Evaluation of a novel palatal suture maturation classification as assessed by CBCT imaging of a pre- and post-expansion treatment cohort

by

Darren Matthew Isfeld

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Medical Sciences – Orthodontics

University of Alberta

Abstract:

Introduction: Evaluated is the novel midpalatal suture maturation classification and methodology proposed by Angelieri et al. (2013). Reliability testing was performed, followed by a retrospective observational longitudinal (cohort) study to evaluate the reliability and usefulness of this novel classification system to predict success of RME treatment.

Methods: Reliability testing focused on a total of sixteen patients aged 9.5 -17 years old with early mixed to full permanent dentition, representing all proposed palatal maturation stages, with accessible pre-expansion CBCTs. The retrospective observational longitudinal (cohort) study evaluated 63 pre-adolescent and adolescent patients aged 11-17 years old with full permanent dentition treated with tooth-borne RME appliances who have CBCTs records taken at T_1 pre- and T_2 post-expansion. CBCT 3D landmarking produced skeletal and dental widths and dental angulations utilized to evaluate the extent of skeletal and/or dental expansion as it relates to the T1 palatal suture classification of each subject.

Results: There was almost perfect intra-examiner agreement and slight to poor inter-examiner agreement, differing from previously reported reliability, affected by necessary operator calibration and the degree of post-acquisition image sharpness and clarity. Results of the cohort study were wholly unsupportive of the efficacy of the proposed palatal suture maturation classification. Further evaluation of its scientific basis determined that the classification was ill-founded.

Conclusion: Clinicians should not consider this proposed classification as being factual, and halt employing its use to drive clinical decision making which will have real-world patient implications and outcomes.

Preface:

This thesis is an original work by Darren M. Isfeld. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board and the Northern Alberta Clinical Trials and Research Centre (NACTRC), Project Name "Reliability and predictive efficacy of a novel methodology for midpalatal suture maturation classification for individual assessment prior to maxillary expansion", No. Pro00060813, December 5, 2015. Chapter 1 of this thesis has been published as Novel methodologies and technologies to assess mid-palatal suture maturation: A systematic review. Manuel Lagravere; Darren Isfeld; Vladimir Leon-Salazar; Carlos Flores-Mir. Head face med. 2017 June 14:13(1). Doi:10/1186/s13005-017-01442-2. I was responsible for the data collection, analysis and manuscript composition. All other authors were involved with concept formation, manuscript composition and edits.

Acknowledgements:

I would like to thank all committee members for their time, invaluable input, expertise and guidance in this project. Special and limitless thanks to my thesis supervisor, Dr. Manuel Lagravere, for allowing me to take on this project so last minute. I am certainly grateful for his patience with me, support throughout this study and writing of this thesis. Dr. Lagravere's critical thinking skills, attention to detail, as well as, being able to see the big picture is inspiring and is something I will certainly try to embolden in myself as I move forward from academia into private practice.

With Deepest Gratitude,

Darren Matthew Isfeld

Contents:

List of Tables **vi**

List of Figures **vii**

Chapter 1: Introduction and Systematic Review of Literature **1**

1.1 Statement of Problem **2**

1.1.1 Research Questions **3**

1.2 Systematic Review of Literature: Novel methodologies and technologies to assess mid-palatal suture maturation **4**

1..2.1 Introduction **4**

1.2.2 Material and Methods **6**

1.2.3 Results **10**

1.2.4 Discussion **24**

1.2.5 Conclusion **34**

1.2.6 Literature Cited **36**

1.2.7 Appendices 40

Chapter 2: Reliability testing of a novel palatal suture maturation classification **44**

2.1 Introduction **45**

2.2 Methods and Materials **48**

2.3 Statistical Analysis **52**

2.4 Results **54**

2.5 Discussion **59**

2.6 Conclusion **67**

2.7 Appendices **69**

2.8 Literature Cited **75**

Chapter 3: Evaluation of a novel palatal suture maturation classification as assessed by CBCT imaging of a pre- and post-expansion treatment cohort **77**

3.1 Introduction **78**

3.2 Material and Methods **86**

3.3 Statistical Analysis **95**

3.4 Results **97**

3.5 Discussion **99**

3.6 Conclusion, Limitations & Future Recommendations **110**

3.7 Appendices **114**

3.8 Literature Cited **126**

Chapter 4: General Discussion **131**

4.1 Discussion **132**

4.2 Literature Cited **137**

Bibliography **139**

List of Tables:

Table 1.1: Summary of articles that met final inclusion criteria **11**

Table 1.2: Results and conclusions of articles meeting final inclusion criteria **21**

Table 2.1: intra-examiner, inter-examiner, and rater to ground truth agreement from classification of 16 patients palatal suture maturation **54**

Table 2.2: Frequency table of classification of palatal stage maturation by CF in comparison to ground truth. Bolding denotes the percent frequency of CF correctly classifying the appropriate palatal stage to ground truth. **55**

Table 2.3: Frequency table of classification of palatal stage maturation by ML in comparison to ground truth. Bolding denotes the percent frequency of ML correctly classifying the appropriate palatal stage to ground truth. **56**

Table 3.1: Maxillary landmarks defined and shown on cross sectional images identified in each subject's pre- and post-expansion CBCT volumes. **89**

Table 3.2: Dental and skeletal linear & angular measurements generated. **90**

Table 3.3: Dependent variables – difference in skeletal and dental distances and dental angles from T2-T1. **94**

Table 3.4: percent change in dependent variables from pre- to post-expansion (T2-T1). **98**

Table 3.5: Summary of histological studies cited as source of findings used to define the proposed palatal suture maturational stages (A-E). **105**

List of Figures:

Figure 1.1: Flow diagram of the literature search. **9**

Figure 2.1: Diagrammatic representation of the developed novel palatal suture maturation classification identifying key radiological morphological characteristics specific to each maturity level. **46**

Figure 2.2: (a) Representation on proper standardization of head position for image analysis and. **(b)** axial cross section of the mid-palatal suture generated from this protocol. **51**

Figure 2.3: Bland Altman plot evaluating Intra-examiner reliability comparing DAI reading **57**

Figure 2.4: Bland Altman plot evaluating Inter-examiner reliability comparing classifications by ML to CF. **57**

Figure 2.5: Bland Altman plot evaluating reliability comparing classifications by CF to ground truth (DAI classification session 1). **58**

Figure 2.6: Bland Altman plot evaluating reliability comparing classifications by ML to ground truth (DAI classification session 1). **58**

Figure 3.1: Diagrammatic representation of the developed novel palatal suture maturation classification identifying key radiological morphological characteristics specific to each maturity level **85**

Figure 3.2: Orientation of the Cartesian plane coordinate system in 3 planes; X (red), Y (green) and Z (blue). **87**

Figure 3.3: Spherical markers representing the 3D landmarks of interest visualized in the x, y and z planes within the Avizo software version 7.0 **88**

Figure 3.4: Unbalanced distribution of palatal stages across the sample (n=63) as classified according to Angelieri et al[1]. **92**

Figure 3.5: Improved balanced distribution of palatal stages across the sample (n=63) after implementation of modified Angelieri et al. classification. **92**

Chapter 1: Introduction and systematic review of literature

1.1 Statement of Problem

There is significant variation in the timing of skeletal maturation amongst individuals[2, 3] as the palatal suture fusion is poorly correlated with patient age and sex[1]. Failure to properly identify key clinical signs and provide individual assessment to identify a patient's ideal expansion treatment option can lead to iatrogenic side effects and co-morbidities[1, 2]. Conversely prematurely committing a patient to surgically assisted expansion ascribes a patient to a potential significant burden of treatment including increased cost, pain and healing time. To minimize sequalae of failed rapid maxillary expansion (RME) and/or avoid the co-morbidities of surgically assisted RME, a reliable method to classify midpalatal suture maturation with predictive ability to drive clinical decision making, towards non-surgical or surgical RME, in adolescent and young adult patients is needed.

1.1.1 Research Questions

Three research questions were identified:

Primary research question –

What is the reliability of the Angelieri et al midpalatal suture maturation classification system?

Secondary research question –

How useful is this novel classification system to predict success of RME treatment?

Tertiary research question -

What alteration(s) or modification(s) to the Angelieri et al. methodology can be suggested to improve reliability and/or predictive ability of this classification system?

1.2 Systematic Review of Literature: Novel methodologies and technologies to assess mid-palatal suture maturation

Novel methodologies and technologies to assess mid-palatal suture maturation: A systematic review. Manuel Lagravere; Darren Isfeld; Vladimir Leon-Salazar; Carlos Flores-Mir. Head face med. 2017 June 14:13(1). Doi:10/1186/s13005-017-01442-2

1.2.1 Introduction:

Rapid maxillary expansion (RME) is indicated for a number of clinical situations namely when a posterior crossbite exists (unilateral or bilateral) or limited buccal overjet in patients with constricted maxillary base[4]. Maxillary transverse deficiency may be skeletal, dental or both skeletal and dental in origin [1, 3, 4]. Expansion in the transverse dimension has not only been used to improve interdigitation of the occlusion and improved function, but also utilized to increase arch perimeter to resolve maxillary crowding[3]. Recently contemporary orthodontics has focused on smile esthetics with emphasis on transverse arch dimensions and minimizing buccal corridor visibility[4, 5]. Those patients with dentofacial deformity or cleft lip and palate with constricted maxillary segments are candidates for RME or possible surgical expansion[3] dependent upon the time of treatment intervention. Additionally, there has been increased interest in the use of RME to increase nasal airway volume and/or function[3, 4].

Treatment options available to clinicians for maxillary expansion include tooth-borne expanders with or without an acrylic support[3, 6], bone-borne maxillary expansion devices supported by temporary (skeletal) anchorage devices (TSADs/TADs)[6], as well

as surgically assisted rapid palatal expansion (SARPE)[1, 4]. The treatment of choice is dependent on numerous clinical indications including; the extent of correction required, whether skeletal or dentoalveolar correction is indicated, and perceived efficacy of expansion based on timing of treatment[2].

The amount of skeletal or dentoalveolar effect of the RME is directly correlated with the stage of skeletal maturation of the palatal suture. Treatment timing of transverse deficiencies is recommended relatively early up to peak skeletal growth velocity[2]; however there is significant variation in the timing of skeletal maturation amongst individuals[2, 3] as the palatal suture fusion is poorly correlated with patient age and sex[1]. Failure to properly identify key clinical signs and provide individual assessment to identify a patient's ideal expansion treatment option can lead to iatrogenic side effects and co-morbidities[1, 2]. Common side effects of poorly timed and failed conventional RME therapy include acute pain[3], gingival recession, dehiscence formation, palatal mucosa necrosis, buccal dentoalveolar tipping and poor long term expansion stability[1, 2]. Conversely prematurely committing a patient to surgically assisted expansion ascribes a patient to a potential significant burden of treatment including increased cost, pain and healing time.

Numerous methodologies have been proposed to discern the architecture and degree of palatal suture fusion including animal and human histologic studies, evaluation of occlusal radiographs, and computed tomography (CT) of both autopsy material and animal specimens[1]. Such methodologies presented inherent difficulties in assessing the

degree of palatal suture fusion. As defined previously, histological evaluation is the reference standard to evaluate mid palatal suture maturation, unfortunately implementation on active orthodontic patients would require an invasive biopsy, precluding its use[7, 8]. Conversely, serial occlusal radiographic assessment is limited in diagnostic quality due to superimposition of nearby anatomical structures[1]. Cone-beam CT (CBCT) allows for 3D rendering of the maxillofacial complex without superimposition of nearby anatomy, and delivers a lower absorbed dose of radiation to the patient than medical CT[1]. To date however, there has been no validated and clinically accepted non-ionizing method to assess palatal suture maturation.

The objectives of this systematic review are to systematically describe and evaluate the contemporary technologies and methodologies capable of assessing mid-palatal suture maturation.

1.2.2 Material and Methods:

The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement checklist was followed; however, several points did not apply to this systematic review. This is a review of both *in vitro* and *in vivo* studies rather than solely *in vivo* studies, convoluting the direct comparison of results amongst these types of studies and their possible clinical inferences. No protocol registration was done.

Eligibility Criteria

Both *in vitro* and *in vivo* studies will be included to identify all diagnostic modalities of

palatal suture maturation. The intervention(s) will be any diagnostic method that is designed to evaluate the degree of ossification and/or interdigitation of the mid palatal suture (the outcome). Comparison will be to other diagnostic interventions designed to evaluate the same outcome variable.

The "participants" will be any human subjects or human specimens being investigated for the degree of mid-palatal suture maturation. No animal studies were considered as their applicability in humans would be questionable.

Information Sources

A computerized database search was conducted using Medline, Pubmed, Embase and Scopus to search the literature from 1984 up until October 5, 2016. A supplemental hand search was completed of references from retrieved articles that met the final inclusion criteria

Search

Terms and their respective truncations used in the literature search (Appendix 1) were specific to each database. Searches were conducted with the help of a senior librarian who specializes in the Health Sciences. The selection process was carried out together by 2 researchers (DAI and HE). All references were managed by reference manager software EndNote to eliminate duplicates.

Study Selection

The inclusion criterion ''*Diagnostic methods to evaluate cranial suture*

ossification/maturation" was utilized to initially identify possible articles from the published abstract results of the database search. If an abstract was not available, the full text was reviewed for appropriateness of inclusion. Any disagreement on the inclusion of a study was resolved by discussion amongst the reviewers.

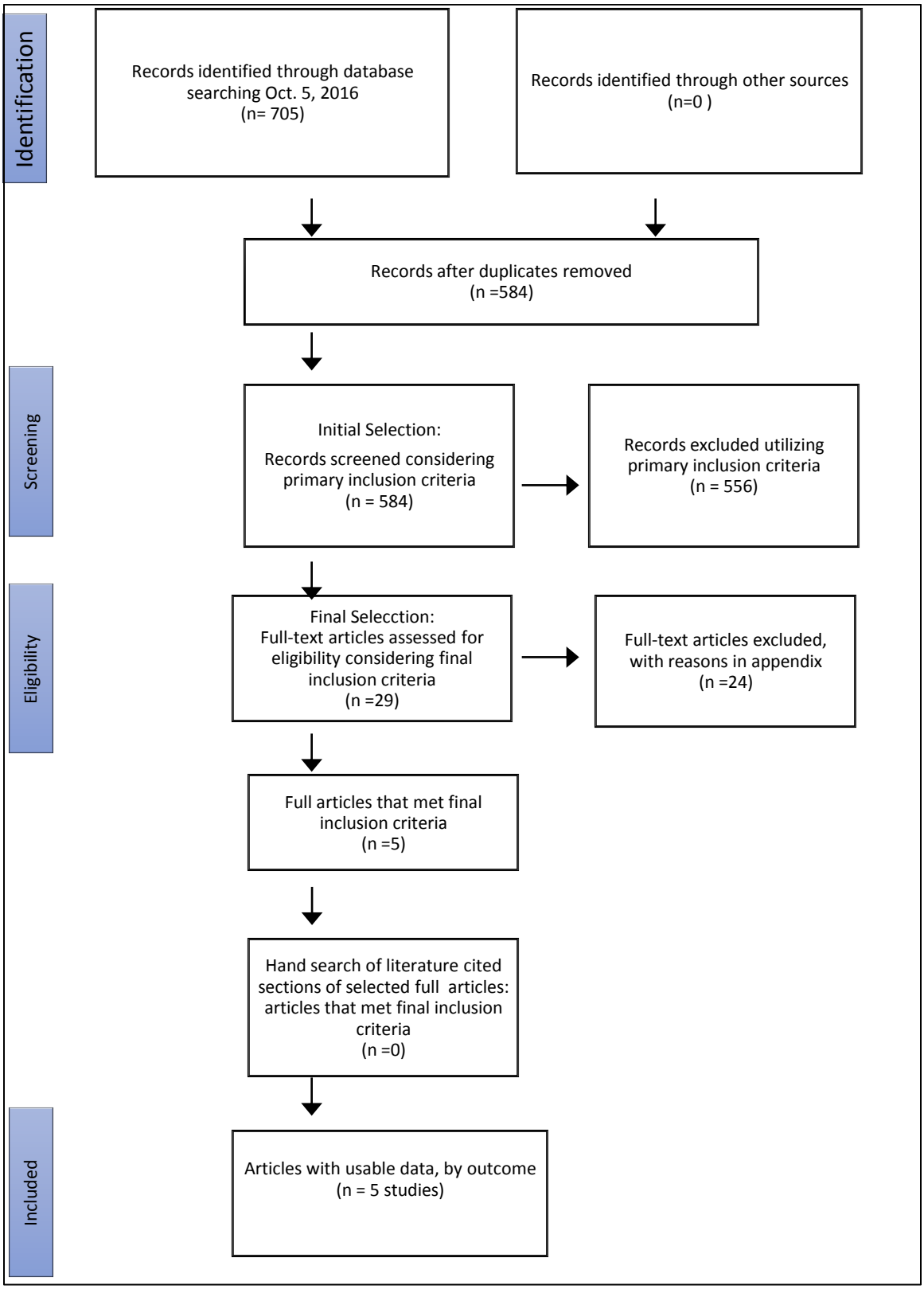
Once these abstracts were selected, full articles were retrieved and inclusion in the systematic review was dependent of fulfilling a final inclusion criterion. The final selection criterion was as follows: ''*In vitro and in vivo human subject studies that describe a novel diagnostic method or technology to assess mid-palatal suture maturation/ossification over time".* Once more, any disagreement on the inclusion of a study following this final criterion was resolved by discussion amongst the reviewers. The references cited in the finally selected articles were also screened for any applicable references missed in the electronic database search.

Studies describing diagnostic methodologies applied to theoretical models without practical application were excluded. Restrictions for language were only applied when resources for translation services were not available.

Data Collection Process

Data extraction was performed and collected by a researcher (DAI). (Appendix 1 and Figure 1)

Figure 1.1: Flow diagram of the literature search.



Data Items

The variables collected included a description of the type of study, type and number of subjects, study objectives, inclusion criteria, imaging modality used, region(s) investigated, methodology to evaluate degree of ossification/maturation of mid-palatal suture.

Summary measures

The outcome measures included quantitative and/or qualitative results attained with applicable units to describe bone density, ossification or maturation of the palatal suture.

Synthesis of results

In the event that data was considered homogenous enough a meta-analysis was planned.

1.2.3 Results

Study Selection

Twenty-nine abstracts met the initial inclusion criteria. Following retrieving of the full articles, only five met the final inclusion criteria. Reasons for exclusion due to final inclusion criteria are stated in Appendix 2. A hand-search of the reference lists from the articles that met the final inclusion criteria identified no new articles. Therefore, a total of five articles were finally considered. (Figure 1.1)

Study Characteristics

The methodology of each selected article was summarized in Table 1.1 and results in Table 1.2. Study parameters, including the type of study, imaging modality used,

methodology to determine the ossification/maturation of the palatal suture and the number of subjects amongst other variables were vastly different amongst the studies meeting the final inclusion criteria. (Table 1.2)

Table 1.1: Summary of articles that met final inclusion criteria

Author(s)	Franchi L. et al.[11]	Sumer AP. Et al.[9]	Korbmacher H. et al. [10]	Angelieri et al.[3]	Kwak KH et al.[12]
Type of Study	Prospective study	Prospective study	In-vitro study	Cross-sectional	Cross-sectional
Human Subjects or Material	Human subjects	Human subjects	Human autopsy material	Human Subjects	Human subject
Study Objective(s)	Assess the midpalatal suture density via lowdose computed tomography (CT) prior to RME (T0), at the end of active RME (T1), and following a 6month retention period (T2).	Evaluate the efficacy of ultrasonography (US) to generate a qualitative assessment of ossicifcation post-SARME.	Quantification of sutural morphology via micro-CT and its association with age.	To validate and present a novel classification system for the individual assessment of midpalatal suture morphology using CBCT.	Evaluate the correlation of fractal patterning to ossification of the palatal suture via CBCT evaluation and determine whether fractal analysis of the midpalatal suture can be used to assess the maturation of the suture.
# of Subjects and Inclusion Criteria (if applicable)	17 patients, 7 male, 10 female, mean age of 11.2 years old, range of 8-14 years old. Inclusion criteria: patients with constricted maxillary arches with or without unilateral or bilateral posterior crossbite, and within cervical vertebral maturation (CS1-CS3)	3 patients, bilateral transverse maxillary deficiencies requiring SARME. Age, sex and developmental characteristics of subjects not given.	28 human-palate specimens, (11 female, 17 male) aged 14-71. The palatal specimens were categorized by the donor's age into age groups (< 25 years, 25 years to < 30 years, ≥ 30 years).	140 subjects (86 female, 56 male), age range from 5.6 to 58.3 years old, Inclusion criteria: patients who are undergoing initial records for orthodontic treatment and who have received no previous orthodontic treatment.	131 subjects, (69 men and 62 women), mean age mean age of 24.1 ± 5.9 years (male subjects 23.1 ± 5.8 years, female subjects 25.2 ± 5.9 years) Age range of18.1-53.4 years old. No specific inclusion criteria noted
Study's Expansion Modality, Expansion protocol, Average amount of Expansion (mm)	**Modality:** butterfly palatal expander **Protocol:** standard protocol – activated twice per day (0.25mmper turn) for 14 days. Retention period of 6 months than appliance removed. **Amount of**	**Modality:** SARME (tooth borne Hyrax). **Protocol:** 0.8-0.9mm expansion/day in two daily activation steps until desired expansion achieved, ~14 days. Retention period of 6 months, than	Not applicable, no expansion performed.	Not applicable, no expansion performed.	Not applicable, no expansion performed

	expansion: 7mm in all subjects	hyrax removed. **Amount of expansion:** not specified but based on clinical needs of patient.			
Imaging Modality	Multi-slice low-dose Computed tomography (brand information not given). Standardized axial CT images parallel to the palatal plane and passing through the furcation of maxillary right first molar, scans acquired and magnified (3x) with Light-Speed 16 softwar (General Electric Medical System, Milwaukee, Wis).	Color-coded Ultrasonography duplex scanner (Aplio 80, Toshib, Tokyo, Japan) with 7.5-MHz linear-array transducer	Scanco Micro-CT 40 (Scanco Medical, Bassersdorf, Switzerland) 70 kV, 114 μA. Isotropic voxel size 37 μm. Maximum scanning time of 200 minutes/specimen. Data analyzed using V4.4A software (Scanco Medical, Bassersdorf, Switzerland). 3D reconstruction via AMIRA 3.00 software (TGS, Mercury Computer Systems, San Diego, CA, USA). Bone volume and quantification via Image Tool 3.00 software (UTHSCSA, San Antonio, TX, USA),	iCAT cone-beam 3-dimensional imaging system (Imaging Sciences International, Hatfield, Pa). 11cm Minimum FOV. Scan time from 8.9 to 20 seconds resolution of 0.25 to 0.30 mm. Image analysis using Invivo5 (Anatomage, San Jose, Calif). A standardized protocol to isolate axial maxiallary cross-sections of the palate was presented.	Cone Beam Computed Tomography (CBCT) (Zenith 3D; Vatech Co., Gveonggi-do, Korea) Field of view 20 × 19 cm; voltage 90 kVp; current 4.0 mA; scan time 24 s). Images were assessed using CT software (Ez3D 2009; Vatech Co.),
Region(s) Investigated	Mid palatal suture and maxilla. 4 regions of interest (ROIs); 1. Anterior sutural ROI (AS ROI): located on the suture 5 mm anterior to nasopalatine 2. Posterior sutural ROI (PS ROI): on suture 5 mm posterior to the nasopalatine duct 3. Anterior bony ROI (AB ROI): control ROI on maxillary bone 3 mm to the right of laterally AS ROI 4. Posterior bony ROI (PB ROI): control ROI on maxillary bone 3 mm right of PS ROI	Midpalatal suture	Midpalatal suture	Axial central cross-sectional slices generated and used for assessment of the midpalatal suture	Axial central cross-sectional slices generated and used for assessment of the midpalatal suture. A long and narrow region of interest within the final axial slice highlighting only the suture was considered for fractal analysis, such that the incisive canal was not incorporated, but rather the ROI extended from posterior to the incisive canal to just anterior to the posterior nasal spine.
Method of Measurements (units)	1 trained and blinded operator (R.L.) calculated	Ultrasonography findings rated via a semi-	Quantification of 3D Suture Morphology in frontal plane	Definition of the proposed palatal suture maturational	1 principal investigator trained in the Angelieri et al.

	bone density values in Hounsfield units (HU). RL performed measurements and repeated all measurements 1 month later. Bone density changes from T0 through T2 at AS ROI and PS ROI contrasted with the Friedman repeated measures ANOVA on ranks fand Tukey post-hoc test (SigmaStat 3.5, Systat Software, Point Richmond, Calif).	quantitative bone fill score (0-3). 0= complete through-transmission of the ultrasound waves, clear gap margins, and no echogenic material; 1 = partial through-transmission of the ultrasound waves, identifiable gap margins, and less than 50% echogenic material; 2 = partial through-transmission of the ultrasound waves, partially obscured gap margins, and greater than 50% echogenic material; 3 = no through transmission of the ultrasound waves, invisible gap margins, and 100% echogenic material. Scores were not supported by histology or CT.	measured: calculated Obliteration index [%], and mean obliteration index [%]. Quantification of 3D Suture Morphology in Axial plane: measured suture length [μm]: linear sutural distance [μm]: interdigitation index;	stages (A-E) determined by two operators. The definition of each palatal suture maturational stage derived from the histological appearance of suture described in previous histologic studies.	method categorized the midpalatal sutures of the patients, and the findings were considered the "ground truth" not "gold standard". Images were reclassified 2 days later two other operators classified 30 images to determine interxaminer reliability. For Fractal analysis, image software (Photoshop CS6 Extended; Adobe Systems, San Jose, CA, USA) was utilized to perform Gaussian blurring and subtract this blurred image from the original, followed by skeletonizing of the binary image, and utilizing the box counting method to determine the fractal dimension. Weighted kappa coefficient was calculated to determine inter- and intra-examiner reliability using MedCalc version 12.3.0 (MedCalc Software, Oostende, Belgium). Fractal dimension at each maturation stage determined by Scheffe's ANova test. Spearman's correlation coefficient was calculated to determine the correlation between the fractal analysis and maturation stage. Utilized IBM SPSS Statistics version 21.0 software (IBM Co., Armonk, NY, USA) $p < 0.05$ was considered

					statistically significant.
Measurement time points	Three time points; Before RME (T0), at the end of RME (T1), and after the 6month retention period	5 time points; after RME, at 2 and 4 months during the expansion period, 6 months later where appliance removed and 2 months post appliance removal. Note opening of midpalatal surture confirmed by plain radiograph after active expansion.	One time point evaluated	Single time point evaluated prior to RME. Palatal maturational stage reclassified 2 days later for each patient.	Single time point Palatal maturational stage reclassified 2 days later for each patient.) .

The studies varied significantly in the number of subjects evaluated and quality of evidence. The studies ranged from having three human subjects in a prospective study[9] to 140 human subjects in a cross-sectional study[1]. The types of studies ranged across the hierarchy of evidence from an *in-vitro* study[10] to prospective *in vivo* studies[9, 11].

The only study characteristic common to all studies was the region of interest (ROI) investigated, generally speaking, the maxilla. Four of the five studies[1, 9, 10, 12] had a single common region of interest being the palatal suture. One study[11] evaluated 4 regions of interests in the palatal suture and surrounding hard tissue.

All studies but one utilized computed tomography (CT) in some form. The types of CT scanners utilized in the four studies included multi-slice low-dose computed tomography (brand information not given)[11], dental cone beam computed tomography (CBCT) [1] and the extremely high resolution Micro-CT[10]. One study[9] utilized a less invasive

modality of ultrasonograpy (US), specifically using color-coded ultrasonography duplex scanner (Aplio 80, Toshib, Tokyo, Japan).

To measure the degree of maturation/ossification at the palatal suture, one of three types of evaluations were utilized amongst the five studies: quantitative, semi-quantitative and qualitative.

Franchi [11] performed a quantitative evaluation of the palate using one blinded operator to calculate the radiodensity (Hounsfield units [HU]) of the ossification at the palatal suture from T0 (pre-expansion) and T2 (at 6 months retention).

Korbmacher [10] also performed a quantitative evaluation of sutural maturation by measuring the maturation of the palate cadaver specimens at one time point. In the coronal plane, an obliteration index (%) and mean obliteration index (%) was calculated by comparison of the total length of the suture to the length that has ossified (evaluated every 370 μ m). The degree of interdigitation of the palatal suture in the axial plane was assessed by calculating the interdigitation index, a comparison of the sutural distance (μm) to linear sutural distance (μm).

Angelieri [1] developed a novel qualitative methodology for individual evaluation of midpalatal suture maturation. Two evaluators defined the maturational stages (A-E) via comparison of the morphological description of the palatal suture found in previous histologic studies[13-15] to the appearance of the suture in the axial plane generated from a standardized CBCT protocol of 140 subjects during initial records[1]. To assess the reliability of defining the maturational stages (A-E) a validation study utilizing 30

random axial CBCT cross-sections of the mid-palatal suture was performed by 3 evaluators and weighted kappa coefficients calculated[1].

