Improving root stage assessment in dental age estimation

Amir Hazwan Bin Abdul Rahim

Centre for Oral Bioengineering

Barts and The London, School of Medicine and Dentistry

Queen Mary University of London, UK

Submitted in partial fulfilment of the requirements of the Degree of Doctor of Philosophy

This work was supported by the Ministry of Higher Education, Government of Malaysia

STATEMENT OF ORIGINALITY

I, Amir Hazwan Bin Abdul Rahim, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that Queen Mary University of London has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: -SIGNED-

Date: 29/05/2023

Details of collaboration and publications: A.H. Abdul Rahim, J.A. Davies, H.M. Liversidge. Reliability and limitations of permanent tooth staging techniques. *Forensic science international*, *346*, 111654. https://doi.org/10.1016/j.forsciint.2023.111654

ABSTRACT

Aim: Tooth staging techniques are subjective and prone to bias. The aims of this thesis were to assess intra-observer reliability of permanent tooth staging techniques (Nolla, Moorrees et al. and Demirjian et al.) and explore crown-root ratios as an alternate method.

Methodology: The reliability sample was archived cross-sectional panoramic radiographs of 200 children (100 males, 100 females) aged 6-15. Seven permanent teeth on the left side were scored twice using Nolla, Moorrees, and Demirjian methods. Weighted Kappa and percentage agreement were calculated. The sample for crown-root ratio was panoramic radiographs of a 100 male and 100 female subjects aged 16-25. Mean crown-root ratios (crown height/root length) for each tooth type (left side) were calculated using ImageJ. Student t-test was used to compare between tooth type and sexes. Moorrees root fractions derived from mean root lengths were defined in terms of crown-root ratios. A separate sample of 62 radiographs of individuals (aged 9-24) with developing roots were assessed comparing Moorrees' staging and crown-root ratio. Accuracy was defined as percentage of teeth with crown root ratio within defined limits.

Result: Results showed excellent reliability with Kappa values of 0.918, 0.922 and 0.938 for Demirjian (N=2682), Nolla (N=2698) and Moorrees (N=2674) respectively. Results for mean crown root ratios from 3019 teeth by tooth type ranged from 0.49-0.68 with third molars having highest ratios. The ratio for some tooth types differed significantly ($P<0.01$) between sexes. The accuracy of using crown-root ratio over root fractions showed an increasing accuracy with root stage, however, the sample included few early or mid-root fractions.

Conclusion: These findings showed that reliability of permanent teeth using Demirjian scoring is marginally more reliable than Nolla or Moorrees. Crown-root ratio has

potential as a less subjective approach to assess root growth stages than root fractions in dental age estimation.

ACKNOWLEDGMENTS

I would like to acknowledge:

God. Praise be to Him the Almighty.

My sponsor, Ministry of Higher Education, Government of Malaysia.

My dearest supervisors – Prof. Helen Liversidge, for her knowledge, expertise, wisdom and support for the past five years. Thank you for her endless patience, for always cheering me on and believing in me. I would also like to thank Dr. Janet Davies for her guidance and advice, and the many words of encouragement. I am forever indebted to them both.

Prof. Dr. Zamri Radzi, Dean of the Faculty of Dentistry, University of Malaya and Head of Department, A/P Dr. Kathreena Kadir. My warmest thanks to my colleagues and mentors – Prof. Dr. Norliza Ibrahim, Dr. Rabi'ah Al-Adawiyah Rahmat and of course, Prof. Dr. Phrabhakaran Nambiar.

I would like to express my very special gratitude to my friends, the beautiful Halim sisters, Jihan and Zati (Taufiq and Keng too), Muz and Adam, and the 'London gang' for their friendship, support and love. "*Insaflah ye sebelum mati*"

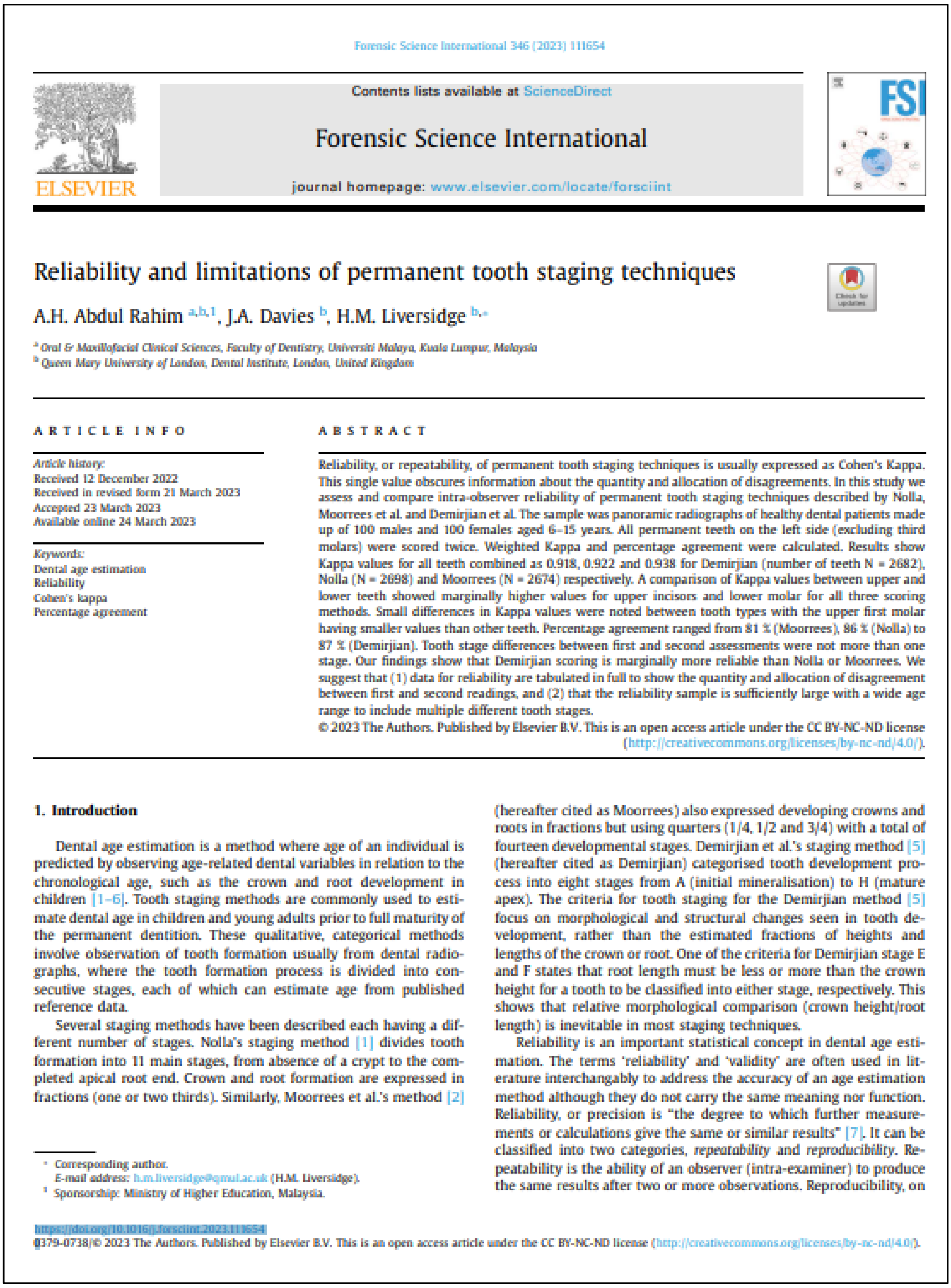
Special thanks and much love to Sarah Sabrina and Melissa Ooi Wan Yun – without whom none of this would even be possible in the first place.

My dearest and beloved Mak, whom I can never thank enough, and Ayah, for their infinite and unconditional love. I would also like to thank my sisters, Akak and Eshah for their love and support for the last 36 years.

Last but not least, myself. Who deserves love and kindness much more than he has given himself.

LIST OF PUBLICATIONS

Rahim, A. A., Davies, J. A., & Liversidge, H. M. (2023). Reliability and limitations of permanent tooth staging techniques. *Forensic science international*, *346*, 111654. https://doi.org/10.1016/j.forsciint.2023.111654



Forensic Science International 346 (2023) 111654

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Reliability and limitations of permanent tooth staging techniques

A.H. Abdul Rahim [a,b,1], J.A. Davies [b], H.M. Liversidge [b,*]

[a] Oral & Maxillofacial Clinical Sciences, Faculty of Dentistry, Universiti Malaya, Kuala Lumpur, Malaysia
[b] Queen Mary University of London, Dental Institute, London, United Kingdom

ARTICLE INFO

Article history:
Received 12 December 2022
Received in revised form 21 March 2023
Accepted 23 March 2023
Available online 24 March 2023

Keywords:
Dental age estimation
Reliability
Cohen's kappa
Percentage agreement

ABSTRACT

Reliability, or repeatability, of permanent tooth staging techniques is usually expressed as Cohen's Kappa. This single value obscures information about the quantity and allocation of disagreements. In this study we assess and compare intra-observer reliability of permanent tooth staging techniques described by Nolla, Moorrees et al. and Demirjian et al. The sample was panoramic radiographs of healthy dental patients made up of 100 males and 100 females aged 6–15 years. All permanent teeth on the left side (excluding third molars) were scored twice. Weighted Kappa and percentage agreement were calculated. Results show Kappa values for all teeth combined as 0.918, 0.922 and 0.938 for Demirjian (number of teeth N = 2682), Nolla (N = 2698) and Moorrees (N = 2674) respectively. A comparison of Kappa values between upper and lower teeth showed marginally higher values for upper incisors and lower molar for all three scoring methods. Small differences in Kappa values were noted between tooth types with the upper first molar having smaller values than other teeth. Percentage agreement ranged from 81 % (Moorrees), 86 % (Nolla) to 87 % (Demirjian). Tooth stage differences between first and second assessments were not more than one stage. Our findings show that Demirjian scoring is marginally more reliable than Nolla or Moorrees. We suggest that (1) data for reliability are tabulated in full to show the quantity and allocation of disagreement between first and second readings, and (2) that the reliability sample is sufficiently large with a wide age range to include multiple different tooth stages.

1. Introduction

Dental age estimation is a method where age of an individual is predicted by observing age-related dental variables in relation to the chronological age, such as the crown and root development in children [1–6]. Tooth staging methods are commonly used to estimate dental age in children and young adults prior to full maturity of the permanent dentition. These qualitative, categorical methods involve observation of tooth formation usually from dental radiographs, where the tooth formation process is divided into consecutive stages, each of which can estimate age from published reference data.

Several staging methods have been described each having a different number of stages. Nolla's staging method [1] divides tooth formation into 11 main stages, from absence of a crypt to the completed apical root end. Crown and root formation are expressed in fractions (one or two thirds). Similarly, Moorrees et al.'s method [2] (hereafter cited as Moorrees) also expressed developing crowns and roots in fractions but using quarters (1/4, 1/2 and 3/4) with a total of fourteen developmental stages. Demirjian et al.'s staging method [5] (hereafter cited as Demirjian) categorised tooth development process into eight stages from A (initial mineralisation) to H (mature apex). The criteria for tooth staging for the Demirjian method [5] focus on morphological and structural changes seen in tooth development, rather than the estimated fractions of heights and lengths of the crown or root. One of the criteria for Demirjian stage E and F states that root length must be less or more than the crown height for a tooth to be classified into either stage, respectively. This shows that relative morphological comparison (crown height/root length) is inevitable in most staging techniques.

Reliability is an important statistical concept in dental age estimation. The terms 'reliability' and 'validity' are often used in literature interchangably to address the accuracy of an age estimation method although they do not carry the same meaning nor function. Reliability, or precision is "the degree to which further measurements or calculations give the same or similar results" [7]. It can be classified into two categories, *repeatability* and *reproducibility*. Repeatability is the ability of an observer (intra-examiner) to produce the same results after two or more observations. Reproducibility, on

* Corresponding author.
E-mail address: h.m.liversidge@qmul.ac.uk (H.M. Liversidge).
[1] Sponsorship: Ministry of Higher Education, Malaysia.

https://doi.org/10.1016/j.forsciint.2023.111654

TABLE OF CONTENTS

4. CHAPTER FOUR: RESULTS

1. INTRODUCTION

1.1 Overview

Growth indicators such as biological, skeletal and dental are effective in predicting the age of an individual. Age estimation, as well as age discrimination for the minority and majority age groups are especially important for the implications they carry in the legal system. In certain jurisdictions, these legal implications would result in either a whole life order or as consequential as the death penalty. Compared to other growth indicators, dental age estimation is favourable due to its accurate, inexpensive, quick and simple nature.

Generally, dental age estimation is carried out through the radiological examination of developing permanent teeth categorised into stages based on qualitative and quantitative observations, known as the staging method. Examples of these observations include root formation assessment and root canals maturation. This method is non-invasive, and allows assessment of developing teeth with minimal risks, as conventional dental radiographs are generally low in radiation exposure. On the other hand, the staging method is limited by its subjective and arbitrary qualities. For example, mature root lengths and stages may be interpreted in various ways by different observers as interpretations are influenced by knowledge, experience and expertise. Subsequently, reliability of the method is compromised and further reduces its accuracy. Therefore, defining these subjective attributes in an objective manner seems to be the sensible approach to better this method.

The aims of this study were to explore and develop a quantitative and objective method (expressed in crown-root ratio) and to assess developing teeth particularly root fractions. Reliability of present staging methods were analysed, and mean crown-root ratios obtained from a sample of mature permanent teeth comprised of the Malaysian population were measured. Accuracy of this quantitative method was tested against a sample of developing permanent teeth of similar group.

1.2 Objectives

a) To assess and compare reliability of present tooth staging techniques.

b) To measure and describe crown-root ratio of mature mandibular and maxillary teeth in young adults (100 males and 100 females).

c) To define crown-root ratios of the average R1/4, R1/3, R1/2, R2/3, and R3/4 from average values of mature teeth, Rc.

d) To test accuracy of crown-root ratios as a method to assign root fractions.

2. LITERATURE REVIEW

2.1 Age Estimation in Forensic Science

Forensic odontology is a branch of forensic science that "deals with the proper handling and examination of dental evidence and the proper evaluation and presentation of dental findings" [1]**.** Among the many roles of a forensic odontologist, one that is mainly sought after is to provide an accurate age estimation when required by the government authority, or the court of law for instance. This process is particularly important should an individual fail to produce proper legal documentations, or when the legality of one's age becomes doubtful. According to the United Nations High Commissioner for Refugees (UNCHR) there were 13.6 million newly displaced people in 2018 [2]. Among these numbers, 580 800 were first-time asylum seekers in Europe [3], and 138 600 were unaccompanied and separated children [2], a 20% decrease from 2017 [4]. An asylum seeker is "an individual who has sought international protection and whose claim for refugee status has not yet been determined" [5]. There are several reasons as to why European countries appeal to immigrants. Italy, for example, is easily reached by her Mediterranean neighbours especially in the summer [6]. A country's change in policy on the other hand, such as border opening in Germany, has also led to an inflow of immigrants in 2015 [7]. For cases such as these, accurate classification of age for young children is consequential as minors are able to have access to education and may be granted permit, while adults face deportation or detention in jail [8].

Identification of unidentified bodies and skeletonised remains also benefit from age estimation methods, especially in mass disasters where casualties are usually of large scales. Estimation of age may assist identification process when there is absence of ante-mortem data, such as the unfortunate event of 2004 tsunami when thousands lost their lives [9]. In 2011's Christchurch earthquake that killed 185, many of whom were children, the need to age victims precisely became increasingly crucial [10]. When bodies are badly charred, and visual recognition of bodies becomes impossibly difficult, such as the case of a bus – car collision on a major highway in southern Spain, 4 out of 28 victims were positively identified through dental eruption and mineralisation assessment alone [11].

2.2 Skeletal Age Estimation

As humans grow, other biological structures such as the bones develop and mature. This growth acts as one of the indicators of an individual's skeletal age [12] and shown to be strongly correlated with chronological age [13,14]. Changes in the epiphysis of tubular bones from the earliest ossification up to the fusion with the diaphysis was staged and classified in the Greulich & Pyle Atlas [15,16]. Therein it compiled a comprehensive collection of the left hand-wrist radiographs of children aged up to 19 years, which visualises development of the carpal bones until maturity is reached. Age is estimated by comparing the hand-wrist radiograph of a subject to the atlas for the closest correspondence. Due to its simplicity, this method is more commonly used as compared to the Tanner Whitehouse method [17]. For this method, score maturity is used rather than a visual comparison. 20 selected regions of interest of the hand bones are staged accordingly, and these stages are given individual scores. Eventually, these scores are added up, and translated into a skeletal age. Despite the more complex nature, it has been demonstrated to be more accurate and reproducible as compared to the Greulich & Pyle Atlas [18]. Another qualitative method for skeletal age estimation is by assessing the cervical vertebrae maturity (CVM). Introduced by Lamparski in 1972, it assessed the developmental stage of C2-C6 by describing the size and shape [19]. The interior borders of the cervical vertebrae go through changes in their shape from being flat (10 years) to deepening concavities (15 years) in females. According to Lamparski, maturity is reached by males approximately one year after their counterparts. Researchers believed skeletal maturity assessed on hand-wrist radiographs as the best indicator for maturity, and most closely related to growth spurt [20-22].

2.3 Dental Age Estimation

Dental development is another biological structure that develops throughout gestation to adulthood that may be correlated with chronological age. Initial calcification of teeth starts as early as 13 weeks in utero [23], which develops as a set of deciduous teeth. These teeth then exfoliate and are replaced by a set of adult or also known as permanent teeth. The first permanent teeth to form are the first mandibular molars, which occurs at birth [24], and the last to mature are the roots of the third molars at 25 years of age [25].

There are several known reasons that could factor in the delay of tooth formation. Certain skeletal diseases, such as cleidocranial dysplasia [26], and cleft lip/palate are known to prolong dental development in children for 0.3 – 0.7 years [27]. Nutrition has also been reported to play a role in tooth formation process. Initially, it was revealed that overweight children showed a significant advancement in tooth development as compared to their malnourished counterparts [28]. However, a study of more than 2000 North Sudanese has shown that the effect of extreme malnutrition on the timing of dental development was negligible [29]. Other local and systemic diseases, such as odontogenic and non-odontogenic tumours, cerebral palsy, and endocrine disorders [30-34] have also been known to cause delayed tooth eruption. However, assessment of tooth formation is seen as a more favourable tool in predicting age as compared to tooth eruption as it is less likely to be influenced by other external factors, such as crowding, and ankylosis of deciduous teeth [35]. Apart from these factors, dental age remains as a reliable method for age estimation in otherwise healthy individuals as it has been shown to be strongly correlated with chronological age [13][36].

