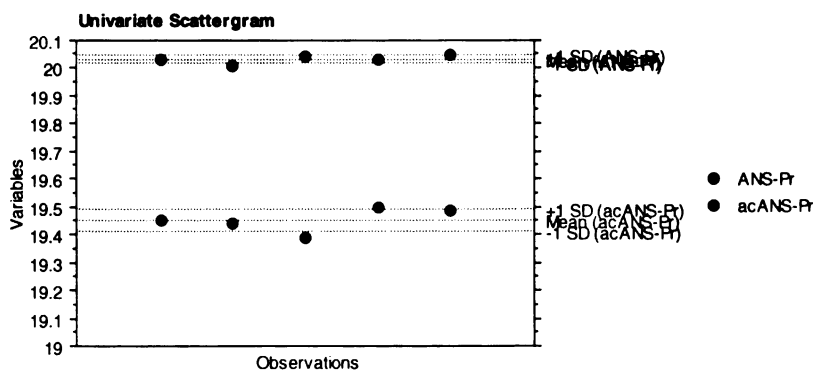
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
IOF, acIOF	.617	4	10.757	.0004	.458	.777



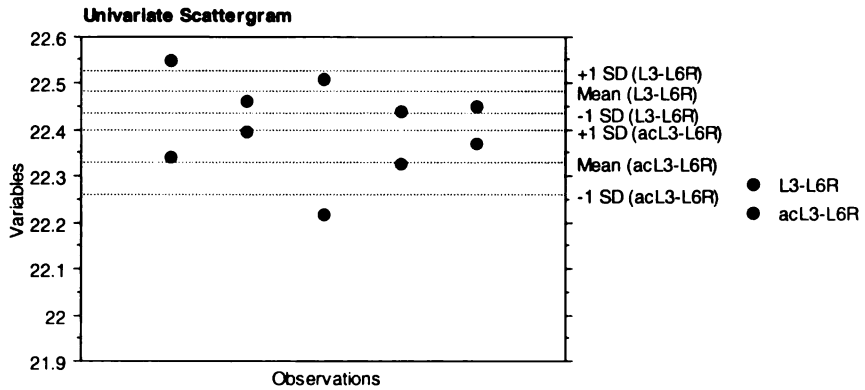
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
ANS-A, acANS-A	.388	4	16.305	<.0001	.322	.454



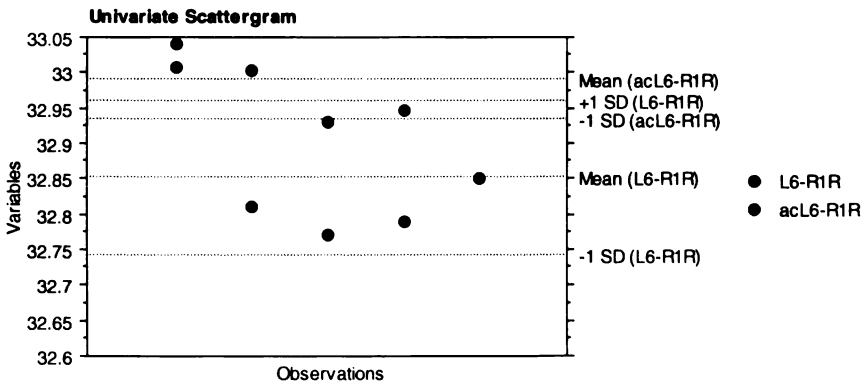
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
ANS-Pr, acANS-Pr	.579	4	30.512	<.0001	.526	.631



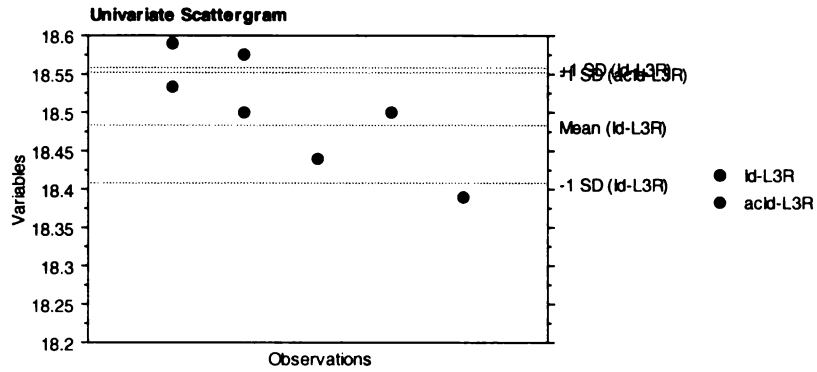
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L3-L6R, acL3-L6R	.152	4	3.541	.0240	.033	.271