Kwak et al.[12] utilized an objective and quantitative method of fractal analysis, a methodology established previously for the evaluation of mammalian cranial sutures[16], to be used for the first time in conjunction with CBCT imaging to evaluate the maturity of the mid-palatal suture[12]. The cross-sectional study involved 131 subjects (69 men and 62 women) with a mean age of 24.1 ± 5.9 years. Each subject underwent CBCT imaging, followed by significant image processing to evaluate CVM stage, palatal stage of maturation (A-E, as defined by Angelieri et al.[1]) and isolation of a region of interest (ROI) for the calculation of the fractal dimension of the palatal suture. To assess the intra- and inter-reliability of defining the maturational stages (A-E) 30 random axial CBCT cross-sections of the mid-palatal suture were staged by two other evaluators under controlled conditions and weighted kappa coefficients calculated, analogous to the study by Angelieri et al[1]. Statistical analysis included utilizing Scheffe's ANOVA to compare the fractal dimension for each individual maturation stage (A-E) and subsequent Spearman's coefficient calculation to ascertain the correlation between fractal dimension and maturation stage. The generation of a receiver operating characteristic (ROC) curve was used to develop an optimal fractal dimension cut-off value and sensitivity, specificity, false positive rate, false negative rate, positive predictability, and negative predictability calculated. For all statistical analysis, results were considered statistically significant at $p<0.05$ [12].

Sumer[9] utilized ultrasonography to evaluate palatal sutural mineralization in 3 patients at 5 different time points; once after the 14 day surgically-assisted RME (SARME)

expansion protocol, 2 months post-expansion, 4 months post-expansion, at time of removal of the tooth-borne expander (6 months post- expansion) and 2 months after appliance removal. The authors report that the ultrasound probe was used intra-orally on the skin that overlies the palatal suture, obtaining axial scans with the probe directed perpendicular to the length of the suture[9]. The authors assigned semi-quantitative bone fill scores (0-3). A bone fill score = 0 was characterized by open suture with clean gap margins and 0% echogenic material. A bone fill score = 1 was characterized by partial ultrasound transmission, localization of gap margins, and reduced echogenic material of $\leq$50%. A bone fill score = 2 was characterized by partial ultrasound transmission, marginally visible gap margins, and increased echogenic material of >50%. A bone fill score =3 was characterized by no ultrasound transmission, 100% echogenic readings, and unidentifiable gap margins. The bone filling trends were qualitatively supported by comparison to conventional occlusal radiography[9].

Results of Individual studies

Franchi et al.[11] utilized the Houndsfield quantitative scale to evaluate the radiodensity of 4 previously mentioned regions of interest in the maxilla, 2 sutural and 2 bony areas. Pre-expansion (T0) statistical analysis noted a significant difference between the anterior and postural sutural regions (563.3 $\pm$ 183.29 HU and and 741.7 $\pm$ 167.1 HU) and anterior and posterior bony areas (1057.5$\pm$ 129.4 HU and 1102.8 $\pm$ 160.9 HU) ($p<0.05$) (Table 1.2). Further statistical analysis yielded a significant difference between the anterior sutural and posterior sutural landmarks at T0 ($p<0.05$), but no significant differences of these

sutural areas at T1 or T2 ($p>0.05$, Mann-Whitney). A significant difference between the radiodensity of the anterior and postural sutural ROIs between T0 and immediately post-expansion (T1), but no difference between their radiodensities when comparing pre-treatment (T0) and the post-expansion retention phase (T2) readings ($p<0.05$). (Table 1.2)

Korbmacher et al.[10] evaluated sutural interdigitation via calculation of an obliteration index for each of the 28 human palate specimens in frontal and axial plane. In the frontal plane demonstrated no age dependent difference in the mean obliteration index between specimens ($p=0.244$). The specimens were classified into one of three age groups (<25yo, ≥25 to <30 yo and ≥30yo) and results demonstrate that the frontal plane obliteration index varied across age groups between a minimum index of 0% to a maximum interdigitation of 7.3% (44yo patient) (Table 1.2). Although the ≥25 to <30 yo age group consistently had a higher obliteration index in the frontal plane compared to other age groups, the results were not significant. Across all age groups, each subject had at least one frontal sutural cross-section that was devoid of interdigitation (mean obliteration index of 0%), with the oldest patient exhibiting a frontal plane mean obliteration index of 0% being a 71yo female. Investigation into the degree of interdigitation in the axial plane demonstrated no significant age-dependent differences in the calculated interdigitation index ($p=0.633$). The authors did report a large standard deviation in the interdigitation index in the axial plane in the youngest and oldest age groups, and considerably less variation in the calculated index in the middle (<25 yo group and >30 yo) group[10] (Table 1.2).

Ultrasonography findings in the Sumer et al. study[9] demonstrated that immediately post-expansion all subjects had a bone fill score =0. (Table 1.2) Two of the three subjects at 2

and 4 months post-expansion were identified as having a bone score = 1, while the remaining subject was determined to have a bone fill score =2 for these same time periods. Following the removal of the tooth-borne appliance at 6 months and 2 months subsequent to that during continued fixed appliance therapy, the bone scores for two of the subjects demonstrated increased mineralization and identification of echogenic material, having bone fill scores =2. The remaining patient received a bone fill score=3 due to incomplete transmission of the waves and 100% echogenicity measured at these respective time points[9] (Table 1.2). It should be noted that no statistics were reported by the authors.

As it relates to the proposed maturational stages of the palatal suture proposed by Angelieri et al.[1], a validation study performed reported a weighted Kappa statistic for intra- and inter-examiner reliability to be κ =0.75 (95% CI, 0.64-0.99) and be κ =0.79 (95% CI, 0.60-0.97) (no p-value reported), respectively. Due to a lack of an histologic or micro-ct gold standard, the authors also reported examiner reliability compared to the "ground truth", a descriptor used to represent consensus among examiners with the principal investigator' radiographic evaluations or other interpretations. Examiner reliability with ground truth ranged from κ =0.82 (95% CI, 0.64-0.99) to κ =0.93 (95% CI, 0.86-1.00) (no p-value reported)[1].

Fractal dimension intra- and inter-reliability results from the Kwak et al. study[12] demonstrated agreement with calculated weighted kappa coefficients of 0.84 (95% confidence interval [CI] 0.74-0.93) and 0.67 (95% CI 0.38–0.95) to 0.72 (95% CI 0.48–0.97), respectively (Table 2). The CVM index inter- and intra-examiner reliability

demonstrated agreement with weighted kappa coefficients from 0.69 (95% CI 0.53–0.86) and 0.71 (95% CI 0.56−0.86), respectively. The authors reported that none of the patients investigated possessed a CVM 1-IV nor was any subject classified as having palatal suture maturational stage A. It was found that 13 of 21 subjects with CVM V were classified as having maturational stage B or C (61.9%; males 77.8%, females 50.0%). Additionally, 42 of 110 subjects with CVM VI were classified as having maturational stage B or C (38.2%; males 41.6%, females 34.0%). Post-hoc analysis demonstrated that maturational stages B, C, D and E were related to differences in mean fractal dimension ($p< 0.05$). A negative correlation existed between fractal dimension and maturation stage (-0.623, $p < 0.001$). Male and Female correlation coefficients determined to be -0.649 ($p < 0.001$) and -0.569 ($p < 0.001$) respectively. A receiver operating characteristic (ROC) curve was generated and determined the boundary between dichotomous maturation stages A–C and D or E, allowing for fractal dimension to be used to identify midpalatal suture fusion. Predictive statistical analysis noted that fractal dimension is a statistically significant indicator capable of predicting dichotomous maturation stages ((A, B, & C) vs. (D or E) (area under ROC curve [AUC] = 0.794, $p < 0.001$)[12] (table 1.2).

Table 1.2: Results and conclusions of articles meeting final inclusion criteria

Author(s)	Franchi L. et al.[11]	Sumer AP. Et al.[9]	Korbmacher H. et al. [10]	Angelieri et al.[3]	Kwak KH et al.[12]
Result(s)	Bone density in the AS ROI and the PS ROI at T0 (563.3 6 183.2 HU and 741.7 6 167.1 HU, respectively) were significantly smaller than values in the AB ROI and the PB ROI at T0 (1057.5 6 129.4 HU and 1102.8 6 160.9 HU, respectively). At T0 there was a significant difference in bone density at AS and PS ROIs, but no difference at T1 and T2. AS and PS ROIs showed significant decreases in density from T0 to T1, significant increases from T1 to T2, and no statistically significant differences from T0 to T2.	No statistics reported. Immediately post expansion all 3 patients had a bone fill score = 0. At 2 and 4 months of expansion there was low echogenicity in the suture (US bone fill score =1) for 2 of 3 subjects. The remaining patient had a bone fill score =2 at 2 and 4 months respectively. At 6 months post expansion and 2 months after expander removal, 2 of the 3 patients showed a qualitative increase in echogenic material in the suture was seen but less than 100% therefore had a bone fill score =2, and the remaining patient demonstrated 100% echogenic material, bone fill score =3. All trends in scores over time were qualitatively confirmed with plain radiographic images.	**Frontal plane:** No age dependent significance was found for the mean obliteration index ($p = 0.244$). The mean obliteration index was low, varying in all groups (minimum 0%; maximum 7.3%). Middle-aged group's mean obliteration index tended to be higher than that of either the younger or older age groups but no significant difference was calculated. The highest mean obliteration index (of 7.3%) was found in a 44-year-old male. The oldest individual with a mean obliteration index of 0% was a 71-year-old female. At least one frontal slice per palate – even in the oldest age group – exhibited a suture completely open cranio-caudally. **Axial plane:** No significant differences detected in all age groups regarding means and standard deviations for suture length, linear sutural distance, and interdigitation index. Interdigitation index computed revealed no significant age-dependent differences ($p = 0.633$). High standard deviation values for suture length, linear sutural distance and	The intra- and inter-examiner reproducibility values demonstrated agreement, with weighted kappa coefficients from 0.75 (95% confidence interval [CI], 0.57-0.93) to 0.79 (95%CI, 0.60-0.97), and the reproducibility of examiners with the ground truth demonstrated agreement with weighted kappa coefficients from 0.82 (95% CI, 0.64-0.99) to 0.93 (95% CI, 0.86-1.00). From the 140 subject sample, stage A was observed in children from 5 to approximately 11 years of age, a 13 year old boy was the sole exception. Should be noted there was no fusion of the palatal suture in subjects aged 5 to almost 11years old. Stage B was observed primarily up to 13 years of age but also 6 of 32 subjects (23% of boys, 15.7% of girls) aged 14 to 18 years old. Stage C primarily depicted from 11 to 18 years of age, with exception being two 10-year-old girls (8.3% of	The intra- and inter-examiner reliability analysis demonstrated agreement for fractal dimension, with a weighted kappa coefficient of 0.84 (95% confidence interval [CI] 0.74–0.93) and 0.67 (95% CI 0.38–0.95) to 0.72 (95% CI 0.48–0.97) respectively. No subjects had a CVM of 1-IV nor maturational stage A present. 13 of 21 subjects with CVM V were found to have maturational stage B or C (61.9%; males 77.8%, females 50.0%). 42 of 110 subjects with CVM VI were found to have maturational stage B or C (38.2%; males 41.6%, females 34.0%). Post-hoc analysis demonstrated that maturational stages B, C, D and E were related to differences in mean fractal dimension ($p< 0.05$). A negative correlation existed between fractal dimension and maturation stage (-0.623, $p < 0.001$). Male and Female correlation coefficients determined to be -0.649 ($p < 0.001$) and -0.569 ($p < 0.001$) respectively. A receiver operating characteristic (ROC) curve determined the boundary between maturation stages A–C and D or E. Fusion of palatal suture was

			interdigitation index were seen in the <25 yo group and >30 yo group, while the 25-30 yo group had far less variation Mean error of measurement amounted to 0.12% for the obliteration index, 2.4% for the suture length, and 0.41% for the linear sutural distance.	girls) and 4 of 32 adults (15.7% of girls, 7.7% of boys). Stage D was observed in 1 of 24 girls aged 11-<14 years old, and 3 of 19 girls aged 14-18 years old, as well as in 3 of 13 males aged 14-18 years old and > 18 years old respectively. Stage E was observed in 5 of 24 females aged 11-<14 years old and , 8 of 19 females aged 14-18 years old and 8 of 19 females aged >18 years old. Stage E was observed in far less males, approximately 9 of 13 males aged >18 years old only.	determinable as a fractal dimension. Fractal dimension is a statistically significant indicator capable of predicting dichotomous maturation stages ((A, B, & C) vs. (D or E) (area under ROC curve [AUC] = 0.794, $p < 0.001$). At optimal fractal dimension cut-off value of 1.0235, statistical analysis to evaluate the predictive ability of fractal analysis to determine maturation stage ((A, B, & C) vs. (D or E)), noted the following values; specificity 86.6%, Sensitivity 64.9%, false positive rate 35.1%, false negative rate 13.4%, positive predictability 80.3%, and negative predictability 74.6%.
Conclusion(s)	Prepubertal subjects demonstrated a lower bone density at the mid palatal suture as compared to the lateral control ROIs on ossified maxillary bone. The post-expansion low bone density at the sutural ROIs supported findings that prepubertal RME effectively opens the suture.	Ultrasound bone fill scores increased directly with the duration of time post active expansion (authors referred to this as part of the expansion period) Non-invasive Ultrasonography can yield accurate information regarding bone formation at the midpalatal suture in patients undergoing	Authors note Micro CT analysis disproves the hypothesis of progressive closure of the suture directly related to patient age. Skeletal age and/or calculation of an obliteration index is not useful in terms of diagnostic criteria to drive clinical decision making regarding the perceived efficacy of non-surgical RME. Micro-CT	Utilizing CBCT to assess the midpalatal suture avoids any overlapping of soft and hard tissues. Authors note that their proposed methodology may be useful in reliably driving clinical decision making as it relates to pursing a non-surgical (RME) or surgical expansion intervention (SARME).	Adult patients posses a greater proportion of non-fused palatal sutures than what is assumed. Therefore age of the patient should not drive SARME initiation. Authors report a significant correlation between fractal dimension and degree of maturation of the midpalatal suture Determination of the fractal dimension cut-off value could be used as a reference to pursue RME vs.

	Six months of retention following RME allows reorganization and ossification of the midpalatal suture with sutural bone density values similar to pre-RME values. .	SARME.	Quantification of the mid palatal suture yields very low obliteration and age-independent interdigitation in the coronal plane. All calculated parameters demonstrated substantial inter-individual and intra-sutural variation.		SARME Fractal analysis can be utilized to evaluate the degree of maturation at the palatal suture.

Synthesis of results

Due to high methodological heterogeneity among the included studies a meta-analysis was not supported.

Risk of bias across studies

Each proposed technology or methodology to assess the maturation of the palatal suture lacked validation with a reference standard, namely histological evaluation. There was a lack of homogeneity in the quality of evidence amongst all five studies, ranging from an in-vitro study on human autopsy material[10] to human subject prospective studies[9, 11]. Sample sizes across all studies varied greatly, from 3 subjects[9] to a high of 140 subjects in a human subject cross-sectional study[1].

Additional analysis

Not applicable due to lack of meta-analysis.

1.2.4 Discussion:

Summary of Evidence

Modality #1 - multi-slice low-dose CT and quantitative bone density measurements (Hounsfield Units).

A technique to assess palatal suture maturation includes the use of multi-slice low-dose CT to capture axial slices of the maxilla and quantitatively measure the bone density at a particular region of interest in Hounsfield units (HU)[11]. It is known that CT is an excellent modality to evaluate the localized architecture of cancellous and cortical bone of the jaws[17]; however less is known regarding the quantitative measurement of bone density, the Hounsfield unit scale. Hounsfield units were first utilized in dentistry to evaluate the pre-surgical bone density of implant sites[17-19]. The Hounsfield unit scale is a linear transformation of tissue attenuation coefficients where air is defined as -1000 HU, distilled water at standardized conditions equal to 0 HU and very dense bone defined as $\geq$1000HU[17]. Consequently the authors considered and utilized the calculated HU's as an applicable unit of measurement to quantitatively assess mineralization at the palatal suture[11].

Franchi et al.[11] utilized the Houndsfield quantitative scale to evaluate the radiodensity of 4 previously mentioned regions of interest in the maxilla, 2 sutural and 2 bony areas. Pre-expansion (T0) statistical analysis noted a significant difference between the anterior and postural sutural regions (563.3 $\pm$ 183.29 HU and and 741.7 $\pm$ 167.1 HU) and anterior and posterior bony areas (1057.5$\pm$ 129.4 HU and 1102.8 $\pm$ 160.9 HU) ($p<0.05$) (table 1.2). Further statistical analysis yielded a significant difference between the anterior sutural

and posterior sutural landmarks at T0 ($p<0.05$), but no significant differences of these sutural areas at T1 or T2 ($p>0.05$, Mann-Whitney). A significant difference between the radiodensity of the anterior and postural sutural ROIs between T0 and immediately post-expansion (T1), but no difference between their radiodensities when comparing pre-treatment (T0) and the post-expansion retention phase (T2) readings ($p<0.05$). (Table 1.2)

Throughout the course of the study, trends in bone density measurements at the suture and its comparison to lateral bony sites followed conventional expectations of successful RME. Pre-expansion the measured HU at the anterior sutural region was significantly smaller than that of the posterior sutural site, and the applied expansion protocol introduced differential sutural opening with greatest opening at the anterior sutural region consistent with the pre-expansion HU scores. Additionally, the results measured at T2, at the end of the 6 month RME retention protocol, where congruent with previous histologic findings, namely post-expansion evidence of reorganization and sutural interdigitation[20].

An inherent advantage of using a low-dose CT protocol, where the voltage was decreased to 80 kilovolts (KV), is subjecting the patient to a lower absorbed dose required for children undergoing radiologic evaluation[21]. Additionally, when the kilovoltage is reduced, image contrast of anatomical structures increases while still acceptable for assessing bone quality via this protocol[21]. Future areas of interest relating to the findings and protocol of this study would include further studies to define an anterior sutural HU: postural sutural HU ratio that best predicts the success of RME treatment. Conversely, further studies could elucidate specific ratios comparing sutural radiodensity to maxillary

bony radiodensity that may predict an improved expansion prognosis.

Modality #2 - Micro-CT quantification of 3D palatal suture in the frontal and axial planes.

Korbmacher et al.[10] proposed assessing palatal suture maturation via micro-CT scanning and calculation of a number of developed indices, namely the obliteration index (%) and mean obliteration index (%) in the frontal plane, as well as, suture length [μm], linear sutural distance [μm] and interdigitation index in the axial plane.

Korbmacher et al.[10] evaluated 28 human palate specimens in frontal and axial plane. In the frontal plane demonstrated no age dependent difference in the mean obliteration index between specimens (p=0.244). The specimens were classified into one of three age groups (<25yo, ≥25 to <30 yo and ≥30yo) and results demonstrate that the frontal plane obliteration index varied across age groups between a minimum index of 0% to a maximum interdigitation of 7.3% (44yo patient) (Table 1.2). Although the ≥25 to <30 yo age group consistently had a higher obliteration index in the frontal plane compared to other age groups, the results were not significant. Across all age groups, each subject had at least one frontal sutural cross-section that was devoid of interdigitation (mean obliteration index of 0%), with the oldest patient exhibiting a frontal plane mean obliteration index of 0% being a 71yo female. Investigation into the degree of interdigitation in the axial plane demonstrated no significant age-dependent differences in the calculated interdigitation index (p=0.633). The authors did report a large standard deviation in the interdigitation index in the axial plane in the youngest and oldest age groups, and considerably less variation in the calculated index in the middle (<25 yo

group and >30 yo) group[10] (Table 1.2).

Results indicated a generally low obliteration index amongst all subjects as well as an age-independent degree of interdigitation in the axial plane; however across all measured indices there was significant intra-sutural and inter-subject variation[10]. This was the first time micro-CT was used on human samples and although this methodology was not implemented as part of an active expansion study, its principles can still be important to evaluate the pre-expansion maturity of the palatal suture. Additionally, it could be applied during mid-expansion protocol to evaluate the efficacy of treatment via calculation of the above noted indices and evaluation of the sutural architecture.

A limiting feature of the Korbmacher et al.[10] modality is the fact cadaver specimens and not living patients were used, making direct translation of this study's findings poorly applicable to clinical practice[15]. Considering the limitations of the gantry size of the micro-CT unit, and maximum scanning time used (200 minutes), micro-CT is best used on *ex-vivo* samples, and very small *in-vivo* samples to avoid an excessive absorbed dose emitted to patients [22]. Consequently, the use of of micro-CT for *in-vivo* radiologic evaluation of the palate is impractical at this time. Therefore, continued improvements to micro-CT technology including decreasing the emitted radiation while maintaining superior resolution, is necessary prior to implementation of such a technique on active RME patients.

An area of interest is the development of a CT-based strain assessment of peri-sutural and maxillary tissues, the development of which the authors believe will help facilitate predicting the success of RME treatment[10].

Modality #3 - Ultrasonography (US) and assignment of semi-quantitative bone fill scores (0-3).

Sumer et al. [9] utilized ultrasonography to evaluate sutural mineralization at 5 time points during the SARME and retention protocol for 3 patients, scoring each patient's palatal suture calcification via assignment of a semi-quantitative bone fill scores (0-3).

Ultrasonography findings in the Sumer et al. study[9] demonstrated that immediately post-expansion all subjects had a bone fill score =0. (Table 1.2) Two of the three subjects at 2 and 4 months post-expansion were identified as having a bone score = 1, while the remaining subject was determined to have a bone fill score =2 for these same time periods. Following the removal of the tooth-borne appliance at 6 months and 2 months subsequent to that during continued fixed appliance therapy, the bone scores for two of the subjects demonstrated increased mineralization and identification of echogenic material, having bone fill scores =2. The remaining patient received a bone fill score=3 due to incomplete transmission of the waves and 100% echogenicity measured at these respective time points[9]. (Table 1.2) It should be noted that no statistics were reported by the authors.

The results of this study follow those of a similar animal study[23], such that there was a statistically significant increase in bone fill scores that were directly related to the length of time the patient has been in retention post expansion. A major advantage to US is it's low cost and non-invasiveness[9, 23, 24], as well as improved usability compared to other methodologies, with the ability to perform real-time chair side evaluations with smaller hand held units. Additionally, ultrasonography is a reliable method to image early bone

formation as demonstrated by previous studies involving distraction osteogenesis [9, 23, 24]. A study comparing US to normal panoramic radiography, demonstrated that the efficacy of US to measure an osteotomy gap during distraction osteogenesis is equal to that of the conventional radiography[9, 25]. Ultrasonography also demonstrated increased reliability compared to panoramic radiography to evaluate the maturation of early bone formation[9, 25] in the distraction gap. A disadvantage to US is its inability to penetrate cortical bone[9]. However, following SARME or successful RME the osteotomy gap and its margins are easily visualized[9]. An area of significant future interest is to ascertain whether this technology can penetrate an immature mid-palatal suture prior to the start of RME treatment, and allow the clinician to perform a chair side subjective evaluation of the bone maturity and interdigitation along the whole length of the suture. Limitations to this study included a very small sample size of three patients, and lack of a gold standard (histology) or CT to validate the findings. Consequently, an area of future research is the use of this technology and bone fills scores in a similar larger sample size study in conjunction with a gold standard methodology to support the findings[9].

Modality #4 - Cone Beam Computed Tomography (CBCT) and proposed maturation stages.

Angelieri et al.[1] utilized a standardized methodology to capture axial CBCT cross-sections of the palatal suture to provide individual staging of midpalatal suture maturation from the author's proposed maturation stages (A-E).

As it relates to Angelieri et al.[1], a validation study performed reported a weighted Kappa statistic for intra- and inter-examiner reliability to be $\kappa = 0.75$ (95% CI, 0.64-0.99) and be

κ =0.79 (95% CI, 0.60-0.97) (no p-value reported), respectively. Due to a lack of an histologic or micro-ct gold standard, the authors also reported examiner reliability compared to the "ground truth", a descriptor used to represent consensus among examiners with the principal investigator' radiographic evaluations or other interpretations. Examiner reliability with ground truth ranged from κ =0.82 (95% CI, 0.64-0.99) to κ =0.93 (95% CI, 0.86-1.00) (no p-value reported)[1].

Results of the validation study demonstrated "almost perfect" inter-examiner reliability with the "ground truth", however the authors did not report appropriate p-values with their statistics. As was mentioned before, there was no reference standard utilized during the validation study, but rather utilized what the authors termed the "ground truth"[1], the professional opinion of the principal investigator when utilizing their own proposed maturation stages to classify each patients' sutural maturation. Due to the lack of a gold standard, nor listed p-values, the results of the validation should be cautioned. An additional limitation of this methodology is the proposed novel palatal suture maturation classification system itself. The authors developed the stages (A-E) based on comparison of CBCT axial cross-sections of the palatal suture to the perceived likeness of this radiographic morphology to the histological morphology of the suture as determined by previous studies[13-15]. Theoretically direct comparison of the histological morphology to the CBCT morphology of the suture is incompatible due to the histological assessment being on the microscopic scale as compared to the macro or eye level scale of sutures depicted in the CBCT axial slices. Consequently, any inference or direct translation of the sutural histological appearance and subsequent development of CBCT based sutural maturation stages is not possible. Therefore, the findings and developed maturational

stages should be used with caution, and should not drive clinical decision making. Rather, at best, this maturational staging may be used as part of an extended protocol to subjectively assess palatal suture maturity during the treatment planning process. Future studies to thoroughly validate the proposed maturation stages to an available gold standard is advised.

Modality #5 - CBCT and fractal analysis to quantitatively ascertain degree of sutural maturation per proposed maturation stages of Anglieri et al.[1].

Kwak et al.[12] utilized CBCT imaging in conjunction with quantitative fractal analysis to ascertain if this analysis can be correlated to the maturational stage of each subjects palatal suture. Conceptually fractal analysis is based on the observation that cranial sutures can be visualized as a fractal pattern[16], the dimensions of which are directly related to localized stresses experienced[12]. Additionally, the closer the approximation of two articulating bones, the more complex sutural morphology[12] suggestive of a more mature suture. Conceptually sound, fractal analysis has demonstrated its applicability in various areas dental research[26].

Fractal dimension intra- and inter-reliability results from the Kwak et al. study[12] demonstrated agreement with calculated weighted kappa coefficients of 0.84 (95% confidence interval [CI] 0.74-0.93) and 0.67 (95% CI 0.38–0.95) to 0.72 (95% CI 0.48–0.97), respectively (table 1.2). The CVM index inter- and intra-examiner reliability demonstrated agreement with weighted kappa coefficients from 0.69 (95% CI 0.53–0.86) and 0.71 (95% CI 0.56−0.86), respectively. The authors reported that none of the patients investigated possessed a CVM 1-IV nor was any subject classified as having palatal

suture maturational stage A. It was found that 13 of 21 subjects with CVM V were classified as having maturational stage B or C (61.9%; males 77.8%, females 50.0%). Additionally, 42 of 110 subjects with CVM VI were classified as having maturational stage B or C (38.2%; males 41.6%, females 34.0%). Post-hoc analysis demonstrated that maturational stages B, C, D and E were related to differences in mean fractal dimension ($p < 0.05$). A negative correlation existed between fractal dimension and maturation stage (-0.623, $p < 0.001$). Male and Female correlation coefficients determined to be -0.649 ($p < 0.001$) and -0.569 ($p < 0.001$) respectively. A receiver operating characteristic (ROC) curve was generated and determined the boundary between dichotomous maturation stages A–C and D or E, allowing for fractal dimension to be used to identify midpalatal suture fusion. Predictive statistical analysis noted that fractal dimension is a statistically significant indicator capable of predicting dichotomous maturation stages ((A, B, & C) vs. (D or E) (area under ROC curve [AUC] = 0.794, $p < 0.001$)[12](table 1.2).

The study notes a significant correlation between fractal patterning and degree of maturation of the midpalatal suture, and consequently the authors feel that fractal analysis can provide an objective and quantitative methodology to assess palatal suture maturity[12].

Disadvantages of this methodology include requiring significant training and proficiency in classifying the maturation stage of palatal sutures as proposed by Angelieri et al.[1]. Another disadvantage is requiring the clinician to have significant familiarity with image processing and possessing necessary software. Consequently, the time, cost and resources to do so may be prohibitive to clinicians. Additionally, this modality relies on complex statistical analyses to determine the variable (optimal cut-off value) to predict the

dichotomous maturation stage of the patient's palatal suture. Kwak et al. [12] argue that if an individual's fractal dimensions can be compared, it may provide a straightforward and clinically viable method to assess the maturation of the palatal suture and aid in clinical decision making as it relates to the modality of expansion at the diagnostic record visit[12]. Conversely, the authors do note a variety of methods to calculate fractal dimensions, and the fact these varying techniques produce different fractal dimension values. Consequently, Kwak et al.[12] argue for a more agreed upon method for its calculation to be utilized clinically.

Performing and interpreting these analyses requires significant advanced knowledge of statistics. Ultimately it is the view of the authors that this methodology is impractical in terms of time, cost, resources and knowledge required to complete this methodology for each patient as part of their diagnostic work up in day-to-day clinical practice. Furthermore, as was stated previously in the discussion, utilization of the crudely proposed maturational staging as defined by Angelieri et al.[1] should be used with caution, and lacks validation to a reference standard as does this study as mentioned by Kwak et al.[12]. Further areas of interest include the development of a ratio comparing the fractal dimensions of a mature coronal suture to that of the mid-palatal suture[12]. Additionally, improvement in the accuracy of the methodology may be gained by refinement and minimization of the number of actions needed to determine fractal dimensions[12].

Limitations

As mentioned before significant methodological differences were identified (sample size, *in vitro* vs. *in viv*o, imaging technique used, lack of adequate reference standard). The results were non-homogenous consequently a meta-analysis could not be performed, nor direct comparison of the studies possible, limiting any major conclusions regarding these newer contemporary methodologies to assess mid-palatal sutural maturation. Selection bias may have been introduced due to withdrawing one article for lack of translation services (Appendix 2). Overall, these studies did not present solid evidence of their validity for the accurate determination of the maturation of the palatal suture. As a consequence of this weak body of evidence, it is of utmost importance that clinicians use a multitude of diagnostic criteria to properly direct clinical decision making as it pertains to the maturity of the mid palatal suture and appropriate modality of expansion, namely RME or SARME.

