7-8-2016

The Influence of the Quality of Radiographs and Training on the Reproducibility of the Cervical Vertebrae Maturation Method

Anfal Khajah
khajah@uchc.edu

The Influence of the Quality of Radiographs and Training on the Reproducibility of the Cervical Vertebrae Maturation Method

Anfal Khajah

B.M.Sc. (DENT), Bachelor of Medical Sciences in Dentistry, Kuwait University, 2006

B.D.M., Bachelor of Dental Medicine, Kuwait University- Faculty of Dentistry, 2009

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Dental Science

At the

University of Connecticut

2016

APPROVAL PAGE

Masters of Dental Science Thesis

The Influence of the Quality of Radiographs and Training on the Reproducibility of the Cervical Vertebrae Maturation Method

Presented by

Anfal Khajah, B.M.Sc. (DENT), B.D.M.

Major Advisor ______________________________
Flavio Uribe, D.D.S., M.D.S.

Associate Advisor ______________________________
Ravindra Nanda, B.D.S., M.D.S., Ph.D.

Associate Advisor ______________________________
Aditya Tadinada, B.D.S., M.Dent.Sc.

Program Director ______________________________
Flavio Uribe, D.D.S., M.D.S.

University of Connecticut
2016

Acknowledgements

I would like to thank my advisors and committee members. In particular, I would like to express my gratitude to my major advisor Dr. Uribe for being a great mentor and providing me with guidance and encouragement during the entire project. I am thankful to Dr. Tadinada for his help and expertise in radiology and CBCT imaging. I am especially grateful to Dr. Nanda for all the support over the past three years. I am also very thankful to all faculty members and residents who participated in this research. Without them this research could not have been completed.

I am thankful to Dr. Derek Sanders for providing us with the CBCTs from his private practice in Florida. I am also thankful to the American Association of Orthodontists Foundation for providing us with the lateral cephalograms of Iowa Facial Growth Study that were selected from their website to be used in this study.

Finally and most importantly, I am deeply grateful to my family for their endless inspiration, support and love. This accomplishment would not have been possible without them.

Table of Contents

Abstract

Aims & Objectives: The objective of this study was to assess the reproducibility of cervical vertebral maturation method based on training and quality of radiographic images.

Materials & Methods: Ten evaluators (5 orthodontic residents and 5 orthodontists) were randomly divided into two groups: training group and non-training group. All the participants evaluated 80 radiographic images previously acquired in four different formats: 2 dimensional digital images, 2 dimensional digitized hard copy images, 2-dimensional digital images reconstructed from a 3- dimensional radiograph, and 3-dimensional images (CBCT). They had to evaluate the morphology of the cervical vertebrae and to determine a CVM stage of each radiographic image.

Results: The overall interobserver agreement level between the ten evaluators in cervical vertebrae maturation staging was 65.25%. The interobserver agreement between the trained evaluators was higher than the non-trained evaluators. Overall, the kappa values for assessing the curvature of the vertebrae were highly variable ranging from no agreement to perfect agreement. The interobserver agreement for determining both the shape of the body of the vertebrae and CVM staging was fair among all the different x-rays groups. The majority (81.5%) of the disagreements were one cervical stage apart. 15.6% of the disagreements were two cervical stages apart.

Conclusion: Our findings suggest that training seems to have some influence on the reproducibility of the cervical vertebrae maturation method. The quality of radiographs does not seem to have a major influence on the reproducibility of the method.

Introduction

Background:

Growth prediction of the human face has always been an area of interest in orthodontic research. It is thought that the ability to predict the amount, timing and direction of facial growth will enable more efficient treatment planning, selection of more appropriate treatment regimens and optimization of treatment timing. Changes in the human face due to growth are important during orthodontic treatment where a skeletal imbalance is present; a favorable growth pattern may facilitate treatment, whereas, an adverse growth pattern may make treatment more difficult or even impossible. Such growth patterns, if anticipated, may result in altered treatment plans. Patterns of facial growth may have an impact on the stability of completed orthodontic treatment. Thus, knowledge of future growth may influence the selection of post-orthodontic treatment retention protocols. An indication of the maturation level of an individual is necessary to predict future growth.

A number of different methods are available to evaluate the maturation stage of an individual such as chronologic age, height, weight, sexual maturation characteristics, dental development, and skeletal development[1]. According to Fishman in 1979, skeletal age rather than chronological age would be a more accurate parameter on which to base facial growth prediction[2]. Frontal sinus morphology, hand wrist radiograph analysis, and the morphology of the cervical vertebrae are three major areas that have been studied to assess skeletal age and skeletal maturation. Skeletal maturation staging from radiographic analysis is a widely used approach for predicting the timing of pubertal growth and for estimating growth velocity and the proportion of growth remaining.

The gold standard method of evaluating skeletal maturity has been a hand-wrist radiograph[1]. There are two general approaches for assessment of the hand-wrist radiograph[3]. The first method consists is a comparative measure to reference radiographs as described by Greulich and Pyle. Greulich and Pyle method uses an atlas of typical radiographs of hands taken at 6 monthly intervals of chronological age. An average of the development of all hand bones is used to deliver an approximate skeletal age. The second method of assessment of the hand-wrist radiograph uses specific indicators to relate skeletal maturation to the pubertal growth curve. A number of different indicators have been used, including the onset of calcification of the sesamoid bone, and developmental staging of the middle phalanx of the third finger. Changes in the morphology of the cervical vertebrae were initially studied in 1972 by Lamparski who was able to correlate these morphological changes to the timing of peak height velocity and stages of hand-wrist ossification[4]. His method analyzed size and shape changes in the bodies of five cervical vertebrae (from the second one to the sixth) [5]. Further refinements to the technique were undertaken such that only the upper four cervical vertebrae were used for the analysis. This was done by Hassel and Farman who reviewed lateral cephalometric and left hand wrist radiographs from the Bolton Brush Growth Study to develop an index based on the lateral profiles of the second, third, and fourth cervical vertebrae[5, 6]. Their sample consisted of 11 groups of 10 males and 10 females (220 subjects) aged from 8- 18 years. Six categories of cervical vertebrae skeletal maturation were defined and the observations of each category were described as follows and as shown in **Figure 1**:

- **Category 1** is *initiation* stage: The inferior borders of C2, C3, and C4 vertebral body are flat and the superior vertebral borders are tapered posterior to anterior. Very significant amount of adolescent growth is expected (85- 100% of adolescent growth is remaining).

- **Category 2** is *acceleration* stage: Concavities are developing in lower borders of C2 and C3 but lower border of C4 vertebral body is flat. C3 and C4 are more rectangular in shape. Significant amount of adolescent growth is expected (65- 85% of adolescent growth is remaining).
- **Category 3** is *transition* stage: Distinct concavities in lower borders of C2 and C3 are seen. Concavity developing in lower border of body of C4. C3 and C4 are rectangular in shape. Moderate amount of adolescent growth is expected (25- 65% of adolescent growth is remaining).
- **Category 4** is *deceleration* stage: Distinct concavities in lower borders of C2, C3, and C4 are seen. C3 and C4 are nearly square in shape. Small amount of adolescent growth is expected (10- 25% of adolescent growth is remaining).
- **Category 5** is *maturation* stage: Accentuated concavities of inferior vertebral body borders of C2, C3, and C4 are seen. C3 and C4 are square in shape. Insignificant amount of adolescent growth is expected (5- 10% of adolescent growth is remaining).
- **Category 6** is *completion* stage: Deep concavities are present at the inferior vertebral body borders of C2, C3, and C4. C3 and C4 heights are greater than widths. Adolescent growth is completed.

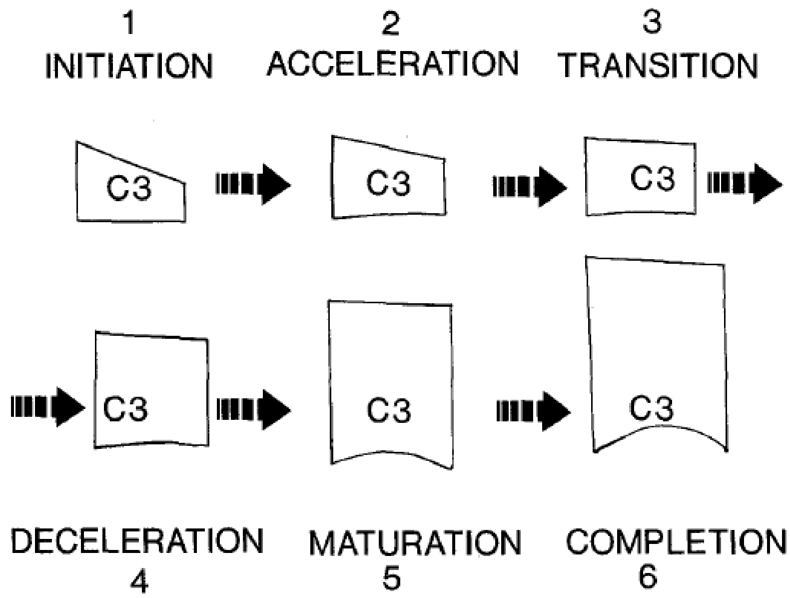


Figure 1: Cervical vertebrae maturation indicators using C3 as a guide

In 2005, Baccetti et al introduced a new clinically improved cervical vertebral maturation method that is comprised of six maturational stages (cervical stage 1 through cervical stage 6)[7]. The features of each cervical stage were summarized and correlated with the peak in mandibular growth as follows and as illustrated by the schematic representation in **Figure 2** by Baccetti et al[7]:

- ***Cervical stage 1 (CS1):*** The lower borders of all the three vertebrae (C2-C4) are flat. The bodies of both C3 and C4 are trapezoid in shape (the superior border of the vertebral body is tapered from posterior to anterior). The peak in mandibular growth will occur on average 2 years after this stage.
- ***Cervical stage 2 (CS2):*** A concavity is present at the lower border of C2 (in four of five cases, with the remaining subjects still showing a cervical stage 1). The bodies of both C3 and C4 are still trapezoid in shape. The peak in mandibular growth will occur on average 1 year after this stage.