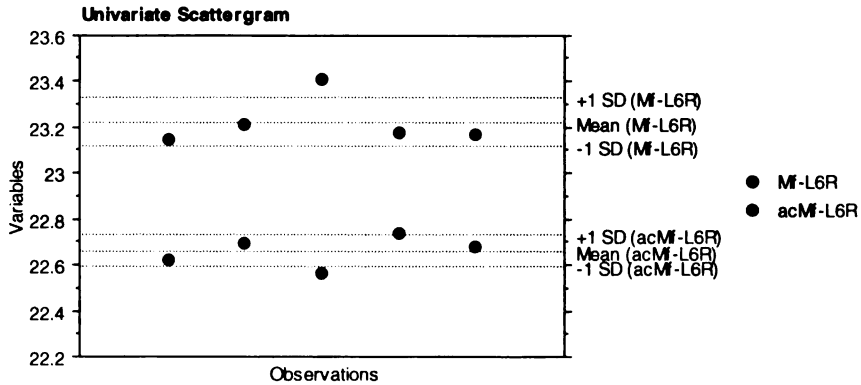
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L6-R1R, acL6-R1R	-.140	4	-3.126	.0353	-.265	-.016



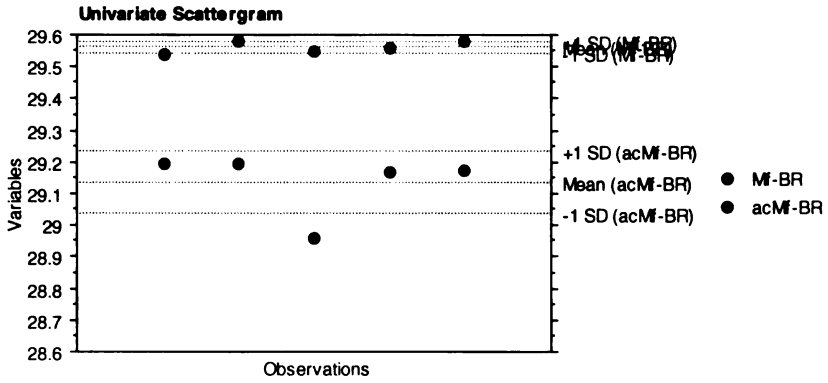
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
ld-L3R, acld-L3R	-.197	4	-2.437	.0714	-.421	.027



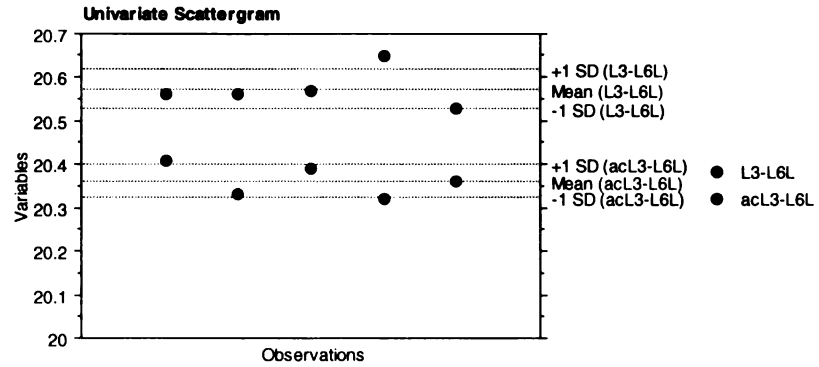
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Mf-L6R, acMf-L6R	.563	4	7.794	.0015	.363	.764



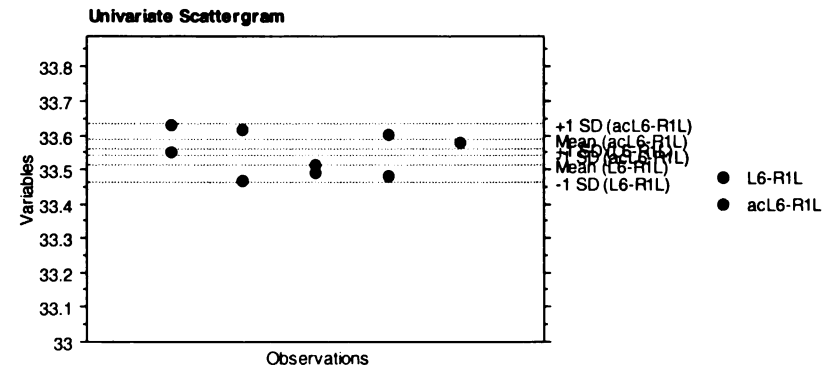
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Mf-BR, acMf-BR	.424	4	10.026	.0006	.307	.542