1.2.5 Conclusions:

- Only a weak body of evidence exists to support the newest technologies and proposed methodologies that evaluate the extent of mid palatal suture maturation.
- All discussed novel methodologies lack validation with histological reference/gold standard. Consequently, it is advised that clinicians use a multitude of diagnostic criteria to subjectively assess palatal suture maturation and drive clinical decision-making as it relates to the appropriate treatment of maxillary skeletal transverse deficiency in late adolescents and young adults (RME vs. SARME).
- Future considerations in the imaging and assessment of the mid-palatal sutural

maturation will likely include some form of invasive CT technology, and proposed methodologies should follow appropriate ALARA radiation safety protocols.

- Non-invasive imaging technologies such as ultrasound present a promising and biologically safe alternative to assess mid-palatal sutural ossification.

Funding

The authors report no funding or conflicts of interest.

1.2.6 Literature Cited

1. McNamara, J.A., *Maxillary transverse deficiency.* Am J Orthod Dentofacial Orthop, 2000. **117**(5): p. 567-70.
2. Bishara, S.E. and R.N. Staley, *Maxillary expansion: clinical implications.* Am J Orthod Dentofacial Orthop, 1987. **91**(1): p. 3-14.
3. Angelieri, F., et al., *Midpalatal suture maturation: classification method for individual assessment before rapid maxillary expansion.* Am J Orthod Dentofacial Orthop, 2013. **144**(5): p. 759-69.
4. Isiksal, E., S. Hazar, and S. Akyalcin, *Smile esthetics: perception and comparison of treated and untreated smiles.* Am J Orthod Dentofacial Orthop, 2006. **129**(1): p. 8-16.
5. Lin, L., et al., *Tooth-borne vs bone-borne rapid maxillary expanders in late adolescence.* Angle Orthod, 2015. **85**(2): p. 253-62.
6. Baccetti, T., et al., *Treatment timing for rapid maxillary expansion.* Angle Orthod, 2001. **71**(5): p. 343-50.
7. Melsen, B., *Palatal growth studied on human autopsy material. A histologic microradiographic study.* American Journal of Orthodontics, 1975. **68**(1): p. 42-54.
8. Persson, M. and B. Thilander, *Palatal suture closure in man from 15 to 35 years of age.* American Journal of Orthodontics, 1977. **72**(1): p. 42-52.

9. Sumer, A.P., et al., *Ultrasonography in the evaluation of midpalatal suture in surgically assisted rapid maxillary expansion.* Journal of Craniofacial Surgery, 2012. **23**(5): p. 1375-7.

10. Korbmacher, H., et al., *Age-dependent three-dimensional microcomputed tomography analysis of the human midpalatal suture.* Journal of Orofacial Orthopedics, 2007. **68**(5): p. 364-76.

11. Franchi, L., et al., *Modifications of midpalatal sutural density induced by rapid maxillary expansion: A low-dose computed-tomography evaluation.* American Journal of Orthodontics & Dentofacial Orthopedics, 2010. **137**(4): p. 486-8; discussion 12A-13A.

12. Kwak, K.H., et al., *Quantitative evaluation of midpalatal suture maturation via fractal analysis.* Korean Journal of Orthodontics, 2016. **46**(5): p. 323-330.

13. Persson, M., B.C. Magnusson, and B. Thilander, *Sutural closure in rabbit and man: a morphological and histochemical study.* J Anat, 1978. **125**(Pt 2): p. 313-21.

14. Cohen, M.M., Jr., *Sutural biology and the correlates of craniosynostosis.* Am J Med Genet, 1993. **47**(5): p. 581-616.

15. Sun, Z., E. Lee, and S.W. Herring, *Cranial Sutures and Bones: Growth and Fusion in Relation to Masticatory Strain.* Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology, 2004. **276**(2): p. 150-161.

16. Yu, J.C., et al., *A fractal analysis of human cranial sutures.* Cleft Palate Craniofac J, 2003. **40**(4): p. 409-15.

17. Shapurian, T., et al., *Quantitative evaluation of bone density using the Hounsfield index.* Int J Oral Maxillofac Implants, 2006. **21**(2): p. 290-7.

18. Duckmanton, N.A., et al., *Imaging for predictable maxillary implants.* Int J Prosthodont, 1994. **7**(1): p. 77-80.

19. Norton, M.R. and C. Gamble, *Bone classification: an objective scale of bone density using the computerized tomography scan.* Clin Oral Implants Res, 2001. **12**(1): p. 79-84.

20. Cleall, J.F., et al., *EXPANSION OF THE MIDPALATAL SUTURE IN THE MONKEY.* Angle Orthod, 1965. **35**: p. 23-35.

21. Ballanti, F., et al., *Immediate and post-retention effects of rapid maxillary expansion investigated by computed tomography in growing patients.* Angle Orthod, 2009. **79**(1): p. 24-9.

22. Perilli, E., I.H. Parkinson, and K.J. Reynolds, *Micro-CT examination of human bone: from biopsies towards the entire organ.* Ann Ist Super Sanita, 2012. **48**(1): p. 75-82.

23. Thurmuller, P., et al., *Use of ultrasound to assess healing of a mandibular distraction wound.* J Oral Maxillofac Surg, 2002. **60**(9): p. 1038-44.

24. Hughes, C.W., et al., *Ultrasound monitoring of distraction osteogenesis.* Br J Oral Maxillofac Surg, 2003. **41**(4): p. 256-8.

25. Bruno, C., et al., *Gray-scale ultrasonography in the evaluation of bone callus in distraction osteogenesis of the mandible: initial findings.* Eur Radiol, 2008. **18**(5): p. 1012-7.

26. Sanchez, I. and G. Uzcategui, *Fractals in dentistry.* J Dent, 2011. **39**(4): p. 273-92.

1.2.7 Appendices:

Appendix 1: Database search and results from 1984 – October 2016

Database	Keywords	New
Medline	(1) Palate.mp or exp palate/ or ex palate, Hard/ (2) Cranial suture. mp or exp cranial sutures (3) Maturation.mp (4) Interdigitation.mp or exp cranial sutures (5) Ossification.mp 1 and 2 and (3 or 4 or 5)	221
Pubmed	(("palate") AND "cranial sutures") AND (("maturation" OR "interdigitation" OR "ossification"))	31
Embase	(1) Palate.mp or exp palate/ or ex primary palate/ or exp secondary palate or exp, hard palate/ (2) Cranial suture. mp or exp cranial suture; (3) Exp maturation/ or exp bone maturation/ or maturation.mp (4) Interdigitation.mp (5) Ossification.mp or exp ossification/ 1 and 2 and (3 or 4 or 5)	31
Scopus	Palate AND cranial sutures AND (maturation OR interdigitation OR ossification) Subjects limited to Medicine, biochemistry, genetics and molecular biology, dentistry Documents limited to articles only	422

Appendix 2: Articles excluded from final inclusion criteria and reasons

Author(s)	Reason Excluded
Fricke-Zech et al. [27]	Animal study (porcine model) using flat-panel volume computed tomography (fpVCT)
Gao et al. [28]	Animal study (dog) using transmission electron microscope (TEM)
Hahn et al. [29]	Animal study (pig) using fpVCT and multislice-computed tomography (MSCT)
Beauthier et al. [30]	Methodology is not novel, used in forensics for age of death using maxilla
Cheung et al. [31]	Animal study (Rhesus monkeys) using conventional radiography and microcomputerized scanning
De Melo Mde et al. [32]	Study used conventional digital radiography
Gurgel jde et al. [33]	Study used conventional radiography that was latter digitized
Kjaer et al. [34]	Study used conventional radiography
Knaup et al. [35]	Study used established histological analysis
Lee et al. [36]	Study used conventional radiography and histological analysis
Leonardi et al. [37]	Study evaluated spheno-occipital synchondroses displacement during RME not mid-palatal suture using multislice multidetector CT
Leonardi et al. [38]	Study evaluated circumaxillary sutures during RME not mid-palatal suture using multislice multidetector CT
Kjaer et al. [39]	Study used conventional radiography and histological analysis
Agrawal et al. [40]	Study used conventional radiography to evaluate patency of cranial suture following surgical treatment of cranial synostosis
Bradley et al. [41]	Animal study (mouse) that utilized conventional histological analysis
Captier et al. [42]	Study evaluated sphenofrontal suture with light microscopy
Lauridsen et al. [43]	Study used established histological analysis
Corega et al. [44]	Study assessed coronal cranial sutures using Micro-CT
Corega et al. [45]	Study assessed coronal cranial sutures using Micro-CT
Bjork et al. [46]	Article in German, no translation services available.
De Araujo Grugel et al. [47]	Study used conventional radiography
Sannomiya et al. [48]	Study used conventional radiography
Takenouchi et al. [49]	Animal study (rats) utilizing in vivo micro-computed tomography (mCT).
Acar et al. [50]	Study used previously described novel technique using Hounsfield units to assign bone density at the palatal suture.

References for Appendix

1. Fricke-Zech, S., et al., *Measurement of the midpalatal suture width.* Angle Orthodontist, 2012. **82**(1): p. 145-50.
2. Gao, Q.W., et al., *[An ultrastructure study on the palatomaxillary suture of dog expanded by NiTi-SMA].* Zhonghua Zheng Xing Wai Ke Za Zhi, 2009. **25**(4): p. 277-9.
3. Hahn, W., et al., *Imaging of the midpalatal suture in a porcine model: flat-panel volume computed tomography compared with multislice computed tomography.* Oral Surgery Oral Medicine Oral Pathology Oral Radiology & Endodontics, 2009. **108**(3): p. 443-9.
4. Beauthier, J.P., et al., *Palatine sutures as age indicator: a controlled study in the elderly.* Journal of Forensic Sciences, 2010. **55**(1): p. 153-8.
5. Cheung, L.K. and Q. Zhang, *Radiologic characterization of new bone generated from distraction after maxillary bone transport.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics, 2003. **96**(2): p. 234-242.
6. de Melo Mde, F., et al., *Digital radiographic evaluation of the midpalatal suture in patients submitted to rapid maxillary expansion.* Indian Journal of Dental Research, 2013. **24**(1): p. 76-80.
7. Gurgel Jde, A., M.F. Malmstrom, and C.R. Pinzan-Vercelino, *Ossification of the midpalatal suture after surgically assisted rapid maxillary expansion.* European Journal of Orthodontics, 2012. **34**(1): p. 39-43.
8. Kjaer, I., *Human prenatal palatal shelf elevation related to craniofacial skeletal maturation.* European Journal of Orthodontics, 1992. **14**(1): p. 26-30.
9. Knaup, B., F. Yildizhan, and H. Wehrbein, *Age-related changes in the midpalatal suture. A histomorphometric study.* Journal of Orofacial Orthopedics, 2004. **65**(6): p. 467-74.
10. Lee, S.K., et al., *Prenatal growth pattern of the human maxilla.* Acta Anatomica, 1992. **145**(1): p. 1-10.
11. Leonardi, R., A. Cutrera, and E. Barbato, *Rapid maxillary expansion affects the spheno-occipital synchondrosis in youngsters. A study with low-dose computed tomography.* Angle Orthodontist, 2010. **80**(1): p. 106-10.
12. Leonardi, R., et al., *Early post-treatment changes of circumaxillary sutures in young patients treated with rapid maxillary expansion.* Angle Orthodontist, 2011. **81**(1): p. 36-41.
13. Kjaer, I., *Prenatal skeletal maturation of the human maxilla.* Journal of Craniofacial Genetics & Developmental Biology, 1989. **9**(3): p. 257-64.
14. Agrawal, D., P. Steinbok, and D.D. Cochrane, *Reformation of the sagittal suture following surgery for isolated sagittal craniosynostosis.* Journal of Neurosurgery, 2006. **105 PEDIATRICS**(SUPPL. 2): p. 115-117.
15. Bradley, J.P., et al., *Studies in cranial suture biology: IV. Temporal sequence of posterior frontal cranial suture fusion in the mouse.* Plastic and Reconstructive Surgery, 1996. **98**(6): p. 1039-1345.
16. Captier, G., et al., *Prenatal organization and morphogenesis of the sphenofrontal suture in humans.* Cells Tissues Organs, 2003. **175**(2): p. 98-104.
17. Lauridsen, H., et al., *Histological investigation of the palatine bone in prenatal trisomy 21.* Cleft Palate-Craniofacial Journal, 2001. **38**(5): p. 492-7.

18. Corega, C., et al., *Three-dimensional cranial suture morphology analysis.* Romanian Journal of Morphology and Embryology, 2010. **51**(1): p. 123-127.
19. Corega, C., et al., *Cranial sutures and diploae morphology.* Romanian Journal of Morphology and Embryology, 2013. **54**(4): p. 1157-1160.
20. Bjork, A. and V. Skieller, *[Growth and development of the maxillary complex]*. Informationen aus Orthodontie und Kieferorthopadie, 1984. **16**(1): p. 9-52.
21. De Araujo Gurgel, J., M.F.V. Malmstrom, and C.R.M. Pinzan-Vercelino, *Ossification of the midpalatal suture after surgically assisted rapid maxillary expansion.* European Journal of Orthodontics, 2012. **34**(1): p. 39-43.
22. Sannomiya, E.K., et al., *Evaluation of optical density of the midpalatal suture 3 months after surgically assisted rapid maxillary expansion.* Dento-Maxillo-Facial Radiology, 2007. **36**(2): p. 97-101.
23. Takenouchi, H., et al., *Longitudinal quantitative evaluation of the mid-palatal suture after rapid expansion using in vivo micro-CT.* Archives of Oral Biology, 2014. **59**(4): p. 414-23.
24. Acar, Y.B., M. Motro, and A.N. Erverdi, *Hounsfield Units: a new indicator showing maxillary resistance in rapid maxillary expansion cases?* Angle Orthodontist, 2015. **85**(1): p. 109-16.

Chapter 2: Reliability testing of a novel palatal suture maturation classification

2.1 Introduction:

Maturation of the midpalatal suture (MPS or PS) occurs by intramembranous ossification[13]; however there is significant variation in the skeletal maturation timing amongst individuals[2] as the fusion is poorly correlated with patient age and sex[1]. Consequently, RME treatment has unpredictable outcomes in still growing but older patients, namely late adolescents and young adults. Failure to properly predict the degree of midpalatal fusion in these patients can lead to choosing the least favorable and potentially damaging expansion modality. Iatrogenic effects such as acute pain, recession, mucosal necrosis, buccal tipping and poor stability[1, 5], can result if tooth-borne RME is incorrectly selected for a patient with a near or fully fused palate. Conversely prematurely committing a patient to surgically assisted rapid palatal expansion (SARPE) creates a burden of increased cost, pain and healing time to the patient. Current gold standard methodologies (biopsy) to assess the degree of palatal suture fusion present are invasive and for obvious reasons not performed on live patients. The advent of accessible and often times, in-office, cone-beam computed tomography (CBCT) offers a low cost, low radiation modality for individual visualization of the palatal suture morphology[1]. Angelieri et. al[1]developed a novel classification system to image through CBCT the palatal suture and determine the palatal suture maturation stage using an ordinal scale to assign a particular maturation stage (A-E) to the degree of predicted palatal fusion in a patient (Figure 1) where stage A represents the most immature suture and stage E represents a mature and fused MPS. Since successful RME treatment is dependent upon the degree of palatal fusion and properly chosen modality, the authors suggested that their methodology and classification system can direct the

appropriate treatment decision making to avoid significant comorbidities of an incorrectly chosen expansion modality.

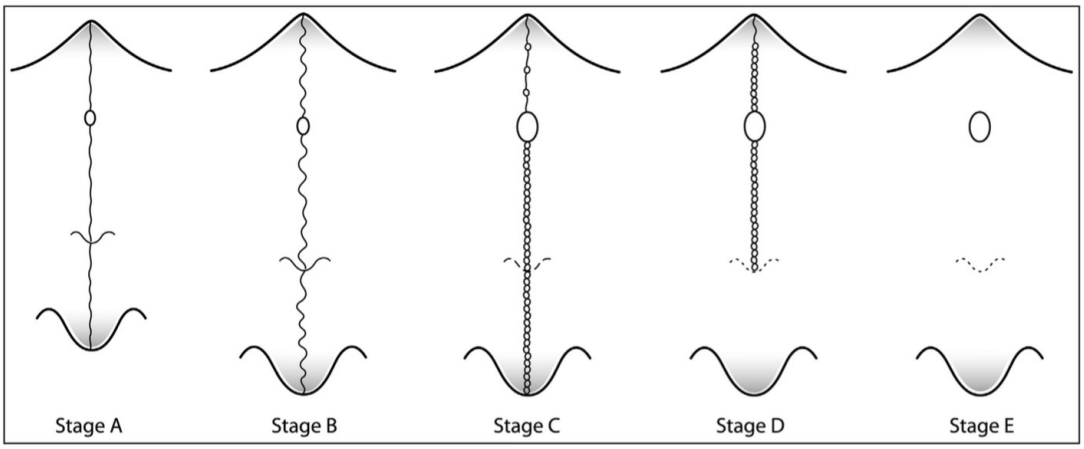


Figure 2.1: Diagrammatic representation of the developed novel palatal suture maturation classification[1] identifying key radiological morphological characteristics specific to each maturity level. Stage A is defined by a single non-scalloped line, stage B is defined by a single scalloped line, stage C is defined by two parallel, scalloped, high density radiopaque lines with small radiolucent spaces between the lines, stage D is defined by maturation of the 2 high density lines in the palatine bone, stage E is defined by complete maturation of the palatal suture and no visible lines in at least a segment of the maxillary bone.

Reliability testing evaluates the degree of observer reproducibility and agreement of measurement outcomes[51]. Inter-examiner reliability refers to the degree to which different observers agree, whereas intra-examiner reliability refers to the degree to which the same rater agree with him or herself over multiple measurement sessions [52]. Due to the nature of the ordinal scale used in the developed sutural maturation classification (A-E), Angelieri et. al.[1] utilized weighted kappa statistic to evaluate intra-examiner agreement and inter-examiner agreement.

In addition to calculating inter- and intra-examiner reliability, in the radiographic study performed by Angelieri et. al.[1], the agreement of the examiners to what the authors called, "ground truth" was evaluated. The authors recognize that the gold standard to evaluate sutural maturation, namely biopsy and histological evaluation of palatal suture, is impractical. *In lieu of* an unfeasible gold standard, the authors opted to use a *ground truth* (expert opinion) that equates to a consensus of radiographic interpretation or reliable interpretation. Consequently, a single principal investigator evaluated the maturity of the palatal sutures visible on constructed radiographic images, and this investigator's interpretation was considered reliable to a degree such that it was considered the *ground truth*[1].

The results of the Angelieri et. al.[1] validation study noted significant intra- and inter-examiner agreement, ranging from $\kappa = 0.75$ (95% confidence interval [CI], 0.57-0.93) to 0.79 (95% CI, 0.60-0.97) [no p-values given]. Additionally, the agreement of observers to the previously mentioned ground truth noted even greater agreement, ranging from $\kappa = 0.82$ (95% CI, 0.64-0.99) to 0.93 (95% CI, 0.86-1.00) [no p-values given]. It should be noted no p-values were given, nor a description of the weighting used to calculate the weighted kappa statistics. Consequently, there is no understanding of the magnitude of the differences in agreement amongst raters or readings.

Although the methodology of the Angelieri et. al.[1] study is promising, further validation of this novel technique is necessary to establish its efficacy and reliability to clinical practice. The primary research question of this study is *what is the reliability of the Angelieri et al midpalatal suture maturation classification system*?

2.2 Methods and Materials

Reliability testing focused on a total of sixteen patients aged 9.5 -17 years old with early mixed to full permanent dentition who received routine pre-treatment CBCTs taken during preliminary record appointments. The 16 patients were chosen at random and represented all previously defined[1] maturational stages (A-E). Two researchers were blinded (DAI and CF) to all 16 subjects presented during the reliability testing. The remaining reliability testing clinician (ML) supplied at random twelve of the sixteen reliability test subjects from a database of patients from a previous expansion study[53]. Clinician ML had access to these twelve patients demographic information as well as pre- and post-RME treatment outcomes if he so chose to evaluate these. However, these patients were numerically coded with the master key never being utilized by ML. Consequently, for the purpose of reliability testing ML was also considered to be blinded Prior to reliability testing, calibration of the involved clinicians was performed. Calibration and reliability testing was carried out by 3 experienced clinicians [DAI, ML and CF], each of which has at minimum 2.5 years of experience in the diagnostic interpretation of CBCT for clinical and/or research purposes. Throughout the course of the calibration and reliability testing, evaluations performed by DAI were considered the *ground truth*, as previously defined in the introduction. Additionally, for the purpose of clinician calibration, reliability testing and completion of this study, the developed novel classification developed by Angelieri et. al[1] (Figure 2.1) will be used to identify the appropriate palatal maturational stage of evaluated subjects.

Clinician calibration was performed via explanation of each palatal stages radiographic morphology according to Angelieri et al.[1], as well as visually depicting each palatal

stage in a Power Point presentation (Microsoft Office PowerPoint 2007; Microsoft, Redmond, Wash) with a white background[1]. During the calibration procedure clinicians ML and CF were asked to evaluate seven palatal sections separate from the reliability testing sample, representative of maturational stages (A-E), in dimly lit settings without changes to contrast, brightness or other visual modifications. Clinicians ML and CF were able to openly discuss with each other and DAI the appropriate grading of the palatal maturation of these subjects. Any discrepancies or clarification of the stages were open to full discussion and further understanding as needed.

Following calibration, DAI delivered a second presentation (Microsoft Office PowerPoint 2007; Microsoft, Redmond, Wash) regarding image analysis, detailing the previously developed protocol[1] to isolate palatal axial cross-sections from patient 3D volumes for the purpose of reliability testing. Significant variations or adaptations from this previously described protocol[1] will be italicized.

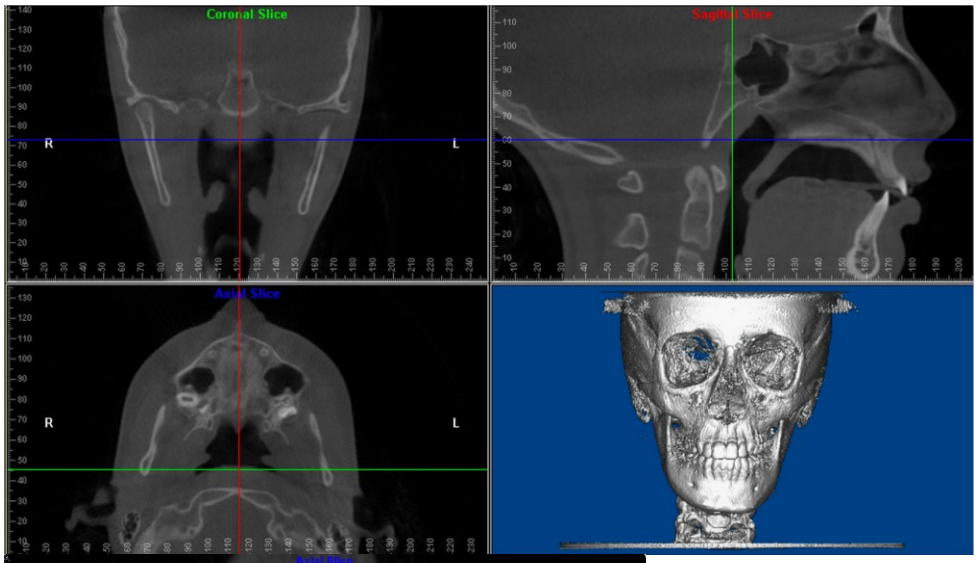
Angelieri et al[1] Image Analysis Methodology:

Image analysis was performed using *dolphin software (Chatsworth, Calif).* The following stepwise procedures were performed to isolate palatal axial cross-sections for the evaluation of each subject's palatal stage maturation stage.

1. Subject Head Orientation. In the sagittal plane the palate was made to be perpendicular to a true vertical reference plane from the position indicating cross-hairs. In both the coronal and axial plane the true vertical reference plane from the position indicating crosshairs were placed in the mid-sagittal position (figure 2.2a).

2. Standardization of palatal axial cross-sections. In the sagittal plane the horizontal reference plane of the position indicating crosshairs was then placed to bisect the superior-inferior width of the subject's palate (figure 2.2a). The axial cross-sections developed at this superior-inferior height was used to assess the maturational stage of the subject in the axial cross-section (figure 2.2b). The slice thickness of the axial cross-section was the default thickness of 1.0mm. To evaluate the maturational level of a subject with a curved palate, two axial slices were developed (anterior and posterior) according to the protocol defined above and evaluated in concert to identify the appropriate stage. A subject is considered to have a curved palate when only one aspect of the palate, either anterior or posterior, can be viewed in the axial slice while attempting to standardize the palatal axial slice. If a subject's palate was considered to be thick, 2 axial slices were generated by bisecting the superior-inferior height of the palate at two different vertical heights (oral and nasal). A thick palate is one in which the suture can be identified in a minimum of 3 different axial slices at different superior-inferior heights (oral, central and nasal).

(a)



(b)

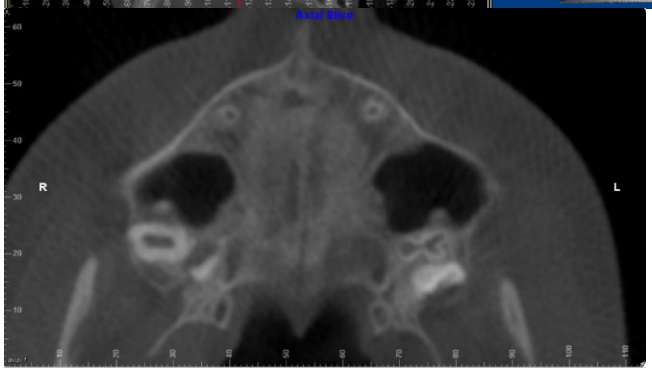


Figure 2.2: (a) Representation on proper standardization of head position for image analysis and. **(b)** axial cross section of the mid-palatal suture generated from this protocol.

Reliability testing protocol

Clinicians CF and ML were asked to evaluate the maturational stages of the 16 reliability testing subjects in one viewing session. The clinician representing the *ground truth* (DAI) performed three rounds of evaluations of these patients all 48 hours apart without randomization. All clinicians used the image analysis procedure above to isolate the axial cross-sections of the 16 subjects involved in the reliability testing. After isolating the 16 palatal axial cross-sections the clinicians were asked to define the maturational stage in

dimly lit settings without changes to contrast, brightness or other visual modifications. Reclassification procedures occurred at the intervals described above.

2.3 Statistical Analysis:

Statistical analysis

Intra-examiner reliability (DAI to DAI) and inter-examiner reliability (ML to CF), as well as examiner to ground truth (ML to DAI & CF to DAI) were investigated by Cohen's Kappa statistic.

Hypotheses

Intra-examiner reliability: Principal investigator DAI (ground truth) performed three different readings (#1, #2, #3) of palatal suture classifications. To assess DAI intra-examiner reliability, Microsoft Excel (2007) random number generator function was utilized to randomly determine which reading session would be used in the calculation. The random number generator assigned reading session #1 and #3 to be used for intra-examiner reliability calculation.

null hypothesis: the intra-examiner agreement between classification sessions for rater DAI is no different than chance agreement. H_0: $\kappa = 0$

Alternative hypothesis: the intra-examiner agreement between classification sessions for rater DAI is different than chance agreement. H_A: $\kappa \neq 0$

Inter-examiner reliability examiner ML and CF:

null hypothesis: the agreement between raters is no different than chance agreement. H_0: $\kappa = 0$

Alternative hypothesis: the agreement between raters is different than chance

agreement. H_A: $\kappa \neq 0$

Agreement between examiner ML and Ground Truth (DAI), as well as CF and Ground Truth (DAI): To assess agreement between examiner and ground truth, Microsoft Excel (2007) random number generator was utilized to randomly determine which reading session by DAI would be used in the in the calculation. The random number generator assigned reading session 1 to be used for the gold standard calculations.

null hypothesis: the agreement between rater and ground truth is no different than chance agreement. H_0: $\kappa = 0$

Alternative hypothesis: the agreement between rater and ground truth is different than chance agreement. H_A: $\kappa \neq 0$

For all calculations a p-value of $p < 0.05$ was considered significant. The Statistical Package for the Social Sciences (SPSS) 23.0v for Windows (SPSS Inc., Chicago, IL, USA) was utilized for statistical analyses. The following guidelines were used to interpret the calculated Kappa coefficients[51, 54]:

- Kappa <0.00 denotes poor agreement
- 0.00 ≤ Kappa ≤ 0.20 denotes slight agreement
- 0.21 ≤ Kappa < 0.40 denotes fair agreement
- 0.41 ≤ Kappa < 0.60 denotes moderate agreement
- 0.61 ≤ Kappa < 0.80 denotes substantial agreement
- 0.81 ≤ Kappa ≤ 1.00 denotes almost perfect agreement

2.4 Results:

Kappa coefficient (κ), was calculated to determine intra-examiner, inter-examiner and rater to ground truth agreement. Results have been tabulated (table 2.1) and raw data, descriptive statistics and 95% confidence interval sample calculation presented [appendix 2.1 -2.5].