Dental age estimation is a method where the age of an individual is predicted by observing age-related dental variables in relation to the chronological age, such as the crown and root development in children [24][37-41]. In these young individuals, age estimation may be carried out through visual observation of tooth eruption, or emergence, such as the timing for the third molars emergence as described by Olze [42] and assessment of tooth tissues mineralisation on radiographs [37-40]. The timing and sequence of tooth eruption have been described and visualised in the past through diagrams and atlases [41][43-45]. Schour and Massler atlas [44,45] was one of the most well-known as it was one of the first of its kind. Ubelaker [43] attempted to overcome the limitations that the Schour and Massler presented such as unclear tooth stages and small ranges [46,47]. The London Atlas is a more recent and comprehensive age estimation tool as compared to the aforementioned ones. Comprising data collection of 72 prenatal and 104 postnatal teeth, it aimed to aid age prediction for individuals aged between 28 weeks in utero and 23 years.

Tooth staging technique is another commonly used method to estimate dental age in children. This qualitative, categorical method involves observation of odontogenesis, where the tooth formation process is classified into several different stages. These stages correspond individually to a certain dental age, depending on the level of dental maturity. Several staging techniques have been described and for each technique, various number of stages were established.

2.3.1 Qualitative Method

2.3.1.1 Nolla's Staging Technique

Nolla's staging technique [37] was first introduced in 1960 when it was developed by using data of 25 boys and 25 girls from the files of Child Development Laboratories of the University of Michigan School. Different types of radiographs were taken, such as the left and right lateral, maxillary and mandibular occlusal, and left and right intraoral periapical radiographs for the left and right maxillary teeth. In total, there were over 3400 radiographs taken for this study. Nolla found that there was no significant difference between the left and right side of the jaw, and as a result of tooth development observation of these children, illustration similar to Figure 1 was produced.

Nolla’s staging technique employed a comprehensive scoring system whereby the tooth development is categorised into 11 main stages, between absence of crypt and apical end of root completed. For each developing crown and root that has not reached maturity, the stages are expressed in fractions (1/3 and 2/3). Each permanent tooth on the mandible and/or maxilla is staged from stage 0 – 10, which also correspond to a score, and later summed to a total that is translated into a dental age based on the sex of the subject. Alternatively, a dental age can also be obtained by translating individual tooth scores instead.

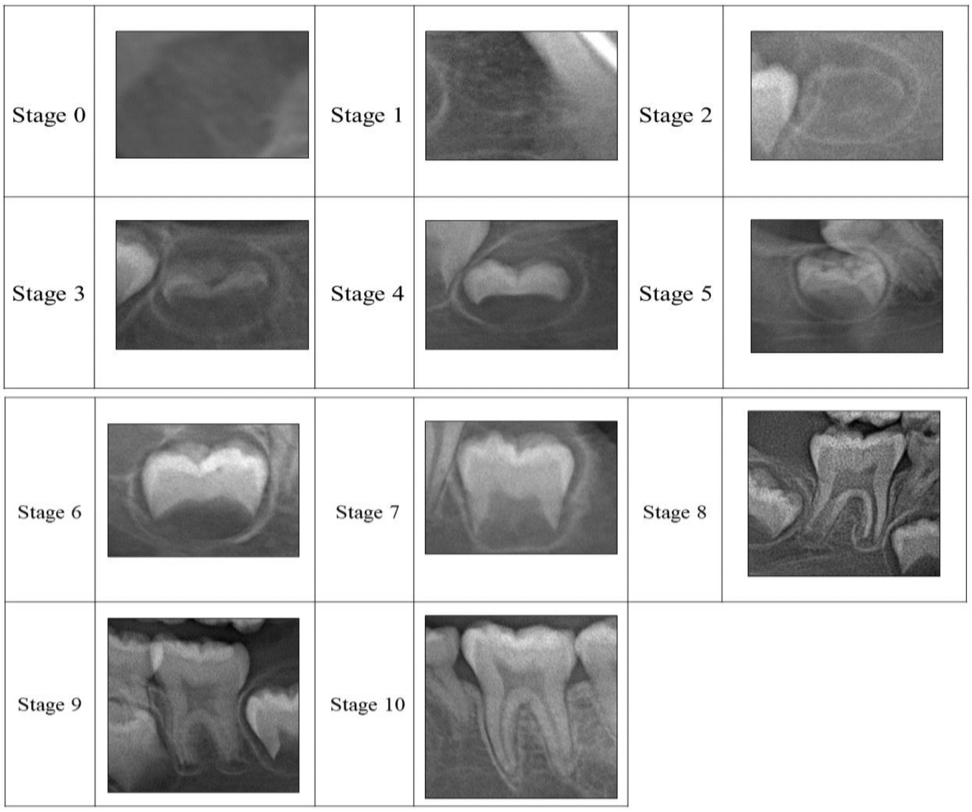


Figure 1: Adapted from Nolla’s staging technique [37] by Ghougassian and Ghafari [48]

In order to improve the staging of a tooth development, Nolla introduced a point system in her technique. For example, when the crown development is fully complete (stage 6) a score 6.0 is given. However, when it is observed that development is halfway through between crown completed (stage 6) and one-third of root completed (stage 7), a value 0.5 is added, giving a score of 6.5. Scores 6.2 and 6.7 however, are given when there is a slight development than stage 6 but has *not yet* reached 6.5, and when a tooth has *almost* reached 7.0 but more developed than a score of 6.5, respectively.

While this grading system may seem reasonable as tooth development is a continuous and gradual process, it is arguably confusing especially since there is inconsistency in the classification of the stages, and whether certain stages need to be classified as a stage at all. For instance, stage 5 states that '*crown almost completed*' when the other developing crown stages are expressed in terms of fractions (*one-third of crown completed, two-thirds of crown completed*). By the definition of the point-system, this stage could possibly be classified as 4.5 instead.

2.3.1.2 Moorrees' Staging Technique

Moorrees et al. [24], however, had a slightly different approach than Nolla's [37]. Based on a previously established technique by Gleiser and Hunt [38], it was developed through the analysis of longitudinal data of over 200 children. Intraoral and lateral oblique radiographs of these children were observed and each developing permanent tooth was given 14 different stages between initial cusp formation (Ci) to apical closure complete (Ac) as described in **Table 1**. A cumulative percentage frequency for each stage was calculated for children who had attained or passed these stages, and mean age for each stage was calculated. Norms that include the means and standard deviations attainment of age for each tooth of the permanent teeth formation were published in this study for boys and girls.

Initial cusp formation	Ci	Initial cleft formation	Cli
Coalescence of cusps	Cco	Root length 1/4	R1/4
Cusp outline complete	Coc	Root length 1/2	R1/2
Crown 1/2 complete	Cr1/2	Root length 3/4	R3/4
Crown 3/4 complete	Cr3/4	Root length complete	Rc
Crown complete	Cr.c	Apex 1/2 closed	A1/2
Initial root formation	Ri	Apical closure complete	Ac

Table 1: Tooth formation stages as described by Moorrees [24] and their abbreviations

According to this method, age of an individual is estimated by first staging the developing permanent teeth. The line drawings of mean age are then referred to, and the mean attainment of age that corresponds to the particular stage and tooth will be the estimated age of the individual. Similar to Nolla's staging technique [37], Moorrees [24] also expressed developing crowns and roots in fractions. However, they were divided into quarters (1/4, 1/2 and 3/4) with considerably more stages (fourteen). In comparison to Nolla's, Moorrees' staging technique is simpler to execute as it does not include point-system and more consistent in classification. However, it may be difficult to stage a tooth when stages initial root formation (Ri), initial cleft formation (Cli) and root length 1/4 (R1/4) are involved. This is due to the fact that a root that has substantially developed and may have possibly reached 1/4 of its length, sometimes is presented with initial cleft formation. When faced with this situation, which stage should the tooth be classified as?

This ambiguous stage, without the knowledge of the actual length of the root in its mature form, and inconsistencies in stages classification as previously recognised for Nolla's staging technique [37], may reduce reliability resulting in inaccuracies when predicting age.

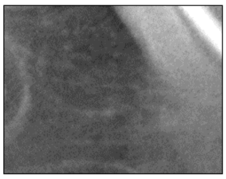
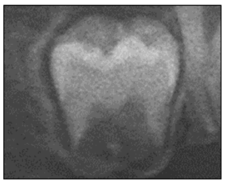
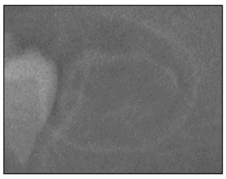
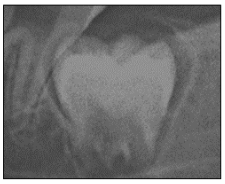
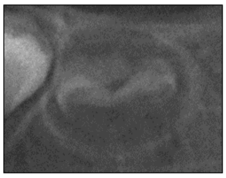
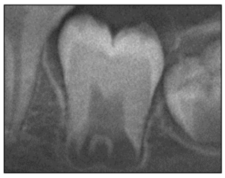
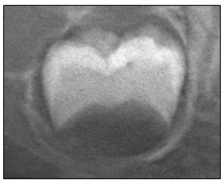
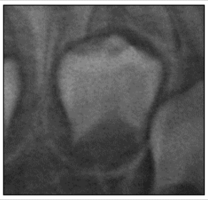
Ci Initial cusp formation		**Cli** Initial cleft formation	
Cco Coalescence of cusps		**R1/4** Root length 1/4	
Coc Cusp outline complete		**R1/2** Root length 1/2	
C1/2 Crown 1/2 complete		**R3/4** Root length 3/4	
C3/4 Crown 3/4 complete		**Rc** Root length complete	
Crc Crown complete		**A1/2** Apex 1/2 closed	
Ri Initial root formation		**Ac** Apical closure complete	

Figure 2: Adapted from Moorrees' staging technique [24]

2.3.1.3 Demirjian's Staging Technique

Demirjian et al. [40,49] staging technique is one of the most widely used techniques since it was established in 1973. This study of almost 3000 French Canadian children adapted the earlier studies on tooth staging as were previously mentioned [24,37] as well as skeletal age estimation methods [16,17].

Based on these techniques, Demirjian et al. [40] derived a system where seven permanent lower left teeth were divided into eight stages, A-H (**Figure 3**). Stage A represents beginning of calcification at the superior level of the crypt while the final stage, H represents completion of root development where the apical end has closed. Similar to Tanner's scoring technique [17], each stage is given a self-weighted score depending on the sex of the individual. After seven permanent left mandibular teeth (excluding the third molar) are staged, a sum of these values, called maturity score is obtained and this score is then converted into dental age. These scores are unique to each sex.

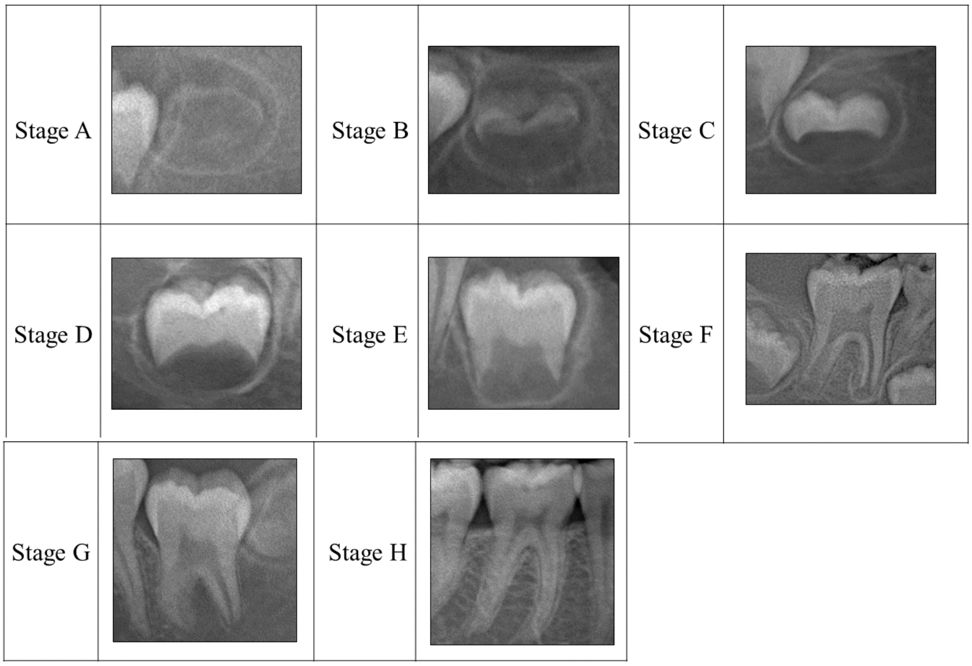


Figure 3: Adapted from Demirjian's staging technique [40] by Liversidge [50]

Compared to Nolla's [37] and Moorrees' staging techniques [24], the criteria for tooth staging for Demirjian technique [40] focus more on morphological and structural changes seen in tooth development, rather than the heights and lengths of the crown or root. It is worth noting, however, one of the criteria in Stages E and F states that root length must be less or more than the crown's height (**Table 2**) for a tooth to be classified into either stage, respectively. This shows that relative morphological comparison (crown height/root length) is inevitable in many staging techniques.

Stage	Description
A	In both uniradicular and multirooted teeth, a beginning of calcification is seen at the superior level of the crypt in the form of an inverted cone or cones. There is no fusion of these calcified points.
B	Fusion of the calcified points forms one or several cusps which unite to give a regularly outlined occlusal surface.
C	a. Enamel formation is complete at the occlusal surface. Its extension and convergence towards the cervical region is seen. b. The beginning of a dentinal deposit is seen. c. The outline of the pulp chamber has a curved shape at the occlusal border.
D	a. The crown formation is complete down to the cemento–enamel junction. b. The superior border of the pulp chamber in the uniradicular teeth has a definite curved form, being concave towards the cervical region. The projection of the pulp horns, if present, gives an outline shaped like an umbrella top. In molars the pulp chamber has a trapezoidal form. c. Beginning of root formation is seen in the form of a spicule.
E	Uniradicular teeth: a. The walls of the pulp chamber now form straight lines, whose continuity it broken by the presence of the pulp horn, which is larger than in the previous stage. b. the root length is less than the crown height. Molars: a. Initial formation of the radicular bifurcation is seen in the form of either a calcified point or a semi-lunar shape. b. the root length is still less than the crown height.
F	Uniradicular teeth: a. The walls of the pulp chamber now form a more or less isosceles triangle. The apex ends in a funnel shape. b. The root length is equal to or greater than the crown height. Molars: a. The calcified region of the bifurcation has developed further down from its semi-lunar stage to give the roots a more definite and distinct outline with funnel shaped endings. b. The root length is equal to or greater than the crown height
G	The walls of the root canal are now parallel and its apical end is still partially open (distal root in molars).
H	a. The apical end of the root canal is completely closed (distal root in molars). b. The periodontal membrane has a uniform width around the root and the apex.

Table 2: Detailed description for each stage as described by Demirjian [40]

Demirjian [40] recognised techniques that are based on the measurements of crowns and roots, as well as changes in morphology are prone to subjectivity. A complete guideline that describes criteria that should be fulfilled before a tooth is classified into a certain stage was provided in order to aid the staging process (**Table 2**).

According to Demirjian [40], when there is only one criterion presented such as stage A and B, it *must* be met for a tooth to be classified as such. When there are two criteria, such as stage E and F, it is sufficient only for the first one to be met. For Stage C and D, for example, which are presented with 3 criteria, the first two criteria *must* be fulfilled. Earlier stage is assumed in borderline cases.

2.3.2 Quantitative Method

Quantitative or metric methods are an alternative way to assess tooth development. Cameriere et al. [51] studied this process by measuring tooth height and open apices in children. They concluded that all open apices ratios have a significant correlation with age, and that this method is highly reliable and reproducible, evidenced by its insignificant intra-observer differences. Thevissen et al. [52] investigated whether measurement of third molars was a good age predictor as compared to scoring system and concluded that measurements and ratios of third molars calculated in their study added no clinical value in age prediction. This may be due to the fact that third molars are more variable in morphology as compared to the other molars [53]. Inclusion of other permanent or developing teeth, although was not suggested, might have a positive influence in this study and should be considered.

The study of crown-root ratio is not a unfamiliar subject outside the field of forensic odontology. Rowlands et al. [54] correlated crown-root ratios with mandibular second premolar's time of eruption to predict eruption timing, hence determining whether early intervention is needed in orthodontic cases. Hölttä et al. [55] tested this method on 435

Finnish children to obtain mean crown-root ratios of mature permanent teeth, excluding the third molars. They also examined the reproducibility of this technique and found it to be highly acceptable (Pearson correlation coefficient of 0.87 and 0.83 for intra and inter-examiner reliability, respectively). While they produced useful data for fully developed teeth, studies on developing ones were not available. Othman et al. [56] also tested this method to determine the normal mean value of crown-root ratios in Malay subjects. They concluded in their paper that crown-root ratio is a reproducible method with an intraclass correlation coefficient of 0.81.

2.4 Reliability Test in Forensic Science

Often in literature the terms 'reliability' and 'validity' are used interchangeably to address the accuracy of an age estimation method. Authors and scientists have the responsibility to use these terms appropriately as they are not defined nor do they function similarly.

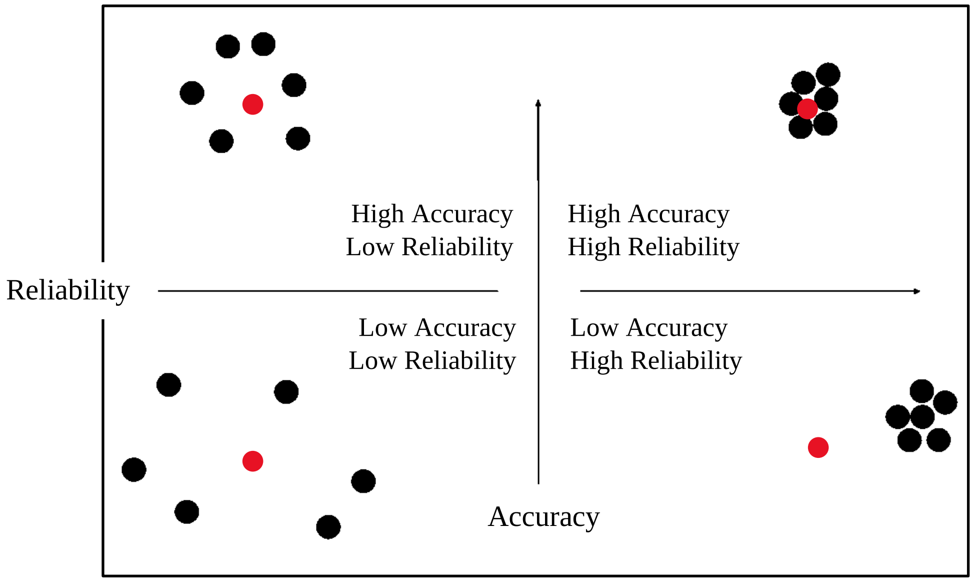


Figure 4: Analogy of accuracy and precision. Red=indicator for accuracy. The further the black balls away, the less accurate.