- ***Cervical stage 3 (CS3):*** Concavities at the lower borders of both C2 and C3 are present. The bodies of C3 and C4 may be either trapezoid or rectangular horizontal in shape. The peak in mandibular growth will occur during the year after this stage.
- ***Cervical stage 4 (CS4):*** Concavities at the lower borders of C2, C3, and C4 are present. The bodies of both C3 and C4 are rectangular horizontal in shape. The peak in mandibular growth has occurred within 1 or 2 years before this stage.
- ***Cervical stage 5 (CS5):*** The concavities at the lower borders of C2, C3, and C4 are still present. At least one of the bodies of C3 and C4 is squared in shape. If not squared, the body of the other cervical vertebra still is rectangular horizontal. The peak in mandibular growth has ended at least 1 year before this stage.
- ***Cervical stage 6 (CS6):*** The concavities at the lower borders of C2, C3, and C4 are still evident. At least one of the bodies of C3 and C4 is rectangular vertical in shape. If not rectangular vertical, the body of the other cervical vertebra is squared. The peak in mandibular growth has ended at least 2 years before this stage.

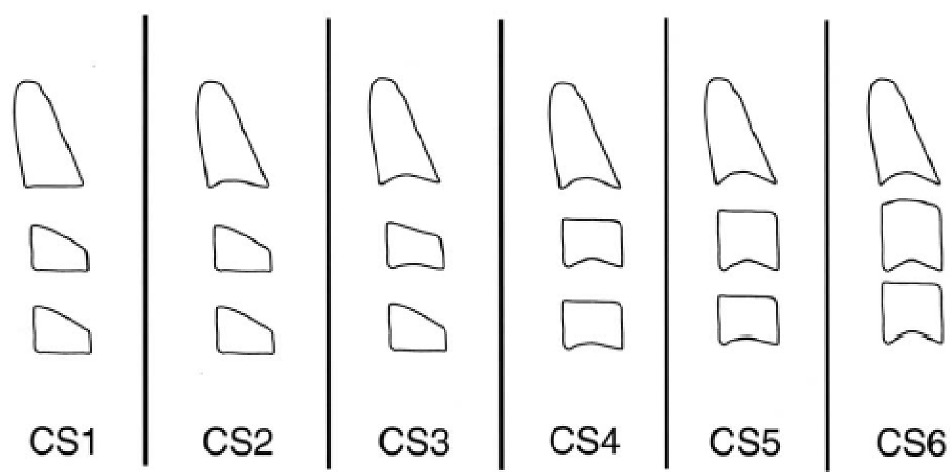


Figure 2: Schematic representation of the stages of cervical vertebrae

CS1 and CS2 are pre-peak stages; the peak in mandibular growth occurs between CS3 and CS4. CS6 is recorded at least 2 years after the peak and active growth is virtually completed when this stage is attained[7].

The cervical vertebral maturation (CVM) method has proved to be effective to assess the adolescent growth peak both in body height and mandibular size[5]. Good correlation between cervical maturation and skeletal age is reported, and this technique may be useful clinically[5, 8] as it enables the clinician to identify an optimal treatment timing of different dentoskeletal malocclusions in all three planes of space[7]. Several clinical studies have shown that the greatest response to functional jaw orthopedics tends to occur during the circumpubertal growth period. Thus, the use of a reliable biologic indicator to detect the pubertal spurt in mandibular growth represents a crucial diagnostic tool for a rational treatment planning in Class II subjects with mandibular deficiencies. It has been reported that Class II treatment is most effective when it includes the peak in mandibular growth and that when the intervention to treatment includes CS3-CS4 interval (growth spurt), the net growth of the mandible in treated samples versus untreated controls ranges from 2.4 mm to 4.7 mm[7]. Furthermore, methods such as CVM can be useful to detect periods of reduced growth rate in the timing of orthognathic surgery or for the long-term evaluation of treatment outcomes[5]. In order for a biologic indicator of skeletal maturity to be reliable and reproducible, interpretation of its data has to be consistent, and the inter-examiner error in the appraisal of the defined stage should be as low as possible[7].

In the literature, majority of the studies have reported that the reproducibility of the CVM method exceeds 90%. However, Kucukkeles et al. reported a reproducibility of 45% and 65%[9, 22]. Similarly, Gabriel et al. evaluated the reproducibility of the CVM method from ten private practice orthodontists trained in the CVM method. Specifically, the authors evaluated 30

individual and 30 pairs of cephalometric radiographs in two sessions to determine the CVM stage. The interobserver agreement was below 50% and the intraobserver agreement was 62%. They concluded that CVM method cannot be recommended as a strict clinical guideline for the timing of orthodontic treatment as has been suggested in the literature. Another recent study by the same group aimed to expand on the reproducibility of the CVM method by determining the reasons for the poor reproducibility. The study aim was to determine which of the individual CVM vertebral patterns could be classified reliably and which could not[10]. The same methods were used except that the evaluation of the 30 cephalometric radiographs was done using questions based on the CVM method. Interobserver agreement was high for assessment of the lower borders of C2, C3, and C4 being either flat or curved, but interobserver agreement was low for assessment of the shape of vertebral bodies of C3 and C4. According to the authors, this led to the overall poor reproducibility of the CVM method and thus the authors stated that the use of this method as a strict clinical guideline for the timing of orthodontic treatment was not supported. A more recent study evaluated the accuracy and repeatability of the CVM method[11] and the influence of operator training. Ten operators underwent training sessions in visual assessment of CVM staging using a series of cases analyzed cephalometrically. The operators were asked to assign CVM stage in a different set of cases two times in two sessions four weeks apart. The outcomes of these sessions were compared with a reference standard for diagnostic accuracy. The overall agreement with the reference standard was about 68% for both sessions and 76.9% for intrarater repeatability. It was concluded that visual assessment of the CVM stages is accurate and repeatable to a satisfactory level and that disagreement is generally limited to one stage and is mostly seen in stages 4 and 5.

For almost a century, two-dimensional (2D) radiographic imaging and cephalometry has been used in orthodontics for diagnosis, treatment planning, evaluation of growth and development, and assessment of treatment progress and outcomes. However, 2D imaging has its limitations. As a result of the recent advances in the field of radiology, the availability of multi-slice computed tomography (CT) has increased and cone beam computed tomography (CBCT) has become more popular as an investigating tool for orthodontic patients[13] and therefore visualization of the cervical vertebrae in three dimensions is now feasible. Many studies in the literature have investigated the use of CBCT for skeletal maturity assessment and if the cervical vertebrae maturation method can be applied with CBCT as it is applied with lateral cephalograms.

A study evaluated the application of the cervical vertebrae maturation method in cone-beam computer tomographic (CBCT) images to bring forth assessment of skeletal maturation in three dimensions[12]. Ninety-eight lateral cephalometric radiographs and CBCT scans were collected from orthodontic patients between 11 to 17 years of age over an 18-month period. CBCT scans were examined in seven sagittal slices based on cervical vertebral maturation staging (CVMS) method proposed by Baccetti et al in 2005. Collected CVMS values were compared with those from corresponding lateral cephalometric radiograph. CVMS measured from CBCT and lateral cephalometric radiographs were the same on average. However, they were not consistent with each other and scored interclass correlation coefficient of 0.155 in a validity test. Interoperator reliability was weak (0.581). The authors of this study concluded that adaptation of cervical vertebrae maturation staging in CBCT requires further clarifications or modifications to become consistent with lateral cephalometric examinations and to become a reliable method. They also

concluded that as an alternative, a completely new method may be developed consisting of maturational indicators or landmarks unique to CBCT imaging.