	Intra-examiner Reliability (DAI1vs3)	Inter-examiner Reliability (MLvsCF)	ML to Ground Truth (DAI_1)	CF to Ground Truth (DAI_1)
Kappa (κ)	k=0.915 (95% CI, 0.752 to 1.078), p <0.005	k=0.040 (95% CI, -0.209 to 0.289), p =0.733	k=0.470 (95% CI, 0.141 to 0.799), p =0.001	k= -0.015 (95% CI, -0.25 to 0.22), p =0.896

Table 2.1: intra-examiner, inter-examiner, and rater to ground truth agreement from classification of 16 patients palatal suture maturation

There was almost perfect intra-examiner agreement between the rater DAI's palatal suture maturation classification at time 1 and 3, κ=0.915 (95% CI, 0.752 to 1.00, p <0.005. Consequently, there is convincing evidence that agreement between classification sessions for rater DAI is far greater than chance agreement.

There was slight inter-examiner agreement between rater ML and CF's staging of the patient's palatal suture, κ =0.040 (95% CI, -0.209 to 0.289), p =0.733. Given the p>0.05, there is weak to no evidence to support this slight agreement is greater than chance agreement.

There was moderate agreement between rater ML and the ground truth for classifying the patient's palatal suture maturation, κ=0.470 (95% CI, 0.141 to 0.799), p =0.001. Consequently, there is convincing evidence that the agreement between rater ML and ground truth is moderately greater than chance agreement.

There was poor agreement between rater CF and the ground truth staging of the patient's palatal suture maturation, $\kappa = -0.015$ (95% CI, -0.25 to 0.22), $p = 0.896$. Given the $p > 0.05$, there is weak to no evidence to support this poor agreement is greater than solely chance agreement.

Cross-tabulation tables (Appendix 2.2-2.5) were generated and used to construct frequency tables (Tables 2.2 and 2.3) to evaluate trends in either over- or under classifying palatal stages by CF and ML in comparison to the ground truth examiner.

DAI raw count (n) and palatal classification	CF percent Frequency (%f) Palatal Stage A	CF percent Frequency (%f) Palatal Stage B	CF percent Frequency (%f) Palatal Stage C	CF percent Frequency (%f) Palatal Stage D	CF percent Frequency (%f) Palatal Stage E
Palatal stage A n = 3	**33.33%**	33.33%	33.33%	0%	0%
Palatal stage B n = 1	0%	**0%**	0%	100%	0%
Palatal stage C n = 7	0%	42.85%	**14.29%**	28.57%	14.29%
Palatal stage D n = 3	0%	66.67%	0%	**33.33%**	0%
Palatal stage E n = 2	0%	0%	100%	0%	**0%**

Table 2.2: Frequency table of classification of palatal stage maturation by CF in comparison to ground truth. Bolding denotes the percent frequency of CF correctly classifying the appropriate palatal stage to ground truth.

DAI raw count (n) and palatal classification	ML percent Frequency (%f) Palatal Stage A	ML percent Frequency (%f) Palatal Stage B	ML percent Frequency (%f) Palatal Stage C	ML percent Frequency (%f) Palatal Stage D	ML percent Frequency (%f) Palatal Stage E
Palatal stage A n = 3	**100%**	0%	0%	0%	0%
Palatal stage B n = 1	0%	**100%**	0%	0%	0%
Palatal stage C n = 7	0%	0%	**71.42%**	14.29%	14.29%
Palatal stage D n = 3	0%	33.33%	33.33%	**33.33%**	0%
Palatal stage E n = 2	0%	0%	100%	0%	**0%**

Table 2.3: Frequency table of classification of palatal stage maturation by ML in comparison to ground truth. Bolding denotes the percent frequency of ML correctly classifying the appropriate palatal stage to ground truth.

For ease of statistical analysis raw reliability data was transformed from alphabetical ordinal data (palatal stages A-E) to ascending numerical ordinal variables where palatal stage A =1, B=2, C=3, D=4 and E = 5, (Appendix 2.3) where an integer of one reflects a difference in palatal suture maturational level by one discrete stage. Following this Bland Altman analyses was performed and presented for intra-examiner agreement (Figure 2.3), inter-examiner agreement (Figure 2.4), agreement of CF to ground truth (Figure 2.5) and agreement of ML to ground truth (Figure 2.6). Intra-examiner agreement (Figure 2.3) demonstrated the smallest average of difference with bias = -0.06, whereas, the largest average difference being observed in the evaluation of ML to ground truth, bias = -0.25 (Figure 2.6). The smallest confidence interval was found in the plot of intra-examiner agreement (Figure 2.3), with a 95% confidence interval ranging from -0.55 to 0.43. The largest confidence interval range was found in the plot of CF to ground truth agreement (Figure 2.5), with a 95% confidence interval ranging from -3.10 to 2.82. Two outliers existed, one in the evaluation of intra-examiner agreement (Figure 2.3), as well as, ML to ground truth (Figure 2.6).

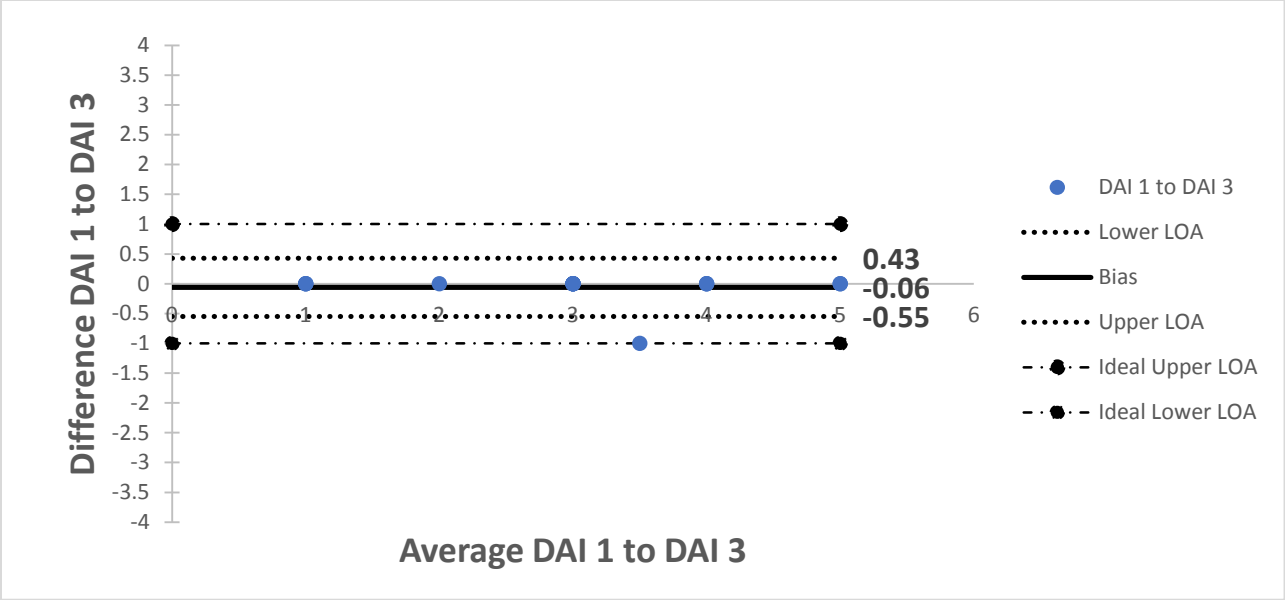


Figure 2.3: Bland Altman plot evaluating Intra-examiner reliability comparing DAI reading sessions at time 1 and 3.

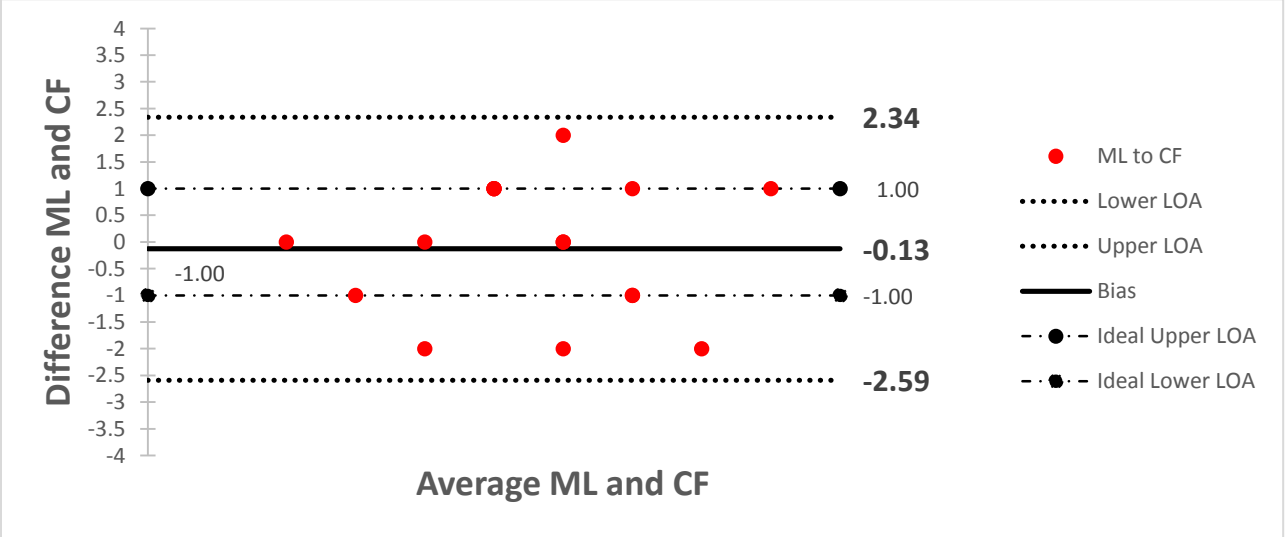


Figure 2.4: Bland Altman plot evaluating Inter-examiner reliability comparing classifications by ML to CF.

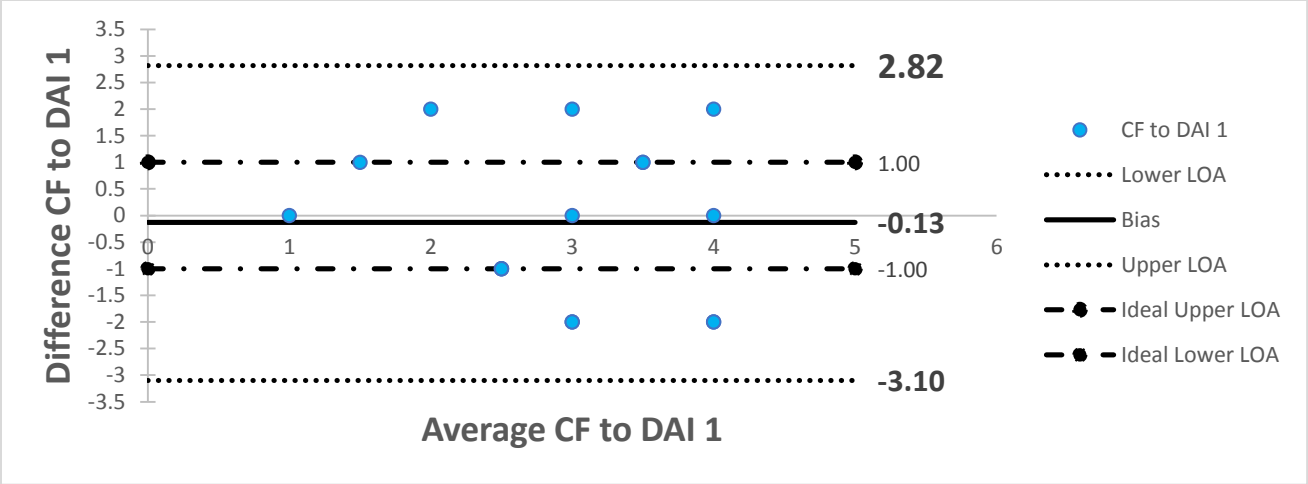


Figure 2.5: Bland Altman plot evaluating reliability comparing classifications by CF to ground truth (DAI classification session 1).

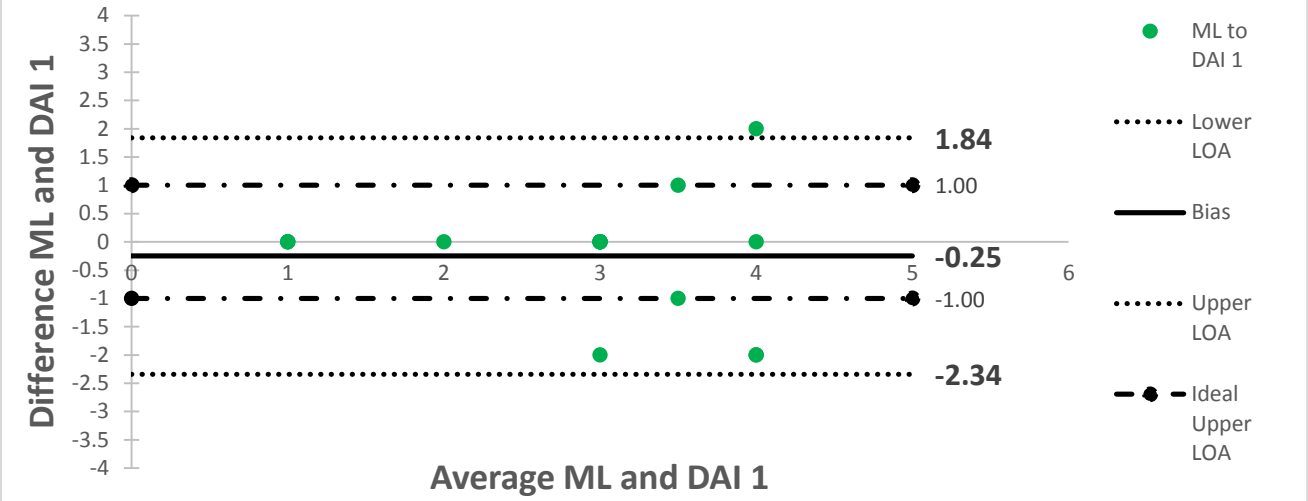


Figure 2.6: Bland Altman plot evaluating reliability comparing classifications by ML to ground truth (DAI classification session 1).

2.5 Discussion:

This study properly utilized the correct measure of agreement for ordinal data, where chance agreement has been removed, namely Cohen's Kappa. Additionally, this was close to the same measurement of agreement utilized by Angelieri et al. who utilized a weighted kappa statistic in their measurement of agreement.[1] Cohen's Kappa statistic (κ) is a measure to evaluate inter-examiner and intra-examiner reliability and validity when assessing nominal data [55] while accounting for the amount of chance observer agreement[52]. A weighted kappa allows for differences in agreement amongst raters or readings to be represented. A weighted kappa attaches increased emphasis on larger disagreements on ratings of ordinal data than smaller disagreements[54]. Cohen's kappa can range of -1 to +1, where $\kappa = 0$ equates to agreement solely based on chance alone, and perfect agreement $\kappa = 1$. A negative kappa is a rare event and can be interpreted as a distinct difference between observers, where agreement is considered to occur less than chance agreement itself disagreements[54]. Although utilizing a weighted kappa was considered for this project, ultimately the assignment of weights is complicated due to lack of a true gold standard, consequently we do not know the true classification of the palate, and ultimately it was deemed impossible to assign appropriate weighting. It could be argued to replicate the weighting used in the previous study[1] for ease of comparison however, the authors failed to provide the readers the weighting utilized, their rationale for the weighting utilized, nor p-values in regards to their agreement results. Consequently, it is advisable to cautiously accept the agreement results found in the previous study[1]. Additionally, readers should have guarded confidence in the ability to truly compare the calculated agreement amongst these two studies.

The primary investigator DAI who was identified as the ground truth had an intra-examiner reproducibility of κ=0.915 (95% CI, 0.752 to 1.078), p <0.005. Examiner DAI demonstrated near perfect intra-examiner agreement when assigning palatal suture maturation classifications at time 1 and 3, with convincing evidence that agreement between classification sessions for rater DAI is far greater than chance agreement. Primary investigator DAI exhibited near perfect intra-examiner agreement. This result can be explained by the fact examiner DAI was the principal investigator in the study and therefore, it could be assumed that DAI held a relative greater understanding of the methodology and classification compared to other examiners. Additionally, relatively speaking DAI was far more practiced in his ability to isolate and interpret axial cross-sections. Given that the classification sessions were performed only 48 hours apart, the high degree of intra-examiner agreement, may have been confounded by memory bias. Nonetheless, the results reflect this greater competence and consistency to assess palatal sections relative to examiners ML and CF, as demonstrated by the high degree of intra-examiner reliability (Table 2.1).

There was very slight inter-examiner agreement between CF and ML and no evidence to suggest that any such agreement was greater than chance alone k=0.040 (95% CI, -0.209 to 0.289), p =0.733 (table 2.1). Additionally, examiner CF showed less agreement with the ground truth than ML, having a CF-to-ground truth agreement of k= -0.015 (95% CI, -0.25 to 0.22), p =0.896 (table 2.1). Examiner CF demonstrated poor agreement when assigning palatal suture maturation classifications in comparison to the ground truth. The kappa statistic being effectively zero indicates that both inter-examiner agreement, and the agreement between CF and the ground truth was solely based on chance. The large p-

values are reflective of no convincing evidence that any inter-examiner agreement or agreement between CF and ground truth is greater than chance agreement. When focusing on both the low kappa statistic for inter-examiner agreement (ML and CF) and CF to ground truth, is suggestive of a systematic error, or a distinct difference in approach to the methodology by CF in comparison to examiners ML and DAI, such that any agreement was by chance alone. The small kappa value and systematic difference in methodological approach can be explained by a couple reasons. Firstly, during the reliability testing the initial task was to isolate the axial slice following the methodology presented. This methodology is not intuitive and requires strict adherence to its protocol, if not followed the morphology of the palatal suture developed can look vastly different than if the proper protocol followed. Refinement of the protocol in the presence of systematic error would in theory improve the agreement between CF and DAI. Secondly, clarity of the developed axial slices of the palate is affected by the voxel definition at the visual level. Undoubtedly, voxel size affects the clarity of the scanned image, most notably as voxel size increases, 3D volume clarity deteriorates[56]. In fact, when using the default slice thickness of 1.0mm, poor voxel size was a limiting factor affecting the clarity of the generated axial cross-sections, introducing reading error throughout the study. This affected the accuracy and precision to which appropriate subject classifications could be made, especially as it pertains to finer radiological differences needing to be ascertainable to delineate between palatal stages A, B and C. The specific radiological subtleties include whether or not two high-density lines were present (stage C), or whether a single hi-density line was in fact scalloped (stage B) or straight (stage A). Consequently, this lack of clarity in regards to these diagnostic criteria impeded the

proper classification of palatal stages A, B and C. The difficulty delineating these diagnostic criteria, and subsequent difficulty assigning a subject to palatal stages A, B and C is supported by the findings in the frequency table (Table 2.2) for CF. When the ground truth noted that palatal stage A was observed, CF evenly distributed the classification to palatal stages to A, B and C, while palatal stages D and E were not represented (Table 2.2). Diagnostically, considering the methodology, delineation between palatal stage C, D and E is relatively simple in theory; two-high density lines along the maxillary and palatine bone denotes palatal stage C, continued maturation and visible loss of the suture in the palatine bone is relevant to palatal stage D, while, complete or near complete lack of a suture in both the maxillary bone and palatine bone is indicative of palatal stage E. However, the decreased ability to discern between palatal stages A, B and C thus complicates this delineation in clinical applications, making the assessment between stages A and D, as well as, stages B and D, increasingly as difficult to ascertain as the difference between C and D. This is supported in the findings when the ground truth noted the presence of palatal stage B, examiner CF 100% of the time classified the subject(s) as having palatal stage D (Table 2.2). Conversely, although ML demonstrated moderate agreement greater than chance with the ground truth (k=0.470 (95% CI, 0.141 to 0.799), p =0.001), when the ground truth noted the presence of palatal stage D, examiner ML evenly distributed the classification to palatal stages to B, C and D. This finding reiterates that the differentiation between stages A, B, C, and the differences of each of these stages to stage D were clinically more difficult than anticipated given the clear criteria set out in the classification[1]. Interestingly, both examiner CF and ML 100% of the time classified a patient as having stage C, with two

high-density lines present in the palatine and maxillary bone, when the ground truth reported the patient as having stage E, a fully mature palate with minimal to no visible suture (Tables 2.2 and 2.3). Once more this could be explained by poor adherence to the methodology to isolate that particular patient's axial slice by both operators and/or is a true sign of the difficulty in clinically observing even the greatest of radiological differences between the stages (two hi density lines –stage C vs. none- stage E) given the image clarity and voxel size.

Bland Altman analysis is typically used in the evaluation of two quantitative methods of measurement, say for example if a new diagnostic test has been introduced to replace a pre-existing proven methodology. Bland Altman analysis quantifies agreement via evaluation of the mean and standard deviations of difference between methods of measurement and the construction of 95% upper and lower limits of agreement (LOA). Following this, a plot of the mean difference of the measurements (y-axis) is plotted against the average of the differences on the x-axis[57].

By nature, any instrument of measurement possesses some amount of inherent error and limitations to precisely analyze a variable. Thus, imprecision of two methods, instruments or individuals generates variability. If the variability of the differences observed are relegated to imprecision in analytical measurement, the mean difference between the methods or examiners should theoretically be null[57].That said, intra-examiner agreement (Figure 2.3) demonstrated the smallest average of difference with bias = -0.06 and denotes high accuracy in the staging of subjects palatal sutures. This minimal average difference between readings was less than disagreement in assigning a subject to a full palatal suture maturation above or below (+/- 1.00) the respective stage. In fact, inter-

examiner agreement, CF to ground truth and ML to ground truth, also demonstrated minimal average difference in readings, with the largest average difference being observed in the evaluation of ML to ground truth, bias = -0.25 (Figure 2.6). These small average differences were not large enough to be clinically important.

In regards to trends, any plot involving CF noted a consistently large variability throughout the scatter plots with large confidence intervals (Figure 2.4 and 2.5). The intra-examiner reliability plot (Figure 2.3) had the most consistent average differences and consistently narrow variability between the two classification sessions. When evaluating the plot of ML to ground truth (Figure 2.6) a trend developed where the difference in ML to DAI increased as the average increased.

The Bland Altman analysis performed utilized upper and lower limits of agreement such that 95% of the data should be present within $_{+/-}$ 2 standard deviations of the absolute mean difference[57]. A sample calculation of the upper and lower limits for the Bland Altman analyses has been provided (Appendix 2.2). The smallest 95% confidence interval observed for intra-examiner agreement ranged from -0.55 to 0.43 (Figure 2.3) and exemplifies the reliable precision to which DAI staged the patients as the ground truth during reliability testing. The largest 95% confidence interval depicted in the plot of CF to ground truth agreement (Figure 2.5), ranged from -3.10 to 2.82. Bland Altman analysis does not directly assess agreement between two methods or examiners, rather it evaluates the bias, upper and lower limits of agreement and the overall range of agreement where 95% of the differences in measurement are found. The decision as to how wide is a 95% confidence interval is in fact a clinical decision[57]. For our purposes, a confidence interval whose absolute value is equal to 5, then the whole palatal suture

maturation range (A-E alphabetically, 1-5 numerically) would be present in the interval and gives too much latitude to examiners. Appropriately a narrower and defined upper and lower LOAs should be adhered to. Consequently, it is proposed that a narrower confidence interval range with an absolute value of 2 would allow for mean differences to extend to +/- 1.00, representing the ability of an examiner to incorrectly score a patient's degree of palatal maturation by one whole stage, either being less mature (-1.00, ideal lower LOA) or more mature (+1.00, ideal upper LOA). Utilizing these proposed ideal upper and lower LOAs and corresponding ideal confidence intervals, it is apparent that the plots produced for inter-examiner agreement (Figure 2.4), CF to ground truth (Figure 2.5) and ML to ground truth (Figure 2.6) all possessed larger than clinically accepted confidence intervals and therefore any agreement should be taken with serious caution.

It is expected that 95% of the differences be between 2 standard deviations as represented by the upper and lower LOAs. In fact, 2 outliers existed, one in the evaluation of intra-examiner agreement (Figure 2.3), as well as, ML to ground truth (Figure 2.6). The outlier present in the plot of intra-examiner agreement can be accounted for by the fact that DAI was so reliable in staging patients between readings, with such a small range of upper and lower LOAs that any major distinct difference in a single reading, made the difference fall outside of this 95% confidence interval. Interestingly, If the proposed confidence interval with upper and lower LOAs of +/- 1.00 were utilized, this data point and would no longer be considered an outlier. The outlier present in the plot of ML to ground truth (Figure 2.6) is present in a data set with such a large range of upper and lower LOAs. Consequently, the outlier's presence outside of these limits most likely represents a true outlier. If further statistical analyses were to be utilized these data points could be kept

and then more robust statistical measures applied as needed, and then remove the outlier and repeat the barrage of tests to see if there was a difference. If no difference observed the outlier could be kept.

Limitations:

There were some limitations present in the study such the sample size, memory bias and the lack of a gold standard for comparison. In comparison to the Angelieri et al[1] study, the number of subjects utilized in our study for reliability testing was roughly half, n=16 compared to n= 30 subjects respectively. The small size of our study may have decreased the statistical power, such that outcomes that are deemed statistically significant have decreased probability of actually representing a true difference[58]. Repeated viewing sessions by the ground truth performed only 48 hours apart may contribute to the increased intra-examiner agreement via the introduction of memory bias.

If repeated, the reliability testing would be better served with an increased sample size equal to or greater than the study in question[1] and have any evaluators performing multiple classifications sessions perform them 2-4 weeks apart to minimize memory bias. The lack of a true gold standard, poor voxel definition and limitations of the proposed methodology and classification system[1] will all be addressed in the following chapter as these limitations permeated throughout the course of the study.

Voxel size and difficulty in clinically observing radiological anatomical differences between the stages was identified as a limiting factor during reliability testing and throughout the study. Utilizing only the default slice thickness of 1.0mm, was a limitation as altering the slice thickness could have increased the clarity of the generated axial images. Frequency tables developed illustrated that differentiation between stages A, B,

C, and the differences of each of these stages to stage D were clinically more difficult than anticipated given the clear criteria set out in the classification[1].

Small sample size affected the statistical power of the chosen analyses, and should the reliability testing be repeated, a sample size equal to or greater than that of the study in question[1] is advised.

2.6 Conclusion:

The kappa statistic was correctly employed in replacement of a weighted kappa statistic due to a lack of a true gold standard making the assignment of weights impossible. There was almost perfect intra-examiner agreement between the rater DAI's palatal suture maturation classification at time 1 and 3, κ=0.915 (95% CI, 0.752 to 1.078), $p < 0.005$. Consequently, there is convincing evidence that agreement between classification sessions for rater DAI is far greater than chance agreement.

Agreement of CF to ground truth, Inter-examiner agreement (ML to CF) and agreement of ML to ground truth ranged from poor, to slight, to moderate agreement respectively. Additionally, there is weak to no evidence to support that any agreement with examiner CF was greater than chance agreement alone. Proposed was the concept of severe systematic differences in the approach of CF to the assessment of the palatal stages. Reasons for this departure possibly stem from poor voxel definition for this operator and not a specific attempt to purposely not follow the methodology. Ultimately, reliability testing disagrees with that of the original study, indicating that this classification system is not as reliable as previously presented[1]. This study indicates that the proposed

methodology is in fact non-intuitive, requires operator calibration and heavily influenced by the degree of post-acquisition image sharpness and clarity.

2.7 Appendices:

Appendix 2.1 Raw counts from Reliability testing

Subject	MLE Classification	CFE Classification	DAI-1E Classification	DAI-2E Classification	DAI-3E Classification
1	D	B	D	D	D
2	D	C	C	C	D
3	C	B	C	C	C
4	C	D	C	C	C
5	C	E	C	C	C
6	C	B	C	C	C
7	C	C	E	E	E
8	**E**	D	C	C	C
9	C	B	C	C	C
10	C	D	D	D	D
11	B	D	B	D	B
12	C	C	E	E	E
13	A	B	A	A	A
14	**B**	**B**	**D**	D	D
15	A	A	A	A	A
16	A	C	A	A	A

Appendix 2.2 Inter-examiner reliability (CF and ML) descriptive statistics, crosstabulation table, symmetric measures and sample calculation of kappa statistic

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
ML * CF	16	100.0%	0	0.0%	16	100.0%

ML * CF Crosstabulation

Count

		CF					
		A	B	C	D	E	Total
ML	A	1	1	1	0	0	3
	B	0	1	0	1	0	2
	C	0	3	2	2	1	8
	D	0	1	1	0	0	2
	E	0	0	0	1	0	1
Total		1	6	4	4	1	16

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.040	.127	.341	.733
N of Valid Cases		16			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Kappa statistic **(κ)** = $(P_{observed} - P_{chance}) / (1-P_{chance}) = 0.040$

<u>Confidence interval sample calculation at 95%:</u>

Standard error = 0.127

z-score at 95% confidence, two tailed = z = 1.96

z-score x standard error = 1.96 x0.127 = 0.249

→ lower interval 0.040-0.249 = -0.209

→ upper interval 0.040+0.249 = +0.289

(κ)=0.040(95% CI, -0.209 to 0.289), p =0.733

Appendix 2.3 Reliability ML to ground truth (DAI time 1) descriptive statistics, crosstabulation table and symmetric measures (sample calculation provided in appendix 2.2).

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
ML * DAI_1	16	100.0%	0	0.0%	16	100.0%

ML * DAI_1 Crosstabulation

Count

		DAI_1					
		A	B	C	D	E	Total
ML	A	3	0	0	0	0	3
	B	0	1	0	1	0	2
	C	0	0	5	1	2	8
	D	0	0	1	1	0	2
	E	0	0	1	0	0	1
Total		3	1	7	3	2	16

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.470	.168	3.426	.001
N of Valid Cases		16			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

<u>95% Confidence interval:</u>

1.96x0.168 = 0.329

→ lower interval 0.470-0.329 = 0.141

→ upper interval 0.470+0.329 = 0.799

(κ)=0.470 (95% CI, 0.141 to 0.799), p =0.001

Appendix 2.4 Reliability CF to ground truth (DAI time 1) descriptive statistics, crosstabulation table and symmetric measures (sample calculation provided in appendix 2.2).