Ferrante and Cameriere [57] in their paper had defined these terms clearly and how they differ from each other. Reliability, or precision is defined as "the degree to which further measurements or calculations give the same or similar results". It can be classified into two categories, *repeatability* and *reproducibility*. Repeatability is the ability of an observer (intra-examiner) to produce the same results after two or more observations. Reproducibility, on the other hand, is the likelihood or agreement between two or more observers (inter-examiners) to produce the same results when carrying out the same observation. Accuracy, or similarly known as validity on the other hand, is defined as “the degree of conformity of a measured or calculated quantity to its actual true value”. A method can be highly accurate and low in precision (close to the true value but with huge error), but it can also be highly precise and low in accuracy (small error but further from its true value). Ideally, when an age estimation method is devised, high accuracy (without bias) and precision (small error) are desired.

Reliability of a method is influenced by several factors. Sun stated in the meta-analysis of Cohen’s kappa that study settings, test properties, and subject characteristics influence reliability of a method, and that it may vary when setting of a study is altered [58]. Observer’s characteristics such as knowledge, experience, and expertise have also been known to affect reliability [59]. Staaf et al. reported that precision of a technique may be influenced by three factors, correct interpretation of radiographs, biological variability, and the quality of sample [60].

Cohen’s coefficient kappa is one of the most widely accepted statistical tests to evaluate reliability for qualitative or categorical methods [61]. It measures agreement between two observers or observations and takes into account and adjusts the amount of agreement that occurs by chance. The formula to calculate kappa, κ is

$$\kappa = \frac{\text{observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}} = \frac{\rho_o - \rho_e}{1 - \rho_e} = 1 - \frac{(1 - \rho_o)}{(1 - \rho_e)}$$

Degrees of agreement and kappa classification are represented in the table below:

Classification	Kappa value
Very good agreement	0.80 ≥ 1.00
Good agreement	0.60 ≥ 0.80
Moderate agreement	0.40 ≥ 0.60
Poor agreement	0.00 ≥ 0.40

Table 3: Kappa coefficient classification

Cohen's kappa takes into account agreement between observers, but not the *degree* of agreement itself [62]. In order to improve this limitation, Cohen's weighted kappa was introduced [63]. The weighted kappa is measured by giving weights to disagreements. The higher the disagreement, more weight is given, as simplified in the table below:

Score	A	B	C
A	0	1	2
B	1	0	1
C	2	1	0

Table 4: Scoring and weights as per Cohen's weighted kappa [61]

When two scores are rated similarly (A-A, B-B, C-C), no weight is given. When two scores differ, but only by one score (A-B, B-C) value of weight, or 'penalty' given is 1. The bigger the difference between two scores (A-C), more weight is given (2). Hence, the kappa coefficient value will be influenced by how much weight or disagreement is given.

Cohen's kappa, however, may not accurately reflect the true level of agreement between observers [64] as true agreement may be attributed to chance agreement when it is not supposed to [65].

Reliability of staging methods is understudied as compared to accuracy or validity, where a lot of emphasis is given. Attempts to test reproducibility have been documented in the past [66], but little has been done to evaluate *repeatability* (intra-observer reliability) and limitations of tooth staging techniques. These limitations have to be recognised and studied, so reliability of these methods can be improved.

All of the previously mentioned staging techniques utilise fractions to represent certain stage of root development (one-third of root completed and two-third of root completed for Nolla [37] and R1/4, R1/2, R3/4 for Moorrees [24]), and this contributes to some major drawbacks, such as subjective definitions by observers. For example, tooth 37 on **Figure 5** may be classified as R1/2 by observer A and R3/4 by observer B, depending on their own independent interpretations, knowledge, and experience. For another instance, observer A may also classify this tooth differently on two different occasions, reducing its intra-observer reliability.

Figure 5: A developing second permanent molar

Ferrante and Cameriere [57] discussed that reliability is influenced by two other factors, such as differences among the observers and small changes in the morphological variable itself. These subjective interpretations may be owed to the fact that the information on the real root/tooth length is not fully known and that morphological variations (i.e., tooth length) exist in human dentition. Hence, these fractions assigned by the observers are guesses or hypotheticals. Discrepancies such as this, and inaccurate interpretation of the stages contribute to bias and reduce precision of the predicted age.

Another limitation of staging techniques is the difficulty in classifying a tooth when it is in an ambiguous stage. Demirjian [40] discussed that when such occurrence happens, the lesser stage should be considered. Even though increasing the number tooth formation stages, such as fourteen stages by Moorrees' [24] seems like a logical way to overcome this, Dhanjal et al. stated that "increasing tooth formation stages might improve accuracy, but too many reduces precision" [66].

As previously recognised, reliability of a method is influenced by the observer's expertise and knowledge. Thus, training and calibration among observers may be advantageous and improve the limitations presented by these staging techniques. However, as long as subjectivity persists, precision will remain a challenge.

The rationale of this study was to improve this limitation of subjectivity by developing a quantitative method to aid tooth staging in dental age estimation.

3. METHODOLOGY

3.1 Materials

This study was divided into three parts. Part 1 (reliability of staging methods), Part 2 (crown-root ratio definition), and Part 3 (testing crown-root ratio accuracy). This study included a collection of panoramic radiographs of healthy patients attending various clinics at the Faculty of Dentistry, University of Malaya in Kuala Lumpur, Malaysia. Subjects were made up of the Malaysian population, mainly Malay, Chinese, and Indian by ethnicity, and categorised by sex. Only subjects aged 6 to 25 were included in this study. Other details such as hospital number, date of birth, and date of radiographic exposure were recorded on an Excel spreadsheet (*Microsoft365 (Office), version18*), secured, coded, and made accessible only to the principal investigator to preserve anonymity and confidentiality. Ethics approval was submitted on 27th May 2019 and approved by the Faculty of Dentistry Medical Ethics Committee, University of Malaya on 16th July 2019. Ethics reference number was DF OS 1909/0042(P).

Rating	Quality	Basis
1	Excellent	No errors of patient preparation, exposure, positioning, processing or film handling
2	Diagnostically acceptable	Some errors of patient preparation, exposure, positioning, processing or film handling, but which do not detract from the diagnostic utility of the radiograph
3	Unacceptable	Errors of patient preparation, exposure, positioning, processing, or film handling, which render the radiograph diagnostically unacceptable

Table 5: Quality ratings for dental radiographs

Table 5 described the criteria used during extraction of data and radiographs were (1) subjects with no pathology and facial trauma and (2) only radiographs with excellent or diagnostically acceptable quality as outlined by National Radiological Protection Board [67] were selected.

3.1.1 Part 1 – Reliability of Staging Methods

A sample of 200 subjects (100 males and 100 females) aged 6 – 15 was selected as detailed in **Table 6**. A total of 2698 teeth (*n*=2698) were scored. Mean age for males was 10.95 (SD=2.93/SE=0.29) and females was 10.81 (SD=2.90/SE=0.29).

Age (years)	Male	Female
6	10	10
7	10	10
8	10	10
9	10	10
10	10	10
11	10	10
12	10	10
13	10	10
14	10	10
15	10	10
TOTAL	**100**	**100**

Table 6: Distribution of subjects for PART 1

3.1.2 Part 2 – Crown-Root Ratio Definition

Age (years)	Male	Female
16	10	10
17	10	10
18	10	10
19	10	10
20	10	10
21	10	10
22	10	10
23	10	10
24	10	10
25	10	10
TOTAL	**100**	**100**

Table 7: Distribution of subjects for PART 2

A sample of 200 subjects (100 males and 100 females) aged 16-25 was selected. A total of 3019 teeth (*n*=3019) were measured. Mean age for males was 21.18 (SD=2.91/SE=0.29) and females was 21.18 (SD=2.93/SE=0.29). Distribution of subjects was represented in **Table 7**.

Only teeth that were fully mature (defined as having attained stage 10 for Nolla [37], Ac for Moorrees [24] or stage H for Demirjian [40]) were included in this study and those still in development were excluded. Teeth with existing pathologies, extensive restorations that may compromise crown heights, presence of apical lesions, buccally and lingually tilted teeth, and dental anomalies were also excluded.

3.1.3 Part 3 – Crown-Root Ratio Accuracy

Age (years)	Male	Female
9	0	3
10	0	2
11	0	6
12	0	8
13	1	3
14	1	3
15	2	2
16	6	1
17	2	1
18	1	1
19	2	3
20	1	2
21	1	3
22	0	3
23	0	1
24	0	3
TOTAL	**17**	**45**

Table 8: Distribution of subjects for PART 3

A sample of 20 subjects (8 males and 12 females) attending multiple clinics at the Faculty of Dentistry, University of Malaya was collected. Only subjects with two or more panoramic radiographs taken at different clinical sessions (between May 2012 and August 2017) were included. Collectively, a total of 62 subjects and panoramic radiographs aged between 9 and 24 years were selected. 818 teeth (*n*=818) were scored and subsequently measured. Mean age for males was 17.16 (SD=2.23/SE=0.54) and females was 15.84 (SD=4.76/SE=0.71) years. Distribution of subjects for each radiograph was shown in **Table 8**.

Teeth with existing pathologies, extensive restorations that may compromise crown heights, presence of apical lesions, buccally and lingually tilted teeth, and dental anomalies were also excluded.

3.2 Methods

3.2.1 Part 1 – Reliability of Staging Methods

All seven permanent teeth on the left and right side of the mandible and maxilla (central incisor, lateral incisor, canine, first premolar, second premolar, first molar and second molar) were staged according to stages as described by Nolla [37], Moorrees [24], and Demirjian [40]. Third molars were excluded in this study. Alphanumeric notation was adopted, and defined in **Table 9**.

Mandibular	Notation	Maxillary	Notation
First Incisor	LL1	First Incisor	UL1
Second Incisor	LL2	Second Incisor	UL2
Canine	LL3	Canine	UL3
First Premolar	LL4	First Premolar	UL4
Second Premolar	LL5	Second Premolar	UL5
First Molar	LL6	First Molar	UL6
Second Molar	LL7	Second Molar	UL7
Third Molar	LL8	Third Molar	UL8

Table 9: Alphanumeric notation and its description

3.2.1.1 Nolla's Staging Technique

For the first method of staging, teeth were staged 0 – 10 as described by Nolla [37]. Teeth with absence of crypt, presence of crypt, initial calcification, one-third of crown completed, two-third of crown completed, crown almost completed, crown completed, one-third of root completed, two-third of root completed, root almost completed – open apex, and root completed were staged as 0 – 10 respectively as shown in **Table 10.**

Stage	Description
0	Absence of crypt
1	Presence of crypt
2	Initial calcification
3	One-third of crown completed
4	Two-third of crown completed
5	Crown almost completed
6	Crown completed
7	One-third of root completed
8	Two-third of root completed
9	Root almost completed – open apex
10	Root completed

Table 10: Stages as described by Nolla [37]

3.2.1.2 Moorrees' Staging Technique

Stage	Abbreviation	Code
Initial cusp formation	Ci	1
Coalescence of cusps	Cco	2
Cusp outline complete	Coc	3
Crown 1/2 complete	Cr1/2	4
Crown 3/4 complete	Cr3/4	5
Crown complete	Cr.c	6
Initial root formation	Ri	7
Initial cleft formation	Cli	8
Root length 1/4	R1/4	9
Root length 1/2	R1/2	10
Root length 3/4	R3/4	11
Root length complete	Rc	12
Apex 1/2 closed	A1/2	13
Apical closure complete	Ac	14

Table 11: Stages as described by Moorrees [24] and alternative numbering system

Teeth were staged into fourteen stages as described by Moorrees [24]. The stages were initial cusp formation (Ci), coalescence of cusps (Cco), cusp outline complete (Coc), crown 1/2 complete (Cr1/2), crown 3/4 complete (Cr3/4), crown complete (Cr.c), initial root formation (Ri), initial cleft formation (Ci), root length 1/4 (R1/4), root length 1/2 (R1/2), root length 3/4 (R3/4), root length complete (Rc), apex 1/2 closed (A1/2), and apical closure complete (Ac). Alternatively, teeth were also staged with a digital system that corresponded to respective stages as detailed in **Table 11**.

3.2.1.3 Demirjian's Staging Technique

Stage	Description
A	Beginning of calcification
B	Complete fusion of calcified cusps
C	Complete formation of enamel and deposition of dentine
D	Complete formation of the crown
E	Walls of pulp chamber form a straight line, and the root length is less than the crown height
F	The walls of pulp chamber form isosceles triangle and root length is equal or greater than the crown height
G	Walls of root canal are parallel, and apex is still open
H	Apical end of root canal completely closed

Table 12: Stages as described by Demirjian [40]

Teeth were staged as described by Demirjian [40] into eight stages based on their description in **Table 12** for the third method of scoring. The stages were stage A (beginning of calcification), stage B (complete fusion of calcified cusps), stage C (complete formation of enamel and deposition of dentine), stage D (complete formation of the crown), stage E (walls of pulp chamber form a straight line, and the root length is less than the crown height), stage F (the walls of pulp chamber form isosceles triangle and root length is equal or greater than the crown height), stage G (walls of root canal are parallel and apex is still open), and stage H (apical end of root canal completely closed).

Two independent series of scoring were conducted for each staging method, with a minimum break of two weeks. Reliability of teeth, each jaw (mandibular and maxillary teeth), as well as both jaws combined (mandibular and maxillary teeth) was assessed for each method by using weighted Kappa. Scores were crosstabulated to illustrate raw agreements (highlighted in yellow) as well as disagreements between both observations.

Agreement was defined as stages scored similarly and disagreement was stages that were scored differently in both observations.

Observation 1

Observation 2

Stage	A	B	C	Total
A	AA	BA	CA	AA+BA+CA
B	AB	BB	CB	AB+BB+CB
C	AC	BC	CC	AC+BC+CC
Total	AA+AB+AC	BA+BB+BC	CA+CB+CC	Grand Total

Raw agreements were expressed in percentage agreement and calculated as below:

$$\text{Stage A percentage agreement} = \frac{AA}{(AA + AB + AC + BA + CA)} \times 100\ \%$$

$$\text{Stage B percentage agreement} = \frac{BB}{(BB + BA + BC + AB + CB)} \times 100\ \%$$

$$\text{Stage C percentage agreement} = \frac{CC}{(CC + CA + CB + AC + BC)} \times 100\ \%$$

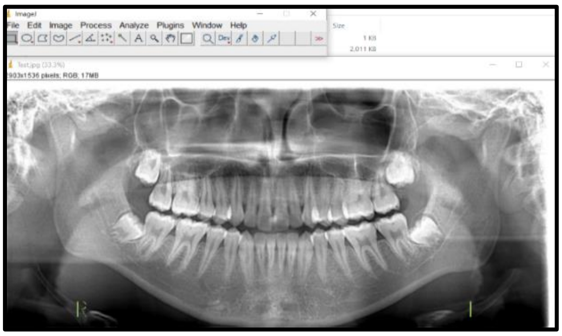
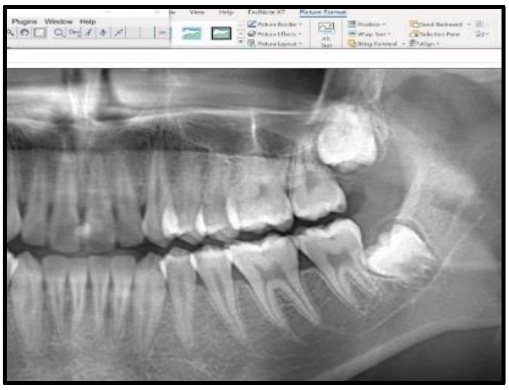
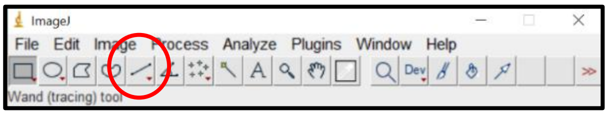
Data for male and female subjects were pooled as this study was focused on reliability of staging technique, and not accuracy of the technique itself.