Another study compared three methods of assessing skeletal maturity[13]. The first method used skeletal maturity indicators from hand wrist radiographs. The second method used cervical vertebrae maturity index (CVMI) from lateral cephalograms. In the third method, the cervical vertebrae maturity index was assessed from sagittal sections of CBCT.

The study material consisted of 100 subjects (51 female, 49 male) who had CBCT, lateral cephalograms, and hand wrist radiographs. The age range of the study group was 3-35 years. The mean chronologic age of the subjects was 11 years with a standard deviation of 5.57. The cervical vertebrae maturation of the sample was evaluated by Hassel and Farman's method. In this study, a very good correlation was found between the CBCT- CVMI and cephalograms-CVMI as well as between CBCT- CVMI and skeletal maturity indicators from hand wrist radiography. This study also proved that chronological age is a poor indicator of maturity. The results of this study suggest the use of CBCT to assess skeletal maturity whenever CBCT is used as a diagnostic tool for orthodontic patients.

Hypothesis/ Goal/ Specific Objectives

Hypothesis:

- The quality of radiographic image has an influence on the reproducibility of the CVM method.
- Training has an influence on the reproducibility of the CVM method.

Goal:

- To assess the reproducibility of cervical vertebral maturation method based on training and quality of radiographic images.

Specific Objectives:

- To assess the influence of the quality of the radiographic image on the reproducibility of the CVM method.
- To assess the influence of training on the reproducibility of the CVM method.

Study Design/ Procedures/ Methods

This study involved an evaluation of the reliability of determining the skeletal maturation status of the cervical vertebrae from radiographic images of different quality, previously acquired in different formats: 1) hard copy in 2 dimensions, 2) digital image in 2 dimensions, 3) 2-dimensional digital image acquired from a 3- dimensional radiograph 4) a 3-dimensional radiograph (CBCT). Additionally, the influence of training on the evaluators/ subjects in determining the cervical maturation stage was assessed. All the radiographs (total of 80) were evaluated by ten evaluators/ subjects (residents/orthodontists). Five of the evaluators were orthodontic residents and the other five were orthodontic faculty members from the Division of Orthodontics at the School of Dental Medicine in the University of Connecticut. All the evaluators completed a questionnaire which contains questions about the shape of the vertebrae in each radiograph. In order to assess the reliability of the cervical vertebrae maturation method,

the results obtained were compared between the different evaluators. The radiographs evaluated by the evaluators were de-identified. All radiographs (total of 80) were cropped to include only cervical vertebra C2 to C4 to eliminate any additional information such as stage of the development of the dentition that might generate bias during the evaluation. The study was approved by the IRB of the UCONN School of Dental Medicine (IRB# 16-008-1).

The radiographs were retrieved from a total of 80 records, which were obtained from three different sources. Forty records (lateral cephalometric radiographs) were selected from the longitudinal growth records from the Iowa Facial Growth Study through the American Association of Orthodontists Foundation website. Twenty records (lateral cephalometric radiographs) were selected from the electronic health record (HER) system "axium" of patients treated at UCONN School of Dental Medicine Orthodontic Clinic. Another twenty records (CBCT scans) were obtained from a private orthodontic office in Florida. The CBCT scans were viewed as a 3- dimensional as well as a 2-dimensional image (one CBCT record provided two different image modalities, one in two dimensions and one in three dimensions). A CBCT reconstruction software (Invivo 5) was used to acquire a two dimensional radiograph from the three dimensional scan. While viewing the three dimensional scans, evaluators had full control of the scan volume to scroll through all the three orthogonal planes (Axial, Sagittal and Coronal). The evaluators also had the ability to control the histogram and could make any changes to the contrast and density to help them best evaluate the image.

The influence of training in the reliability of assessment of the shape and stage of the cervical vertebrae was assessed by randomly dividing both groups of evaluators (5 orthodontists/ 5 residents) into 2 groups. The orthodontists group had 2 evaluators receiving training and the

other 3 did not receive any training. The resident group had 3 evaluators receiving training and the other 2 did not receive training.

The training session was before the main evaluation session (the evaluation session is described in page 15). The training session included a detailed explanation of the rules to be followed for using cervical vertebrae maturation method and assigning CVM stages. Twenty lateral cephalograms already selected from the longitudinal growth records from the Iowa Facial Growth Study through the American Association of Orthodontists Foundation website were used for this training session. Those twenty lateral cephalograms were not used for the main evaluation session. The same questionnaire (see 6 questions below) was given to the evaluators who answered the same six questions for each radiograph. After completing the questionnaires, a discussion was carried out with the study coordinator and any conflict about the result was discussed immediately and clarified. The session was considered successful only if at least 80% of the cases are correctly identified. Subjects who were unable to reach this result underwent a second session of retraining one week later. Those subjects were ready for the main evaluation session as explained above two weeks after the second training session. The same radiographs were used for the second training session but were viewed by the subjects in a different order than the previous session.

Methods:

A. Subjects and Recruitment:

The group of evaluators of the radiographic records included a total of ten subjects. Among these subjects were five orthodontic residents and five orthodontic faculty members in the School of Dental Medicine at the University of Connecticut. The subjects did not participate in the design or construction of the study. The subjects were recruited by sending an email to all the orthodontic residents and faculty members in the

School of Dental Medicine at the University of Connecticut. The first five residents and the first five faculty members who responded to the email were included in the study. All the subjects included were provided with a cover letter/ information sheet that briefly describes the study.

B. Survey Instrument:

A hard copy survey that includes six multiple choice questions was distributed to all the evaluators. The same six questions were asked for each radiograph. The answers to the six questions provided information about the cervical vertebrae morphology and the CVM stage of the radiographic image. The questionnaire did not contain any identifiers of the subjects. It took around 45- 60 minutes to evaluate the images and complete the questionnaire.

The radiographic sample used in this study included 80 records selected randomly from three different sources. The records from the American Association of Orthodontists Foundation website (AAOF) are from the longitudinal growth records of untreated subjects and were selected by the study coordinator. The records from the electronic record system “axium” of UCONN School of Dental Medicine are for subjects who were seeking orthodontic treatment and have been treated at the orthodontic clinic in the University of Connecticut. These records were selected by the study coordinator. Although these “axium” records are fully identifiable, the study coordinator recorded only age and gender of the patients’ radiographic image and no additional information were recorded. The records (CBCT scans) were obtained from a private orthodontic office in Florida. These records are from individuals who are starting orthodontic treatment at this orthodontic office. These records were provided to the study coordinator in a de- identifiable manner. The only information that was provided is the age and

gender of the patients' radiographic image. The following inclusion and exclusion criteria were applied to all the records:

Inclusion Criteria:

1. Subjects age range between 10- 16 years
2. Radiographs with clear and visible C2 to C4 vertebra
3. Absence of anomalies of the vertebra

Exclusion Criteria:

1. Radiographs in which the subject is wearing a thyroid collar that interferes with the visualization of C2 to C4
2. Radiographs without a clear and visible C2 to C4 vertebra
3. Low quality radiographs

Procedures:

After the radiographs were collected, a total of ten evaluators were recruited to participate in the study to evaluate the radiographic images. Among the evaluators there were five orthodontic residents and five orthodontic faculty members at the University of Connecticut. Since one of the objectives of this study is to evaluate the influence of training on the reproducibility of the cervical vertebrae maturation method, the evaluators were further divided into two groups within each group. Three subjects from the resident group and two subjects from the orthodontist group were part of a training session and an evaluation session. The other five members did not undergo a training session and were only part of an evaluation session. The evaluators were randomly assigned to either group. Five dark covered envelopes per group contained "training" (2 for the orthodontist group and 3 for the resident group) and "no training" (3 for the orthodontist group and 2 for the resident group) slips for evaluator allocation. Each of the

evaluators was asked to pick one of the envelopes and was assigned to a group based on the slip inside the envelope.