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
CF * DAI_1	16	100.0%	0	0.0%	16	100.0%

CF * DAI_1 Crosstabulation

Count

		DAI_1					
		A	B	C	D	E	Total
CF	A	1	0	0	0	0	1
	B	1	0	3	2	0	6
	C	1	0	1	0	2	4
	D	0	1	2	1	0	4
	E	0	0	1	0	0	1
Total		3	1	7	3	2	16

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	-.015	.120	-.131	.896
N of Valid Cases		16			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

<u>95% Confidence interval:</u>

1.96x0.120= 0.235

→ lower interval -0.015 -0.235 = -0.25

→ upper interval -0.015+0.235 = 0.22

(κ)=0-0.015 (95% CI,)-0.25 to 0.22), p =0.896

Appendix 2.5 Intra-examiner reliability (DAI 1 and 3) descriptive statistics, crosstabulation table, and symmetric measures (sample calculation provided in appendix 2.2).

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
DAI_1 * DAI_3	16	100.0%	0	0.0%	16	100.0%

DAI_1 * DAI_3 Crosstabulation

Count

		DAI_3					
		A	B	C	D	E	Total
DAI_1	A	3	0	0	0	0	3
	B	0	1	0	0	0	1
	C	0	0	6	1	0	7
	D	0	0	0	3	0	3
	E	0	0	0	0	2	2
Total		3	1	6	4	2	16

Symmetric Measures

		Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Agreement	Kappa	.915	.083	6.620	.000
N of Valid Cases		16			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

<u>95% Confidence interval:</u>

1.96x0.083 = 0.163

→ lower interval 0.915-0.163 = 0.752

→ upper interval 0.915+0.163 = 1.078

k=0.915 (95% CI, 0.752 to 1.078), p <0.005

Appendix 2.6 Converted raw counts of reliability testing to numerical ordinal data, such that Stage A=1, B=2, C=3, D=4 and E=5. Additionally, sample calculation for upper and lower limits of the Bland Altman analyses using raw counts.

Subject	ML Classification	CF Classification	DAI 1 Classification	DAI 2 Classification	DAI 3 Classification
1	4	2	4	4	4
2	4	3	3	3	4
3	3	2	3	3	3
4	3	4	3	3	3
5	3	5	3	3	3
6	3	2	3	3	3
7	3	3	5	5	5
8	5	4	3	3	3
9	3	2	3	3	3
10	3	4	4	4	4
11	2	4	2	4	2
12	3	3	5	5	5
13	1	2	1	1	1
14	2	2	4	4	4
15	1	1	1	1	1
16	1	3	1	1	1

Bland Altman analysis upper and lower limits of agreement calculation for ML and CF.

ML	CF	dif ML to CF	Mean of ML and CF
4	2	2	3
4	3	1	3.5
3	2	1	2.5
3	4	-1	3.5
3	5	-2	4
3	2	1	2.5
3	3	0	3
5	4	1	4.5
3	2	1	2.5
3	4	-1	3.5
2	4	-2	3
3	3	0	3
1	2	-1	1.5
2	2	0	2
1	1	0	1
1	3	-2	2

Bias = average of dif (difference) of ML to CF = -0.125

Standard deviation of dif (difference) of ML to CF = -1.258

Lower limit = bias – (1.96 x std dev) = -0.125 – (1.96*1.258) = -2.59

upper limit = bias + (1.96 x std dev) = -0.125 + (1.96*1.258) = 2.34

2.8 Literature Cited:

1. Persson, M., B.C. Magnusson, and B. Thilander, *Sutural closure in rabbit and man: a morphological and histochemical study.* J Anat, 1978. **125**(Pt 2): p. 313-21.
2. Baccetti, T., et al., *Treatment timing for rapid maxillary expansion.* Angle Orthod, 2001. **71**(5): p. 343-50.
3. Angelieri, F., et al., *Midpalatal suture maturation: classification method for individual assessment before rapid maxillary expansion.* Am J Orthod Dentofacial Orthop, 2013. **144**(5): p. 759-69.
4. Isiksal, E., S. Hazar, and S. Akyalcin, *Smile esthetics: perception and comparison of treated and untreated smiles.* Am J Orthod Dentofacial Orthop, 2006. **129**(1): p. 8-16.
5. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.
6. McHugh, M.L., *Interrater reliability: the kappa statistic.* Biochem Med (Zagreb), 2012. **22**(3): p. 276-82.
7. Lagravere, M.O., et al., *Transverse, vertical, and anteroposterior changes from bone-anchored maxillary expansion vs traditional rapid maxillary expansion: a randomized clinical trial.* Am J Orthod Dentofacial Orthop, 2010. **137**(3): p. 304.e1-12; discussion 304-5.
8. Viera, A.J. and J.M. Garrett, *Understanding interobserver agreement: the kappa statistic.* Fam Med, 2005. **37**(5): p. 360-3.

9. Maclure, M. and W.C. Willett, *Misinterpretation and misuse of the kappa statistic.* Am J Epidemiol, 1987. **126**(2): p. 161-9.

10. Cooper, D., et al., *Effect of voxel size on 3D micro-CT analysis of cortical bone porosity.* Calcif Tissue Int, 2007. **80**(3): p. 211-9.

11. Giavarina, D., *Understanding Bland Altman analysis.* Biochem Med (Zagreb), 2015. **25**(2): p. 141-51.

12. Button, K.S., et al., *Power failure: why small sample size undermines the reliability of neuroscience.* Nat Rev Neurosci, 2013. **14**(5): p. 365-76.

Chapter 3: Evaluation of a novel palatal suture maturation classification as assessed by CBCT imaging of a pre- and post-expansion treatment cohort

3.1 Introduction:

The primary palate (median palatine process) and secondary palate (2 lateral palatine processes) develop and fuse during the 5 to12 weeks of human intrauterine development[59, 60]. Specifically, in the 5th developmental week the primary palate forms from the fusion of two medial nasal prominences. The secondary palate develops in the 7th developmental week from the anteroinferior dissent and diminution in size of the tongue, with concurrent elongation and reorientation of the lateral palatine processes into a superior and horizontal position relative to the tongue[59, 60]. This fusion of the palatal shelves occurs prior to ossification of the cartilaginous cranial base ossification[34]. During developmental weeks 7 to12, fusion along the two lateral palatine processes occurs at the incisive foramen and proceeds posteriorly[60], fusing superiorly with the nasal septum and anteriorly with the primary palate[59, 60]. The primary palate and bones of the secondary palate namely, the maxillary and palatine bone ossify via intramembranous ossification.

The midpalatal suture (MPS) is defined by three specific regions from most anterior to posterior as the interpremaxillary, maxillary and interpalatine regions[59]. The interpremaxillary suture is the first to form, and is established by 47 days in utero. The maxillary and palatine aspects of the secondary palate show signs of sutural formation by 10.5 weeks old, and established by week twelve of life when soft tissues of the soft palate fuse posterior to the hard palate[59, 60]. Melsen[7, 61] investigated autopsy material to histologically stage the maturation of the MPS. From birth up to 10 years of age, Melsen identified the “infantile” stage to have a MPS characterized by being broad and smooth. From 10-13 years old, during the “juvenile” developmental stage Melsen identified the

suture as being a more organized squamous sutural morphology with areas of overlap exists. From 13-14 years old, Meslen identified the "adolescent" stage to be characterized by a greater degree of corrugation and interdigitation[7]. A characterization of an "adult" stage of the MPS was carried out by Melsen and Melsen (1982)[61] noting this mature suture is defined by increased interdigitation and multiple bony articulations of osseous margins[61].

Excluding the articulation with itself at the MPS, the maxilla articulates with the rest of the craniofacial complex via a series of cranial and circummaxillary sutures namely the frontonasal, zygomaticotemporal, frontomaxillary, frontozygomatic, internasal, nasomaxillary, zygomaticomaxillary, intermaxillary, and pterygomaxillary sutures[62]. The role of the these cranial and circummaxillary sutures is to be a site of growth, connect bones, act as a site of articulation allowing for minute movement and absorb applied forces[62].

As it relates to timing and force delivery to create maxillary skeletal expansion a histological investigation[7] supported clinically by an additional study[63] inferred that skeletally mature MPS's are directly related to increased effort and difficulty in maxillary skeletal expansion with decreased transverse stability, especially in patients treated past 12 years of age or older[7, 61]. These findings were supported by Bacetti (2001)[2] whose study found that rapid expansion of the maxilla was capable of producing more long-term transverse skeletal changes when treatment is applied prior to the patient's pubertal peak of skeletal growth, whereas application of rapid maxillary expansion (RME) past this pubertal peak growth lead to an increasing proportion of expansion from dentoalveolar adaptations[2].

A recent study[62] investigated the effect of RME loading on the circummaxillary sutures. The results noted in an average patient population aged of 12.3 +1.9 years old, significant transverse expansion across intermaxillary, internasal, maxillonasal, frontomaxillary and frontonasal sutures was observed, with no significant changes to the frontozygomatic, zygomaticomaxillary, zygomaticotemporal and pterygomaxillary suture. The results suggested that these circummaxillary sutures absorb and are affected by the forces generated from RME treatment. Specifically, the results noted that the most anterior sutures, especially the MPS and interfaces at the frontonasal maxillary complex demonstrated the greatest transverse change compared to more posterior cranial and circummaxillary sutures located further from the site of force application[62]. Ultimately, the study concluded that efficacy of maxillary skeletal expansion is dependent upon the pre-treatment patency of the MPS and affiliated circummaxillary sutures or the degree of sutural opening due to applied orthopedic forces during RME treatment, with decreased opening related to decreased clinical transverse expansion outcomes [62]. Ultimately, the findings highlight the predictability of skeletal maxillary expansion prior to or at the time of pubertal growth spurt while inferring a decreased predictability and prognosis of skeletal expansion with an increase in dentoalveolar adaptations when transverse expansion is applied past the pubertal growth spurt.

Literature notes great individual variability in skeletal maturation. Consequently, this variability is also observed in the maturation of the MPS and circummaxillary sutures[2]. Circummaxillary sutures have been reported to fuse shortly after the initiation of puberty, approximately age 11-13 for females and 13-15 for males[64]. Meanwhile multiple other studies have shown greater variability in the fusion of some individuals MPS ranging

from 15-19 years old[8], upwards to 71 years[10], with the greatest degree of sutural obliteration occurring in the third decade of life[8]. These studies demonstrate that chronologic age of the patient is an unreliable and poor indicator of degree of fusion of the maxillary sutures[1]. The great variability regarding onset of sutural fusion and thus difficulty predicting the pre-treatment degree of MPS maturation creates a unique challenge in treatment planning and delivery of reliable skeletal expansion outcomes in late stage adolescent and young adult patients[1]. Improper prediction of the pre-treatment degree of MPS maturation may lead to the least favourable expansion modality being chosen.

If the MPS is nearly fused in these older patients and tooth-born RME attempted, dentoalveolar adaptations will predominate with minimal to no skeletal expansion, putting the patient at higher risk for iatrogenic effects such as acute pain, periodontal recession, mucosal necrosis, severe buccal tipping and associated bite opening, as well as poor transverse and occlusal stability[1, 5]. Conversely, if the MPS and circummaxillary sutures are in fact more patent than predicted, then committing a patient to a surgical expansion modality creates an unnecessary burden of increased cost, pain, healing time, iatrogenic and potentially severe surgical complications[1]. Additionally, although surgical expansion may deliver the most predictable outcome in those patients with a mature or nearly fused MPS, surgical expansion via surgically assisted rapid palatal expansion (SARPE) and/or multi-piece maxillary osteotomies has the most unpredictable long term outcomes and greatest degree of relapse compared to other orthognathic surgeries[1]. Given that biopsy is the current gold standard to identify the degree of MPS maturation and is unfeasible to be performed on patients, what is needed is the ability for

clinicians to provide a minimally invasive individual MPS maturation assessment, to drive clinical decision making towards the least harmful, most advantageous, and predictable expansion modality.

Bjork's classic studies involving implants[65, 66] suggest that the pattern of skeletal maxillary transverse development parallels statural growth velocity and distance curves, with comparable onset of growth acceleration and quiescence. Conversely, it was found utilizing digitized 3D study casts and measuring the palatal surface area or volume that transverse development of the palate occurs during the primary dentition into middle mixed dentition[67]. Additional methods to assess and infer the degree of craniofacial skeletal maturity includes hand-wrist radiograph assessment, evaluation of overall dental development, primary and secondary sexual characteristics changes, and identification of cervical vertebral maturation (CVM) stage, with the largest growth spurt of the craniofacial complex being identified in CVM stage 3 and 4[2]. Capturing pre-treatment maxillary occlusal radiographs have been suggested[1, 59] to directly observe the patency of the MPS; however, this presents several issues: Firstly, the superimposition of adjacent hard and soft tissue structures including the vomer and soft tissue nares complicates the radiographic evaluation and may lead to false positive or negative assessments of the degree of patency at the MPS. Additionally, due to the localized nature of this radiograph, evaluation of the circummaxillary sutures is not possible[1]. Use of posterior-anterior (PA) cephalometric radiography to make angular and linear measurements in the maxillary transverse dimension possess the same projection and superimposition concerns as the aforementioned occlusal radiography. However, PA cephalometry is susceptible to the additional concerns that complicate landmark identification in lateral

cephalometry including head positioning errors, structure geometry being curved or straight, landmark clarity with regard to contrast levels between adjacent structures, radiographic quality and clarity, as well as operator inexperience[68, 69]. In fact, two-dimensional (2D) radiography produces low precision and accuracy in landmark reproducibility[70]. In addition, a PA projection of the palatal suture is questionable as the different progressive maturational stages from the front to the back all do superimpose.

The recent development of cone-beam computed tomography (CBCT) offers a viable alternative to 2D radiography. CBCT technology has been made increasingly more accessible, at a lower cost and delivering a lower and acceptable[71] patient exposure alternative to medical computed tomography[1] while providing improved diagnostic ability and long term benefit to case management over two-dimensional radiography[71]. Use of CBCT has circumvented the shortfalls of occlusal radiography and other 2D radiography, namely allowing clinicians to vary the field of view (FOV) to include the circummaxillary sutures, generate three dimensional (3D) volumes [1] free of head positioning and projection errors with true, 1:1 ratio of anatomic linear measurements[72], as well as generate hard tissue cross-sections for clear visualization of areas of interests without superimposition of adjacent structures[1, 62]. These advantages allow CBCT technology to produce more precise and unobstructed landmark identification[70] and as a result, more reliable linear and angular measurements of 3D versus 2D landmarks would be expected. In fact, comparison of direct caliper measurements on anatomic dental models to measurements performed on 3D volumes of dental casts using voxel sizes between 0.40mm and 0.25mm were found to be accurate yielding absolute errors of 0.05$\pm$0.04mm and 0.07$\pm$0.05mm for the 0.40mm and 0.25mm-voxel group respectively

with excellent intraoperator reliability demonstrating an interclass correlation value of 1.00[73]. When pre- and post-treatment CBCTs are taken, the anatomical displacement of structures over time can be investigated, a diagnostic technique used since the inception of lateral cephalometric radiography by Braodbent (1931)[74]. Given that the linear and angular measurements are defined by the landmarks utilized, accurate and reliable detection of 3D landmarks of interest is necessary to minimize measurement errors which could be significant with slight errors in landmark identification and placement[75]. In fact, a recent study[76] evaluated the reliability of 3D landmark identification using CBCT and found excellent intra- and interobserver reliability, suggesting that a well-trained and calibrated clinician can utilize CBCT technology to produce reliable and reproducible landmarking.

As mentioned in the previous chapter, Angelieri et. al [1] developed an imaging protocol and novel, patient specific, classification system dependent upon CBCT imaging of the MPS, assigning an ordinal scale maturation stage (A-E) related to the degree of observed palatal suture fusion (Figure 3.1).

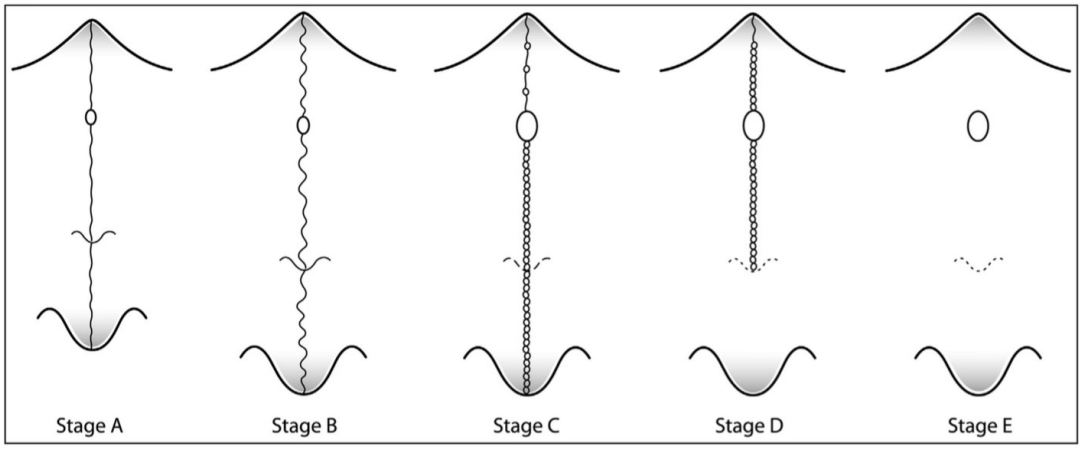


Figure 3.1: Diagrammatic representation of the developed novel palatal suture maturation classification[1] identifying key radiological morphological characteristics specific to each maturity level. Stage A is defined by a single non-scalloped line, stage B is defined by a single scalloped line, stage C is defined by two parallel, scalloped, high density radiopaque lines with small radiolucent spaces between the lines, stage D is defined by maturation of the 2 high density lines in the palatine bone, stage E is defined by complete maturation of the palatal suture and no visible lines in at least a segment of the maxillary bone.

Although intriguing, relating actual pre- and post-treatment clinical RME outcomes and potential co-factors to further evaluate the predictive ability of this novel classification technique is necessary to establish its usefulness to clinical practice and drive RME modality decision making. This fact generated the secondary research question of this study: *How useful is this novel classification system to predict success of RME treatment?* Furthermore the *tertiary* research question will also be addressed, namely: *What alteration(s) or modification(s) to the Angelieri et. al[1] methodology can be suggested to improve reliability and/or predictive ability of this classification system?*

3.2 Material and Methods:

The investigation performed is retrospective observational longitudinal (cohort) study, approved by the University of Alberta research ethics board.

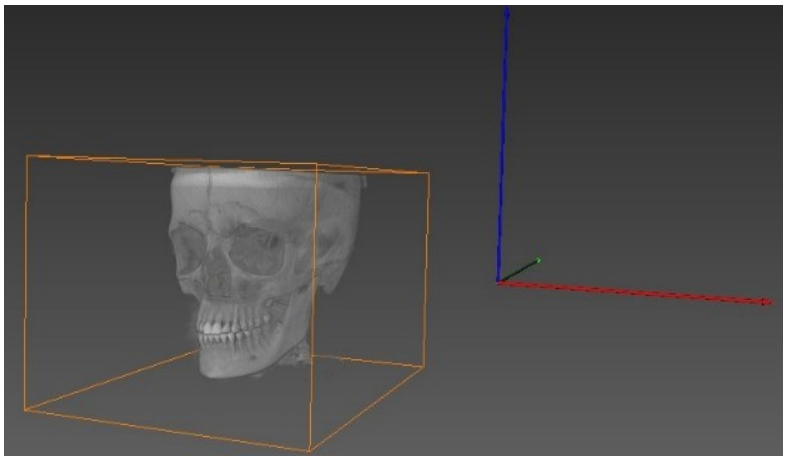
The investigated CBCT data is derived from a database of patients from a previous expansion study[75]. The CBCT data was generated by an I-CAT machine (120kVp, 23.87mAs, 8.9 seconds exposure time, 16x13cm FOV, 0.3mm voxel size). Initial pre- and post-expansion CBCT records were gathered using a standardized I-CAT protocol.

Inclusion Criteria:

A total of 68 pre-adolescent and adolescent patients aged 11-17 years old with full permanent dentition treated with tooth-borne RME appliances who have had CBCTs records taken at two time points, T_1 pre-expansion and T_2 post-expansion treatment were evaluated in the study. The patients were assigned to random numbers as patient identifiers.

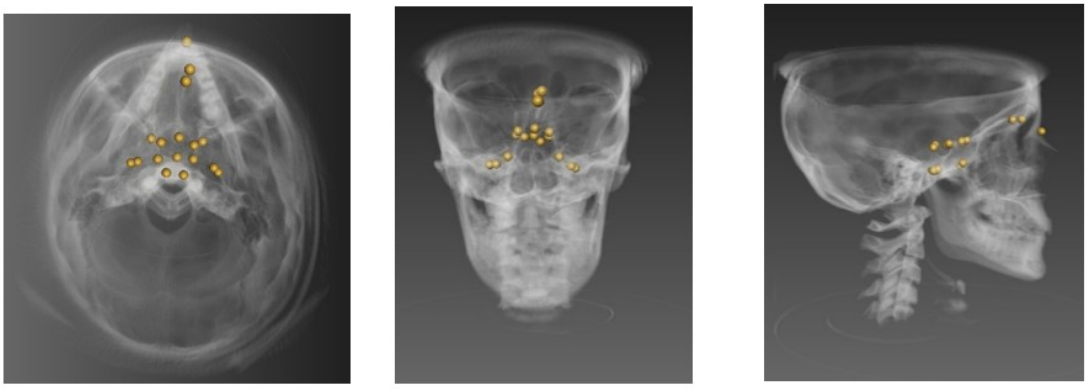
Captured raw data were exported as DICOM files into Avizo version 7.0 software (visualization Sciences Group, Burlington, MA, USA) for further image processing. The software integrated a Cartesian plane coordinate system such that the image could be identified in various planes, namely the x-y, x-z, and y-z planes, representing the axial (right-left), coronal (superior-inferior) and sagittal (anterior-posterior) respectively (Figure 3.2).

Figure 3.2: Orientation of the Cartesian plane coordinate system in 3 planes; X (red), Y (green) and Z (blue).



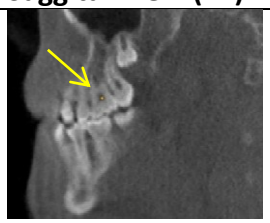
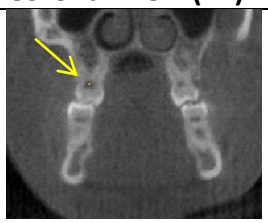
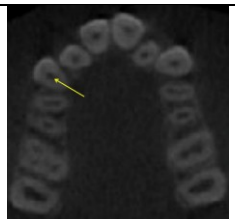
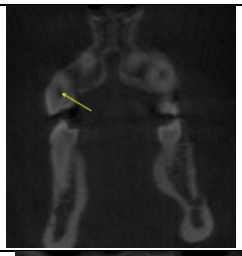
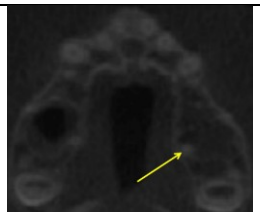
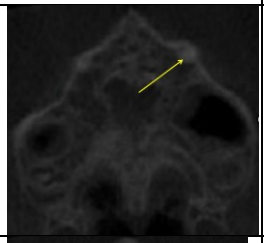
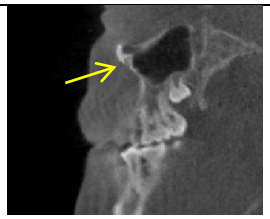
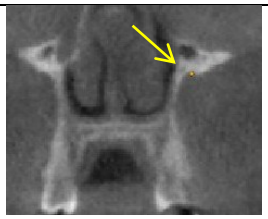
The Avizo software allowed for 3D landmarking within the Cartesian plane coordinate system, where placement of a landmark was characterized by coordinates in all 3 planes. The primary investigator (DAI) identified the anatomical locations of interest and then completed the 3D landmarking by depositing a 0.25mm diameter spherical marker within the Cartesian coordinate plane system (Figure 3.3). Intra-examiner reliability (ICC) has been reported for the x, y, z coordinates of 3D landmarking to be greater than 0.99, with a mean measurement difference less than 0.7mm on all axis and landmarks[53].

Figure 3.3: Spherical markers representing the 3D landmarks of interest visualized in the x, y and z planes within the Avizo software version 7.0



A total of twelve anatomical landmarks were located within the maxilla of each subjects pre- and post-expansion CBCT volumes for further image analysis. (Table 3.1).

Table 3.1: Maxillary landmarks defined and shown on cross sectional images identified in each subject's pre- and post-expansion CBCT volumes.

Landmark	Axial View (XY)	Saggital View (YZ)	Coronal View (XZ)
Pulp chamber (PC)- Upper 1st molar Center of the largest cross sectional pulp chamber area. Defined for both L & R 1st molars: 16 Pulp 26 Pulp			
PC- Upper Canines Center of the largest cross sectional pulp chamber area. Defined for both L & R canines: 13 Pulp 23 Pulp			
Apex (A) - Upper 1st molar Palatal root apex. Defined for both L & R 1st molars: 16 Apex 26 Apex			
A - Upper Canines root apex. Defined for both L & R 1st canines: 13 Apex 23 Apex			
Slope - Point generated at greatest depth of curvature of the palatal slope lingual to the 1st molars. Defined for both L & R slopes.			
Infraorbital Foramen (IOF) Superior most aspect of the infraorbital foramen outer border. Defined for both L & R IOF.			

Three-dimensional landmarking in the Cartesian plane coordinate system allowed for a total of four dental linear measurements, 4 dental angular measurements and 2 skeletal linear measurements (Table 3.2) to be generated using the equation depicted below:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

Table 3.2: Dental and skeletal linear & angular measurements generated.

Dental/skeletal	Linear/ Angular	Measurement	Description
Dental	Linear	intermolar width at the apices	#1.6 palatal root apex - #2.6 palatal root apex
Dental	Linear	intermolar width at the pulp chamber	#1.6 pulp - #2.6 pulp
Dental	Linear	intercanine width at the apices	#1.3 apex - #2.3 apex
Dental	Linear	intercanine width at the pulp chamber	#1.3 pulp - #2.3 pulp cusp
Dental	Angular	right canine angulation	#1.3 angulation
Dental	Angular	Left canine angulation	#2.3 angulation
Dental	Angular	right molar angulation	#16 angulation
Dental	Angular	left molar angulation	#26 angulation
Skeletal	Linear	inferior orbital foramen (IOF) distance	R IOF- L IOF
Skeletal	Linear	palatal width	palatal alveolar process line of best fit from upper right adjacent to #1.6 to upper left adjacent to #2.6.

The distance (d) is measured in millimeters (mm) between the two spherical markers deposited within the coordinate system, and x_1, y_1, z_1 and x_2, y_2, z_2 denote the coordinates of these landmarks making up the linear measurement. Angular measurements were generated using the following trigonometric equation:

$$a = ACOS(d1 * d1 + d3 * d3 - d2 * d2)/(2 * d1 * d3)$$

such that d1, d2, and d3 are the distances representing each side of the triangle, specific to the location of the angle. The angle (a) were defined by radian units and subsequently converted to degrees using Microsoft Excel software (version 15.32, Redmond, Washington). The angle relates to the buccal/lingual axial inclination of the tooth in question when viewed in the coronal plane.

At time 1 (pre-expansion) patient demographics including age and sex were recorded for each patient. Furthermore, each pre-expansion CBCT volume was used to generate a 2D projected lateral cephalogram to evaluate patient Time 1 CVM stage.

Additionally, the image analysis was performed following a mildly modified protocol from Angelieri et. al[1] (as previously described) to isolate MPS containing axial cross-sections to identify the pre-expansion sutural maturation stage according to their proposed novel palatal suture maturation classification. Upon review, pre-expansion palatal staging was unbalanced across the sample (Figure 3.4), and therefore some stages were grouped by reducing the number of categories to create balance across the sample to increase the power of the experimental design (Figure 3.5).

Figure 3.4: Unbalanced distribution of palatal stages across the sample (n=63) as classified according to Angelieri et al[1].