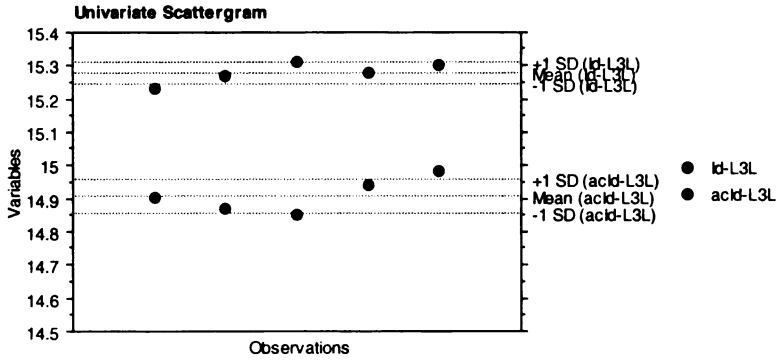
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L3-L6L, acL3-L6L	.212	4	6.556	.0028	.122	.301



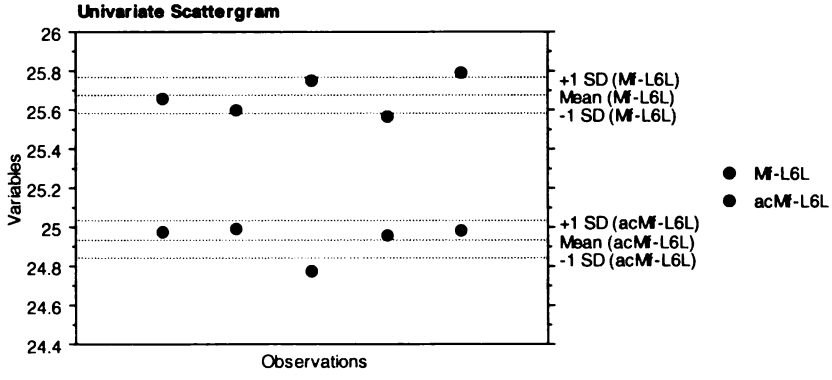
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L6-R1L, acL6-R1L	-.076	4	-2.702	.0540	-.154	.002



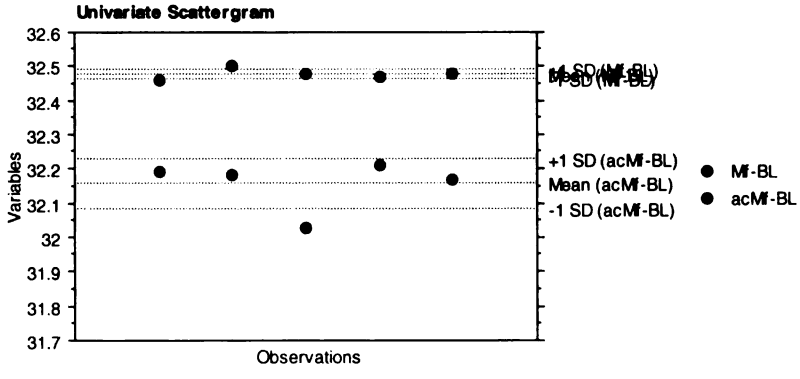
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Id-L3L, acId-L3L	.369	4	13.890	.0002	.296	.443



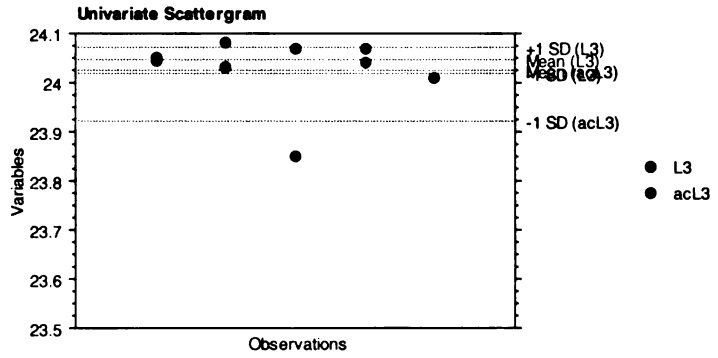
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Mf-L6L, acMf-L6L	.737	4	10.559	.0005	.543	.930