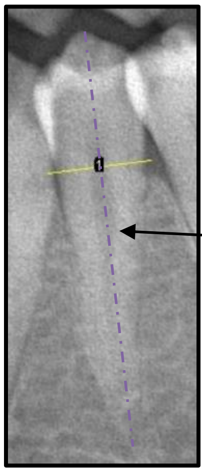
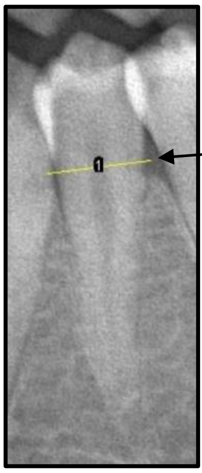
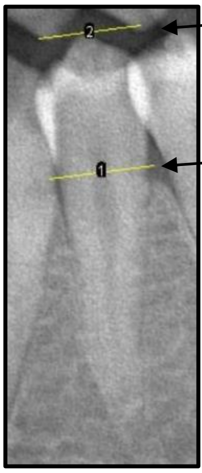
3.2.2 Part 2 - Crown-Root Ratio

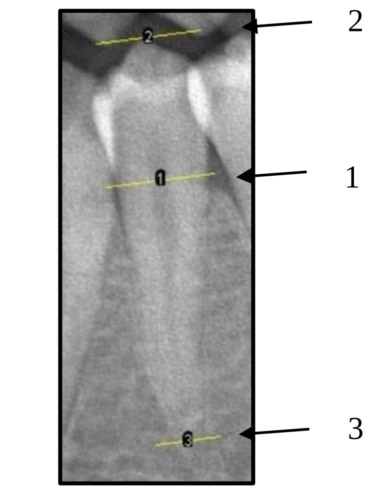
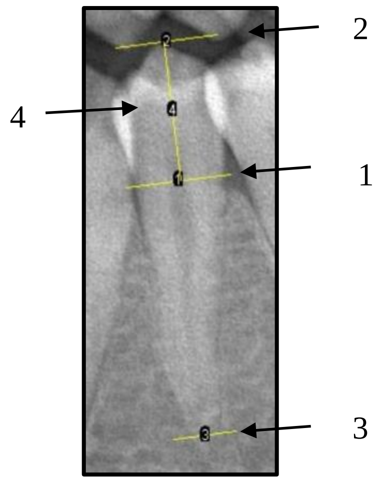
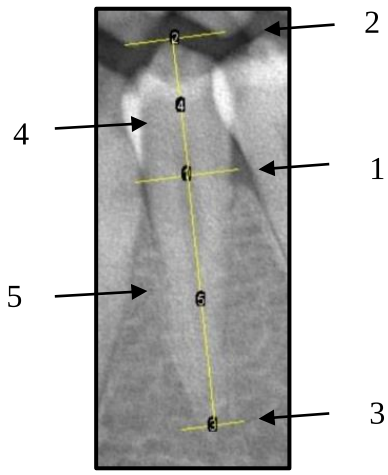
3.2.2.1 Crown-root ratio definition for mature teeth

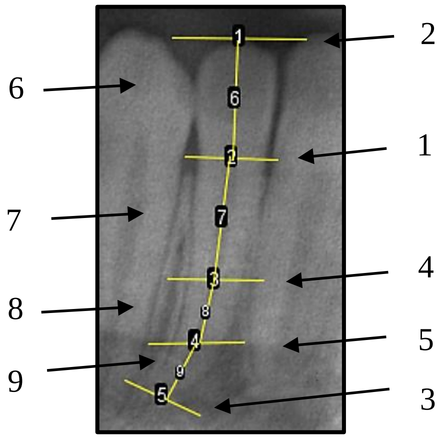
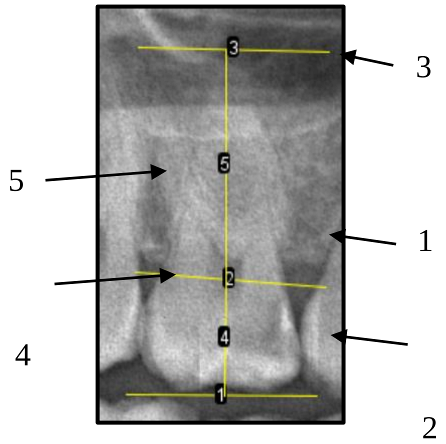
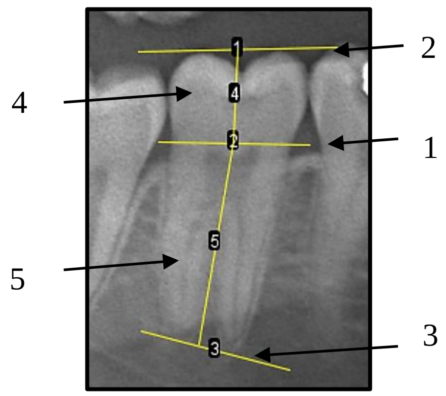
Crown height and root length measurement was done by using an open access image processing program, *ImageJ*, 1.47v. **Table 13** described and illustrated the steps used in order to measure the crown height, root length and analysis of crown-root ratio.

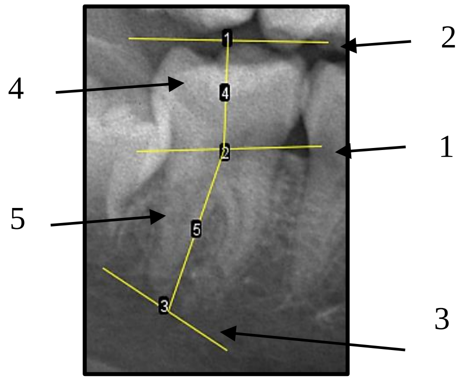
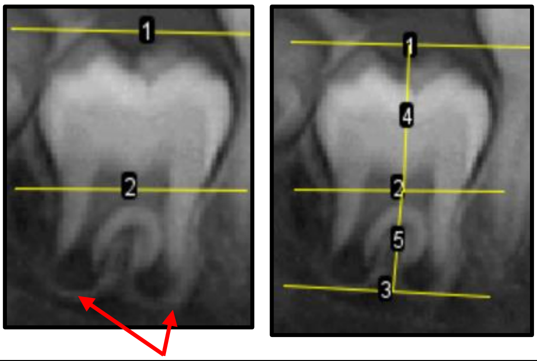
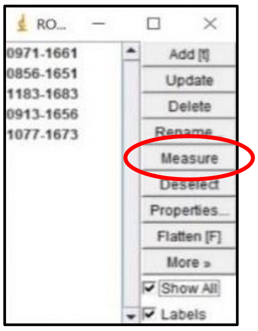
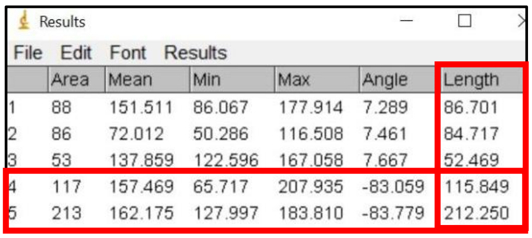
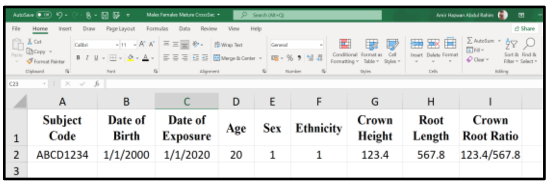
All mature permanent teeth on the left side of the mandible and maxilla (central incisor, lateral incisor, canine, first premolar, second premolar, first molar, second molar and third molar) were measured to obtain the individual average crown-root ratio.

Crown- Root Ratio Measurement	ImageJ/Microsoft Excel
Step 1 – Radiograph, or image was uploaded into the software *ImageJ* by selecting file > open. The intended image was chosen from the folder.	
Step 2 – In order to standardise the magnification of each image uploaded into the software, as too much magnification will result in pixelated image, and too little will affect accuracy of the landmarks positioning, the image was zoomed in to 75% on the desired region (left side of the jaw).	
Step 3 – To start measuring for crown height and root length, several landmarks were identified first such as distal, mesial, occlusal aspect of the tooth, cemento-enamel junction (CEJ), tip of the cusp, and apices. Straight line selection tool was selected (red circle).	

Crown- Root Ratio Measurement	ImageJ/Microsoft Excel
Step 4 – Y–axis of the tooth was identified. Radiolucent appearance of the root canal was an acceptable indicator of the Y–axis.	
Step 5 – CEJ of the tooth was identified and a line was drawn from the mesial to distal aspect of the tooth and marked as '1'. This marking was perpendicular to the Y–axis of the tooth.	
Step 6 – The tip of the cusp was identified and a line was drawn parallel to line '1' and marked as '2'. For posterior teeth where there were more than 1 cusp, the highest cusp was selected.	

Crown- Root Ratio Measurement	**ImageJ/Microsoft Excel**
Step 7 – The apex of the tooth was identified and a line was drawn parallel to line ‘1’ and ‘2’ and marked as ‘3’.	
Step 8 – A line connecting line ‘1’ and ‘2’ was drawn as close as possible to the tooth’s Y–axis and marked as ‘4’. This line represented the crown height.	
Step 9.1 – A line connecting ‘2’ and ‘3’ was drawn as close as possible to the tooth’s Y–axis and marked as ‘5’. This line represented root length for uniradicular teeth with straight roots/no curves.	

Crown- Root Ratio Measurement	ImageJ/Microsoft Excel
Step 9.2 – For uniradicular teeth with a curved root, multiple lines (lines ‘7’, ‘8’ and ‘9’) might be required to represent the root length. Root canal (located in the Y–axis of the tooth/in the mid-section) was used to aid placement of these lines. Lines ‘7’, ‘8’ and ‘9’ were measured as the root length. Anomalies such as dilacerated roots were excluded from the study.	
Step 9.3 – For maxillary molars/teeth with three roots, the longest root was selected as the primary criterion to draw lines ‘3’ and ‘5’ to represent the root length. Where all roots were similar in lengths, *step 9.4* would be adopted. Teeth with more than three roots were excluded from the study.	
Step 9.4 – For teeth with two roots such as the mandibular molars or maxillary premolars, both roots were considered as line ‘3’. The midpoint between these roots was used to mark line ‘5’ and represent the root length. In cases where roots were unequal in lengths, the lengths were averaged by the two apices connected by line ‘3’.	

<table>
<tr><th>Crown- Root Ratio Measurement</th><th>ImageJ/Microsoft Excel</th></tr>
<tr><td>Step 9.5 – For multirooted teeth with one or two curved roots, similar methodology was applied where midpoint of both roots was considered as the average root lengths.</td><td>
</td></tr>
<tr><td>Step 9.6 – For immature teeth, the most apical point of the dental follicle in the apex region was considered as line ‘3’ (red arrows). Similar consideration was taken for both immature uniradicular and multirooted teeth.</td><td>
</td></tr>
<tr><td>Step 10 – The lengths of the lines were measured by selection ‘Measure’ (red circle).</td><td>
</td></tr>
<tr><td>Step 11 – A results box was produced. Only ‘length’ in the seventh column and line ‘4’ and ‘5’ (crown height and root length) on the fifth and sixth row with one decimal value were included.</td><td>
</td></tr>
<tr><td>Step 12 – All results were transferred into a Microsoft Excel spreadsheet that included the individual codes for each subject, date of birth, date of radiographic exposure, age, sex, and ethnicity.</td><td>
</td></tr>
</table>

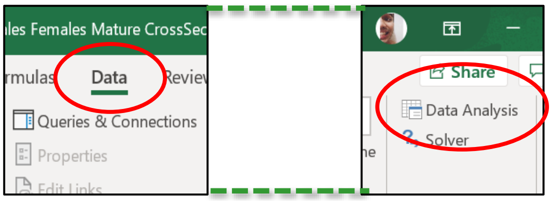
Crown- Root Ratio Measurement	**ImageJ/Microsoft Excel**
Step 13 – Crown-root ratio was calculated by dividing the crown height with the root length.	
Step 14 – Measurements were done for mandibular central incisor, lateral incisor, canine, first premolar, second premolar, first molar, second molar and third molar on the left side.	
Step 15 – Measurements were repeated for maxillary central incisor, lateral incisor, canine, first premolar, second premolar, first molar, second molar and third molar on the left side.	
Step 16 – Data was analysed with *'Data'* → '*Data Analysis*' function.	

Table 13: Crown-root measurement on *ImageJ* and analysis on *Microsoft Excel*

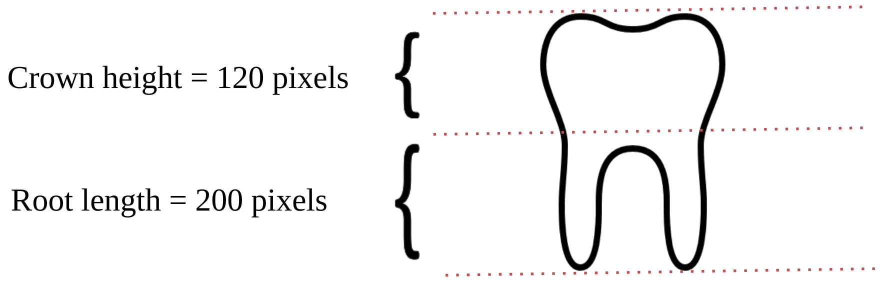
Descriptive analysis was performed to ascertain the means, standard deviations and standard error means for each mandibular and maxillary tooth for males and females. Independent samples t test was also carried out to assess significant differences between male and female subjects for each individual tooth.

3.2.2.2 Crown-root ratio definition for root fraction

Measurements for mature teeth (Rc) crown-root ratio from **3.2.2.1** were used to define the average crown-root ratio for each root fraction described by Nolla's [37] R1/3 and R2/3, and Moorrees' [24] R1/4, R1/2 and R3/4. Formula to calculate the means of each root fraction was:

$$R_{x/y} \text{ crown-root ratio} = \text{crown height}/[\text{root length} \times (x/y)]$$

The following diagram exemplifies the method of calculation for crown-root ratio of each root fraction for mandibular first molar (LL6) according to Moorrees fractions [24]:



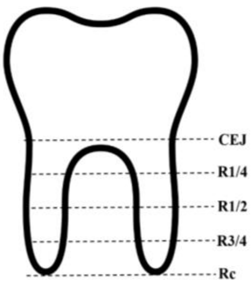


Fig. 6

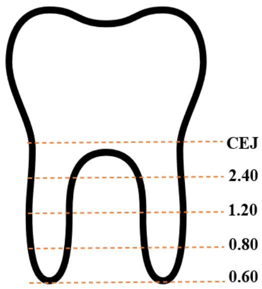


Fig. 7

Figure 6, 7: Crown-root ratio for R1/4, R1/2. R3/4 and Rc of LL6 based on example given

Rc LL6 crown-root ratio $= \text{crown height}/[\text{root length} \times (x/y)] = 120 \text{ pixels}/[200] = 0.60$

R3/4 LL6 crown-root ratio $= \text{crown height}/[\text{root length} \times (x/y)] = 120 \text{ pixels}/[200 \times 3/4] = 0.80$

R1/2 LL6 crown-root ratio $= \text{crown height}/[\text{root length} \times (x/y)] = 120 \text{ pixels}/[200 \times 1/2] = 1.20$

R1/4 LL6 crown-root ratio $= \text{crown height}/[\text{root length} \times (x/y)] = 120 \text{ pixels}/[200 \times 1/4] = 2.40$

Similar formula was used for Nolla's [37] R1/3 and R2/3 root fraction calculations.

3.2.2.3 Upper and lower limits of root fraction

Calculations for mature crown-root ratios in **3.2.2.1** were used to define the average crown-root ratios for each root fraction defined by Nolla's [37] R1/3 and root 2/3, and

Moorrees' [24] R1/4, R1/2 and R3/4. Once the mean crown-root ratio for each mandibular and maxillary root fraction was obtained and analysed, the upper and lower limits were subsequently calculated to discriminate each root fractions. **Figure 8** illustrated root fractions as described by Moorrees [24] and their upper and lower limits.

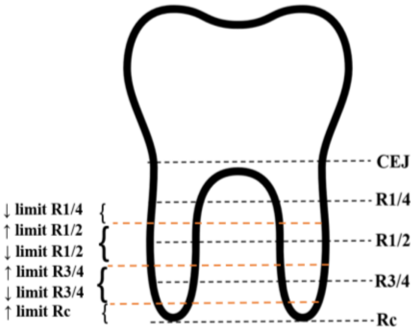


Figure 8: Root fractions as described by Moorrees with definitions of upper and lower limits

Root fractions R1/3 and R2/3 with their respective upper and lower limits were used to represent Nolla's root fractions [37].

3.2.2.3.1 Upper limits of root fraction

Calculation of the upper limit was established by finding the midpoint of two consecutive root fractions. The formula was as follows:

$$Rx_1/y_1 \text{ upper limit} = (Rx_1/y_1 \text{ crown-root ratio} + Rx_2/y_2 \text{ crown-root ratio})/2$$

By using the values obtained in **3.2.2.2**;

R1/2 upper limit = *(R1/2 crown-root ratio+R1/4 crown-root ratio)/2*

= *(1.20+2.40)/2 = 1.80*

R3/4 upper limit = *(R3/4 crown-root ratio+R1/2 crown-root ratio)/2*

= *(0.80+1.20)/2 = 1.00*

Rc upper limit = *(Rc crown-root ratio+R3/4 crown-root ratio)/2*

= *(0.80+0.60)/2 = 0.70*

Similar methodology was applied to calculate Nolla's R1/3 and R2/3 upper limits [37].

3.2.2.3.2 Lower limits of root fraction

Calculation for the lower limit was established by adding the upper limit of the subsequent root fraction by 0.01.

For example, lower limit of R3/4 = upper limit of Rc + 0.01. Formula used was:

$$Rx_1/y_1 \text{ lower limit} = Rx_2/y_2 \text{ upper limit} + 0.01$$

By using the upper limit values obtained in **3.2.2.3.1**,

R1/4 lower limit = *R1/2 upper limit + 0.01*
= *1.80 + 0.01 = 1.81*

R1/2 lower limit = *R3/4 upper limit + 0.01*
= *1.00 + 0.01 = 1.01*

R3/4 lower limit = *Rc upper limit + 0.01*
= *0.70 + 0.01 = 0.71*

Nolla's [37] root fractions' (R1/3 and R2/3) lower limits were also calculated with similar methodology and formula. Summary of root fraction crown root ratios and their upper and lower limits was illustrated in **Figure 9**.

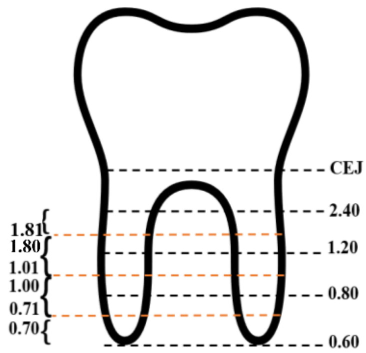


Figure 9: Example of root fraction crown-root ratio values and their upper and lower limits

Stage R1/4 as the most extreme or earliest stage for Moorrees [24] was not limited by, and did not require an upper limit. This is due to the fact that there was no earlier stage than R1/4, and therefore any crown-root ratios measured smaller or less than its lower limit would be considered into stage R1/4. Similar method was applied for Nolla's [37] earliest root fraction stage, R1/3 which was only bounded by the lower limit.

Conversely, the final root stage, stage Rc was limited only by an upper limit, and not a lower limit due to the fact that no more development was expected to take place after stage Rc was achieved. Therefore, crown-root ratio for stage Rc would be the maximum crown-root ratio possible.

3.2.3 Part 3 – Crown-Root Ratio Accuracy

Only Moorrees' [24] staging technique was applied for this stage of the study. All seven permanent teeth on the left side of the mandible and maxilla (central incisor, lateral incisor, canine, first premolar, second premolar, first molar and second molar) were staged as detailed in PART 1 according to stages as defined by Moorrees [24]. Only one series of staging was conducted for this section of the study as reliability of the method was not assessed. Root fraction stages Ac, A1/2 and Rc were combined into one stage, Rc. This is due to the fact that once a root has reached stage root complete Rc, development was already complete and only morphological changes were observed. Thus, all three stages Ac, A1/2 and Rc were categorised as 'mature' or 'root mature', Rc.

Procedure in **Table 13** was repeated to measure the crown height and root length of permanent teeth on the left side of the mandible and maxilla of the same group of test subjects to obtain the crown-root ratio. Descriptive analysis was performed to evaluate the means, standard deviations and standard error means for each mandibular and maxillary tooth for males and females.

Stages were plotted against measured crown-root ratios (as visualised in **Figure 10)** to analyse accuracy of crown-root ratios and their limits as an objective method to predict improve root stage assessment in dental age estimation.