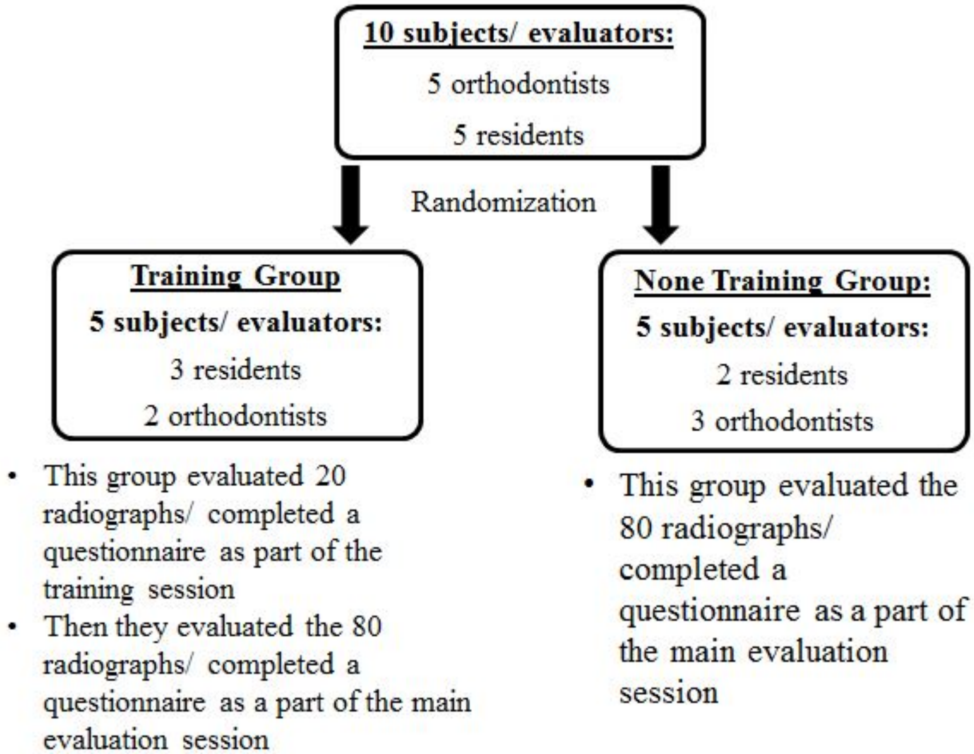
All the groups underwent the same evaluation session as described below:

All the evaluators were provided with a hard copy handout of figures and definitions of the CVM morphology, according to Baccetti et al, to be used at any time during the study (refer to page 2 of cover letter/ information sheet). Then, a high resolution image presentation containing all the selected radiographs was shown to the subjects and they were asked to complete a questionnaire which contains questions about each radiograph being evaluated. For each radiograph, the following six questions regarding cervical vertebrae morphology were asked:

1. Is the lower border of C2 best described as flat or curved?
2. Is the lower border of C3 best described as flat or curved?
3. Is the lower border of C4 best described as flat or curved?
4. Is the vertebral body of C3 best described as trapezoidal, rectangular horizontal, square, or rectangular vertical?
5. Is the vertebral body of C4 best described as trapezoidal, rectangular horizontal, square, or rectangular vertical?
6. What is the CVM stage of the radiographic image?

Each subject was given a questionnaire to answer the same six questions listed above for each image. The subjects had unlimited time to make their evaluations and complete the questionnaires which did not contain identifiers of the subjects. There was no detailed explanation of the cervical vertebrae maturation method or any discussion after the questionnaire completion at the end of the evaluation session.

Summary of Methods and Procedures:



Results

The overall interobserver agreement level between the ten evaluators in cervical vertebrae maturation staging was 65.25%. When looking at the agreement within each x-rays group, the interobserver agreement in staging using the Axium lateral cephalograms was 68.5%. When using the AAOF lateral cephalograms, the interobserver agreement was 60.5%. When 2D generated lateral cephalograms and CBCTs were used, the interobserver agreement between the ten evaluators in determining a CVM stage was 64.5% and 67.5% respectively. (**Figure 3**)

The influence of training on the reproducibility of the cervical vertebrae maturation method was assessed by comparing the interobserver agreement level in determining a CVM stage between the training and non-trained groups. The interobserver agreement between the five trained evaluators was 71.75% while the interobserver agreement between the five none trained evaluators was 62.75%. (**Figure 4**)

The interobserver agreement in staging between the training and none training groups was further analyzed comparing their agreement among the four different x-rays groups. In the Axium lateral cephalograms images group, the agreement of the trained evaluators was 76% while the agreement of the non-trained evaluators was 63%. In the second x-rays group (AAOF lateral cephalograms), the trained evaluators agreed in 70% of the times while the non-trained evaluators agreed in 57% of the times. Looking at the 2D generated lateral cephalograms, the agreement was 71% and 62% in the trained versus the non- trained evaluators respectively. The interobserver agreement level was higher among the trained evaluators compared to the non-trained evaluators in the three x-rays groups except for the CBCT images. The agreement in

staging among the trained evaluators when evaluating the CBCT images was almost similar to the agreement among the non-trained evaluators (70% vs. 96%). (**Figure 5**)

The interobserver agreement in CVM staging was also analyzed between the five faculty members and the five orthodontic residents to assess whether the evaluator's own experience in orthodontics has an influence on the reproducibility of the cervical vertebrae maturation method. The overall interobserver agreement among the faculty members was 68% while the residents' agreement was 75%. (**Figure 6**)

Comparing the interobserver agreement between the faculty members and the residents in each x-ray group, it was noted that the residents' agreement in staging was higher than the faculty members' agreement in the Axium lateral cephalograms and CBCT images. On the other hand, the agreement of the faculty members was higher than the agreement of the residents in the AAOF lateral cephalograms and the 2D generated lateral cephalograms images. (**Figure 7**)

The highest agreement between the ten evaluators in CVM staging was 100% which occurred in only 3.75% of the times. In other words, the evaluators agreed 100% in determining a CVM stage in only three images of the total 80 images. The lowest agreement was 30% which was observed in only one image.

There were 900 interobserver observations for each of the four x-ray groups. There were 616 (68.5%) agreements among the Axuim lateral cephalograms x-rays group. In the second x-rays group (AAOF), there were 545 (60.5%) agreements. The agreements in the 2D generated lateral cephalograms x-rays group were observed in 580 of the cases (64.5%). The agreements in the CBCT images group were 608 of the cases (67.5%). The majority of the Kendall correlation

coefficient values for CVM staging were substantial (0.61- 0.80) among the different radiographic images of the four groups. The interobserver agreement levels between the trained residents were higher compared to the non- trained residents in all the x-rays groups except for the CBCT group in which the agreement of both the trained and non- trained residents was substantial with Kendall correlation values of 0.77 and 0.73 respectively.

The interobserver disagreements in each x-rays group were analyzed to find the number of stages apart for each disagreement. **(Figure 8)** The total number of disagreements was 249. An 81.5% of disagreements were one cervical stage apart. A 15.6% of the disagreements were two cervical stages apart. Three and four cervical stages apart comprised 2% and 0.8% respectively.

For the Axium lateral cephalograms x-rays group, there were 63 disagreements in staging. Majority of the disagreements (90.47%) were one stage apart and almost 10% were two stages apart. For the second x-rays group (AAOF lateral cephalograms), there were 66 disagreements regarding a CVM stage and the evaluators disagreed by a difference of as many as four stages. Forty four of the disagreements were one stage apart, sixteen were two stages apart, four were three stages apart, and only two of the disagreements were four stages apart. For the 2D generated lateral cephalograms, the total disagreements were 65. Majority of these disagreements (87.69%) were one stage apart. Seven disagreements were two stages apart and only one was three stages apart. When analyzing the disagreement in the CBCT x-rays group, there were 55 disagreements and 45 of them were one stage apart.

The reliability of interobserver agreement in assessing the presence of a curvature of the inferior border of the vertebrae and determining the shape of the cervical vertebrae was assessed using the Fleiss kappa statistic. Overall the kappa values for assessing the curvature of the vertebrae

were highly variable ranging from no agreement to perfect agreement **(Figure 9)**. The interobserver agreement for determining both the shape of the body of the vertebrae and CVM staging was fair among all the different x-rays groups **(Figure 10)**. There were 900 interobserver observations for each of the 5 questions in each of the four x-rays groups. For the Axium lateral cephalograms x-rays group, the Fleiss kappa value for question 1, which addresses the presence of a curvature in the lower border of C2, was 0.55. The Fleiss kappa value for question 2, which addresses the presence of a curvature in the lower border of C3, was 0.73. The Fleiss kappa value for question 3, which addresses the presence of curvature in the lower border of C4, was 0.86. The kappa values for determining the shape of C3 and C4 were 0.43 and 0.39 respectively. The kappa value for determining a CVM stage was 0.38 which indicates a fair agreement.

For the second x-rays group (AAOF lateral cephalograms), the level of agreement for question 1 was fair with a kappa value of 0.40. The agreement in determining the presence of a curvature of C3 was moderate and the kappa value was 0.43. The level of agreement for question 3 was substantial with a kappa value of 0.63. The kappa values for questions 4, 5, and 6 were fair with kappa values of 0.25, 0.34, and 0.27.

When assessing the interobserver agreement in the cervical vertebrae morphology using the 2D generated lateral cephalograms, there was a perfect agreement in determining the presence of a curvature in the lower border of C2. There was only a slight agreement between the evaluators in answering question 2 since the kappa value was 0.06. The agreement in answering question 3 was moderate with a kappa value of 0.53. The agreement in determining the shape of C3 and C4 was fair with kappa values of 0.23 and 0.28 respectively. The agreement in CVM staging had a kappa value of 0.29. For the fourth x-rays group (CBCT), there was an almost no agreement

between the evaluators in answering question 1 since the kappa value was -0.005. The agreement in answering question 2 was slight with a kappa value of 0.1. Agreement in determining the presence of a curvature in the lower border of C3 was moderate with a kappa value of 0.52. The kappa value for determining the shape of C3 and C4 was fair as the values were 0.33 and 0.39 respectively. The agreement in CVM staging was fair too with a kappa value of 0.35.