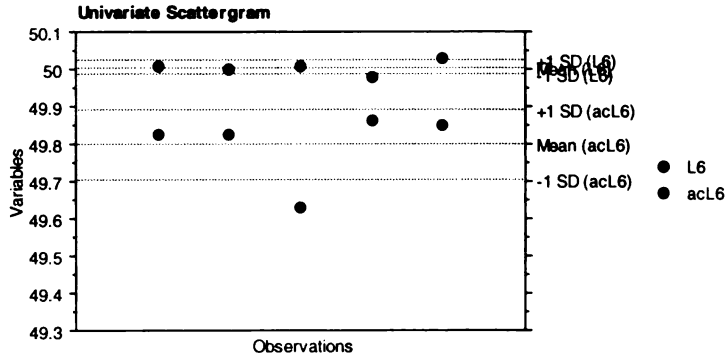
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Mf-BL, acMf-BL	.321	4	9.384	.0007	.226	.415



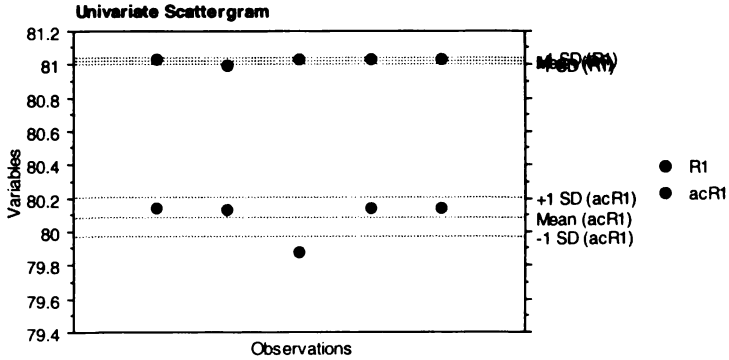
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L3, acL3	.020	4	.354	.7413	-.134	.173



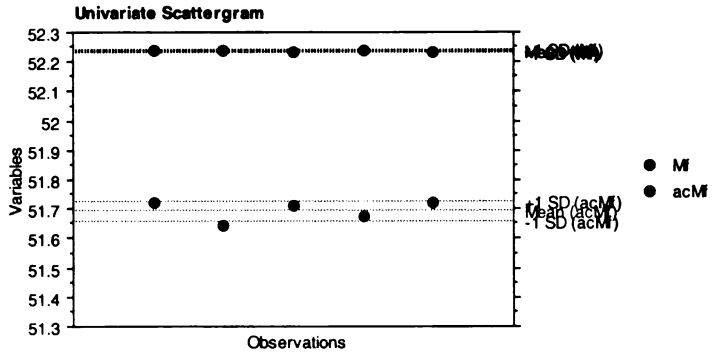
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
L6, acL6	.208	4	4.665	.0096	.084	.331



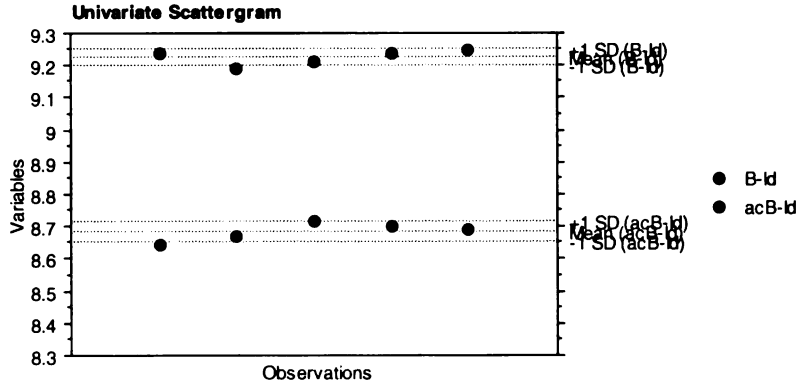
Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
R1, acR1	.936	4	16.914	<.0001	.783	1.090



Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
Mf, acMf	.541	4	32.561	<.0001	.495	.588



Paired Means Comparison
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value	95% Lower	95% Upper
B-ld, acB-ld	.541	4	29.717	<.0001	.491	.592