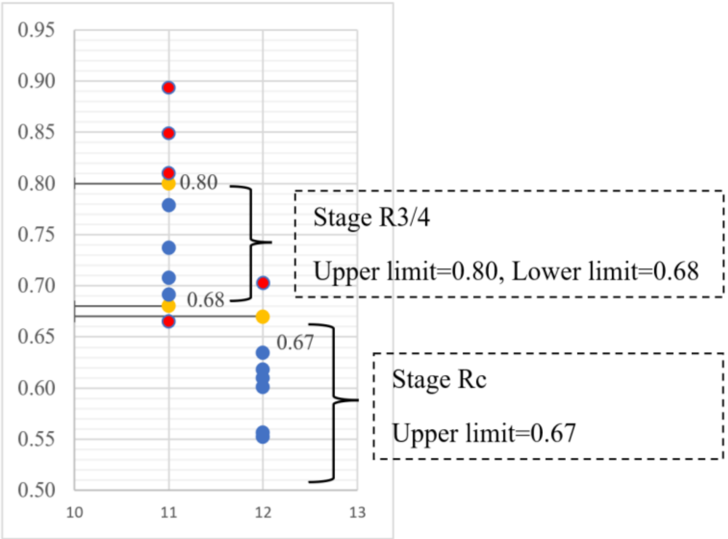


Figure 10: Scatterplot of crown-root ratio against stage. Blue=teeth scored within limits, red=teeth scored outside limits, yellow=upper and lower limits. Stage 10=R1/2, 11=R3/4, 12 and 13 =Rc

Formula to calculate accuracy was as follows:

Rx/y accuracy = Total staged within Rx/y limit(s)/(Total within + outside Rx/y limit(s)) %

The following demonstrates calculation for stage R3/4 (11 on scatterplot) and stage Rc (12 and 13 on scatterplot) accuracy in **Figure 10**:

Stage R3/4 accuracy = *Teeth staged within R3/4 limits / Total teeth in stage R3/4 %*

= *Total of blue / (Total of blue + red) %*

= *4 / (4 + 4) %*

= *50%*

Stage Rc accuracy = *Teeth staged within Rc upper limit / Total teeth in stage Rc %*

= *Total of blue / (Total of blue + red) %*

= *5 / (5 + 1) %*

= *83.3%*

3.2.4 Summary of Methodology

Part 1 – 200 subjects (100 males and 100 females)

Seven permanent teeth on the left side of mandible and maxilla were scored twice to assess reliability of staging techniques

Part 2 – 200 subjects (100 males and 100 females)

Eight permanent teeth on the left side of mandible and maxilla were measured

Average crown-root ratio for mature and root fraction obtained → upper and lower limits established

Part 3 – 62 subjects (17 males and 45 females)

Teeth measured for crown-root ratio and scored according to Moorrees' technique independently

Accuracy of crown-root ratio determined by total number of teeth staged within assigned upper and lower limits

4. RESULTS

4.1 PART 1 –Reliability of Tooth Staging Techniques

4.1.1 Nolla's staging technique – intra observer errors (N=2698)

Intra-observer error or reliability was assessed for each jaw, and both maxilla and mandible combined (**Table 14**) as well as by individual tooth type (**Table 15**, **Figure 11**) by using weighted Kappa coefficient. Reliability for all three readings was excellent. Comparatively, mandibular teeth had slightly higher reliability than their opposing counterparts. Individual assessment of the teeth was also carried out. Maxillary lateral incisor (UL2) was the most reliable with a κ value of 0.944, while maxillary first molar (UL6) proved to be the least reliable with κ =0.832.

			95% Confidence Interval Bound	
Jaw	**Kappa**	**SE**	**Lower**	**Upper**
Maxilla & Mandible	0.922	0.004	0.915	0.930
Maxilla	0.920	0.005	0.910	0.931
Mandible	0.924	0.006	0.913	0.935

Table 14: Kappa values and standard errors for each jaw and combined

			95% Confidence Interval Bound	
Tooth	**Kappa**	**SE**	**Lower**	**Upper**
LL1.1 vs LL1.2	0.913	0.024	0.866	0.960
LL2.1 vs LL2.2	0.875	0.023	0.830	0.920
LL3.1 vs LL3.2	0.902	0.017	0.868	0.936
LL4.1 vs LL4.2	0.927	0.014	0.899	0.954
LL5.1 vs LL5.2	0.910	0.015	0.881	0.940
LL6.1 vs LL6.2	0.939	0.020	0.899	0.978
LL7.1 vs LL7.2	0.905	0.014	0.877	0.932
UL1.1 vs UL1.2	0.919	0.018	0.884	0.955
UL2.1 vs UL2.2	0.944	0.013	0.919	0.970
UL3.1 vs UL3.2	0.915	0.015	0.885	0.944
UL4.1 vs UL4.2	0.915	0.013	0.889	0.940
UL5.1 vs UL5.2	0.929	0.012	0.907	0.952
UL6.1 vs UL6.2	0.832	0.029	0.775	0.890
UL7.1 vs UL7.2	0.889	0.015	0.859	0.919

Table 15: Kappa values and standard errors for mandibular and maxillary teeth

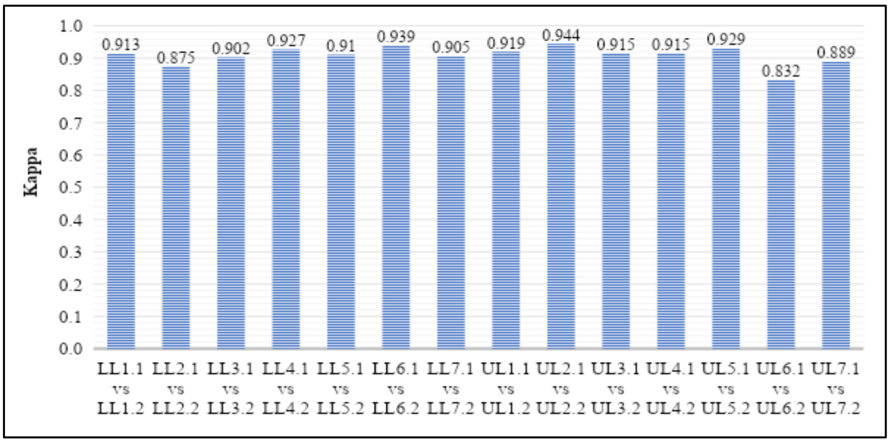


Figure 11: Kappa values for mandibular and maxillary teeth according to Nolla's staging technique

4.1.2 Nolla's staging technique – crosstabulation

Total teeth scored according to Nolla's staging technique was as summarised in **Table 16**. A scoring crosstabulation was constructed to visualise raw agreement between the first and second observations (**Table 17**) and aid percentage agreement calculation and identification of the most and least reliable stages.

	Valid		**Cases Missing**		**Total**	
	N	Percent	N	Percent	N	Percent
Obs1*Obs2	2698	96.40%	102	3.60%	2800	100%

Table 16: Case summary of the first and second observation for all teeth combined based on Nolla's staging technique

		Obs2											
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
Obs1	**0**	10	0	0	0	0	0	0	0	0	0	0	**10**
	1	0	0	0	0	0	0	0	0	0	0	0	**0**
	2	0	0	1	0	0	0	0	0	0	0	0	**1**
	3	0	0	0	2	0	0	0	0	0	0	0	**2**
	4	0	0	0	1	13	3	0	0	0	0	0	**17**
	5	0	0	0	0	20	53	18	0	0	0	0	**91**
	6	0	0	0	0	0	18	259	17	0	0	0	**294**
	7	0	0	0	0	0	0	33	262	19	0	0	**314**
	8	0	0	0	0	0	0	0	42	430	18	0	**490**
	9	0	0	0	0	0	0	0	0	46	278	82	**406**
	10	0	0	0	0	0	0	0	0	0	58	1015	**1073**
	Total	**10**	**0**	**1**	**3**	**33**	**74**	**310**	**321**	**495**	**354**	**1097**	**2698**

Table 17: Crosstabulation for the first and second observation for all teeth combined based on Nolla's staging technique

4.1.3 Nolla's staging technique – percentage agreement

Stage	6	7	8	9	10
Total	259/345	262/373	430/555	278/482	1015/1155
%	75.1	70.2	77.5	57.7	87.9

Table 18: Percentage agreement by stage based on Nolla's staging technique

Readings with fewer than 10% of the total number of teeth staged (stage 0, 1, 2, 3, 4, and 5) were truncated. Stage 10 (apical end of root completed) had the highest percentage agreement (87.9%) while the lowest was displayed by stage 9 (root almost completed – open apex) with 57.7% (**Table 18**, **Figure 12**).

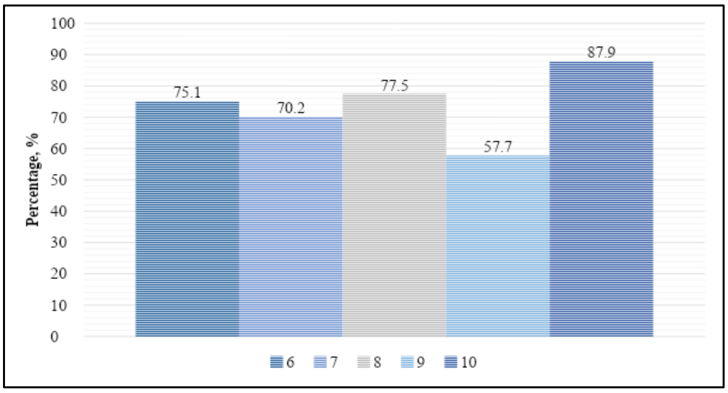


Figure 12: Percentage agreement by stage for all teeth combined based on Nolla's staging technique

4.2.1 Moorrees' staging technique – intra observer errors (N=2674)

			95% Confidence Interval Bound	
Jaw	**Kappa**	**SE**	**Lower**	**Upper**
Maxilla & Mandible	0.938	0.003	0.933	0.943
Maxilla	0.940	0.004	0.933	0.946
Mandible	0.936	0.004	0.928	0.943

Table 19: Kappa values and standard errors for each jaw and combined

Intra-observer error or reliability was assessed for each jaw, and both maxilla and mandible combined (**Table 19**) as well as by individual tooth type (**Table 20**, **Figure 13**) by using weighted Kappa coefficient. In general, Moorrees' technique displayed excellent reliability and better results (κ=0.938) as compared to Nolla's and Demirjian's. Maxillary teeth were also more reliable than that of mandibular, although the difference was minute

(0.004). Individual tooth assessment revealed maxillary lateral incisor (UL2) as the most reliable (κ=0.953) while maxillary first molar (UL6) scored the least with κ=0.817.

			95% Confidence Interval Bound	
Tooth	**Kappa**	**SE**	**Lower**	**Upper**
LL1.1 vs LL1.2	0.864	0.019	0.826	0.901
LL2.1 vs LL2.2	0.902	0.015	0.873	0.931
LL3.1 vs LL3.2	0.917	0.012	0.893	0.941
LL4.1 vs LL4.2	0.943	0.009	0.926	0.960
LL5.1 vs LL5.2	0.946	0.009	0.929	0.963
LL6.1 vs LL6.2	0.891	0.018	0.855	0.928
LL7.1 vs LL7.2	0.934	0.009	0.917	0.951
UL1.1 vs UL1.2	0.938	0.013	0.914	0.963
UL2.1 vs UL2.2	0.953	0.010	0.933	0.972
UL3.1 vs UL3.2	0.945	0.009	0.928	0.962
UL4.1 vs UL4.2	0.943	0.008	0.928	0.959
UL5.1 vs UL5.2	0.944	0.008	0.928	0.960
UL6.1 vs UL6.2	0.817	0.022	0.774	0.860
UL7.1 vs UL7.2	0.932	0.009	0.915	0.949

Table 20: Kappa values and standard errors for mandibular and maxillary teeth

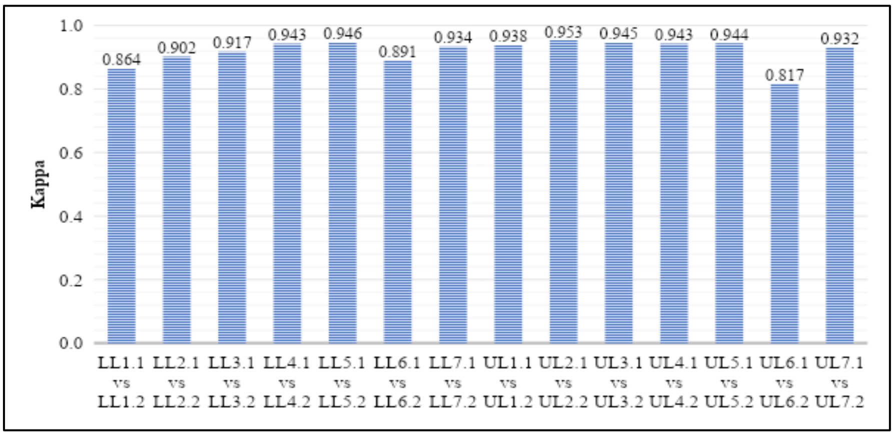


Figure 13: Kappa values for mandibular and maxillary teeth according to Moorrees' staging technique

4.2.2 Moorrees' staging technique – crosstabulation

Total teeth scored according to Moorrees' staging technique was as summarised in **Table 21**. Stages Coc, Cr1/2, Cr3/4, Crc, Ri, and Cli were excluded from the results as the number of scored teeth classified in these stages were fewer than 10% of the total teeth scored. Crosstabulation in **Table 22** showed stage Ac (apical closure complete) as the most reliable where 1002 teeth scored similarly in the first and second scoring while the lowest was stage A1/2 (apex half closed) with 136. Among the root fraction stages, R3/4 (root length 3/4) was the most reliable with N=298, followed by stage R1/4 (root length 1/4) and R1/2 (root length 1/2).

	Valid		**Cases Missing**		**Total**	
	N	Percent	N	Percent	N	Percent
Obs1*Obs2	2674	95.5%	126	4.5%	2800	100%

Table 21: Case summary of the first and second observation for all teeth combined based on Moorrees' staging technique

		Obs2												
		Coc	**Cr1/2**	**Cr3/4**	**Crc**	**Ri**	**Cli**	**R1/4**	**R1/2**	**R3/4**	**Rc**	**A1/2**	**Ac**	**Total**
Obs1	**Coc**	2	0	0	0	0	0	0	0	0	0	0	0	**2**
	Cr1/2	0	3	3	0	0	0	0	0	0	0	0	0	**6**
	Cr3/4	0	2	60	8	0	0	0	0	0	0	0	0	**70**
	Crc	0	0	20	103	19	0	0	0	0	0	0	0	**142**
	Ri	0	0	0	45	113	7	0	0	0	0	0	0	**165**
	Cli	0	0	0	0	4	17	5	0	0	0	0	0	**26**
	R1/4	0	0	0	0	0	3	160	40	0	0	0	0	**203**
	R1/2	0	0	0	0	0	0	33	141	46	0	0	0	**220**
	R3/4	0	0	0	0	0	0	0	17	298	27	0	0	**342**
	Rc	0	0	0	0	0	0	0	0	35	144	59	0	**238**
	A1/2	0	0	0	0	0	0	0	0	0	27	136	38	**201**
	Ac	0	0	0	0	0	0	0	0	0	0	57	1002	**1059**
	Total	**2**	**5**	**83**	**156**	**136**	**27**	**198**	**198**	**379**	**198**	**252**	**1040**	**2674**

Table 22: Crosstabulation for the first and second observation for all teeth combined based on Moorrees' staging technique

4.2.3 Moorrees' staging technique – percentage agreement

Percentage agreement calculated through crosstabulation in **Table 22** and summarised in **Table 23**. Percentages ranged from 42.9% (A1/2) to 91.3% (Ac). Root fraction stages were 66.4% (R1/4), 50.9% (R1/2) and 70.4% (R3/4).

Stage	**R1/4**	**R1/2**	**R3/4**	**Rc**	**A1/2**	**Ac**
Total	160/241	141/277	298/423	144/292	136/317	1002/1097
%	66.4	50.9	70.4	49.3	42.9	91.3

Table 23: Percentage agreement by stage based on Moorrees' staging technique

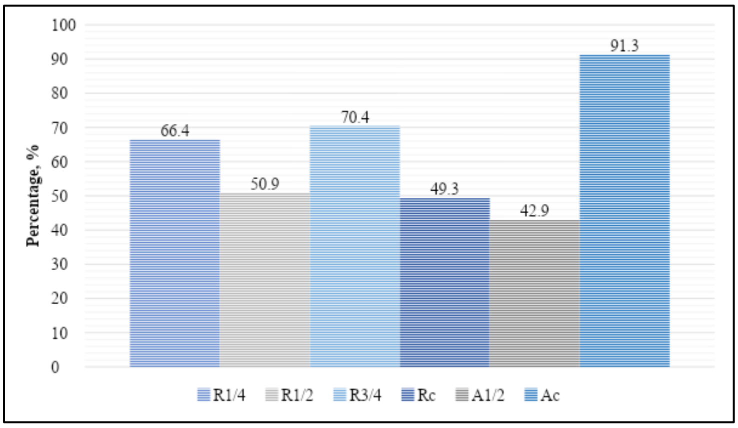


Figure 14: Percentage agreement by stage for all teeth combined based on Moorrees' staging technique

4.3.1 Demirjian's staging technique – intra observer errors - (N=2682)

			95% Confidence Interval Bound	
Jaw	**Kappa**	**SE**	**Lower**	**Upper**
Maxilla & Mandible	0.918	0.004	0.910	0.926
Maxilla	0.910	0.006	0.899	0.922
Mandible	0.927	0.006	0.916	0.938

Table 24: Kappa values and standard errors for each jaw and combined

			95% Confidence Interval Bound	
Tooth	**Kappa**	**SE**	**Lower**	**Upper**
LL1.1 vs LL1.2	0.902	0.025	0.852	0.951
LL2.1 vs LL2.2	0.896	0.022	0.853	0.939
LL3.1 vs LL3.2	0.922	0.016	0.89	0.954
LL4.1 vs LL4.2	0.895	0.017	0.862	0.929
LL5.1 vs LL5.2	0.903	0.015	0.873	0.932
LL6.1 vs LL6.2	0.931	0.022	0.888	0.973
LL7.1 vs LL7.2	0.940	0.012	0.916	0.964
UL1.1 vs UL1.2	0.905	0.021	0.864	0.946
UL2.1 vs UL2.2	0.898	0.019	0.861	0.935
UL3.1 vs UL3.2	0.915	0.016	0.884	0.946
UL4.1 vs UL4.2	0.895	0.015	0.865	0.924
UL5.1 vs UL5.2	0.913	0.014	0.886	0.939
UL6.1 vs UL6.2	0.837	0.030	0.777	0.896
UL7.1 vs UL7.2	0.905	0.015	0.876	0.933

Table 25: Kappa values and standard errors for mandibular and maxillary teeth

Intra-observer error or reliability was assessed for each jaw, and both maxilla and mandible combined (**Table 24**) as well as by individual tooth type (**Table 25**, **Figure 15**) by using weighted Kappa coefficient. Demirjian's staging technique produced the lowest reliability among the other techniques with κ=0.918. However, reliability was still excellent. Mandibular teeth were more reliable than maxillary, with a small difference of 0.017. Mandibular second molar (LL7) was the most reliable compared to the other permanent teeth with κ=0.940 while maxillary first molar (UL6) recorded the lowest Kappa value with 0.837.