Discussion

This study aimed to assess the influence of the quality of radiographs and training on the reproducibility of cervical vertebrae maturation method. Our data showed that the overall interobserver agreement levels for CVM staging of the ten evaluators was 65%. This level of agreement is slightly higher than what was reported by Gabriel et al[9], who found that interobserver agreement for CVM staging among practicing orthodontist was below 50%. They also found that intraobserver agreement was only 62% of the time and ranging from as high as 80% for one clinician to as low as 43% for two clinicians. A more recent study found that the percentage of total interobserver perfect agreement was 42.3% and 46.3% at two different time points [16]. On the other hand, most of the studies in the literature reported interobserver and intraobserver reproducibility levels of greater than 90% [6, 14, 15]. It has been reported that the high levels of reproducibility in these studies is as a result of using traced cervical vertebrae instead of the actual radiograph to determine CVM stages and that the observers performing the tests of reproducibility are often the authors themselves [9,10]. In a study that used an objective analysis of both the concavities and shapes of the cervical vertebrae, it was found that the overall agreement of the ten evaluators with the reference standard was about 68% and intrarater repeatability was 76.9% [11].

Our study also showed that interobserver agreement levels for CVM staging of the five residents was 75% while it was 64% for the faculty members. A study that evaluated the diagnostic accuracy and repeatability of the visual assessment of the CVM stages concluded that the method is accurate and repeatable as long as training is followed and that this accuracy is independent of the rater's own experience in orthodontics [11]. Six of the raters in this study were postgraduate students, two were postdoctoral students, one was assistant professor, and one undergraduate

student. Another recent study examined if the clinical experience has an effect on the reproducibility of cervical vertebrae maturation method [16]. Thirty evaluators were divided into three groups according to the level of their clinical experience: the junior group included 10 recent graduates in dentistry with less than one year of orthodontic experience; the postgraduate group included 10 postgraduate students in orthodontics with clinical experience ranging between 2 and 4 years; and the specialist group including 10 specialist in orthodontics with more than 7 years of orthodontic experience. They found that interobserver agreement was the highest for the junior group in both time intervals and showed an almost perfect agreement (0.87 at T1; 0.86 at T2). On the other hand, the specialist group achieved the lowest Kendall's W values presenting a substantial agreement (0.61 at T1; 0.78 at T2). The increase of the Kendall's W between the two time points was explained as a probably an effect of training. This study concluded that the reproducibility of the method was not improved by the level of orthodontic experience since the group with the lowest level of orthodontic experience had the best performance.

In our study, we assessed the interobserver agreement when evaluating each cervical morphology question among the four different x-rays groups. Our data showed that the interobserver agreement for the assessment of the shape of vertebral bodies of C3 and C4 was low kappa values ranging from 0.23 to 0.39 indicating fair agreement. The interobserver agreement for assessment of the lower borders of C2, C3, and C4 was variable with kappa values ranging from -0.005 to 1 indicating no agreement to perfect agreement. Our finding in relation to the assessment of the shape of C3 and C4 vertebral bodies is consistent with a study by Nestman et al[10], who found a fair interobserver agreement level in determining the shape of the vertebral bodies with kappa values of 0.39 and 0.34. Their interobserver agreement level in

assessing the inferior border of C2 was moderate and the kappa values for assessing the inferior border of C3 and C4 were 0.65 and 0.63 respectively indicating a substantial agreement.

This study also analyzed the number of disagreements in CVM staging and the number of stages apart for each disagreement. The total number of disagreements was 249. Majority of the disagreements (81.5%) were one stage apart among the four x-rays groups. The evaluators' disagreements were of two cervical stages apart in 15.6% of the times. The evaluators disagreed by a difference of as many as four stages in only two of the 66 disagreements in staging the AAOF lateral cephalograms. Our data is similar to many of the studies who examined the spread in terms of the cervical stages when the evaluators disagreed in CVM staging. Gabriel et al[9] study results show that 73.89% of the interobserver disagreements were one cervical stage apart. 18.43% were two cervical stages apart. Three and four cervical stages apart were the least common with a percentage of 6.33 and 1.35 respectively. None of the disagreements were five cervical stages apart. Another study showed that disagreements one cervical stage apart was 23.5% and two cervical stages apart were almost 5%. Only 0.5% of the disagreements were three cervical stages apart[11]. Similar results were reported by Rongo et al[16] who showed that 40% of the disagreements were one cervical stage apart. 12.7% of the disagreements were two stages apart. Disagreements of three and four stages apart were 3.7% and 1.1% respectively. Disagreements that were five cervical stages apart were almost negligible (0.1%).

Conclusion

Our findings suggest that training seems to have some influence on the reproducibility of the cervical vertebrae maturation method. The quality of radiographs does not seem to have a major influence on the reproducibility of the method.

Appendix:

Figure 3: Overall interobserver agreement in CVM staging

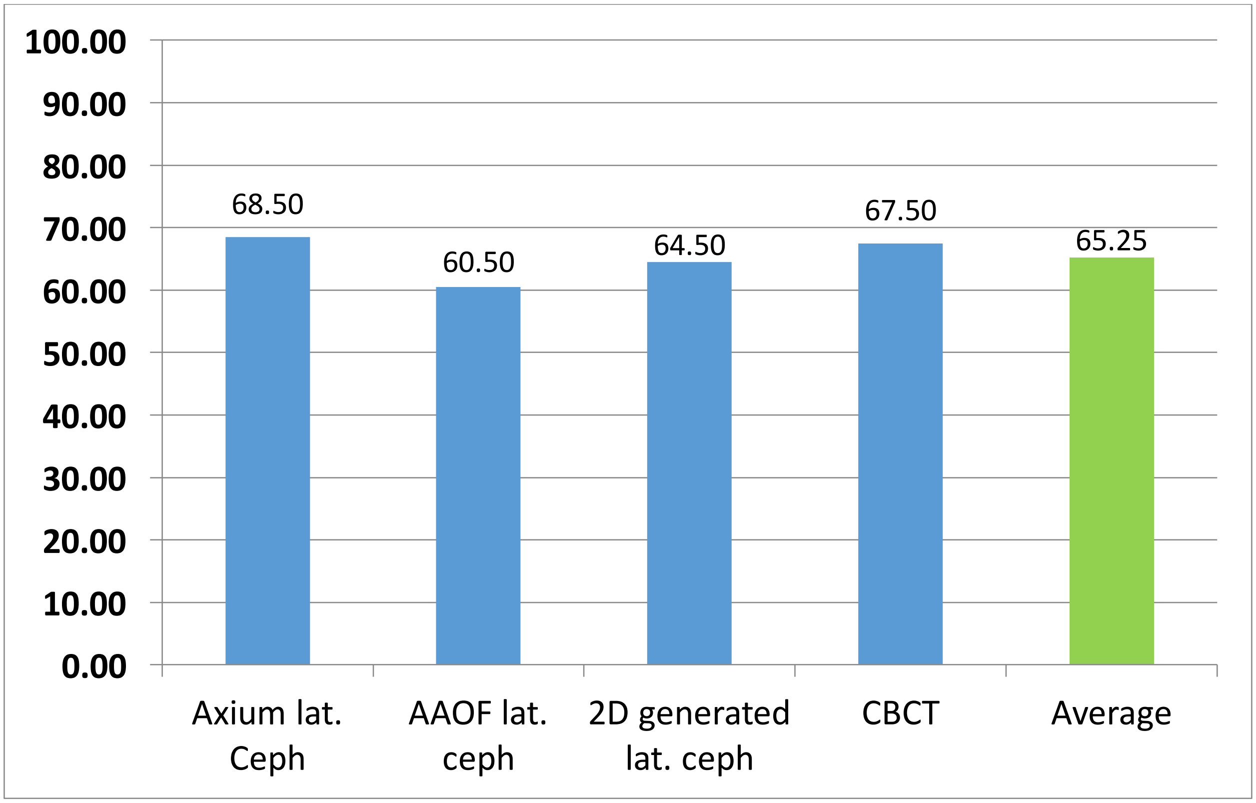


Figure 4: Interobserver agreement in CVM staging between trained and non-trained evaluators