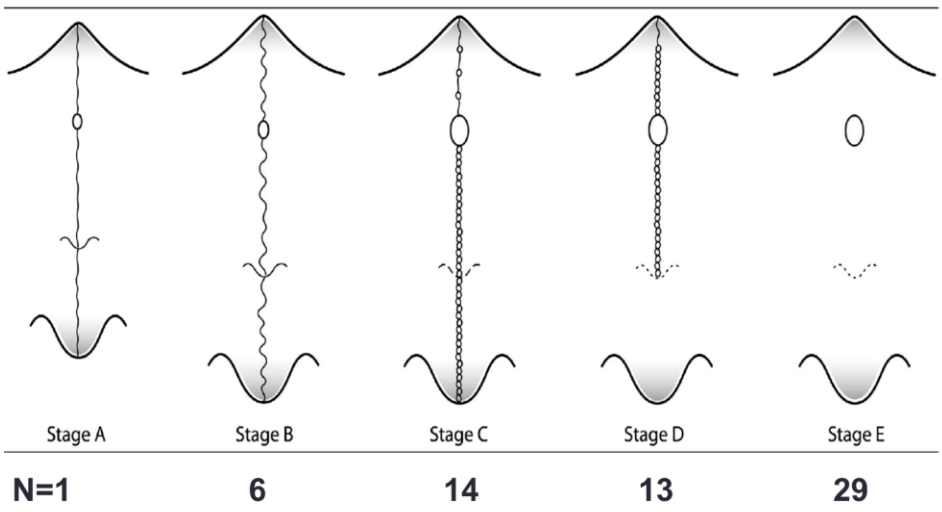
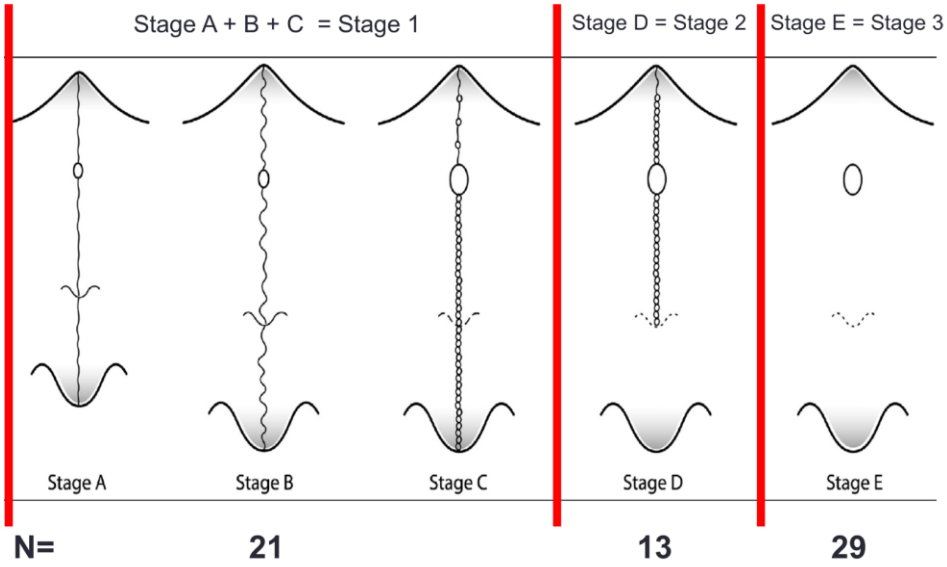


Figure 3.5: Improved balanced distribution of palatal stages across the sample (n=63) after implementation of modified Angelieri et al.[1] classification.



The new classification grouped the previously defined palatal stages A-E[1] based on visibility of the MPS in the maxillary and palatine bone. When the palatal suture was visible in both the maxillary and palatine bone, relating to stages A, B and C, this new cohort was combined into a new palatal stage termed "stage 1" (n =21), and going forward is considered to be representative of the most immature form of the MPS. Palatal Stage D is related to the MPS being visible solely in the maxillary bone and is now termed "stage 2" (n=13) and going forward is considered represent intermediate MPS maturity. Palatal stage E is related to fusion of the MPS such that it is not visible in at least a portion of the maxillary bone if not the whole length, and is now termed "stage 3" and is representative of the most mature form of the MPS (n=29).

Given that pre- and post-expansion CBCT volumes were taken for each subject, dependent variables, namely difference in skeletal and dental distance and dental angles from time 1 (pre-expansion) to time 2 (post-expansion) (Table 3.3) were calculated for each subject. Consequently, the direction and magnitude of transverse skeletal and dental changes from RME treatment were calculated. If the difference in angular measurements from T1 to T2 are measured to be positive, this relates to tooth uprighting, namely via lingual crown torque and/or buccal root torque from T1 to T2.

The evaluator DAI was considered blinded concerning image date when performing pre- and post-expansion 3D landmarking and generation of dental and skeletal linear and dental angular measurements for each subject.

Table 3.3: Dependent variables – difference in skeletal and dental distances and dental angles from T2 – T1.

Dental/skeletal	Linear/ Angular	Measurement	Acronym
Dental	Linear	Difference in intermolar width #1.6 -#2.6 at the apices	Dif_(M_Ap)
Dental	Linear	Difference in intermolar width #1.6 -#2.6 at the pulp chamber	Dif_(M_Pu)
Dental	Linear	Difference in intercanine width #1.3-#2.3 at the apices	Dif_(C_Ap)
Dental	Linear	Difference in intercanine width #1.3-#2.3 at the pulp chamber	Dif_C_Pu
Dental	Angular	Change in right canine #1.3 angulation	Dif_(ANG_13)
Dental	Angular	Change in left canine #2.3 angulation	Dif_(ANG_23)
Dental	Angular	Change in right molar #1.6 angulation	Dif_(ANG_16)
Dental	Angular	Change in left molar #2.6 angulation	Dif_(ANG_26)
Skeletal	Linear	Change in distance from right to left inferior orbital foramen (IOF)	Dif_(IOF)
Skeletal	Linear	Difference in palatal slope widths	Dif_(slope)

A total of five patients were removed from the initial sample of 68 patients. Two patients were removed due to inability to view all necessary vertebrae in the CBCT generated lateral cephalogram and therefore could not assign a pre-expansion CVM stage. The remaining 3 subjects were removed due extremely poor clarity of generated CBCT volumes impeding appropriate landmarking at time 2 (post-expansion). Consequently, a final sample of 63 patients were evaluated.

3.3 Statistical Analysis:

To conduct statistical analysis, standard statistical software package (SPSS version 20 for Mac, Chicago, Ill.) was utilized.

Multivariate analysis of variance (MANOVA) of the absolute difference between the Time 1 (T1: pre-expansion) and Time 2 (T2: post-expansion) methods was used to determine significance and test the null hypothesis that a subject's palatal stage has no effect on any of the of the dependent variables.

The alternative hypothesis is that a subject's palatal stage has an effect on any of the dependent variables. A *p* value of less than 0.05 was considered significant (Appendix 3.1). In respect to angular measurements an absolute difference $\geq 5°$ was considered clinically significant[77, 78], and a change of equal to or greater than 1.0mm[79] for any dental or skeletal width change.

The initial statistical analysis utilized was a two-way multivariate analysis of covariance (MANCOVA) performed with age and CVM stage as covariates, sex and palatal stage as fixed variables and all dependent factors being the difference from T1-T2 for all skeletal and dental widths and angles (Table 3.3). If the covariates of age and CVM stage were deemed to be not statistically significant the statistical analyses would have been changed to a two-way Multivariate analysis of variance (MANOVA). Furthermore, if any one of the fixed factors would have been deemed to be not statistically significant, that fixed factor would have been removed and a one-way MANOVA performed.

To account for differences between subjects regarding age and CVM stages recorded at the time of initial records, age and CVM stage was controlled for by considering these covariates.

Model assumptions were evaluated prior to initiation of statistical analysis (Appendix 3.2).

The model assumptions for MANOVA include:

1. all cases should be independent
2. outcome variables should have a multivariate normal distribution.
3. Relationships among all pairs of Y variables are linearly related
4. variance-covariance matrices equal across groups in the study

A boxplot of Mahalanobis distance (Appendix 3.3) and Q-Q plot (Appendix 3.4) was created to evaluate for outliers and multivariate normality. There were multivariate outliers in the data, as assessed by Mahalanobis distance ($p > .001$). Boxplot of Mahalanobis distance (depicts a right skewedness due to the presence of outliers, namely patient's #50. #8, #4 (extreme outliers, $\geq$3 box-lengths away from the edge of box) and #18, #12 and #7 (1.5 box-lengths away from the edge of box). The outliers and related skewness suggests that multivariate normality is not met. Given the large sample size, outliers will be accepted into our data, and same statistical analysis with these outliers removed will be performed to assess changes in the statistical outcome. Consequently, it has been assumed multivariate normality assumption has been met, as well as linearity. There was homogeneity of variance-covariances matrices, as assessed by Box's test of equality of covariance matrices ($p = .313$) (Appendix 3.5).

The two-way MANCOVA revealed age and CVM stage in relation to the dependent variables was not statistically significant, $F(10, 46) = 0.981$, $p = .473$; Wilks' $\Lambda = .824$; partial $\eta^2 = .176$ and $F(10, 46) = 0.573$, $p = .827$; Wilks' $\Lambda = .889$; partial $\eta^2 = .111$, respectively (Appendix 3.6). Consequently, covariates will not be further analyzed and a two-way MANOVA was performed.

3.4 Results:

Performing the two-way MANOVA revealed palatal stage demonstrated convincing statistical significance [$F(20, 96) = 1.874$, $p = .023$; Wilks' $\Lambda = .517$; partial $\eta^2 = .281$]. Sex, and the interaction between sex and palatal stage in relation to the dependent variables demonstrated suggestive, but inconclusive statistical significance [$F(10, 48) = 1.991$, $p = .055$; Wilks' $\Lambda = .707$; partial $\eta^2 = .293$] and [$F(20, 96) = 1.315$, $p = .189$; Wilks' $\Lambda = .616$; partial $\eta^2 = .215$] respectively (Appendix 3.7). Consequently, the fixed factor of sex was removed and a one-way MANOVA performed using palatal stage as the only fixed factor and same dependent variables. One-way MANOVA demonstrates that the assumption of homogeneity of variance-covariances matrices was failed, as assessed by Box's test of equality of covariance matrices ($p = .027$, Appendix 3.8). Palatal stage demonstrated convincing statistical significance [$F(20, 102) = 1.709$, $p = .044$; Wilks' $\Lambda = .561$; partial $\eta^2 = .251$] (Appendix 3.9a and b).

There was convincing evidence that there is significant difference in Dif_C_Pu and Dif_ANG_16 between subjects with different T1 palatal stages [$F(2, 60) = 3.919$, $p = 0.025$; partial $\eta^2 = .116$] and [$F(2, 60) = 5.588$, $p = 0.006$; partial $\eta^2 = .157$] respectively (Appendix 3.10). Multiple comparisons Bonferonni corrected analyses demonstrated that

at post-expansion there was a mean increase in intercanine width measured at the pulp in (Dif_C_Pu) between subjects in palatal stage 2 and 3 of 1.7022mm (95% CI, -0.021 to 3.43) which was suggestive of statistical significance (p=0.054, Appendix 3.11). Additionally, there was a mean increase the angulation of the right maxillary first molar (Dif_ANG_16) post expansion between subjects in palatal stage 1 and 3 of 5.635 degrees (95% CI, 1.43 to 9.84) which demonstrated statistical significance (p=0.005), however no other group differences were statistically significant (Appendix 3.11). Removal of extreme outliers specifically subject #4, #8 and #50, showed no difference in the findings and consequently it is appropriate to keep these outliers.

Evaluation of the percent changes in intercanine and intermolar pulpal and apical widths, palatal width, left to right infraorbital foramen width and change in left and right canine and molar angulations was calculated (Table 3.4). The greatest percentage change was found in intercanine apical width from pre- to post-expansion, increasing 13.13%. For both intercanine and intermolar widths the greatest percent change occurred at the apical level versus the pulpal level. On average, the percent change in canine angle was two-times that of the percent change experienced at the molar on its respective side. Comparison of the skeletal widths noted that the palatal slope had a greater percent change in width in comparison to the percent change in infraorbital foramen width.

Table 3.4: percent change in dependent variables from pre- to post-expansion (T2-T1)

Percent change % in mm (millimeters)						Percent change % in degrees			
per_C_Ap	per_C_Pu	per_M_Ap	per_M_Pu	Per_slope	per_IOF	per_ANG_13	per_ANG_23	per_ANG_16	per_ANG_26
13.13	7.54	12.89	11.83	9.23	2.83	5.19	5.80	2.20	2.85

3.5 Discussion:

Development of a protocol enabling clinicians to make reliable treatment planning decisions as it pertains to the modality of palatal expansion in late adolescents and young adults has been elusive to date. This retrospective study has offered insight into the efficacy of implementing the Angelieri et al. protocol[1] to predict success of skeletal expansion of the maxilla.

The two-way MANCOVA revealed that the covariates of age and CVM stage in relation to the dependent variables were not statistically significant and thus the covariates were no longer considered in any further statistical analyses. Their removal makes clinical sense since chronological age is a poor predictor of remaining growth, skeletal growth velocity or skeletal maturity[80]. Additionally, cervical vertebrae maturation (CVM) has been suggested in replacement of the hand-wrist radiograph to minimize patient radiation exposure, as no additional radiograph would be needed apart from initial records[80], to individually assess skeletal maturity[2] and assess peak growth spurt (PGS)[81]. However, comparison of the CVM to the Fishman maturation prediction method (FMP) suggest highly variable correlation values between these two methods[80].Taken individually, chronological age or CVM alone do not give a comprehensive assessment of skeletal maturity. Even the hand-wrist radiograph technique, considered to be the standard methodology to assess patient skeletal maturation[81] has limitations. These limitations include the need for additional ionizing radiation exposure[80, 81], sequence of ossification in the hand being subject to polymorphism[81] and the fact this methodology utilizes long bones formed by endochondral ossification to assess the maturity of facial bones formed

by intermembranous ossification, which may additionally show increased plasticity to local environmental conditions[80] and applied forces. Consequently, a comprehensive assessment of a host of biological indicators is best to assess skeletal maturity[2], and these indicators include; hand-wrist skeletal maturation assessment, evaluation of CVM, changes in height, menarche, development of secondary sexual characteristics [2]. Given that the predictive ability of age or CVM taken individually, in isolation of all other indicators, is poor, clinically it makes make sense to disregard these proposed covariates as they did do not affect the relationship between the fixed factors (palatal stage & sex) and the dependent skeletal and dental variables. In other words, the results of this study do not support the use of either age or CVM method to attempt prediction of palatal suture interdigitation.

Results suggested that the fixed variable of sex had no effect any of the dependent variables. This result was interesting considering that it is reported that the circummaxillary sutures fuse earlier in females, shortly after the initiation of puberty, ranging from approximately age 11-13 for females and 13-15 for males[64]. A possible cause of such finding may be that the average age of the 37 females in the study was 13.92 years, similar to that of the 26 male subjects, having an average age of 14.05 years. Given this average male age of 14.05 years, and the age range reported for the fusion of male circummaxillary sutures (13-15 years old), it is statistically possible that all 26 male subjects had fusion of their circumaxillary sutures by 14 years old, as it falls within this noted age range. Taken collectively, it is conceivable that the approximate average age for all subjects in the study of 14 years old corresponds to a timeframe where the degree of circummaxillary suture fusion and subsequent resistance to expansion was on average,

potentially equal amongst all subjects. This possibly accounts for the both the lack of sex effects, as well as, effect of age (as a covariate) on the response variables. Due to the imbalanced sex/age distribution these results should be confirmed or refuted in future projects.

A number of the findings found in Table 3.4 reflect what is anticipated from RME expansion in conjunction with comprehensive full fixed appliance orthodontic treatment. The expansion treatment created a greater percent change of skeletal width measured at the palatal slopes than the more stable and superiorly located IOF. This is anticipated due to the hinge and flexure effect of palatal expansion where the anterior-inferior most aspect of the maxilla separates to a greater extent than the superior-posterior aspect. Consequently, the effects of expansion had little effect on the superiorly located IOFs and greater change experienced at the palatal slopes. The greatest percent change in terms of dental widths was exhibited at the canines measured at the apices, and the least percent change in dental widths experienced at the canines measured at the pulpal space. This reflects a significant degree of labial root movement, which could be accounted for by the following mechanism. In the context of a hypoplastic maxilla requiring expansion, maxillary canines erupt labially with significant buccal proclination and often times remain blocked out, superiorly positioned out of the plane of occlusion, like a number of the individuals in this sample. Subsequent well timed RME treatment increases the arch circumference, maxillary transverse width, and decreases the crowding at the maxillary canines that tend to perpetuate their buccal inclination. The newly developed arch perimeter facilitates uprighting of the maxillary canines, via labial root movement as they fully descend into the plane of occlusion, into a more functional and esthetic axial

inclination. This mechanism would also account for the nearly double amount of percent change in angulation of each maxillary canine in relation to the maxillary molars. The maxillary molars tend not to erupt buccally, even sometimes more lingual than normal among these types of cases, necessitating less uprighting especially if properly controlled and timed expansion can produce bodily expansion at the molars. Any buccal tipping of the maxillary molars produced could later be corrected post-expansion via comprehensive orthodontic treatment using an appropriate bracket prescription on the maxillary first molars (-14 degrees for both Roth and MBT prescriptions). The amount of percent change measured at the molars at both the apices and pulpal space were relatively equal and possibly reflect bodily movement of the maxillary molars on average from the expansion treatment.

The results also suggest that subjects having pre-expansion stage 2 palatal maturity had a 1.7022mm significant increase in intercanine width measured at the pulp after expansion, as compared to those subjects having a stage 3 pre-expansion palatal maturity. These finding suggests that it is more efficacious to increase the intercanine width of a subject via expansion treatment if they possess a more immature palatal stage prior to the expansion mechanics. However, the change in intercanine width does not factor into the type of expansion experienced across the canines, whether it was dental expansion via buccal tipping or true skeletal expansion via separation of the palatal shelves that house the dental alveolus. Lack of significant findings in the variable measuring skeletal expansion namely, the pre- and post-expansion difference in palatal slope widths (DIF_Slope), as well as lack of significant findings in the differences in either of the canine angulations pre- and post-expansion (DIF_ANG_13 and DIF_ANG_23) suggest

that the mechanism of intercanine width development is unaccounted for. It is possible that the intercanine expansion occurred as a combination of both skeletal and dental expansion across the canines, each to a degree that was not statistically significant. Definitive findings in either DIF_Slope or DIF_ANG_13 and DIF_ANG_23 would elucidate the mechanism of expansion across the canines and further suggest how truly efficacious skeletal expansion is for subjects in palatal stage 2 as compared to those in stage 3.

Additionally, the results note subjects having pre-expansion stage 1 palatal maturity had a significant mean increase in the angulation of the right maxillary first molar of 5.635 degrees after expansion, as compared to those who had a stage 3 pre-expansion palatal maturity. The findings explicitly suggest that subjects in palatal stage 1 who undergo expansion have a mean increase in uprighting of tooth #16 to the order of 5.635 degrees in comparison to those subjects who possessed a stage 3 pre-expansion palatal maturity. The findings are suggestive of true skeletal expansion across the suture and uprighting of the posterior segments, levelling the curve of Wilson. Post-expansion CBCT records occurred following the removal of the hyrax appliance, explaining the mechanism of posterior uprighting, namely, via transverse relapse at the cervical level of the maxillary molars due to the stretched fibers of the palatal mucoperiosteum[3]. Contrary to the suggestive findings of skeletal expansion, there was no significant difference in pre- and post-expansion palatal slope widths (DIF_Slope) between subjects with different T1 palatal stages. Additionally, regardless whether expansion be skeletal, dental or a combination thereof, conventional RME occurs bilaterally. Lack of significant findings in regards to the change in angulation of the contralateral maxillary first molar

(DIF_ANG_26) makes the significant findings at the right maxillary first molar circumspect, and these findings should be taken with caution.

Although there was a clinically significant effect of T1 palatal stage on intercanine width measured at the pulp from stage 2 to stage 3 and difference in angulation of 1.6 from palatal stage 1 to 3, there were 8 other dental widths, angles and skeletal widths that were unaffected by the patients palatal stage at T1. Consequently, it can be inferred that palatal stage at T1 truly had little to no effect on these important and clinically relevant angulations and widths throughout the course of expansion treatment and on a qualitative level, we accept our null hypothesis and reject the alternative hypothesis. There was no clinically or statistically significant effect of T1 palatal stage on bony expansion across the suture. The impact of an early compared to a more advanced palatal stage 3 was not analyzed in this study. Maybe there is the need to further subdivide palatal stage 3 in to early, intermediate and advanced interdigitation stages for clinically meaningful findings that reflect the progressive difficulties in maxillary expansion in more mature individuals. The results of this study were so thoroughly unsupportive of this proposed novel classification system[1] that an investigation into the scientific basis for the proposed novel classification is warranted. The authors [1] noted that the characteristics defining each proposed *human, midpalatal suture* maturation stage (A-E), defined explicitly in the *axial* plane, were taken from the findings of previously performed histological evaluations[13-15]. These histological studies and characteristics thereof are outlined in Table 3.5.

Table 3.5: Summary of histological studies cited as source of findings used to define the proposed palatal suture maturational stages (A-E)[1].

Author	Study Design and Objective	Specimens	Plane of histological sections (coronal, sagittal, axial)	Characteristics of histological sections	Magnification	Comments
Persson et al. 1978[13]	Cross-sectional Observation study. Objective was to describe the morphology of the initial stages of sutural closure in human and rabbit subjects, to ascertain the role of cellular components during sutural initial stage obliteration process via histochemical and histological assessment.	**Human Material:** necropsy samples from 24 human subjects aged 15-35yo taken post-mortem. Samples are of intermaxillary and transverse palatine sutures of the palate. **Animal Material:** sagittal and interfrontal sutures of rabbits 25-36 months in age	Unknown - plane of histological sections not communicated	Collagen fiber bundles demonstrate two structural patterns; oriented perpendicular and parallel to sutural margins. Dense bundles of perpendicular fibers were at times observed to be tendon-like. Two patterns of initial obliteration observed: First, presence of one or more bone spicules extending from sutural margins into sutural gap and/or extending to bridge entirety of sutural gap. Spicules tended to be found in more mature specimens. Second (found almost exclusively to human MPS), presence of irregular, heterogeneous, acellular calcified bodies with clearly demarcated margins, existing freely in the sutural gap, or extending from bone spicules from the sutural margins. Multiple bodies may coalesce to form a larger calcified mass. Cyst-like spaces may exist. Additionally, other specimens with woven bone had bridge the sutural gap.	All specimens evaluated at X 250	Distribution of female and male samples not given. The plane of the photomicrographs taken for the study also unknown/not communicated. Study found that suture obliteration took place via intramembranous ossification, where trans-sutural tendon-like tissue is located. There was no radiological evaluation of the sutures of interest nor radio-morphological description thereof.
Cohen, MM (1993)[14]	Review, to correlate known sutural development and biology to the development of craniosynostosis. Heavy emphasis on review of craniosynostosis associated literature. Brief	Not applicable	Not applicable	Review reiterated findings of Persson et al. 1978, above.	Not Applicable	Paper was a review and delivered no advanced histological findings related to the MPS not previously described by Persson et al. 1978, above. No radiological

	review of findings of Persson et al. (1978)[13], above.					evaluation of the MPS or facial sutures nor radio-morphological description thereof.
Sun et al. (2004)[15]	Investigated the effect of masticatory strain in the posterior interfrontal and anterior interparietal sutures. Additionally, growth was quantified at said sutures and adjacent bone surfaces.	**Animal Material:** 14 Hanford miniature pigs (*Sus scrofa, Charles River Labs*) Four 3-month old, four 5-month old and six 7-month old miniature swine	Coronal plane	Interparietal sutures of all subjects were straight/flat or mildly irregular in morphology, being closed only on the ectocranial aspect and patent (yet narrower) on the endocranial aspect. Interfrontal suture displayed complex internal interdigitation. Interfrontal suture of all 3mo old pigs was interdigitated, in the 5- and 7- month old pigs the endocranial aspect of this suture was interdigitated, while the ectocranial aspect was relatively straight/flat. Suture width was greter on the ecto-rather than endocranial aspect of the sutures for both the interfrontal and interparietal sutures. In flat areas sharpey's fibers were positioned perpendicularly into the bony margins, and in the interdigitated aspects the collagen fibers were positioned obliquely to the bony margins.	Calibration bar at 500um and 1000 um	Findings contradict human suture fusion which occurs from the endocranial aspect. There was no radiological evaluation of the sutures of interest nor radio-morphological description thereof.

The most recent cited study was conducted by Sun et al. (2004)[15] and primarily investigated how strain from the muscles of mastication effected the posterior interfrontal and anterior parietal cranial sutures in 14 miniature pigs (*Sus scrofa*) of varying ages

from 3- to 7- months old. Histological sections of these cranial sutures were taken in the *coronal* plane and photomicrographs presented with a calibration bar to the order of 500-1000 μm. Obvious concerns exist in this study to highly question its use to develop the proposed novel MPS maturation classification[1]. Firstly, the study utilized miniature pigs instead of human specimens, and the ages of these subjects were not correlated to human controls. Secondly, the sutures evaluated were cranial sutures, not the MPS. The MPS, like other midline sutures, is considered an end-to-end type suture whereas various off-midline cranial sutures of humans and other species are considered overlapping or beveled type sutures[14]. The forces exerted across midline end-to-end type sutures are equal across the sutural gap. Conversely, beveled type sutures of the cranial complex experience asymmetric forces across the sutural gap, creating a different biological response and morphology in comparison to end-to-end type midline sutures, making these beveled type sutures poor correlates to the MPS. Given the species utilized (not human), their age (not correlated to human controls), plane of histological sections (coronal vs. axial) and sutures identified (off-midline beveled cranial sutures vs midline end-to end MPS), it is difficult to draw conclusions from this study to establish, or even propose any of the five maturational stages (A-E) noted in the proposed axially defined human MPS classification[1].

A second paper[14] cited as being utilized to develop this proposed classification[1] was not an experimental study, but rather a review of known sutural biology as it relates to craniosynostosis, or the premature fusion of cranial sutures. This review focused extensively on the *histological* findings of prematurely fused human *cranial sutures* and offered no new insight into human MPS sutural biology. Consequently, it is difficult to

draw conclusions from this study to establish, or even propose any of the five maturational stages (A-E)[1]. However, an extremely limited portion of this paper was dedicated to a review of the findings from the next paper cited by Anglieri et al. (2013)[1], an experimental study by Persson et al. (1978)[13].

The last cited source is a cross-sectional observational study performed by Persson et al. (1978)[13]. Human autopsy material from 24 human subjects were utilized to evaluate both the intermaxillary and transverse palatine sutures. As with the previous studies outlined, obvious concerns exist in this study to highly question its use to develop the proposed MPS maturation classification[1]. Firstly, the age of the human specimens ranged from 15-35 year old, consequently these samples would not have included patent or immature palatal sutures. Additionally, the authors only described two patterns of histological obliteration, not five stages as outlined by Angelieri et al.[1]. Secondly, nowhere in the study was the plane of the histological sections utilized communicated. In fact, the photomicrographs presented appear to be in the *coronal* plane, rather than, the *axial* plane utilized in the proposed classification[1].

More fundamentally ill-conceived problems exist in using both Sun et al. (2004)[15] and Persson et al. (1978)[13] to formulate the proposed classification[1] based on *axial* cross-sections of the MPS and the observed *radiological morphology*. Outside of the fact, that as previously mentioned, in both studies the plane of histological sections (*coronal*) is perpendicular to those noted in the proposed MPS classification[1] (*axial*), a more fundamentally flawed assertion by Angelieri et al. (2013)[1] from these cited studies[13, 15] is obvious. The histological sections were magnified to the order of x250 for the Persson et al. (1978) study[13], and to the level of 500-1000 μm for the Sun et al. (2004) study[15].

The ability to relate the histological morphology of the sutures in the coronal plane at x250 magnification, and on the micro-meter level scale, to the axially generated cross-sections of the MPS at eye-level, is not possible and to date there has been no scientific precedence for such a correlation. Additionally, no radiological evaluation of the histological sections were presented in either study[13, 15]. Consequently, utilization of these studies [13-15] to draw inferences into the radiological morphology of the MPS could be deemed scientifically flawed, fundamentally impossible, and not plausible.

Undeniably, this proposed palatal maturation classification system[1] has no concrete scientific basis. At its best, it is based on an extremely subjective and questionably conceived morphological interpretation of previous histological studies[13-15] that falls outside the realm of basic science or plausibility. Consequently, it is advisable that until further research supportive of this classification[1] is developed and rigorously tested, that clinicians do not consider this proposed classification[1] as being factual, and halt employing its use to drive clinical decision making which will have real-world patient implications and outcomes.

3.6 Conclusion, Limitations & Future Recommendations:

Multiple limitations were present in the study and inherent to the proposed novel classification[1] itself.

When the subjects were evaluated for their pre-expansion MPS classification, it was found the sample as a whole was poorly balanced amongst all maturation stages (Figure 3.4). Our sampling bias was due to limited available information at the University of Alberta. In fact, the greatest disparity in the distribution was found to be of those subjects with the most immature to intermediately mature palatal stages, especially of palatal stage A which consisted of a single subject (n=1) (Figure 3.4). Taken collectively, stages A-C (n=21) consisted of approximately 2/3rds the amount of subjects in palatal maturation stage E which had the greatest distribution of subjects (n= 29). The distribution of stage E subjects were more than double that of stage D subjects (n= 13) (Figure 3.5).

When stages were modified into stages 1, 2 and 3 (Figure 3.5) there was an increase the balance of the sample distribution as a whole and improved statistical power given that the power of a balanced or equal-allocation sample is greater than the power of a highly-unbalanced sample. It should be noted however that the benefits of the modified staging were limited to the immature palatal stages (A, B and C), collectively termed modified stage 1. In spite of this reallocation and modification of the palatal stages, the ratio of stage E (modified stage 3) to stage D (modified stage 2) subjects was unchanged, and was a limitation of the statistical power of the study.