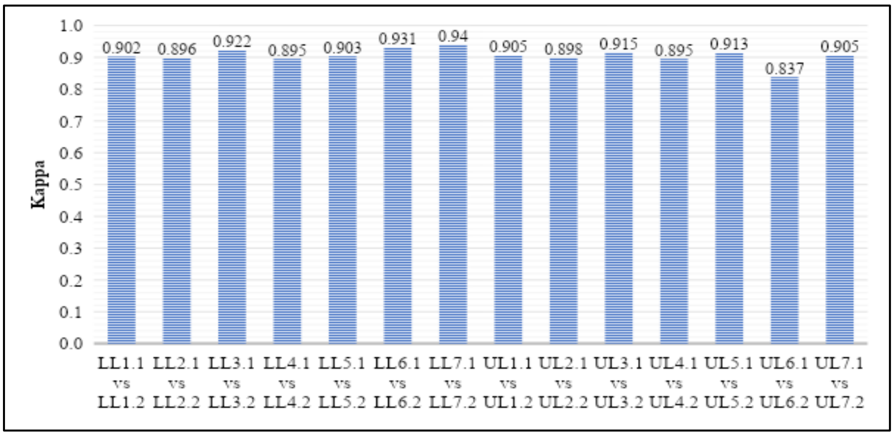


Figure 15: Kappa values for mandibular and maxillary teeth according to Demirjian's staging technique

4.3.2 Demirjian's staging technique – crosstabulation

	Valid		Cases Missing		Total	
	N	Percent	N	Percent	N	Percent
Obs1*Obs2	2682	96.1%	110	3.9%	2792	100%

Table 26: Case summary of the first and second observation for all teeth combined based on Demirjian's staging technique

		Obs2								
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**Total**
Obs1	**A**	0	0	0	0	0	0	0	0	**0**
	B	0	3	1	0	0	0	0	0	**4**
	C	0	0	72	30	0	0	0	0	**102**
	D	0	0	19	109	42	0	0	0	**170**
	E	0	0	0	17	328	36	0	0	**381**
	F	0	0	0	0	23	498	29	0	**550**
	G	0	0	0	0	0	38	337	43	**418**
	H	0	0	0	0	0	0	71	986	**1057**
	Total	**0**	**3**	**92**	**156**	**393**	**572**	**437**	**1029**	**2682**

Table 27: Crosstabulation for the first and second observation for all teeth combined based on Demirjian's staging technique

Total teeth scored according to Demirjian's staging technique was crosstabulated and summarised in **Tables 26** and **27**. As there were fewer than 268 teeth (< 10%) scored in stages B, C and D, these stages were not considered. Stage H (apical end closed) had the highest raw agreement (986) followed by stages F (root length ≥ crown height) with 498, 337 for G (apical end partially open) 337 and E (root length ≤ crown height) with 328.

4.3.3 Demirjian's staging technique – percentage agreement

Stage	E	F	G	H
Total	328/446	498/624	337/518	986/1100
%	73.5	79.8	65.1	89.6

Table 28: Percentage agreement by stage based on Demirjian's staging technique

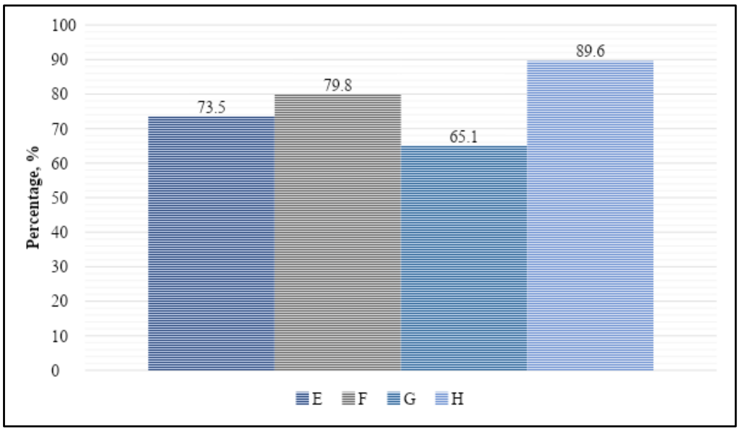


Figure 16: Percentage agreement by stage for all teeth combined based on Demirjian's staging technique

Stage H (apical end closed) ranked first as the most reliable with 89.6% agreement while stage G (apical end partially open) on the other hand was last with 65.1%. Stages E (root length ≤ crown height) and F (root length ≥ crown height) generated good and almost similar agreements with 73.5% and 79.8%, respectively (**Table 28**, **Figure 16**).

4.2 Part 2 - Crown-Root Ratio

4.2.1 Crown-root definition for mature teeth

Descriptive analysis was carried out to define the means of all mature permanent teeth including left maxillary and mandibular central incisor, lateral incisor, canine, first premolar, second premolar, first molar, second molar, and third molar separately for males and females. The results for mean, standard deviation and standard error of each mandibular and maxillary tooth were tabulated in **Table 29** and shown in **Figure 17**.

	Male				**Female**			
Tooth	**N**	**Mean**	**SD**	**SE**	**N**	**Mean**	**SD**	**SE**
LL1	87	0.57	0.080	0.009	84	0.59	0.078	0.009
LL2	92	0.58	0.085	0.009	95	0.60	0.072	0.007
LL3	95	0.55	0.081	0.008	95	0.57	0.076	0.008
LL4	97	0.53	0.080	0.008	98	0.58	0.069	0.007
LL5	99	0.49	0.095	0.010	95	0.52	0.060	0.006
LL6	100	0.57	0.071	0.007	98	0.60	0.069	0.007
LL7	98	0.63	0.085	0.009	100	0.68	0.094	0.009
LL8	93	1.01	0.395	0.042	89	1.10	0.328	0.038
UL1	91	0.58	0.088	0.009	91	0.61	0.080	0.009
UL2	96	0.52	0.080	0.008	93	0.58	0.066	0.007
UL3	98	0.51	0.074	0.008	98	0.53	0.069	0.007
UL4	92	0.59	0.080	0.008	96	0.60	0.076	0.008
UL5	95	0.55	0.079	0.008	95	0.58	0.075	0.008
UL6	98	0.60	0.096	0.010	96	0.60	0.085	0.009
UL7	100	0.62	0.092	0.009	99	0.63	0.076	0.008
UL8	90	0.92	0.475	0.050	76	0.97	0.407	0.047

Table 29: Crown-root ratio definition for mature mandibular and maxillary permanent teeth

Crown height and root length of 3019 mature mandibular and maxillary permanent teeth on the left side were measured. Crown-root ratio was derived by dividing the crown height and root length. According to **Table 29**, the third molars (UL8 and LL8) had the highest crown-root ratio mean for both males and females, alike. Apart from the third molars, crown-root ratio means of the other maxillary and mandibular teeth were relatively uniformed, ranging between 0.49–0.68. Males' mandibular first premolar (LL5) presented as the smallest ratio, 0.49 while the highest crown-root ratio, 0.68 was observed in mandibular second molar (LL7) of the female group. Standard deviations for the third molars of both sex groups were also among the highest (0.33 – 0.48).

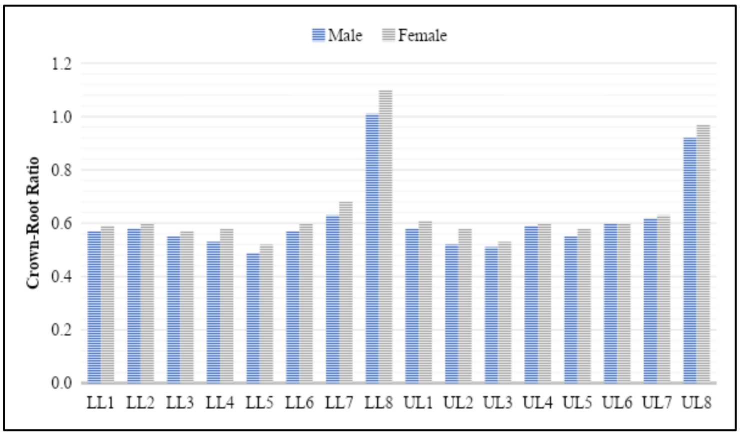


Figure 17: Mean crown-root ratio of mature mandibular and maxillary teeth

Boxplots were produced to visualise the distribution and variation of crown-root ratio for all mandibular and maxillary as seen in **Figures 18, 19, 22** and **23**.

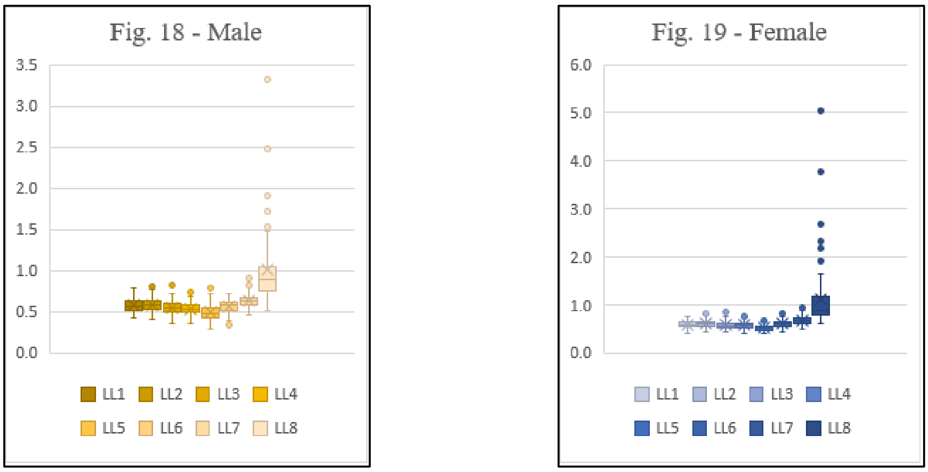


Figure 18 & 19: Boxplot of mandibular teeth crown-root ratio of male and female subjects

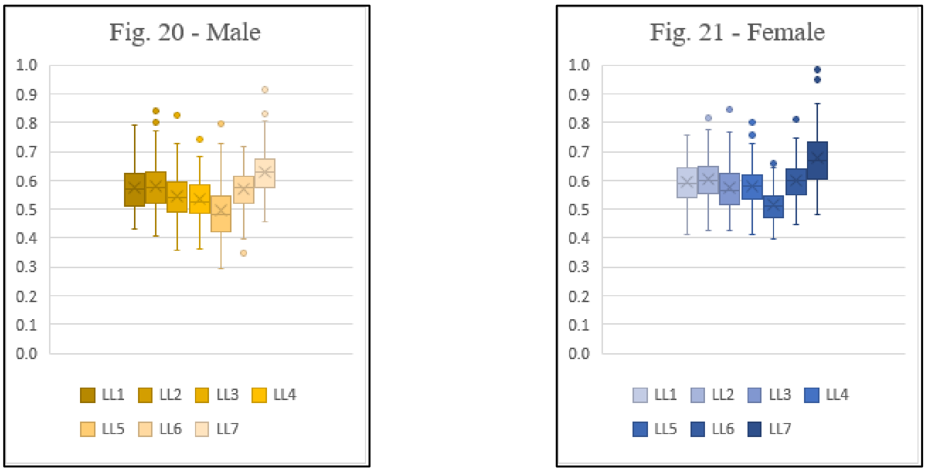


Figure 20 & 21: Boxplot of male and female subjects' mandibular teeth crown-root ratio (third molars excluded)

Extensive size of the mandibular and maxillary third molars boxes and whiskers indicated high variations and this was expressed by both male and female subjects. Comparatively, the boxplots of the other seven permanent teeth were significantly shorter and uniformed demonstrating consistency with small variations. This was evident in **Figures 20, 21, 24** and **25** when third molars were omitted from the data. Collectively, variations between mandibular and maxillary seven permanent teeth (LL1 – LL7, UL1 – UL7) were almost identical, and differences were minute.

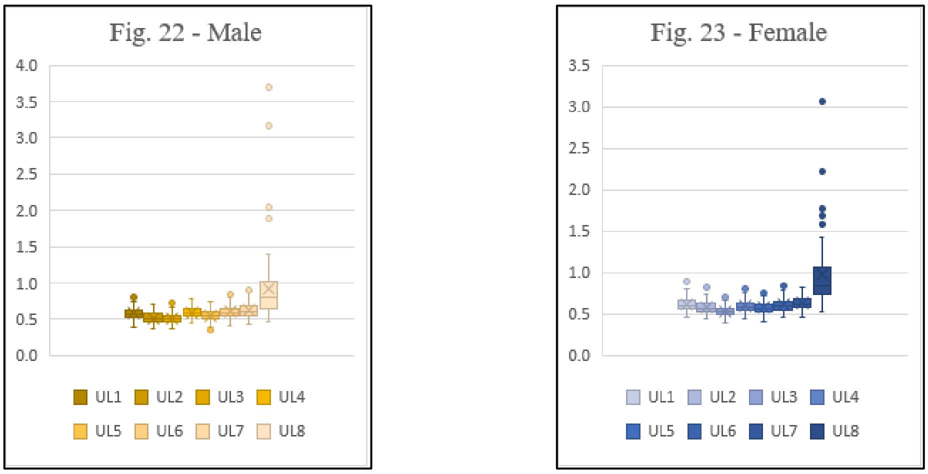


Figure 22 & 23: Boxplot of maxillary teeth crown-root ratio of male and female subjects

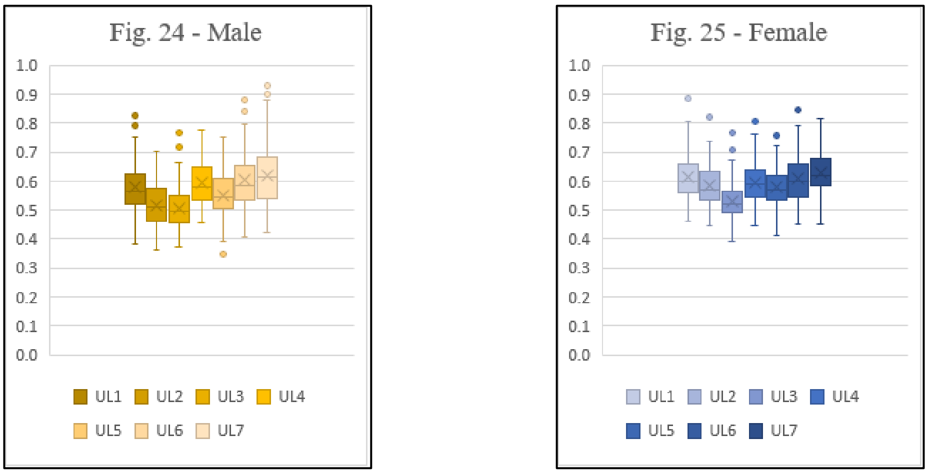


Figure 24 & 25: Boxplot of male and female subjects' maxillary teeth crown-root ratio (third molars excluded)

Medians of each tooth could also be observed and distinguished more clearly in these figures, ranging less than 0.20 between highest and lowest medians. Lowest medians for mandibular and maxillary teeth were represented by LL5 and UL3, respectively. **Figures**

20, 21, 24 and **25** also enabled skewness of each tooth to be analysed and the first seven permanent teeth proved to be normally distributed as compared to the third molars, which were negatively skewed in all cases**.**

4.2.2 Crown height measurement

Crown height was defined as the length between the highest cusp of a tooth and its CEJ. **Table 30** and **Figure 26** provided a detailed summary of mean, standard deviation and standard error of mandibular and maxillary crown heights as measured in pixels for both male and female subjects. Shortest crown height means were LL1 in both sex groups (males=88.5, females=84.3) while the highest on the other hand, was observed in UL3 (males=120.1, females=114.3). Variations ranged between 6.97 (LL7, males) to 14.43 (LL3, males).