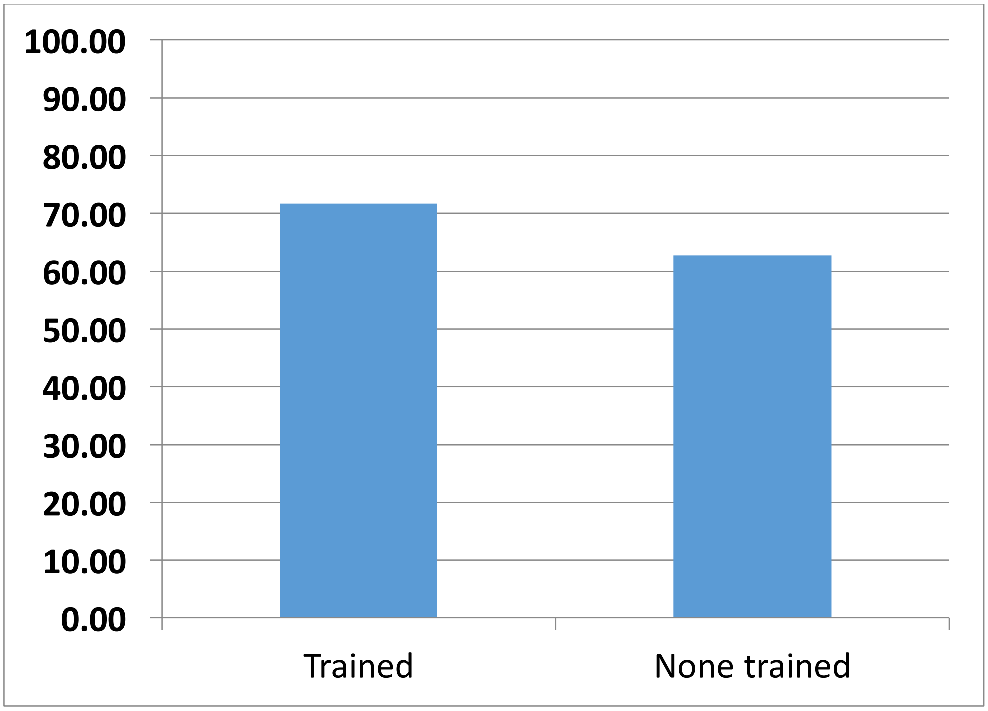
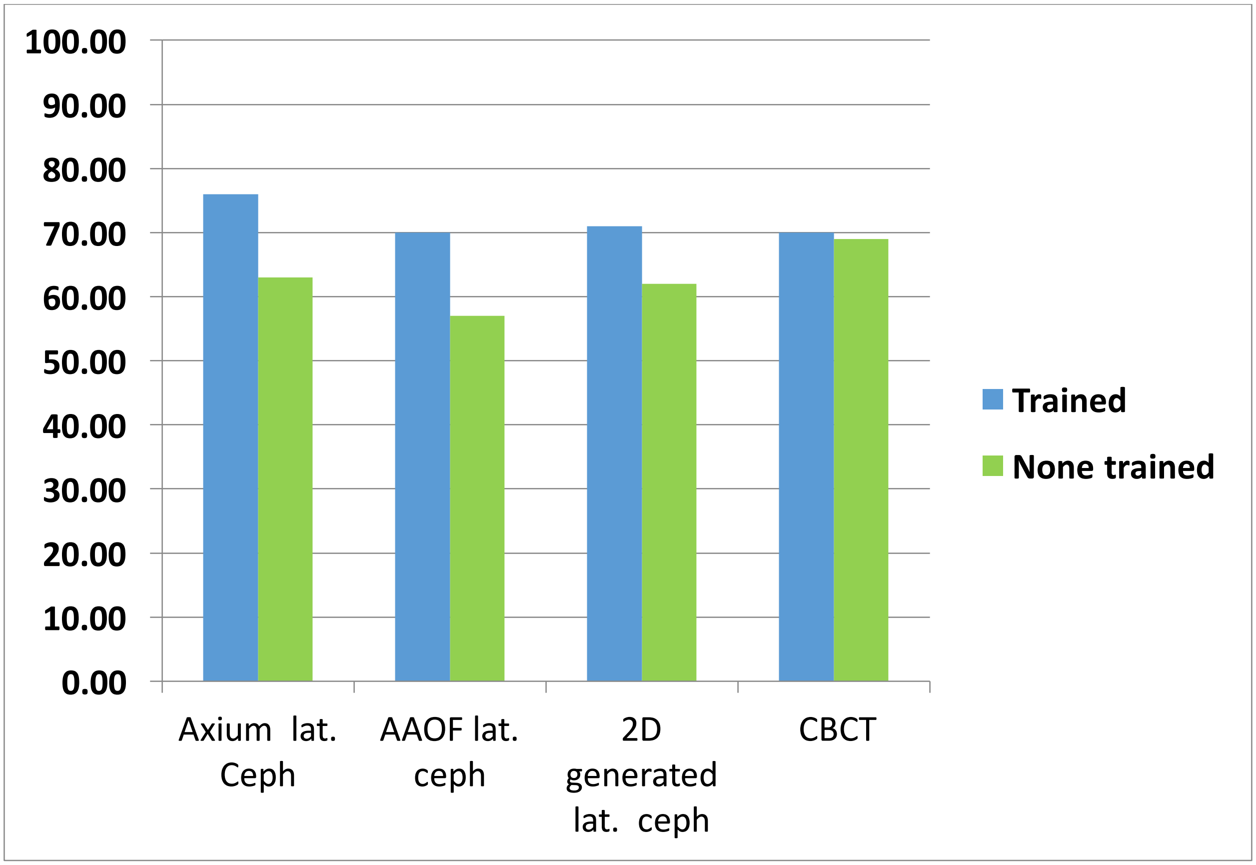


Figure 5:



100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Axium lat. Ceph
AAOF lat. ceph
2D generated lat. ceph
CBCT
Trained
None trained

Figure 6: Interobserver agreement in CVM staging between residents and faculty members

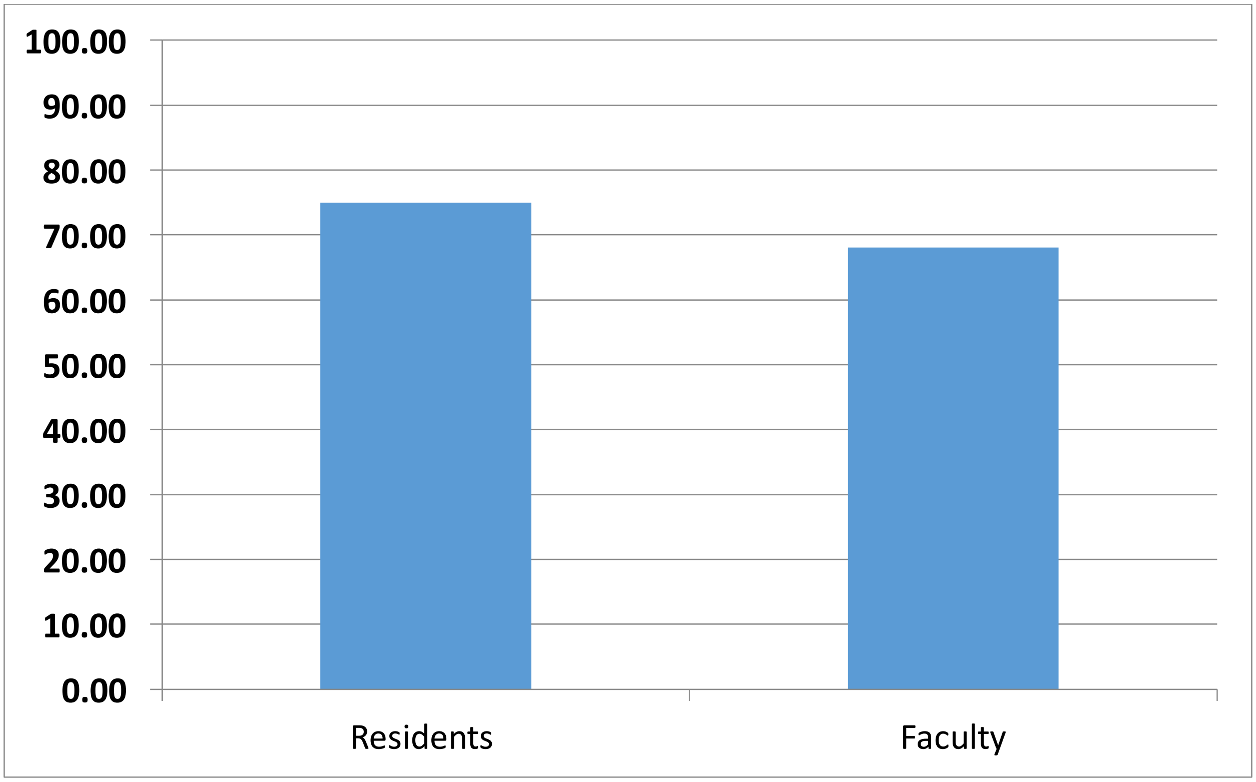


Figure 7:

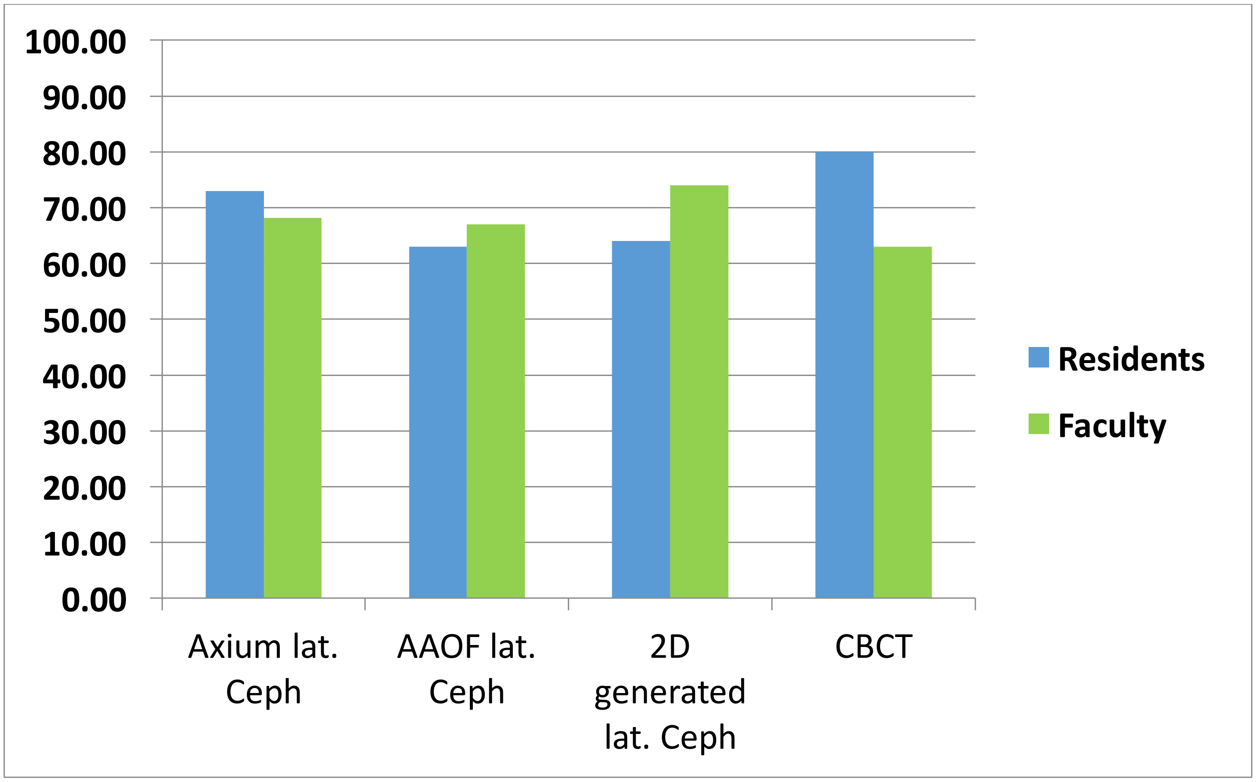

100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Axium lat. Ceph
AAOF lat. Ceph
2D generated lat. Ceph
CBCT
Residents
Faculty

Figure 8: Cervical stage differences for interobserver disagreements

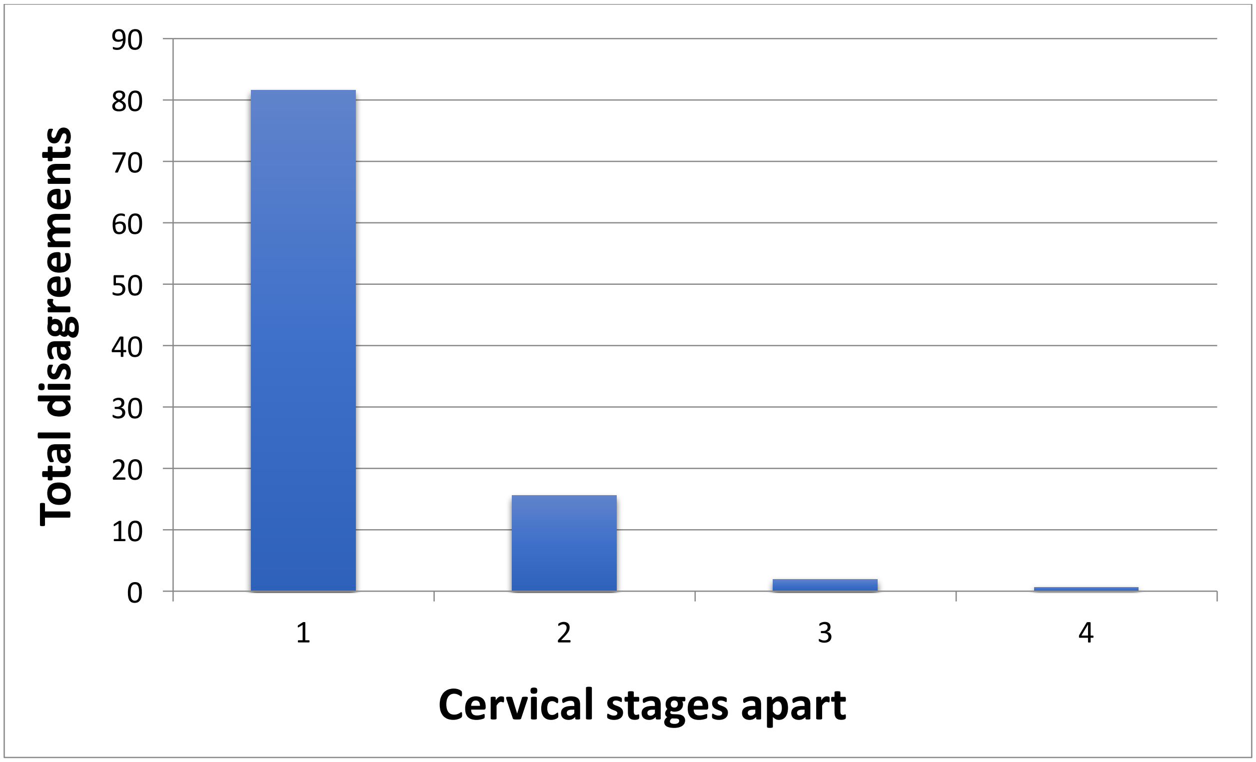


Figure 9: Kappa values for assessing the inferior border of the cervical vertebrae