The proposed classification[1] was based on the qualitative interpretation of the MPS in the axially generated slices, which in itself was a limitation due to the subjective nature of this interpretation. This was further compounded by the large voxel size, constraining the axial cross-section thickness to a default of 1.0mm and associated poor image clarity, which negatively affected the visualization and interpretation of the radio-morphology of the T1 MPS, rendering the classification of subjects increasingly difficult. Poor image quality affected both the reliability testing and assessment of the efficacy of the classification to predict success of RME therapy. These findings are directly attributable to the poor reliability findings (Chapter 2, Table 2.1) that was noted by the remaining raters to be a consistent and significant source of difficulty in staging subjects. Additionally, a work in progress has found that local greyscale within a specific bony region of interest (ROI) is not reliable when scans are taken between days and within the same day. A large variation in greyscale was noted, affecting the relative greyscale distribution in the ROI. The large variation in local greyscale would decrease the efficacy in utilizing greyscale values for quantitative evaluation of the ROI, and conversely confound the qualitative interpretation of this same region. Extrapolation of these findings to this study would offer insight into another aspect of difficulty in staging the T1 palatal sutures. Repeated daily use of the CBCT utilized to generate the T1 records of the participating subjects would have introduced a great variability of relative greyscale distribution in the small ROI of this study (the axially generated MPS cross-section). This would impact image clarity and contrast, compounding the already poor image quality introduced by the large voxel size, furthering the difficulty in the qualitative interpretation and staging of the axially generated T1 MPS cross-sections.

Voxel size can be controlled by the size of the field of view (FOV) investigated[82], slice thickness in millimeters (mm's)[83], and matrix size[84].Given that all scans were full-FOV, large voxel size and poor image clarity was anticipated but not to the extent it presented, hampering the identification of radiological characteristics to be able to more properly classify subject maturation stage. Limiting the field of view (FOV)[82] and matrix size[84], decreases voxel size and creates an image with greater spatial resolution at the expense of increasing the patient dose exposure[84].Thus, a suggestion to improve the protocol would be to ensure the CBCT unit utilized is capable of changing the FOV, and to capture a second scan limited to only the hard palate of the patient, decreasing slice thickness and matrix size. The negative to this protocol would delivering increased radiation to the patient in the form of a second, limited FOV scan of the hard palate. Following the principles of ALARA, the clinical decision to proceed with this secondary scan would be based on the risk/benefit ratio to the patient namely; whether the risk associate with an increased patient dose is worth the benefit of improved diagnostic information in order to identify the modality of expansion (conventional RME or SARME) that offered the lowest burden or treatment, highest predictability and efficacy of expansion. However, the value of this improved protocol is only valuable if the academic foundations of the methodology and proposed classification[1] were deemed scientifically sound.

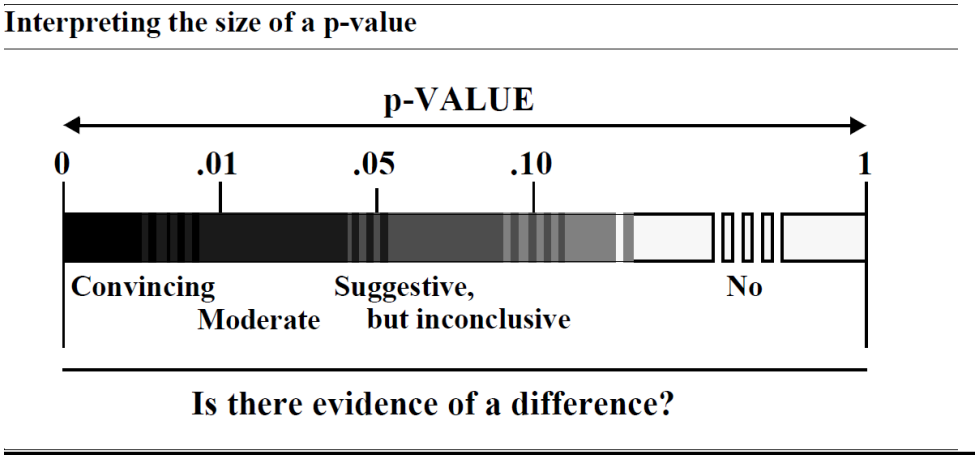
There were significant limitations and concerns inherent to the foundational development of the proposed classification presented by Angelieri et al. (2013)[1]. These concerns directly relate to the tertiary research question: *What alteration(s) or modification(s) to the Angelieri et al..[1] methodology can be suggested to improve reliability and/or*

predictive ability of this classification system? Firstly, the presented classification was not cross-referenced with a gold standard, without which substantiating the proposed classification is not possible. Employing a gold standard in subsequent studies can be suggested however the current gold standard methodology (biopsy) to assess the degree of palatal suture fusion is invasive and not performed on live subjects. Furthermore, if such a proposal were adapted by an ethics review board for implementation in a clinical study, utilizing the histological sections developed from biopsy would still not substantiate the proposed classification.

As previously mentioned correlating the highly magnified histological morphology of the palatal suture to generated radiologic cross-sections of the MPS at eye-level through standard CBCT imaging is not feasible and to date there has been no scientific precedence for such a correlation. As concluded in the discussion, this proposed palatal classification system[1] is significantly flawed in regards to its scientific basis. The same conclusion holds, the proposed classification[1] is based on an extremely subjective and poorly conceived qualitative interpretation of previous histological studies[13-15]. Ultimately, improvement of the proposed classification[1], is not possible. A comprehensive overhaul of its scientific basis would be required and rigorously tested against a gold standard. As it stands today, it is advisable that the proposed classification[1] not be employed to drive clinical decision making in clinical orthodontics.

3.7 Appendix:

Appendix 3.1 Interpretation of p-value

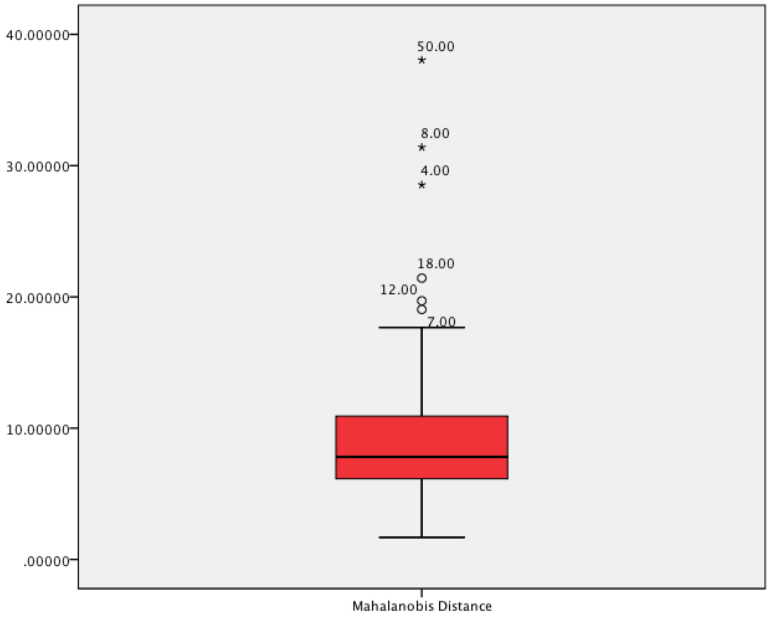


Appendix 3.2

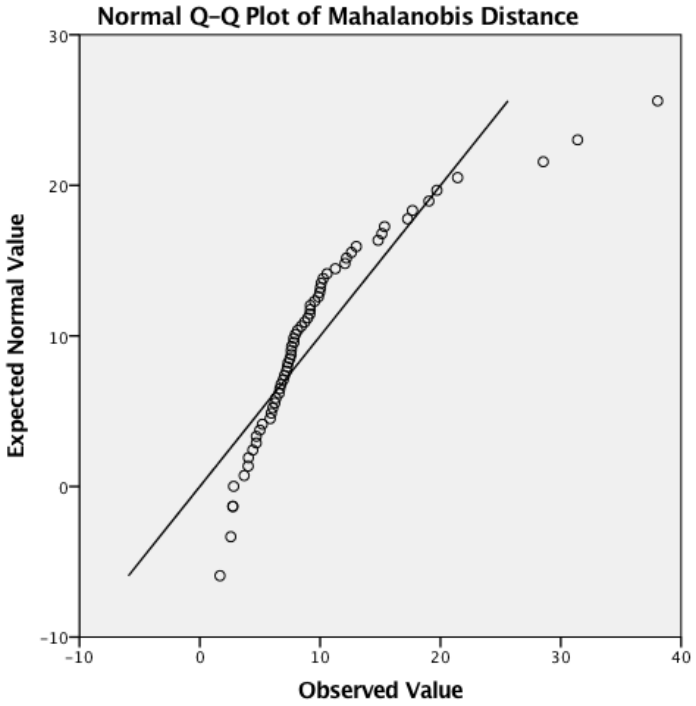
Appendix 3.2 Descriptive statistics of CVM and repeated measures for dental and skeletal widths and dental angulations.

		T1				T2				Difference T2-T1			
Distances	N	**Mean at T1**	Std. Deviation	Minimum	Maximum	**Mean at T2**	Std. Deviation	Minimum	Maximum	**Mean**	Std. Deviation	Minimum	Maximum
CVM	63	**3.94**	1.18	1	6	**-**	-	-	-	**-**	-	-	-
C_Ap	63	**23.77**	2.81	18.02	29.17	**26.83**	3.58	19.80	35.86	**3.06**	2.43	-1.53	8.91
C_Pu	63	**29.07**	3.10	22.03	36.24	**31.15**	3.07	21.06	37.42	**2.08**	2.19	-5.36	8.33
M_Ap	63	**30.33**	3.43	24.57	39.35	**34.17**	4.09	25.52	44.18	**3.85**	2,46	-2.53	8.18
M_Pu	63	**40.93**	3.22	35.52	49.27	**45.68**	3.44	38.20	55.23	**4.74**	2.81	-6.81	11.00
Slope	63	**24.85**	3.00	15.22	31.67	**27.10**	3.08	18.79	34.22	**2.25**	2.11	-7.31	8.19
IOF	63	**45.90**	3.27	39.13	55.14	**47.16**	3.36	40.25	54.24	**1.26**	2.05	-7.10	8.00
ANG_13	63	**126.29**	9.41	101.66	148.61	**132.36**	9.32	105.37	153.21	**6.07**	10.56	-26.37	30.22
ANG_23	63	**128.41**	11.76	96.86	149.49	**135.48**	12.64	105.39	158.59	**7.06**	10.33	-13.28	38.79
ANG_16	63	**113.86**	8.91	93.50	137.18	**116.25**	9.97	99.15	147.79	**2.39**	6.39	-14.39	31.59
ANG_26	63	**114.42**	10.21	96.75	151.88	**117.52**	9.77	96.54	153.40	**3.10**	4.78	-5.99	14.87

Appendix 3.3: Boxplot of Mahalanobis distance calculated from experimental data Anticipated in a sample with no outliers is a box plot with an observed normal distribution. Mahalnobis distance corresponds to a transformed Euclidean distance and therefore is unitless.



Appendix 3.4: Q-Q plot of Mahalanobis distance from experimental data.



Appendix 3.5: Box's test of equality of covariance matrices for 2 way MANCOVA

Box's Test of Equality of Covariance Matrices[a]

Box's M	210.344
F	1.063
df1	110
df2	2613.297
Sig.	.313

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + CVM + AGE + New_Palate + M0_F1 + New_Palate * M0_F1

Appendix 3.6: multivariate tests for 2 way MANCOVA statistical analysis.

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.244	1.486[b]	10.000	46.000	.175	.244
	Wilks' Lambda	.756	1.486[b]	10.000	46.000	.175	.244
	Hotelling's Trace	.323	1.486[b]	10.000	46.000	.175	.244
	Roy's Largest Root	.323	1.486[b]	10.000	46.000	.175	.244
CVM	Pillai's Trace	.111	.573[b]	10.000	46.000	.827	.111
	Wilks' Lambda	.889	.573[b]	10.000	46.000	.827	.111
	Hotelling's Trace	.125	.573[b]	10.000	46.000	.827	.111
	Roy's Largest Root	.125	.573[b]	10.000	46.000	.827	.111
AGE	Pillai's Trace	.176	.981[b]	10.000	46.000	.473	.176
	Wilks' Lambda	.824	.981[b]	10.000	46.000	.473	.176
	Hotelling's Trace	.213	.981[b]	10.000	46.000	.473	.176
	Roy's Largest Root	.213	.981[b]	10.000	46.000	.473	.176
New_Palate	Pillai's Trace	.512	1.616	20.000	94.000	.065	.256
	Wilks' Lambda	.546	1.627[b]	20.000	92.000	.063	.261
	Hotelling's Trace	.728	1.637	20.000	90.000	.061	.267
	Roy's Largest Root	.528	2.484[c]	10.000	47.000	.018	.346
M0_F1	Pillai's Trace	.330	2.264[b]	10.000	46.000	.030	.330
	Wilks' Lambda	.670	2.264[b]	10.000	46.000	.030	.330
	Hotelling's Trace	.492	2.264[b]	10.000	46.000	.030	.330
	Roy's Largest Root	.492	2.264[b]	10.000	46.000	.030	.330
New_Palate * M0_F1	Pillai's Trace	.444	1.342	20.000	94.000	.173	.222
	Wilks' Lambda	.600	1.337[b]	20.000	92.000	.177	.225
	Hotelling's Trace	.591	1.330	20.000	90.000	.182	.228
	Roy's Largest Root	.409	1.921[c]	10.000	47.000	.066	.290

a. Design: Intercept + CVM + AGE + New_Palate + M0_F1 + New_Palate * M0_F1

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Appendix 3.7 multivariate tests for 2 way MANCOA statistical analysis.

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.807	20.099[b]	10.000	48.000	.000	.807
	Wilks' Lambda	.193	20.099[b]	10.000	48.000	.000	.807
	Hotelling's Trace	4.187	20.099[b]	10.000	48.000	.000	.807
	Roy's Largest Root	4.187	20.099[b]	10.000	48.000	.000	.807
New_Palate	Pillai's Trace	.545	1.837	20.000	98.000	.027	.273
	Wilks' Lambda	.517	1.874[b]	20.000	96.000	.023	.281
	Hotelling's Trace	.813	1.909	20.000	94.000	.020	.289
	Roy's Largest Root	.616	3.020[c]	10.000	49.000	.005	.381
M0_F1	Pillai's Trace	.293	1.991[b]	10.000	48.000	.055	.293
	Wilks' Lambda	.707	1.991[b]	10.000	48.000	.055	.293
	Hotelling's Trace	.415	1.991[b]	10.000	48.000	.055	.293
	Roy's Largest Root	.415	1.991[b]	10.000	48.000	.055	.293
New_Palate * M0_F1	Pillai's Trace	.426	1.325	20.000	98.000	.182	.213
	Wilks' Lambda	.616	1.315[b]	20.000	96.000	.189	.215
	Hotelling's Trace	.555	1.304	20.000	94.000	.196	.217
	Roy's Largest Root	.373	1.829[c]	10.000	49.000	.080	.272

a. Design: Intercept + New_Palate + M0_F1 + New_Palate * M0_F1

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Appendix 3.8 : Failed Box's test of equality of covariance matrices for one-way MANOVA.

Box's Test of Equality of Covariance Matrices[a]

Box's M	197.739
F	1.280
df1	110
df2	4648.362
Sig.	.027

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + New_Palate

Appendix 3.9a multivariate tests for one-way MANCOA statistical analysis.

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.806	21.200[b]	10.000	51.000	.000	.806
	Wilks' Lambda	.194	21.200[b]	10.000	51.000	.000	.806
	Hotelling's Trace	4.157	21.200[b]	10.000	51.000	.000	.806
	Roy's Largest Root	4.157	21.200[b]	10.000	51.000	.000	.806
New_Palate	Pillai's Trace	.484	1.659	20.000	104.000	.053	.242
	Wilks' Lambda	.561	1.709[b]	20.000	102.000	.044	.251
	Hotelling's Trace	.702	1.756	20.000	100.000	.037	.260
	Roy's Largest Root	.559	2.907[c]	10.000	52.000	.006	.359

a. Design: Intercept + New_Palate

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Appendix 3.9b Levene's test for equality of means of one-way Manova

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
DIF_C_Ap	.798	2	60	.455
DIF_C_Pu	1.161	2	60	.320
DIF_M_Ap	.062	2	60	.940
DIF_M_Pu	1.638	2	60	.203
DIF_Slope	1.050	2	60	.356
DIF_IOF	.366	2	60	.695
DIF_ANG_13	.636	2	60	.533
DIF_ANG_23	3.757	2	60	.029
DIF_ANG_16	.064	2	60	.938
DIF_ANG_26	.381	2	60	.685

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + New_Palate

Appendix 3.10 Tests of between-subjects effects for one-way Manova

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	DIF_C_Ap	11.323[a]	2	5.662	.957	.390	.031
	DIF_C_Pu	34.448[b]	2	17.224	3.919	.025	.116
	DIF_M_Ap	17.986[c]	2	8.993	1.506	.230	.048
	DIF_M_Pu	7.216[d]	2	3.608	.450	.640	.015
	DIF_Slope	.010[e]	2	.005	.001	.999	.000
	DIF_IOF	6.177[f]	2	3.088	.725	.489	.024
	DIF_ANG_13	493.022[g]	2	246.511	2.304	.109	.071
	DIF_ANG_23	136.446[h]	2	68.223	.631	.535	.021
	DIF_ANG_16	397.050[i]	2	198.525	5.588	.006	.157
	DIF_ANG_26	54.877[j]	2	27.438	1.209	.306	.039
Intercept	DIF_C_Ap	548.947	1	548.947	92.811	.000	.607
	DIF_C_Pu	297.891	1	297.891	67.784	.000	.530
	DIF_M_Ap	857.628	1	857.628	143.639	.000	.705
	DIF_M_Pu	1312.782	1	1312.782	163.844	.000	.732
	DIF_Slope	287.980	1	287.980	62.353	.000	.510
	DIF_IOF	104.288	1	104.288	24.478	.000	.290
	DIF_ANG_13	1902.786	1	1902.786	17.781	.000	.229
	DIF_ANG_23	2984.260	1	2984.260	27.619	.000	.315
	DIF_ANG_16	449.213	1	449.213	12.645	.001	.174
	DIF_ANG_26	485.003	1	485.003	21.375	.000	.263
New_Palate	DIF_C_Ap	11.323	2	5.662	.957	.390	.031
	DIF_C_Pu	34.448	2	17.224	3.919	.025	.116
	DIF_M_Ap	17.986	2	8.993	1.506	.230	.048
	DIF_M_Pu	7.216	2	3.608	.450	.640	.015
	DIF_Slope	.010	2	.005	.001	.999	.000
	DIF_IOF	6.177	2	3.088	.725	.489	.024
	DIF_ANG_13	493.022	2	246.511	2.304	.109	.071
	DIF_ANG_23	136.446	2	68.223	.631	.535	.021
	DIF_ANG_16	397.050	2	198.525	5.588	.006	.157
	DIF_ANG_26	54.877	2	27.438	1.209	.306	.039
Error	DIF_C_Ap	354.882	60	5.915			
	DIF_C_Pu	263.682	60	4.395			
	DIF_M_Ap	358.242	60	5.971			
	DIF_M_Pu	480.743	60	8.012			
	DIF_Slope	277.115	60	4.619			
	DIF_IOF	255.631	60	4.261			
	DIF_ANG_13	6420.571	60	107.010			
	DIF_ANG_23	6482.961	60	108.049			

	DIF_ANG_16	2131.551	60	35.526			
	DIF_ANG_26	1361.422	60	22.690			
Total	DIF_C_Ap	955.133	63				
	DIF_C_Pu	570.195	63				
	DIF_M_Ap	1308.353	63				
	DIF_M_Pu	1903.986	63				
	DIF_Slope	597.324	63				
	DIF_IOF	361.047	63				
	DIF_ANG_13	9232.637	63				
	DIF_ANG_23	9759.695	63				
	DIF_ANG_16	2888.511	63				
	DIF_ANG_26	2020.490	63				
Corrected Total	DIF_C_Ap	366.205	62				
	DIF_C_Pu	298.130	62				
	DIF_M_Ap	376.228	62				
	DIF_M_Pu	487.959	62				
	DIF_Slope	277.125	62				
	DIF_IOF	261.808	62				
	DIF_ANG_13	6913.593	62				
	DIF_ANG_23	6619.407	62				
	DIF_ANG_16	2528.600	62				
	DIF_ANG_26	1416.299	62				

a. R Squared = .031 (Adjusted R Squared = -.001)

b. R Squared = .116 (Adjusted R Squared = .086)

c. R Squared = .048 (Adjusted R Squared = .016)

d. R Squared = .015 (Adjusted R Squared = -.018)

e. R Squared = .000 (Adjusted R Squared = -.033)

f. R Squared = .024 (Adjusted R Squared = -.009)

g. R Squared = .071 (Adjusted R Squared = .040)

h. R Squared = .021 (Adjusted R Squared = -.012)

i. R Squared = .157 (Adjusted R Squared = .129)

j. R Squared = .039 (Adjusted R Squared = .007)

Appendix 3.11: Bonferroni corrected multiple comparison post-hoc testing of One-way Manova.

Multiple Comparisons

Bonferroni

Dependent Variable	(I) New_Palate	(J) New_Palate	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
DIF_C_Ap	1.00	2.00	.5148	.85827	1.000	-1.5990	2.6287
		3.00	.9633	.69686	.516	-.7530	2.6797
	2.00	1.00	-.5148	.85827	1.000	-2.6287	1.5990
		3.00	.4485	.81175	1.000	-1.5508	2.4478
	3.00	1.00	-.9633	.69686	.516	-2.6797	.7530
		2.00	-.4485	.81175	1.000	-2.4478	1.5508
DIF_C_Pu	1.00	2.00	-.3989	.73981	1.000	-2.2210	1.4232
		3.00	1.3033	.60068	.102	-.1761	2.7827
	2.00	1.00	.3989	.73981	1.000	-1.4232	2.2210
		3.00	1.7022	.69971	.054	-.0212	3.4256
	3.00	1.00	-1.3033	.60068	.102	-2.7827	.1761
		2.00	-1.7022	.69971	.054	-3.4256	.0212
DIF_M_Ap	1.00	2.00	.8436	.86232	.996	-1.2803	2.9675
		3.00	1.2080	.70015	.269	-.5164	2.9325
	2.00	1.00	-.8436	.86232	.996	-2.9675	1.2803
		3.00	.3645	.81558	1.000	-1.6443	2.3732
	3.00	1.00	-1.2080	.70015	.269	-2.9325	.5164
		2.00	-.3645	.81558	1.000	-2.3732	1.6443
DIF_M_Pu	1.00	2.00	.1292	.99894	1.000	-2.3311	2.5896
		3.00	.7221	.81107	1.000	-1.2756	2.7197
	2.00	1.00	-.1292	.99894	1.000	-2.5896	2.3311
		3.00	.5928	.94479	1.000	-1.7341	2.9198
	3.00	1.00	-.7221	.81107	1.000	-2.7197	1.2756
		2.00	-.5928	.94479	1.000	-2.9198	1.7341
DIF_Slope	1.00	2.00	-.0324	.75842	1.000	-1.9003	1.8356
		3.00	-.0220	.61579	1.000	-1.5387	1.4946
	2.00	1.00	.0324	.75842	1.000	-1.8356	1.9003
		3.00	.0103	.71731	1.000	-1.7564	1.7771
	3.00	1.00	.0220	.61579	1.000	-1.4946	1.5387

		2.00	-.0103	.71731	1.000	-1.7771	1.7564
DIF_IOF	1.00	2.00	-.7129	.72843	.995	-2.5070	1.0812
		3.00	.0934	.59144	1.000	-1.3633	1.5500
	2.00	1.00	.7129	.72843	.995	-1.0812	2.5070
		3.00	.8063	.68895	.740	-.8906	2.5031
	3.00	1.00	-.0934	.59144	1.000	-1.5500	1.3633
		2.00	-.8063	.68895	.740	-2.5031	.8906
DIF_ANG_13	1.00	2.00	7.1486	3.65064	.165	-1.8428	16.1399
		3.00	5.0320	2.96406	.284	-2.2683	12.3324
	2.00	1.00	-7.1486	3.65064	.165	-16.1399	1.8428
		3.00	-2.1166	3.45275	1.000	-10.6205	6.3874
	3.00	1.00	-5.0320	2.96406	.284	-12.3324	2.2683
		2.00	2.1166	3.45275	1.000	-6.3874	10.6205
DIF_ANG_23	1.00	2.00	1.7568	3.66834	1.000	-7.2781	10.7918
		3.00	3.3428	2.97843	.799	-3.9929	10.6785
	2.00	1.00	-1.7568	3.66834	1.000	-10.7918	7.2781
		3.00	1.5860	3.46949	1.000	-6.9592	10.1312
	3.00	1.00	-3.3428	2.97843	.799	-10.6785	3.9929
		2.00	-1.5860	3.46949	1.000	-10.1312	6.9592
DIF_ANG_16	1.00	2.00	2.2713	2.10344	.854	-2.9094	7.4520
		3.00	5.6353*	1.70785	.005	1.4289	9.8416
	2.00	1.00	-2.2713	2.10344	.854	-7.4520	2.9094
		3.00	3.3640	1.98942	.288	-1.5359	8.2638
	3.00	1.00	-5.6353*	1.70785	.005	-9.8416	-1.4289
		2.00	-3.3640	1.98942	.288	-8.2638	1.5359
DIF_ANG_26	1.00	2.00	2.5881	1.68104	.387	-1.5522	6.7284
		3.00	1.2540	1.36489	1.000	-2.1077	4.6156
	2.00	1.00	-2.5881	1.68104	.387	-6.7284	1.5522
		3.00	-1.3341	1.58992	1.000	-5.2500	2.5818
	3.00	1.00	-1.2540	1.36489	1.000	-4.6156	2.1077
		2.00	1.3341	1.58992	1.000	-2.5818	5.2500

Based on observed means.

The error term is Mean Square(Error) = 22.690.

*. The mean difference is significant at the

3.8 Literature Cited:

1. Revelo, B. and L.S. Fishman, *Maturational evaluation of ossification of the midpalatal suture.* American Journal of Orthodontics & Dentofacial Orthopedics, 1994. **105**(3): p. 288-92.
2. Kosowski, T.R., et al., *Cleft palate.* Semin Plast Surg, 2012. **26**(4): p. 164-9.
3. Kjaer, I., *Human prenatal palatal shelf elevation related to craniofacial skeletal maturation.* European Journal of Orthodontics, 1992. **14**(1): p. 26-30.
4. Melsen, B., *Palatal growth studied on human autopsy material. A histologic microradiographic study.* American Journal of Orthodontics, 1975. **68**(1): p. 42-54.
5. Melsen, B. and F. Melsen, *The postnatal development of the palatomaxillary region studied on human autopsy material.* Am J Orthod, 1982. **82**(4): p. 329-42.
6. Ghoneima, A., et al., *Effects of rapid maxillary expansion on the cranial and circummaxillary sutures.* American Journal of Orthodontics & Dentofacial Orthopedics, 2011. **140**(4): p. 510-9.
7. Wertz, R. and M. Dreskin, *Midpalatal suture opening: a normative study.* Am J Orthod, 1977. **71**(4): p. 367-81.
8. Baccetti, T., et al., *Treatment timing for rapid maxillary expansion.* Angle Orthod, 2001. **71**(5): p. 343-50.
9. Liou, E., *Eric Liou.* Revista Dental Press de Ortodontia e Ortopedia Facial, 2009. **14**: p. 27-37.

10. Persson, M. and B. Thilander, *Palatal suture closure in man from 15 to 35 years of age.* American Journal of Orthodontics, 1977. **72**(1): p. 42-52.

11. Korbmacher, H., et al., *Age-dependent three-dimensional microcomputed tomography analysis of the human midpalatal suture.* Journal of Orofacial Orthopedics, 2007. **68**(5): p. 364-76.

12. Angelieri, F., et al., *Midpalatal suture maturation: classification method for individual assessment before rapid maxillary expansion.* Am J Orthod Dentofacial Orthop, 2013. **144**(5): p. 759-69.

13. Isiksal, E., S. Hazar, and S. Akyalcin, *Smile esthetics: perception and comparison of treated and untreated smiles.* Am J Orthod Dentofacial Orthop, 2006. **129**(1): p. 8-16.

14. Bjork, A., *Sutural growth of the upper face studied by the implant method.* Acta Odontol Scand, 1966. **24**(2): p. 109-27.

15. Bjork, A. and V. Skieller, *Growth of the maxilla in three dimensions as revealed radiographically by the implant method.* Br J Orthod, 1977. **4**(2): p. 53-64.

16. Primozic, J., et al., *Three-dimensional longitudinal evaluation of palatal vault changes in growing subjects.* Angle Orthod, 2012. **82**(4): p. 632-6.

17. Lou, L., et al., *Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2007. **104**(3): p. 402-11.

18. Lagravere, M.O., et al., *Cranial base foramen location accuracy and reliability in cone-beam computerized tomography.* Am J Orthod Dentofacial Orthop, 2011. **139**(3): p. e203-10.

19. Ludlow, J.B., et al., *Precision of cephalometric landmark identification: cone-beam computed tomography vs conventional cephalometric views.* Am J Orthod Dentofacial Orthop, 2009. **136**(3): p. 312.e1-10; discussion 312-3.

20. Ludlow, J.B., et al., *Dosimetry of 3 CBCT devices for oral and maxillofacial radiology: CB Mercuray, NewTom 3G and i-CAT.* Dentomaxillofac Radiol, 2006. **35**(4): p. 219-26.

21. Mah, J. and D. Hatcher, *Three-dimensional craniofacial imaging.* Am J Orthod Dentofacial Orthop, 2004. **126**(3): p. 308-9.

22. Damstra, J., et al., *Accuracy of linear measurements from cone-beam computed tomography-derived surface models of different voxel sizes.* Am J Orthod Dentofacial Orthop, 2010. **137**(1): p. 16.e1-6; discussion 16-7.

23. Broadbent, B.H., *A NEW X-RAY TECHNIQUE and ITS APPLICATION TO ORTHODONTIA.* The Angle Orthodontist, 1931. **1**(2): p. 45-66.