	Male				**Female**			
Tooth	**N**	**Mean**	**SD**	**SE**	**N**	**Mean**	**SD**	**SE**
LL1	87	88.5	13.054	1.400	84	84.3	10.243	1.118
LL2	92	92.6	12.394	1.292	95	89.6	9.283	0.952
LL3	95	111.0	14.429	1.480	95	104.0	9.824	1.008
LL4	97	100.2	11.583	1.176	98	103.5	8.837	0.893
LL5	99	98.9	11.593	1.165	95	100.3	7.340	0.753
LL6	100	104.6	8.280	0.828	98	106.1	7.318	0.739
LL7	98	104.9	6.973	0.704	100	107.7	7.722	0.772
LL8	93	103.6	9.469	0.982	89	104.7	10.196	1.081
UL1	91	113.4	11.187	1.173	91	112.5	9.187	0.963
UL2	96	101.1	9.760	0.996	93	103.6	9.881	1.025
UL3	98	120.1	12.198	1.232	98	114.3	9.920	1.002
UL4	92	113.0	10.150	1.058	96	108.9	9.374	0.957
UL5	95	104.5	10.082	1.034	95	103.9	9.816	1.007
UL6	98	109.1	10.315	1.042	96	108.4	8.516	0.869
UL7	100	103.6	8.337	0.834	99	104.9	9.063	0.911
UL8	90	105.3	13.361	1.408	76	108.1	11.283	1.294

Table 30: Mean crown height (in pixel) for mandibular and maxillary teeth in male and female subjects

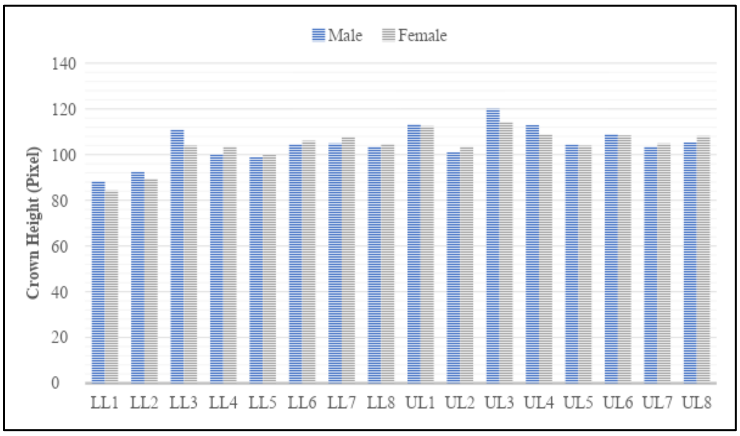
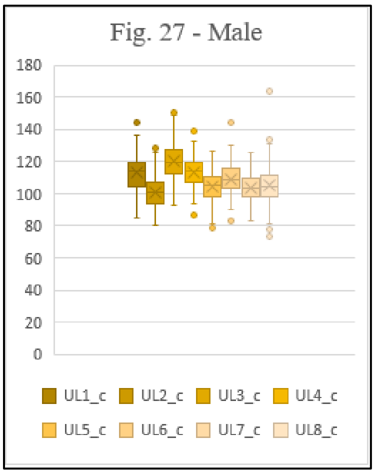


Figure 26: Mean crown height (in pixel) of mature mandibular and maxillary teeth



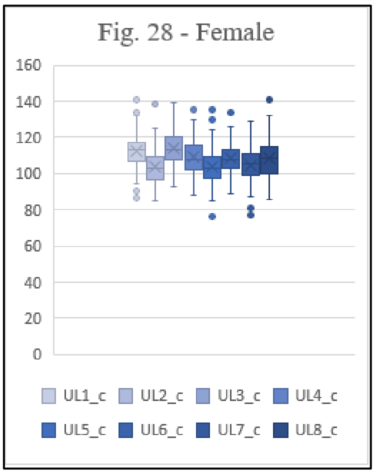
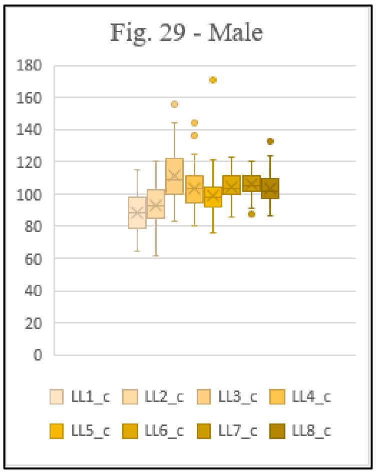


Figure 27 & 28: Boxplot of maxillary crown height (in pixel) of male and female subjects



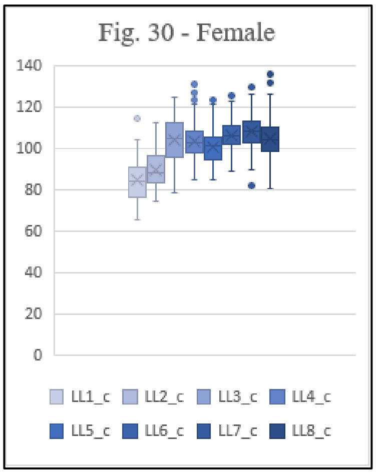


Figure 29 & 30: Boxplot of mandibular crown height (in pixel) of male and female subjects

Overall, lower incisors had considerably shorter crown heights compared to the other permanent teeth. However, differences were minimal and the mean heights of the remaining permanent teeth were homogenous, particularly observed in the upper dentition. Variations were also more consistent in maxillary teeth. Their opposing counterparts however, were more diverse in variation, median and interquartile range (**Figures 27–30**).

4.2.3 Root length measurement

	Male				Female			
Tooth	**N**	**Mean**	**SD**	**SE**	**N**	**Mean**	**SD**	**SE**
LL1	87	155.6	20.373	2.184	84	143.7	20.201	2.204
LL2	92	161.2	22.929	2.391	95	150.4	21.125	2.167
LL3	95	205.0	25.216	2.587	95	183.8	24.658	2.530
LL4	100	189.9	21.230	2.156	98	180.2	21.613	2.183
LL5	99	204.3	28.497	2.864	95	196.9	22.780	2.337
LL6	100	186.1	18.682	1.868	98	178.7	18.346	1.853
LL7	98	170.7	19.432	1.963	100	161.6	20.084	2.008
LL8	93	113.1	27.134	2.814	89	110.4	30.111	3.192
UL1	91	198.6	26.466	2.774	91	185.4	18.792	1.970
UL2	96	199.0	23.469	2.395	93	178.6	17.304	1.794
UL3	98	240.0	28.979	2.927	98	218.2	25.970	2.623
UL4	92	192.4	21.407	2.232	96	184.7	20.675	2.110
UL5	95	191.7	23.143	2.374	95	180.6	19.944	2.046
UL6	98	183.7	22.253	2.248	96	180.1	21.091	2.153
UL7	100	170.3	21.194	2.119	99	167.8	18.960	1.906
UL8	90	128.6	32.343	3.409	76	122.7	30.626	3.513

Table 31: Mean root length (in pixel) for mandibular and maxillary teeth in male and female subjects

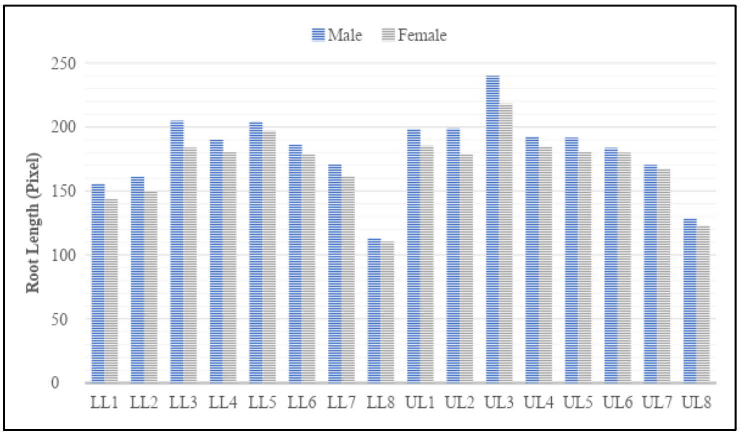
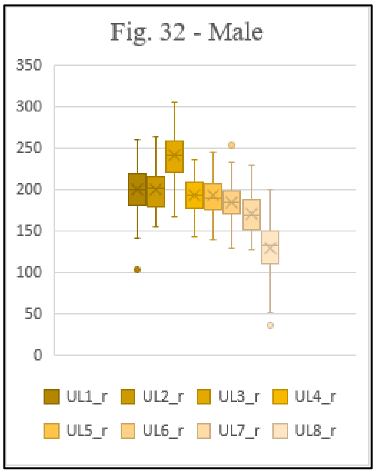


Figure 31: Mean root length (in pixel) of mature mandibular and maxillary teeth

Root length was defined as the length measured between the CEJ and the apex (or the apex of longest root in multirooted teeth). Range of root length means was much wider than that of crown heights where the means fluctuated between 113.1 to 240.0 in males (LL8 and UL3, respectively) and 110.4 to 218.2 in female subjects of similar teeth (**Table 31, Figure 31**). Similar results were observed in the spread of data where magnitude of root length standard deviation was as high as three times than that of crown height's.



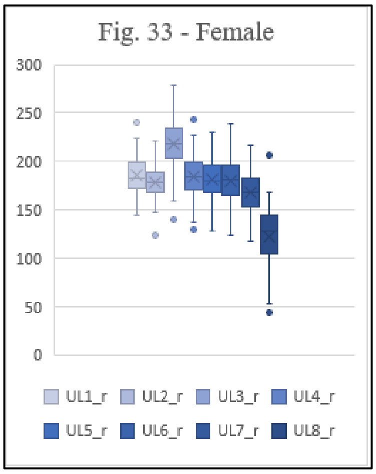
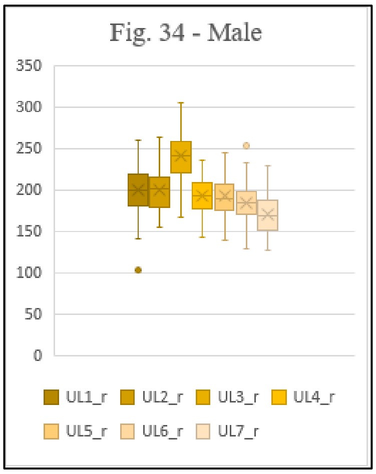


Figure 32 & 33: Boxplot of maxillary root length (in pixel) of male and female subjects



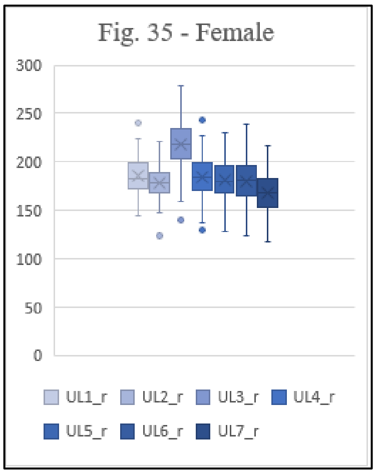
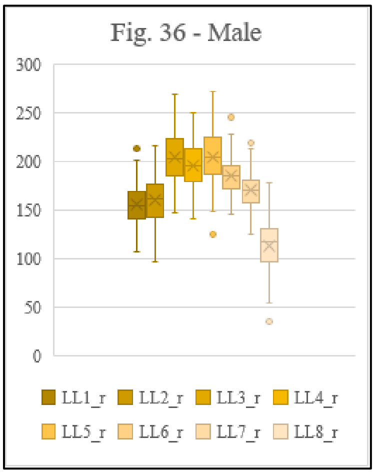


Figure 34 & 35: Boxplot of maxillary root length (in pixel) of male and female subjects (third molars excluded)

UL8 displayed significantly higher variations in both sex classes succeeded by UL3, while UL7 and UL2 presented with the lowest in males and females, respectively (**Figures 32, 33**). Clear differences could also be observed between the highest (UL3) and lowest medians (UL8) in both groups while the remaining maxillary teeth had otherwise consistent values. However, these differences in medians and variations did not greatly affect the homogeneity when UL8 were removed from the data (**Figures 34, 35)**.



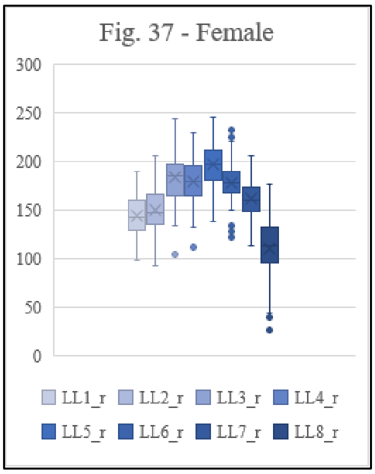
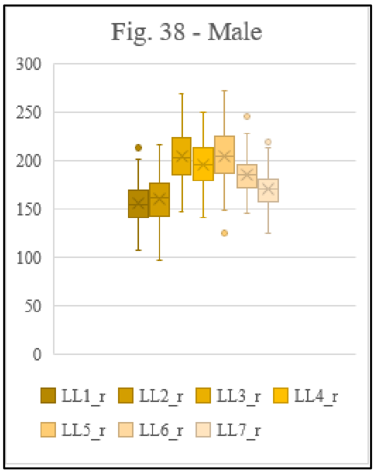


Figure 36 & 37: Boxplot of mandibular crown height (in pixel) of male and female subjects



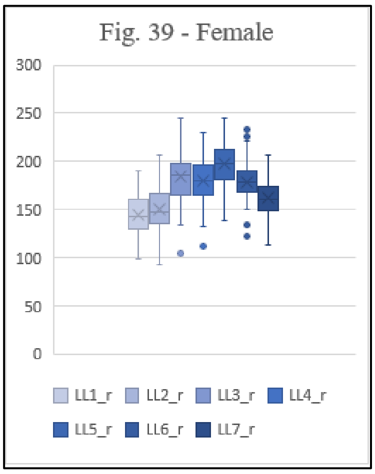


Figure 38 & 39: Boxplot of mandibular crown height (in pixel) of male and female subjects (third molars excluded)

Similar pattern of pits and peaks was displayed by mandibular teeth where LL5 (males) and LL8 (females) exhibited highest variations while LL6 of males and females were the least variable. Although LL8 medians were clearly the lowest, removal from the data had little to no effect to the pattern/s homogeneity.

4.2.4 T–test: paired two sample for means – sex

Two-sample t–test (also known as independent samples t–test) was used to determine whether the means of the two sets of data were statistically different from one another. Each tooth was analysed separately for differences between two sex groups and the results were summarised in **Table 32**. Differences were considered significant when p value (2-tailed) was less than 0.05 ($p<0.05$). Sex differences were more prevalent in mandibular teeth, LL2, LL3, LL4, LL6, and LL7 as compared to maxillary (UL1, UL2 and UL5). However, p values for LL2 and LL3 were considerably higher, indicating less significance.

	Male				Female				
Tooth	**N**	**Mean**	**SD**	**SE**	**N**	**Mean**	**SD**	**SE**	**Sig (2-tailed)**
LL1	84	0.573	0.078	0.009	84	0.582	0.081	0.009	0.070
LL2	92	0.581	0.085	0.009	92	0.604	0.073	0.008	0.022
LL3	95	0.547	0.081	0.008	95	0.574	0.076	0.008	0.027
LL4	97	0.533	0.078	0.008	97	0.581	0.069	0.007	0.000
LL5	95	0.495	0.095	0.010	95	0.515	0.060	0.006	0.074
LL6	98	0.567	0.071	0.007	98	0.599	0.069	0.007	0.002
LL7	98	0.629	0.085	0.009	98	0.678	0.094	0.010	0.000
LL8	89	1.012	0.397	0.042	89	1.102	0.693	0.073	0.087
UL1	91	0.580	0.086	0.009	91	0.613	0.078	0.008	0.003
UL2	93	0.516	0.080	0.008	93	0.584	0.066	0.007	0.000
UL3	98	0.507	0.075	0.008	98	0.530	0.070	0.007	0.051
UL4	92	0.594	0.080	0.008	92	0.597	0.076	0.008	0.802
UL5	95	0.552	0.079	0.008	95	0.581	0.075	0.008	0.005
UL6	96	0.604	0.097	0.010	96	0.610	0.085	0.009	0.668
UL7	99	0.619	0.093	0.009	99	0.631	0.076	0.008	0.334
UL8	76	0.920	0.432	0.050	76	0.969	0.407	0.047	0.135

Table 32: P values between males and females for each mandibular and maxillary tooth

4.2.5 T–test: paired two sample for means – jaw

Independent samples t–test conducted for mandibular-maxillary teeth comparison showed that among male subjects, all five sets of pairing (LL2-UL2, LL3- UL3, LL4-UL4, LL5- UL5, and LL6- UL6) showed very strong distinctions as compared to three in females (LL3- UL3, LL5- UL5, and LL7- UL7). LL2-UL2 also showed significant difference, though marginally lower.

Sex	**LL**	**N**	**Mean**	**SD**	**SE**	**UL**	**N**	**Mean**	**SD**	**SE**	**Sig (2-tailed)**
M	**1**	87	0.573	0.080	0.009	**1**	87	0.582	0.088	0.009	0.476
	2	92	0.581	0.085	0.009	**2**	92	0.517	0.080	0.008	0.000
	3	95	0.547	0.081	0.008	**3**	95	0.505	0.074	0.008	0.000

	4	92	0.533	0.080	0.008	**4**	92	0.594	0.080	0.008	0.000
	5	95	0.495	0.095	0.010	**5**	95	0.552	0.079	0.008	0.000
	6	98	0.567	0.071	0.007	**6**	98	0.603	0.096	0.010	0.001
	7	98	0.629	0.085	0.009	**7**	98	0.618	0.092	0.009	0.300
	8	90	0.972	0.395	0.042	**8**	90	0.915	0.475	0.050	0.380
F	**1**	84	0.593	0.078	0.009	**1**	84	0.615	0.080	0.009	0.053
	2	93	0.604	0.072	0.007	**2**	93	0.584	0.066	0.007	0.036
	3	95	0.574	0.076	0.008	**3**	95	0.532	0.069	0.007	0.000
	4	96	0.581	0.069	0.007	**4**	96	0.596	0.076	0.008	0.135
	5	95	0.515	0.060	0.006	**5**	95	0.581	0.075	0.008	0.000
	6	96	0.601	0.069	0.007	**6**	96	0.610	0.085	0.009	0.345
	7	99	0.677	0.094	0.009	**7**	99	0.631	0.076	0.008	0.000
	8	76	0.964	0.328	0.038	**8**	76	0.969	0.407	0.047	0.937

Table 33: P values of mandibular-maxillary teeth comparison for both sex groups. M= Male, F=Female, LL=Mandibular, UL=Maxillary

4.2.6 T–test: paired two sample for means – crown height

Sex	LL	N	Mean	SD	SE	LL	N	Mean	SD	SE	Sig (2-tailed)
M	**6**	98	104.6	8.34	0.843	7	98	105.9	6.97	0.704	0.141
	6	93	104.4	8.37	0.868	8	93	103.6	9.47	0.982	0.515
	7	93	105.9	7.00	0.726	8	93	103.6	9.47	0.982	0.050
F	**6**	98	106.1	7.32	0.739	7	98	107.8	7.68	0.776	0.101
	6	89	106.6	7.25	0.768	8	89	104.7	10.20	1.081	0.093
	7	89	107.9	7.64	0.810	8	89	104.7	10.20	1.081	0.013

Table 34: P value of mandibular molars crown height comparison (in pixel). M= Male, F=Female, LL=Mandibular, UL=Maxillary

Sex	UL	N	Mean	SD	SE	UL	N	Mean	SD	SE	Sig (2-tailed)
M	**6**	98	104.6	8.34	0.843	7	98	105.9	6.97	0.704	0.141
	6	93	104.4	8.37	0.868	8	93	103.6	9.47	0.982	0.515
	7	93	105.9	7.00	0.726	8	93	103.6	9.47	0.982	0.050
F	**6**	98	106.1	7.32	0.739	7	98	107.8	7.68	0.776	0.101
	6	89	106.6	7.25	0.768	8	89	104.7	10.20	1.081	0.093
	7	89	107.9	7.64	0.810	8	89	104.7	10.20	1.081	0.013

Table 35: P value of maxillary molars crown height comparison (in pixel). M= Male, F=Female, LL=Mandibular, UL=Maxillary

Differences between posterior teeth crown heights were evaluated through a two-sample t–test (**Tables 34, 35**). Mandibular teeth comparisons showed that none of the male combinations demonstrated any significant correlations, while in females, only the difference between LL7 and LL8 was significant. Upon maxillary analysis, UL6 indicated significant differences with UL7 in both sex classes as well as UL8 in males.