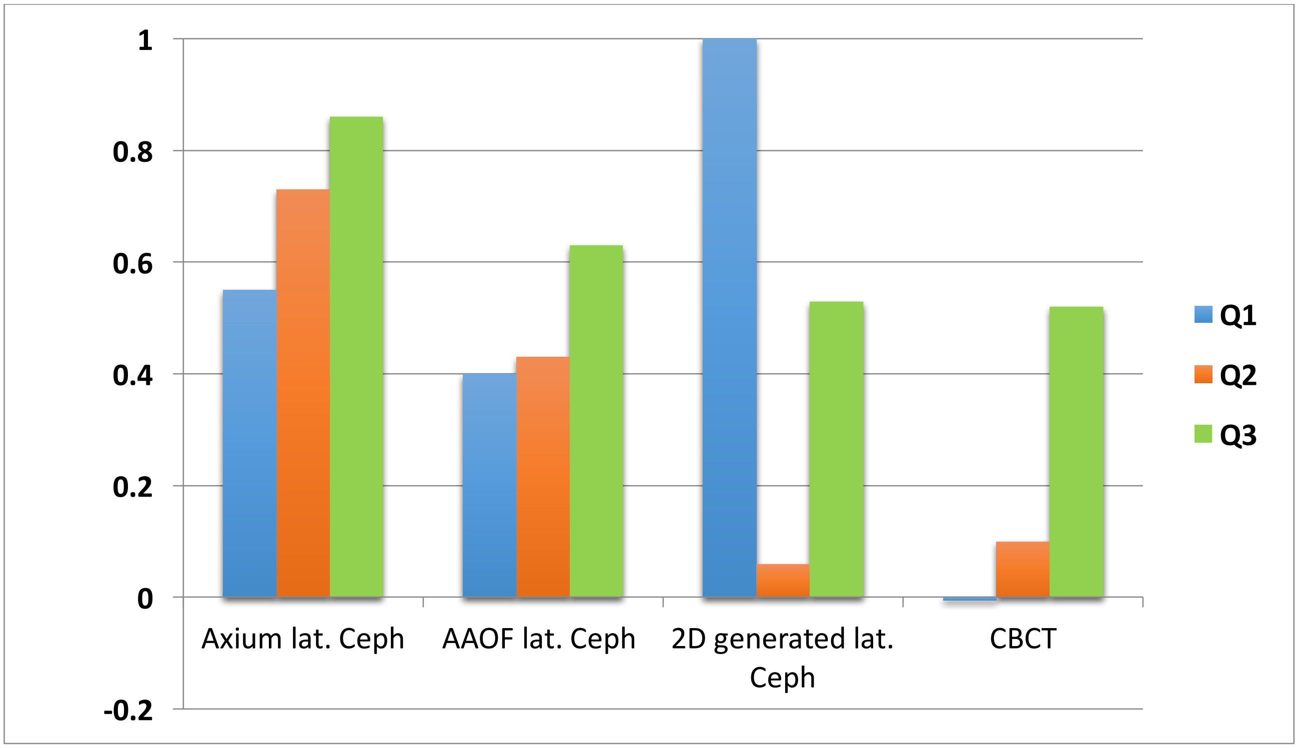
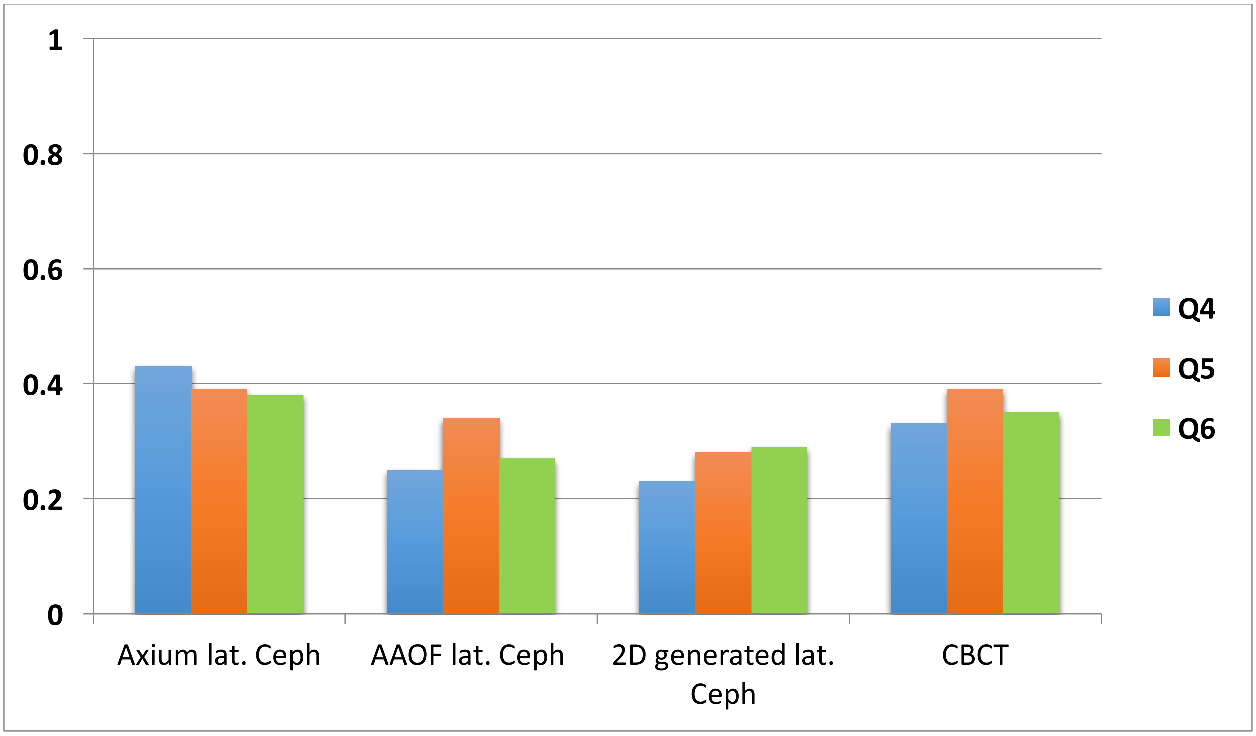


Figure 10: Kappa values for determining the shape of vertebral bodies and a CVM stage



References

1. Uysal T, Ramoglu SI, Basciftci FA, Sari Z. Chronologic age and skeletal maturation of the cervical vertebrae and hand-wrist: Is there a relationship? *Am J Orthod Dentofacial Orthop*. 2006;130(5):622-628. doi: S0889-5406(06)00894-8 [pii].

2. Fishman LS. Chronological versus skeletal age, an evaluation of craniofacial growth. *Angle Orthod*. 1979;49(3):181-189. doi: 10.1043/0003-3219(1979)049<0181:CVSAAE>2.0.CO;2 [doi].

3. Flores-Mir C, Nebbe B, Major PW. Use of skeletal maturation based on hand-wrist radiographic analysis as a predictor of facial growth: A systematic review. *Angle Orthod*. 2004;74(1):118-124. doi: 10.1043/0003-3219(2004)074<0118:UOSMBO>2.0.CO;2 [doi].

4. Chen LL, Xu TM, Jiang JH, Zhang XZ, Lin JX. Quantitative cervical vertebral maturation assessment in adolescents with normal occlusion: A mixed longitudinal study. *Am J Orthod Dentofacial Orthop*. 2008;134(6):720.e1-720.e7; discussion 720-1. doi: 10.1016/j.ajodo.2008.03.014 [doi].

5. Baccetti T, Franchi L, McNamara JA,Jr. An improved version of the cervical vertebral maturation (CVM) method for the assessment of mandibular growth. *Angle Orthod*. 2002;72(4):316-323. doi: 10.1043/0003-3219(2002)072<0316:AIVOTC>2.0.CO;2 [doi].

6. Hassel B, Farman AG. Skeletal maturation evaluation using cervical vertebrae. *Am J Orthod Dentofacial Orthop*. 1995;107(1):58-66. doi: S0889-5406(95)70157-5 [pii].

7. Baccetti T, Franchi L, McNamara Jr JA. The cervical vertebral maturation (CVM) method for the assessment of optimal treatment timing in dentofacial orthopedics. . 2005;11(3):119-129.

8. Franchi L, Baccetti T, McNamara JA,Jr. Mandibular growth as related to cervical vertebral maturation and body height. *Am J Orthod Dentofacial Orthop*. 2000;118(3):335-340. doi: S0889-5406(00)57941-4 [pii].

9. Gabriel DB, Southard KA, Qian F, Marshall SD, Franciscus RG, Southard TE. Cervical vertebrae maturation method: Poor reproducibility. *Am J Orthod Dentofacial Orthop*. 2009;136(4):478.e1-7; discussion 478-80. doi: 10.1016/j.ajodo.2007.08.028 [doi].

10. Nestman TS, Marshall SD, Qian F, Holton N, Franciscus RG, Southard TE. Cervical vertebrae maturation method morphologic criteria: Poor reproducibility. *Am J Orthod Dentofacial Orthop*. 2011;140(2):182-188. doi: 10.1016/j.ajodo.2011.04.013 [doi].

11. Perinetti G, Caprioglio A, Contardo L. Visual assessment of the cervical vertebral maturation stages: A study of diagnostic accuracy and repeatability. *Angle Orthod*. 2014;84(6):951-956. doi: 10.2319/120913-906.1 [doi].

12. Shim J, Heo G, Lagravere M. Assessment of skeletal maturation based on cervical vertebrae in CBCT. *Int Orthod* 2012; 10: 351-362. doi: 10.1016/j.ortho.2012.09.005 [doi].

13. Joshi V, Yamaguchi T, Matsuda Y, Kaneko N, Maki K, Okano T. Skeletal maturity assessment with the use of cone-beam computerized tomography. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2012; 113: 841–9. doi: 10.1016/j.oooo.2011.11.018 [doi].

14. Flores-Mir C, Burgess CA, Champney M, Jensen RJ, Pitcher MR, Major PW. Correlation of skeletal maturation stages determined by cervical vertebrae and hand-wrist evaluations. *Angle Orthod* 2006;76:1-5.

15. Ozer T, Kama JD, Ozer SY. A practical method for determining pubertal growth spurt. *Am J Orthod Dentofacial Orthop* 2006;130:131.e1-6.

16. Rongo R, Valleta R, Bucci R, Bonetti GA, Michelotti A, D'Anto V. Does clinical experience affect the reproducibility of cervical vertebrae maturation method? *Angle Orthod*. 2015; 85: 841-847.

17. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37:360–363.

18. Zhao XG, Lin J, Jiang JH, Wang Q, NG SH. Validity and reliability of a method for assessment of cervical vertebral maturation. *Angle Orthod*. 2012; 82:229-234.

19. Cericato GO, Bittencourt MAV, Paranhos LR. Validity of the assessment method of skeletal maturation by cervical vertebrae: a systematic review and meta-analysis. *Dentomaxillofac Radiol.* 2015; 44:20140270.

20. Ballrick JW, Fields HW, Beck FM, Sun Z, Germak J. The cervical vertebrae staging method's reliability in detecting pre and post mandibular growth. *Orthodontic Waves*. 2013; 72:105-111.

21. Predko AE, Kaminek M, Langova K, Kowalski P, Fudalej PS. Reliability of the cervical vertebrae maturation (CVM) method. *Bratisl Lek Listy*. 2015; 116 (4)222-226.

22. Kucukkeles N, Acar A, Biren S, Arun T. Comparison between cervical vertebrae and hand wrist maturation for assessment of skeletal maturity. *J Clin Pediatr Dent* 1999;24:47-52.