24. Lagravere, M.O., P.W. Major, and J. Carey, *Sensitivity analysis for plane orientation in three-dimensional cephalometric analysis based on superimposition of serial cone beam computed tomography images.* Dentomaxillofac Radiol, 2010. **39**(7): p. 400-8.

25. de Oliveira, A.E., et al., *Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2009. **107**(2): p. 256-65.

26. Lagravere, M.O., et al., *Transverse, vertical, and anteroposterior changes from bone-anchored maxillary expansion vs traditional rapid maxillary expansion: a randomized clinical trial.* Am J Orthod Dentofacial Orthop, 2010. **137**(3): p. 304.e1-12; discussion 304-5.

27. Cao, L., et al., *Effect of maxillary incisor labiolingual inclination and anteroposterior position on smiling profile esthetics.* Angle Orthod, 2011. **81**(1): p. 121-29.

28. Chirivella, P., et al., *Comparison of the effect of labiolingual inclination and anteroposterior position of maxillary incisors on esthetic profile in three different facial patterns.* J Orthod Sci, 2017. **6**(1): p. 1-10.

29. O'Higgins, E.A., R.H. Kirschen, and R.T. Lee, *The influence of maxillary incisor inclination on arch length.* Br J Orthod, 1999. **26**(2): p. 97-102.

30. Flores-Mir, C., et al., *Correlation of Skeletal Maturation Stages Determined by Cervical Vertebrae and Hand-wrist Evaluations.* The Angle Orthodontist, 2006. **76**(1): p. 1-5.

31. Bonfim, M.A., et al., *Cervical vertebrae maturation index estimates on cone beam CT: 3D reconstructions vs sagittal sections.* Dentomaxillofac Radiol, 2016. **45**(1): p. 20150162.

32. Bishara, S.E. and R.N. Staley, *Maxillary expansion: clinical implications.* Am J Orthod Dentofacial Orthop, 1987. **91**(1): p. 3-14.

33. Persson, M., B.C. Magnusson, and B. Thilander, *Sutural closure in rabbit and man: a morphological and histochemical study.* J Anat, 1978. **125**(Pt 2): p. 313-21.

34. Cohen, M.M., Jr., *Sutural biology and the correlates of craniosynostosis.* Am J Med Genet, 1993. **47**(5): p. 581-616.

35. Sun, Z., E. Lee, and S.W. Herring, *Cranial Sutures and Bones: Growth and Fusion in Relation to Masticatory Strain.* Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology, 2004. **276**(2): p. 150-161.

36. Aktan, A.M., et al., *Effects of voxel size and resolution on the accuracy of endodontic length measurement using cone beam computed tomography.* Ann Anat, 2016. **208**: p. 96-102.

37. Pauwels, R., et al., *Effect of exposure parameters and voxel size on bone structure analysis in CBCT.* Dentomaxillofac Radiol, 2015. **44**(8): p. 20150078.

38. Scarfe, W.C. and A.G. Farman, *What is cone-beam CT and how does it work?* Dent Clin North Am, 2008. **52**(4): p. 707-30, v.

Chapter 4: General Discussion

4.1 General Discussion:

Chapter 1- Review of Literature

A systematic review of literature pertaining to novel methodologies and technologies to assess mid-palatal suture maturation in humans was performed. A computerized database search was conducted using Medline, PubMed, Embase and Scopus to search the literature from 2007 up until October 5, 2016. A supplemental hand search was completed of references from retrieved articles that met the final inclusion criteria. The inclusion criterion ‘‘*Diagnostic methods to evaluate cranial suture ossification/maturation*” was utilized to initially identify possible articles from the published abstract results of the database search. Once these abstracts were selected, full articles were retrieved and inclusion in the systematic review was dependent of fulfilling a final inclusion criterion. The final selection criterion was as follows: ‘‘*In vitro and in vivo human subject studies that describe a novel diagnostic method or technology to assess mid-palatal suture maturation/ossification over time”.* Twenty-nine abstracts met the initial inclusion criteria. Following assessment of full articles, only five met the final inclusion criteria. The number of subjects involved and quality of studies varied, ranging from an *in-vitro* study using autopsy material to prospective studies with *in vivo* human patients. Three types of evaluations were identified: quantitative, semi-quantitative and qualitative evaluations. Four of the five studies utilized computed tomography (CT), while the remaining study utilized non-invasive ultrasonography. No methodology was validated against a histological-based reference standard. Weak limited evidence exists to support the newest technologies and proposed methodologies to assess mid-palatal suture maturation. Non-invasive imaging technologies present a promising and biologically safe alternative to ionizing imaging to assess sutural ossification. Due to the lack of reference standard validation, it is advised that clinicians still use a multitude of diagnostic criteria to subjectively assess palatal suture maturation and drive clinical decision-making.

Chapter 2 – Reliability Testing

Angelieri et al.[1] developed a novel classification system to image through CBCT the palatal suture and determine the palatal suture maturation stage using an ordinal scale to assign a particular maturation stage (A-E) (Chapter 2, Figure 2.1). Since successful RME treatment is dependent upon the degree of palatal fusion and properly chosen modality, the authors suggested that their classification system can direct treatment decision making to avoid significant comorbidities of an incorrectly chosen expansion modality. Anglieri et al.[1] performed a validation study that noted significant intra- and inter-examiner agreement, as well as agreement of rater to *ground truth* (previously defined in Chapter 2). Further validation of this novel technique was performed to establish its efficacy and reliability to clinical practice, leading to the primary research question of this study;

What is the reliability of the Angelieri et al.[1] midpalatal suture maturation classification system?

Intra-examiner reliability (DAI to DAI) and inter-examiner reliability (ML to CF), as well as examiner to ground truth (ML to DAI & CF to DAI) were investigated by Cohen's Kappa statistic using specified guidelines (Chapter 2). Results of the reliability testing are found in Chapter 2 (Table 2.1). There was almost perfect intra-examiner agreement between the rater DAI's palatal suture maturation classification at time 1 and 3, κ=0.915 (95% CI, 0.752 to 1.00, p <0.005. There was slight inter-examiner agreement between rater ML and CF's staging of the patient's palatal suture, κ =0.040 (95% CI, -0.209 to 0.289), p =0.733. Given the p>0.05, there is weak to no evidence to support this slight agreement is greater than chance agreement. There was moderate agreement between rater ML and the ground truth for classifying the patient's palatal suture maturation, κ=0.470 (95% CI, 0.141 to 0.799), p =0.001. There was poor agreement between rater CF and the ground truth staging of the patient's palatal suture maturation, κ = -0.015 (95%

CI, -0.25 to 0.22), p =0.896. Given the p>0.05, there is weak to no evidence to support this poor agreement is greater than solely chance agreement.

Ultimately, reliability testing disagrees with that of the original study[1], indicating that this classification system is not as reliable as previously presented. This study indicates that the proposed methodology is in fact non-intuitive, requires operator calibration and heavily influenced by the degree of post-acquisition image sharpness and clarity.

Chapter 3 - Main Investigation

The main study was a retrospective observational longitudinal (cohort) study. The objective of this aspect of the study was to answer the secondary research question: *How useful is this novel classification system[1] to predict success of RME treatment?* As well as, the *tertiary* research question, namely: *What alteration(s) or modification(s) to the Angelieri et al.[1] methodology can be suggested to improve reliability and/or predictive ability of this classification system?*

A total of 63 pre-adolescent and adolescent patients aged 11-17 years old with full permanent dentition treated with tooth-borne RME appliances who have had CBCTs records taken at two time points, T_1 pre-expansion and T_2 post-expansion treatment without comprehensive orthodontic treatment were evaluated in the study. The primary investigator (DAI) identified anatomical locations of interest and completed the 3D landmarking using Avizo version 7.0 software (visualization Sciences Group, Burlington, MA, USA). The 3D landmarks utilized were previously described (Chapter 3, Table 3.1) and produced a number of skeletal and dental widths (mm's), and dental angulations (degrees) (Chapter 3, Table 3.2).

Time 1 (pre-expansion) patient age and CVM stage were considered covariates, sex and palatal stage as fixed variables. The pre-expansion palatal staging was unbalanced across the sample (Chapter 3, Figure 3.4), and therefore a modified classification was developed to improve balance and increase the power of the experimental design (Chapter 3, Figure 3.5). Evaluation of each

subjects pre- and post-expansion CBCT volumes allowed for evaluation of the dependent variables, namely difference in skeletal and dental distances and dental angles from time 1 (pre-expansion) to time 2 (post-expansion) (Chapter 3, Table 3.3).

Multivariate analysis of variance (MANOVA) of the absolute difference between the Time 1 (T1: pre-expansion) and Time 2 (T2: post-expansion) methods was used to determine significance and test the null hypothesis that a subject's palatal stage has no effect on any of the of the dependent variables. The alternative hypothesis is that a subject's palatal stage has an effect on any of the dependent variables. A *p*-value of less than 0.05 was considered significant (Appendix 3.1). In respect to angular measurements an absolute difference ≥ 5° was considered clinically significant[2, 3], and a change of equal to or greater than 1.0mm[4] for any dental or skeletal width change. The initial statistical analysis utilized was a two-way multivariate analysis of covariance (MANCOVA) revealed age and CVM stage in relation to the dependent variables was not statistically significant (Appendix 3.6). Consequently, covariates were not further analyzed.

Two-way MANOVA revealed palatal stage demonstrated convincing statistical significance [$F(20, 96) = 1.874$, $p = .023$; Wilks' $\Lambda = .517$; partial $\eta^2 = .281$]. Sex, and the interaction between sex and palatal stage in relation to the dependent variables demonstrated inconclusive statistical significance [$F(10, 48) = 1.991$, $p = .055$; Wilks' $\Lambda = .707$; partial $\eta^2 = .293$] and [$F(20, 96) = 1.315$, $p = .189$; Wilks' $\Lambda = .616$; partial $\eta^2 = .215$] respectively (Appendix 3.7). Consequently, the fixed factor of sex was removed and a one-way MANOVA performed using palatal stage as the only fixed factor. One-way MANOVA revealed palatal stage demonstrated convincing statistical significance [$F(20, 102) = 1.709$, $p = .044$; Wilks' $\Lambda = .561$; partial $\eta^2 = .251$] (Appendix 3.9a and b).

There was convincing evidence that there is significant difference in Dif_C_Pu and Dif_ANG_16 between subjects with different T1 palatal stages [$F(2, 60) = 3.919$, p =0.025; partial $\eta^2 = .116$] and [$F(2, 60) = 5.588$, p =0.006; partial $\eta^2 = .157$] respectively (Appendix 3.10). Multiple comparisons Bonferonni corrected analyses demonstrated that at post-expansion there was a mean increase in intercanine width measured at the pulp in (Dif_C_Pu) between subjects in palatal stage 2 and 3 of 1.7022mm (95% CI, -0.021 to 3.43) which was suggestive of statistical significance (p=0.054, Appendix 3.11). Additionally, there was a mean increase the angulation of the right maxillary first molar (Dif_ANG_16) post expansion between subjects in palatal stage 1 and 3 of 5.635 degrees (95% CI, 1.43 to 9.84) which demonstrated statistical significance (p=0.005), however no other group differences were statistically significant (Appendix 3.11).

Although there was a clinically significant effect of T1 palatal stage on intercanine width measured at the pulp from stage 2 to stage 3 and difference in angulation of 1.6 from palatal stage 1 to 3, there were 8 other dental widths, angles and skeletal widths that were unaffected by the patients palatal stage at T1. Consequently, it can be inferred that palatal stage at T1 truly had little to no effect on these important and clinically relevant angulations and widths during the expansion treatment and we accept our null hypothesis, rejecting the alternative hypothesis.

A review of the basis for the theoretical development of the novel palatal suture classification[1] due to the highly unsupportive results of this study was performed. Three histological studies were utilized for its development (Chapter 3, Table 3.5) and multiple concerns were identified in using these studies to develop the novel classification. The use of specimens of species other than human autopsy material, evaluation of photomicrographs in the coronal plane rather than axial plane, utilizing an improper age of human specimens that would not include subjects with patent sutures, and correlating the morphology of highly magnified histological photomicrographs to eye level CBCT generated axial cross-sections presented great concerns. Consequently, utilization of

these studies [5-7] to draw inferences into the radiological morphology of the MPS was deemed scientifically flawed, fundamentally impossible, and not plausible. It is advisable that until further research supportive of this classification[1] is developed and rigorously tested, that clinicians do not consider this proposed classification[1] as being factual, and halt employing its use to drive clinical decision making which will have real-world patient implications and outcomes. In response to the tertiary research question, major limitations included lack of a gold standard and poor resolution and clarity to view axial generated slices. Ultimately however; improvement of the proposed classification[1], is not possible. A comprehensive overhaul of its scientific basis would be required and rigorously tested against a gold standard.

4.2 Literature Cited:

1. Angelieri, F., et al., *Midpalatal suture maturation: classification method for individual assessment before rapid maxillary expansion.* Am J Orthod Dentofacial Orthop, 2013. **144**(5): p. 759-69.
2. Cao, L., et al., *Effect of maxillary incisor labiolingual inclination and anteroposterior position on smiling profile esthetics.* Angle Orthod, 2011. **81**(1): p. 121-29.
3. Chirivella, P., et al., *Comparison of the effect of labiolingual inclination and anteroposterior position of maxillary incisors on esthetic profile in three different facial patterns.* J Orthod Sci, 2017. **6**(1): p. 1-10.
4. O'Higgins, E.A., R.H. Kirschen, and R.T. Lee, *The influence of maxillary incisor inclination on arch length.* Br J Orthod, 1999. **26**(2): p. 97-102.

5. Persson, M., B.C. Magnusson, and B. Thilander, *Sutural closure in rabbit and man: a morphological and histochemical study.* J Anat, 1978. **125**(Pt 2): p. 313-21.

6. Sun, Z., E. Lee, and S.W. Herring, *Cranial Sutures and Bones: Growth and Fusion in Relation to Masticatory Strain.* Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology, 2004. **276**(2): p. 150-161.

7. Cohen, M.M., Jr., *Sutural biology and the correlates of craniosynostosis.* Am J Med Genet, 1993. **47**(5): p. 581-616.

Bibliography:

1. Angelieri, F., et al., *Midpalatal suture maturation: classification method for individual assessment before rapid maxillary expansion.* Am J Orthod Dentofacial Orthop, 2013. **144**(5): p. 759-69.
2. Baccetti, T., et al., *Treatment timing for rapid maxillary expansion.* Angle Orthod, 2001. **71**(5): p. 343-50.
3. Bishara, S.E. and R.N. Staley, *Maxillary expansion: clinical implications.* Am J Orthod Dentofacial Orthop, 1987. **91**(1): p. 3-14.
4. McNamara, J.A., *Maxillary transverse deficiency.* Am J Orthod Dentofacial Orthop, 2000. **117**(5): p. 567-70.
5. Isiksal, E., S. Hazar, and S. Akyalcin, *Smile esthetics: perception and comparison of treated and untreated smiles.* Am J Orthod Dentofacial Orthop, 2006. **129**(1): p. 8-16.
6. Lin, L., et al., *Tooth-borne vs bone-borne rapid maxillary expanders in late adolescence.* Angle Orthod, 2015. **85**(2): p. 253-62.
7. Melsen, B., *Palatal growth studied on human autopsy material. A histologic microradiographic study.* American Journal of Orthodontics, 1975. **68**(1): p. 42-54.
8. Persson, M. and B. Thilander, *Palatal suture closure in man from 15 to 35 years of age.* American Journal of Orthodontics, 1977. **72**(1): p. 42-52.
9. Sumer, A.P., et al., *Ultrasonography in the evaluation of midpalatal suture in surgically assisted rapid maxillary expansion.* Journal of Craniofacial Surgery, 2012. **23**(5): p. 1375-7.

10. Korbmacher, H., et al., *Age-dependent three-dimensional microcomputed tomography analysis of the human midpalatal suture.* Journal of Orofacial Orthopedics, 2007. **68**(5): p. 364-76.
11. Franchi, L., et al., *Modifications of midpalatal sutural density induced by rapid maxillary expansion: A low-dose computed-tomography evaluation.* American Journal of Orthodontics & Dentofacial Orthopedics, 2010. **137**(4): p. 486-8; discussion 12A-13A.
12. Kwak, K.H., et al., *Quantitative evaluation of midpalatal suture maturation via fractal analysis.* Korean Journal of Orthodontics, 2016. **46**(5): p. 323-330.
13. Persson, M., B.C. Magnusson, and B. Thilander, *Sutural closure in rabbit and man: a morphological and histochemical study.* J Anat, 1978. **125**(Pt 2): p. 313-21.
14. Cohen, M.M., Jr., *Sutural biology and the correlates of craniosynostosis.* Am J Med Genet, 1993. **47**(5): p. 581-616.
15. Sun, Z., E. Lee, and S.W. Herring, *Cranial Sutures and Bones: Growth and Fusion in Relation to Masticatory Strain.* Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology, 2004. **276**(2): p. 150-161.
16. Yu, J.C., et al., *A fractal analysis of human cranial sutures.* Cleft Palate Craniofac J, 2003. **40**(4): p. 409-15.
17. Shapurian, T., et al., *Quantitative evaluation of bone density using the Hounsfield index.* Int J Oral Maxillofac Implants, 2006. **21**(2): p. 290-7.

18. Duckmanton, N.A., et al., *Imaging for predictable maxillary implants.* Int J Prosthodont, 1994. **7**(1): p. 77-80.

19. Norton, M.R. and C. Gamble, *Bone classification: an objective scale of bone density using the computerized tomography scan.* Clin Oral Implants Res, 2001. **12**(1): p. 79-84.

20. Cleall, J.F., et al., *EXPANSION OF THE MIDPALATAL SUTURE IN THE MONKEY.* Angle Orthod, 1965. **35**: p. 23-35.

21. Ballanti, F., et al., *Immediate and post-retention effects of rapid maxillary expansion investigated by computed tomography in growing patients.* Angle Orthod, 2009. **79**(1): p. 24-9.

22. Perilli, E., I.H. Parkinson, and K.J. Reynolds, *Micro-CT examination of human bone: from biopsies towards the entire organ.* Ann Ist Super Sanita, 2012. **48**(1): p. 75-82.

23. Thurmuller, P., et al., *Use of ultrasound to assess healing of a mandibular distraction wound.* J Oral Maxillofac Surg, 2002. **60**(9): p. 1038-44.

24. Hughes, C.W., et al., *Ultrasound monitoring of distraction osteogenesis.* Br J Oral Maxillofac Surg, 2003. **41**(4): p. 256-8.

25. Bruno, C., et al., *Gray-scale ultrasonography in the evaluation of bone callus in distraction osteogenesis of the mandible: initial findings.* Eur Radiol, 2008. **18**(5): p. 1012-7.

26. Sanchez, I. and G. Uzcategui, *Fractals in dentistry.* J Dent, 2011. **39**(4): p. 273-92.

27. Fricke-Zech, S., et al., *Measurement of the midpalatal suture width.* Angle Orthodontist, 2012. **82**(1): p. 145-50.

28. Gao, Q.W., et al., *[An ultrastructure study on the palatomaxillary suture of dog expanded by NiTi-SMA].* Zhonghua Zheng Xing Wai Ke Za Zhi, 2009. **25**(4): p. 277-9.

29. Hahn, W., et al., *Imaging of the midpalatal suture in a porcine model: flat-panel volume computed tomography compared with multislice computed tomography.* Oral Surgery Oral Medicine Oral Pathology Oral Radiology & Endodontics, 2009. **108**(3): p. 443-9.

30. Beauthier, J.P., et al., *Palatine sutures as age indicator: a controlled study in the elderly.* Journal of Forensic Sciences, 2010. **55**(1): p. 153-8.

31. Cheung, L.K. and Q. Zhang, *Radiologic characterization of new bone generated from distraction after maxillary bone transport.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontics, 2003. **96**(2): p. 234-242.

32. de Melo Mde, F., et al., *Digital radiographic evaluation of the midpalatal suture in patients submitted to rapid maxillary expansion.* Indian Journal of Dental Research, 2013. **24**(1): p. 76-80.

33. Gurgel Jde, A., M.F. Malmstrom, and C.R. Pinzan-Vercelino, *Ossification of the midpalatal suture after surgically assisted rapid maxillary expansion.* European Journal of Orthodontics, 2012. **34**(1): p. 39-43.

34. Kjaer, I., *Human prenatal palatal shelf elevation related to craniofacial skeletal maturation.* European Journal of Orthodontics, 1992. **14**(1): p. 26-30.

35. Knaup, B., F. Yildizhan, and H. Wehrbein, *Age-related changes in the midpalatal suture. A histomorphometric study.* Journal of Orofacial Orthopedics, 2004. **65**(6): p. 467-74.

36. Lee, S.K., et al., *Prenatal growth pattern of the human maxilla.* Acta Anatomica, 1992. **145**(1): p. 1-10.

37. Leonardi, R., A. Cutrera, and E. Barbato, *Rapid maxillary expansion affects the spheno-occipital synchondrosis in youngsters. A study with low-dose computed tomography.* Angle Orthodontist, 2010. **80**(1): p. 106-10.

38. Leonardi, R., et al., *Early post-treatment changes of circumaxillary sutures in young patients treated with rapid maxillary expansion.* Angle Orthodontist, 2011. **81**(1): p. 36-41.

39. Kjaer, I., *Prenatal skeletal maturation of the human maxilla.* Journal of Craniofacial Genetics & Developmental Biology, 1989. **9**(3): p. 257-64.

40. Agrawal, D., P. Steinbok, and D.D. Cochrane, *Reformation of the sagittal suture following surgery for isolated sagittal craniosynostosis.* Journal of Neurosurgery, 2006. **105 PEDIATRICS**(SUPPL. 2): p. 115-117.

41. Bradley, J.P., et al., *Studies in cranial suture biology: IV. Temporal sequence of posterior frontal cranial suture fusion in the mouse.* Plastic and Reconstructive Surgery, 1996. **98**(6): p. 1039-1345.

42. Captier, G., et al., *Prenatal organization and morphogenesis of the sphenofrontal suture in humans.* Cells Tissues Organs, 2003. **175**(2): p. 98-104.

43. Lauridsen, H., et al., *Histological investigation of the palatine bone in prenatal trisomy 21.* Cleft Palate-Craniofacial Journal, 2001. **38**(5): p. 492-7.

44. Corega, C., et al., *Three-dimensional cranial suture morphology analysis.* Romanian Journal of Morphology and Embryology, 2010. **51**(1): p. 123-127.

45. Corega, C., et al., *Cranial sutures and diploae morphology.* Romanian Journal of Morphology and Embryology, 2013. **54**(4): p. 1157-1160.

46. Bjork, A. and V. Skieller, *[Growth and development of the maxillary complex].* Informationen aus Orthodontie und Kieferorthopadie, 1984. **16**(1): p. 9-52.

47. De Araujo Gurgel, J., M.F.V. Malmstrom, and C.R.M. Pinzan-Vercelino, *Ossification of the midpalatal suture after surgically assisted rapid maxillary expansion.* European Journal of Orthodontics, 2012. **34**(1): p. 39-43.

48. Sannomiya, E.K., et al., *Evaluation of optical density of the midpalatal suture 3 months after surgically assisted rapid maxillary expansion.* Dento-Maxillo-Facial Radiology, 2007. **36**(2): p. 97-101.

49. Takenouchi, H., et al., *Longitudinal quantitative evaluation of the mid-palatal suture after rapid expansion using in vivo micro-CT.* Archives of Oral Biology, 2014. **59**(4): p. 414-23.

50. Acar, Y.B., M. Motro, and A.N. Erverdi, *Hounsfield Units: a new indicator showing maxillary resistance in rapid maxillary expansion cases?* Angle Orthodontist, 2015. **85**(1): p. 109-16.

51. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.

52. McHugh, M.L., *Interrater reliability: the kappa statistic.* Biochem Med (Zagreb), 2012. **22**(3): p. 276-82.

53. Lagravere, M.O., et al., *Transverse, vertical, and anteroposterior changes from bone-anchored maxillary expansion vs traditional rapid maxillary expansion: a randomized clinical trial.* Am J Orthod Dentofacial Orthop, 2010. **137**(3): p. 304.e1-12; discussion 304-5.

54. Viera, A.J. and J.M. Garrett, *Understanding interobserver agreement: the kappa statistic.* Fam Med, 2005. **37**(5): p. 360-3.

55. Maclure, M. and W.C. Willett, *Misinterpretation and misuse of the kappa statistic.* Am J Epidemiol, 1987. **126**(2): p. 161-9.

56. Cooper, D., et al., *Effect of voxel size on 3D micro-CT analysis of cortical bone porosity.* Calcif Tissue Int, 2007. **80**(3): p. 211-9.

57. Giavarina, D., *Understanding Bland Altman analysis.* Biochem Med (Zagreb), 2015. **25**(2): p. 141-51.

58. Button, K.S., et al., *Power failure: why small sample size undermines the reliability of neuroscience.* Nat Rev Neurosci, 2013. **14**(5): p. 365-76.

59. Revelo, B. and L.S. Fishman, *Maturational evaluation of ossification of the midpalatal suture.* American Journal of Orthodontics & Dentofacial Orthopedics, 1994. **105**(3): p. 288-92.

60. Kosowski, T.R., et al., *Cleft palate.* Semin Plast Surg, 2012. **26**(4): p. 164-9.

61. Melsen, B. and F. Melsen, *The postnatal development of the palatomaxillary region studied on human autopsy material.* Am J Orthod, 1982. **82**(4): p. 329-42.

62. Ghoneima, A., et al., *Effects of rapid maxillary expansion on the cranial and circummaxillary sutures.* American Journal of Orthodontics & Dentofacial Orthopedics, 2011. **140**(4): p. 510-9.

63. Wertz, R. and M. Dreskin, *Midpalatal suture opening: a normative study.* Am J Orthod, 1977. **71**(4): p. 367-81.

64. Liou, E., *Eric Liou.* Revista Dental Press de Ortodontia e Ortopedia Facial, 2009. **14**: p. 27-37.

65. Bjork, A., *Sutural growth of the upper face studied by the implant method.* Acta Odontol Scand, 1966. **24**(2): p. 109-27.

66. Bjork, A. and V. Skieller, *Growth of the maxilla in three dimensions as revealed radiographically by the implant method.* Br J Orthod, 1977. **4**(2): p. 53-64.

67. Primozic, J., et al., *Three-dimensional longitudinal evaluation of palatal vault changes in growing subjects.* Angle Orthod, 2012. **82**(4): p. 632-6.

68. Lou, L., et al., *Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2007. **104**(3): p. 402-11.

69. Lagravere, M.O., et al., *Cranial base foramen location accuracy and reliability in cone-beam computerized tomography.* Am J Orthod Dentofacial Orthop, 2011. **139**(3): p. e203-10.

70. Ludlow, J.B., et al., *Precision of cephalometric landmark identification: cone-beam computed tomography vs conventional cephalometric views.* Am J Orthod Dentofacial Orthop, 2009. **136**(3): p. 312.e1-10; discussion 312-3.

71. Ludlow, J.B., et al., *Dosimetry of 3 CBCT devices for oral and maxillofacial radiology: CB Mercuray, NewTom 3G and i-CAT.* Dentomaxillofac Radiol, 2006. **35**(4): p. 219-26.

72. Mah, J. and D. Hatcher, *Three-dimensional craniofacial imaging.* Am J Orthod Dentofacial Orthop, 2004. **126**(3): p. 308-9.

73. Damstra, J., et al., *Accuracy of linear measurements from cone-beam computed tomography-derived surface models of different voxel sizes.* Am J Orthod Dentofacial Orthop, 2010. **137**(1): p. 16.e1-6; discussion 16-7.

74. Broadbent, B.H., *A NEW X-RAY TECHNIQUE and ITS APPLICATION TO ORTHODONTIA.* The Angle Orthodontist, 1931. **1**(2): p. 45-66.

75. Lagravere, M.O., P.W. Major, and J. Carey, *Sensitivity analysis for plane orientation in three-dimensional cephalometric analysis based on superimposition of serial cone beam computed tomography images.* Dentomaxillofac Radiol, 2010. **39**(7): p. 400-8.

76. de Oliveira, A.E., et al., *Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2009. **107**(2): p. 256-65.

77. Cao, L., et al., *Effect of maxillary incisor labiolingual inclination and anteroposterior position on smiling profile esthetics.* Angle Orthod, 2011. **81**(1): p. 121-29.

78. Chirivella, P., et al., *Comparison of the effect of labiolingual inclination and anteroposterior position of maxillary incisors on esthetic profile in three different facial patterns.* J Orthod Sci, 2017. **6**(1): p. 1-10.

79. O'Higgins, E.A., R.H. Kirschen, and R.T. Lee, *The influence of maxillary incisor inclination on arch length.* Br J Orthod, 1999. **26**(2): p. 97-102.

80. Flores-Mir, C., et al., *Correlation of Skeletal Maturation Stages Determined by Cervical Vertebrae and Hand-wrist Evaluations.* The Angle Orthodontist, 2006. **76**(1): p. 1-5.

81. Bonfim, M.A., et al., *Cervical vertebrae maturation index estimates on cone beam CT: 3D reconstructions vs sagittal sections.* Dentomaxillofac Radiol, 2016. **45**(1): p. 20150162.

82. Aktan, A.M., et al., *Effects of voxel size and resolution on the accuracy of endodontic length measurement using cone beam computed tomography.* Ann Anat, 2016. **208**: p. 96-102.

83. Pauwels, R., et al., *Effect of exposure parameters and voxel size on bone structure analysis in CBCT.* Dentomaxillofac Radiol, 2015. **44**(8): p. 20150078.

84. Scarfe, W.C. and A.G. Farman, *What is cone-beam CT and how does it work?* Dent Clin North Am, 2008. **52**(4): p. 707-30, v.