4.2.7 T–test: paired two sample for means – root length

Contrary to findings in **4.2.6**, all 12 sets of posterior teeth root length combination of the male and female groups demonstrated exceptionally significant differences where all p values were less than 0.01 ($p<0.01$) (**Tables 36, 37**).

Sex	LL	N	Mean	SD	SE	LL	N	Mean	SD	SE	Sig (2-tailed)
M	6	98	186.3	18.83	1.903	7	98	170.7	19.43	1.963	0.000
	6	93	186.2	19.24	1.995	8	93	113.1	27.13	2.814	0.000
	7	93	170.4	19.25	1.996	8	93	113.1	27.13	2.814	0.000
F	6	98	178.7	18.35	1.853	7	98	161.4	20.19	2.039	0.000
	6	89	178.2	18.22	1.932	8	89	110.4	30.11	3.192	0.000
	7	89	161.5	20.84	2.209	8	89	110.4	30.11	3.192	0.000

Table 36: P value of mandibular molars root length comparison (in pixel). M= Male, F=Female, LL=Mandibular, UL=Maxillary

Sex	UL	N	Mean	SD	SE	UL	N	Mean	SD	SE	Sig (2-tailed)
M	6	98	183.7	22.25	2.248	7	98	170.5	21.01	2.123	0.000
	6	90	182.6	22.64	2.386	8	90	128.6	32.34	3.409	0.000
	7	90	103.6	8.48	0.894	8	90	105.3	13.36	1.408	0.000
F	6	96	180.1	21.09	2.153	7	96	167.5	18.91	1.930	0.000
	6	76	177.9	20.10	2.306	8	76	122.7	30.63	3.513	0.000
	7	76	166.1	18.19	2.086	8	76	122.7	30.63	3.513	0.000

Table 37: P value of maxillary molars root length comparison (in pixel). M= Male, F=Female, LL=Mandibular, UL=Maxillary

4.2.8 Root fraction definition

Table 38, Figures 40 and **41** provided a detailed summary of crown-root ratio for all root fractions and their collective patterns as described by Nolla's and Moorrees' staging methods for both sex groups. Previous data collated in **Table 36** were employed to define the average R1/4, R1/2 and R3/4 for Moorrees', and R1/3, R2/3 for Nolla's crown-root ratio for each mandibular and maxillary tooth.

		Nolla		Moorrees			Both
Tooth	**Sex**	**R1/3**	**R2/3**	**R1/4**	**R1/2**	**R3/4**	**Rc**
LL1	Male	1.72	0.85	2.29	1.15	0.76	0.57
	Female	1.78	0.89	2.37	1.19	0.79	0.59
LL2	Male	1.74	0.87	2.32	1.16	0.77	0.58
	Female	1.81	0.90	2.41	1.21	0.80	0.60
LL3	Male	1.65	0.82	2.19	1.09	0.73	0.55
	Female	1.72	0.86	2.29	1.15	0.76	0.57
LL4	Male	1.60	0.80	2.13	1.07	0.71	0.53
	Female	1.74	0.87	2.32	1.16	0.77	0.58
LL5	Male	1.48	0.74	1.98	0.99	0.66	0.49
	Female	1.55	0.77	2.06	1.03	0.69	0.52
LL6	Male	1.70	0.85	2.27	1.13	0.76	0.57
	Female	1.80	0.90	2.40	1.20	0.80	0.60
LL7	Male	1.89	0.94	2.51	1.26	0.84	0.63
	Female	2.03	1.01	2.71	1.35	0.9	0.68
LL8	Male	3.03	1.51	4.04	2.02	1.35	1.01
	Female	3.30	1.65	4.41	2.20	1.47	1.10
UL1	Male	1.74	0.87	2.32	1.16	0.77	0.58
	Female	1.84	0.92	2.45	1.23	0.82	0.61
UL2	Male	1.55	0.77	2.06	1.03	0.69	0.52
	Female	1.75	0.89	2.34	1.17	0.78	0.58
UL3	Male	1.52	0.76	2.03	1.01	0.68	0.51
	Female	1.59	0.79	2.20	1.06	0.71	0.53
UL4	Male	1.78	0.89	2.38	1.19	0.79	0.59
	Female	1.79	0.89	2.38	1.19	0.79	0.60
UL5	Male	1.66	0.83	2.21	1.10	0.74	0.55
	Female	1.74	0.87	2.32	1.16	0.77	0.58
UL6	Male	1.81	0.90	2.41	1.21	0.80	0.60
	Female	1.83	0.91	2.44	1.22	0.81	0.60
UL7	Male	1.85	0.93	2.47	1.24	0.82	0.62
	Female	1.89	0.95	2.52	1.26	0.84	0.63
UL8	Male	2.75	1.37	3.66	1.83	1.22	0.92
	Female	2.91	1.45	3.88	1.94	1.29	0.97

Table 38: Crown-root ratio definition for mandibular and maxillary root fraction

As carefully explained previously, calculation for crown-root ratio for each fraction is:

$$R_{x/y} \text{ crown-root ratio} = \text{crown height}/[\text{root length} \times (x/y)]$$

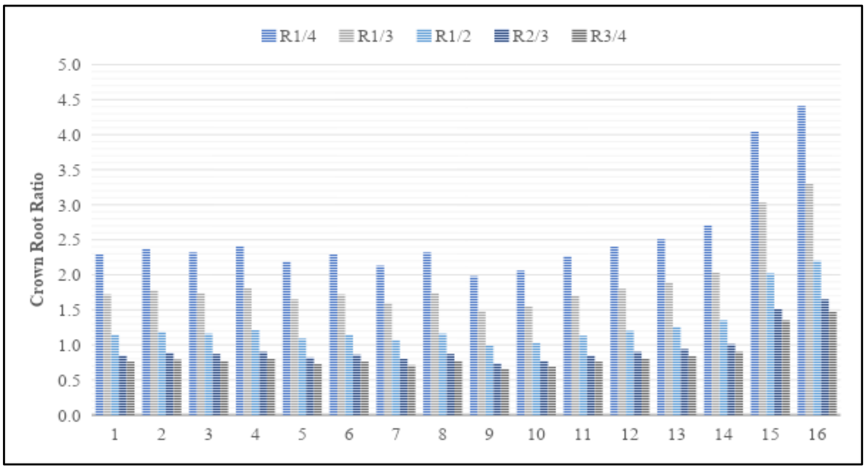


Figure 40: Crown-root ratio definition of mandibular root fraction for male and female subjects

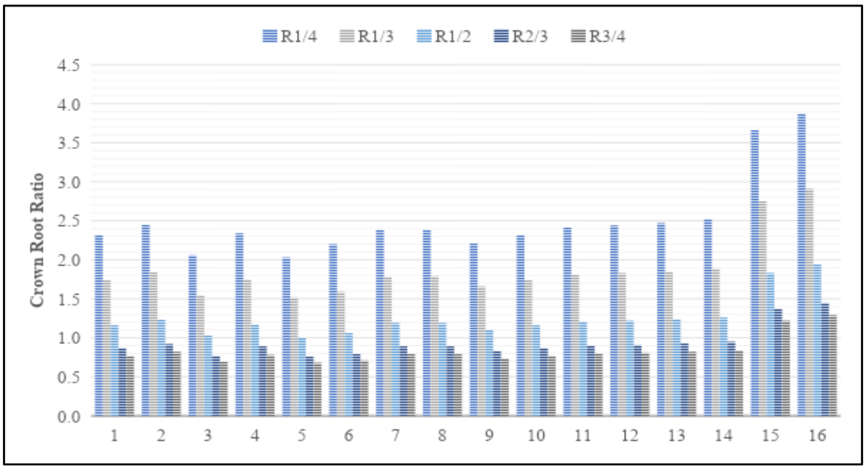


Figure 41: Crown-root ratio definition of maxillary root fraction for male and female subjects

4.2.9 Definition for root fraction – Nolla's staging technique

Through the mathematical formula mentioned previously, information regarding crown-root ratio for root fractions as described by Nolla's staging technique (R1/3. R2/3, and Rc) was extracted from **4.2.8** and displayed in **Table 39.** The data provided was sex-specific, and especially crucial in determining upper and lower limits in the next stage of the study.

Figures 42 – 44 provided examples of crown-root ratio assignments based on type of tooth and sex of the subject.

Tooth	Sex	R1/3	R2/3	Rc
LL1	Male	1.72	0.85	0.57
	Female	1.78	0.89	0.59
LL2	Male	1.74	0.87	0.58
	Female	1.81	0.90	0.60
LL3	Male	1.65	0.82	0.55
	Female	1.72	0.86	0.57
LL4	Male	1.60	0.80	0.53
	Female	1.74	0.87	0.58
LL5	Male	1.48	0.74	0.49
	Female	1.55	0.77	0.52
LL6	Male	1.70	0.85	0.57
	Female	1.80	0.90	0.60
LL7	Male	1.89	0.94	0.63
	Female	2.03	1.01	0.68
LL8	Male	3.03	1.51	1.01
	Female	3.30	1.65	1.10
UL1	Male	1.74	0.87	0.58
	Female	1.84	0.92	0.61
UL2	Male	1.55	0.77	0.52
	Female	1.75	0.89	0.58
UL3	Male	1.52	0.76	0.51
	Female	1.59	0.79	0.53
UL4	Male	1.78	0.89	0.59
	Female	1.79	0.89	0.60
UL5	Male	1.66	0.83	0.55
	Female	1.74	0.87	0.58
UL6	Male	1.81	0.90	0.60
	Female	1.83	0.91	0.60
UL7	Male	1.85	0.93	0.62
	Female	1.89	0.95	0.63
UL8	Male	2.75	1.37	0.92
	Female	2.91	1.45	0.97

Table 39: Crown-root ratio definition for mandibular and maxillary root fraction according to Nolla's staging technique

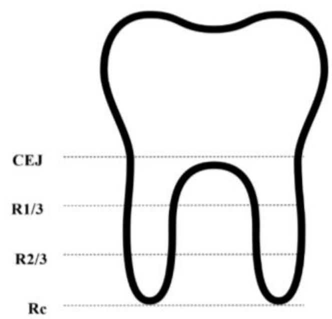


Figure 42: Mandibular tooth and root fractions as defined by Nolla's staging technique

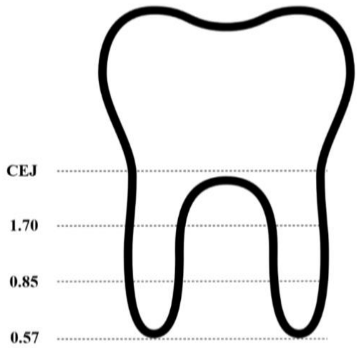


Fig. 43 - Male

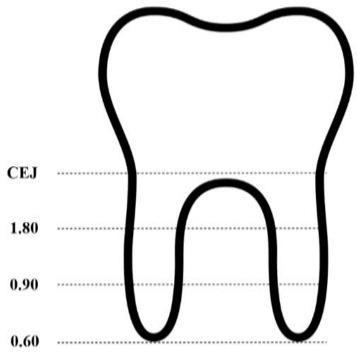


Fig. 44 - Female

Figure 43 & **44**: LL6 crown-root ratio of root fractions R1/3, R2/3 and Rc as described by Nolla's staging technique specific to their respective sex groups.

Tooth	Sex	R1/3	R2/3	Rc
LL6	Male	1.70	0.85	0.57

Table 40: Crown-root ratio for LL6 rootfractions in male subjects as extracted from **Table 39**

Tooth	Sex	R1/3	R2/3	Rc
LL6	Female	1.80	0.90	0.60

Table 41: Crown-root ratio for LL6 root fractions in female subjects as extracted from **Table 39**

4.2.10 Definition for root fraction – Moorrees' staging technique

Through the same mathematical formula applied in **4.2.9**, the sex-specific data regarding crown-root ratio for root fractions as described by Moorrees' staging technique was as summarised in **Table 42. Figures 45 – 47** provided examples of crown-root ratio assignment based on type of tooth and sex of the subject for Moorrees' staging technique.

	Sex	R1/4	R1/2	R3/4	Rc
LL1	Male	2.29	1.15	0.76	0.57
	Female	2.37	1.19	0.79	0.59
LL2	Male	2.32	1.16	0.77	0.58
	Female	2.41	1.21	0.80	0.60
LL3	Male	2.19	1.09	0.73	0.55
	Female	2.29	1.15	0.76	0.57
LL4	Male	2.13	1.07	0.71	0.53
	Female	2.32	1.16	0.77	0.58
LL5	Male	1.98	0.99	0.66	0.49
	Female	2.06	1.03	0.69	0.52

LL6	Male	2.27	1.13	0.76	0.57
	Female	2.40	1.20	0.80	0.60
LL7	Male	2.51	1.26	0.84	0.63
	Female	2.71	1.35	0.90	0.68
LL8	Male	4.04	2.02	1.35	1.01
	Female	4.41	2.20	1.47	1.10
UL1	Male	2.32	1.16	0.77	0.58
	Female	2.45	1.23	0.82	0.61
UL2	Male	2.06	1.03	0.69	0.52
	Female	2.34	1.17	0.78	0.58
UL3	Male	2.03	1.01	0.68	0.51
	Female	2.20	1.06	0.71	0.53
UL4	Male	2.38	1.19	0.79	0.59
	Female	2.38	1.19	0.79	0.60
UL5	Male	2.21	1.10	0.74	0.55
	Female	2.32	1.16	0.77	0.58
UL6	Male	2.41	1.21	0.80	0.60
	Female	2.44	1.22	0.81	0.60
UL7	Male	2.47	1.24	0.82	0.62
	Female	2.52	1.26	0.84	0.63
UL8	Male	3.66	1.83	1.22	0.92
	Female	3.88	1.94	1.29	0.97

Table 42: Crown-root ratio definition for mandibular and maxillary root fraction according to Moorrees' staging technique

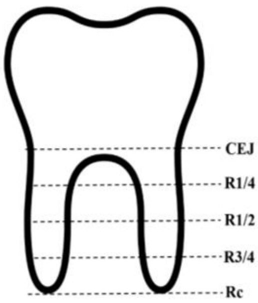


Figure 45: Diagram of mandibular tooth and its root fractions as described by Moorrees' staging technique

Tooth	Sex	R1/4	R1/2	R3/4	Rc
LL6	Male	2.27	1.13	0.76	0.57

Table 43: Crown-root ratio for LL6 root fractions in male subjects as obtained from **Table 42**

Tooth	Sex	R1/4	R1/2	R3/4	Rc
LL7	Female	2.71	1.35	0.90	0.68

Table 44: Crown-root ratio for LL7 root fractions in female subjects as obtained from **Table 42**

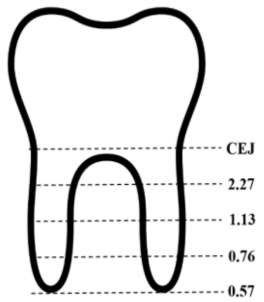


Fig. 46 - Male

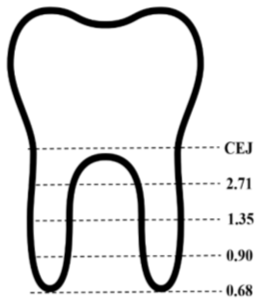


Fig. 47 - Female

Figure 46 & **47**: Crown-root ratio of LL6 and LL7 root fractions R1/4, R1/2, R3/4 and Rc as described by Moorrees' staging technique specific to their respective sex groups

4.2.11 Upper and lower limits for root fraction – Nolla's staging technique

	Sex	R1/3	R1/3 ↓ limit	R2/3 ↑ limit	R2/3	R2/3 ↓ limit	Rc ↑ limit	Rc
LL1	Male	1.72	1.29	1.28	0.85	0.71	0.70	0.57
	Female	1.78	1.34	1.33	0.89	0.74	0.73	0.59
LL2	Male	1.74	1.31	1.30	0.87	0.73	0.72	0.58
	Female	1.81	1.36	1.35	0.90	0.75	0.74	0.60
LL3	Male	1.65	1.24	1.23	0.82	0.69	0.68	0.55
	Female	1.72	1.29	1.28	0.86	0.72	0.71	0.57
LL4	Male	1.60	1.20	1.19	0.80	0.67	0.66	0.53
	Female	1.74	1.31	1.30	0.87	0.73	0.72	0.58
LL5	Male	1.48	1.11	1.10	0.74	0.62	0.61	0.49
	Female	1.55	1.16	1.15	0.77	0.65	0.64	0.52
LL6	Male	1.70	1.28	1.27	0.85	0.71	0.70	0.57
	Female	1.80	1.35	1.34	0.90	0.75	0.74	0.60
LL7	Male	1.89	1.42	1.41	0.94	0.79	0.78	0.63
	Female	2.03	1.52	1.51	1.01	0.85	0.84	0.68
LL8	Male	3.03	2.27	2.26	1.51	1.26	1.25	1.01
	Female	3.30	2.48	2.47	1.65	1.38	1.37	1.10
UL1	Male	1.74	1.31	1.30	0.87	0.73	0.72	0.58
	Female	1.84	1.38	1.37	0.92	0.77	0.76	0.61
UL2	Male	1.55	1.16	1.15	0.77	0.65	0.64	0.52
	Female	1.75	1.32	1.31	0.89	0.74	0.73	0.58
UL3	Male	1.52	1.14	1.13	0.76	0.64	0.63	0.51
	Female	1.59	1.19	1.18	0.79	0.66	0.65	0.53
UL4	Male	1.78	1.34	1.33	0.89	0.74	0.73	0.59
	Female	1.79	1.34	1.33	0.89	0.75	0.74	0.60
UL5	Male	1.66	1.25	1.24	0.83	0.69	0.68	0.55
	Female	1.74	1.31	1.30	0.87	0.73	0.72	0.58

UL6	Male	1.81	1.36	1.35	0.90	0.75	0.74	0.60
	Female	1.83	1.37	1.36	0.91	0.76	0.75	0.60
UL7	Male	1.85	1.39	1.38	0.93	0.78	0.77	0.62
	Female	1.89	1.42	1.41	0.95	0.79	0.78	0.63
UL8	Male	2.75	2.06	2.05	1.37	1.15	1.14	0.92
	Female	2.91	2.18	2.17	1.45	1.21	1.20	0.97

Table 45: Crown-root ratio of mandibular and maxillary root fraction according to Nolla's staging technique, as well as their upper and lower limits

Once the mean crown-root ratio for each root fraction was determined, the limits of each fraction were then established. Through data obtained in **Table 39** and formulae stated in **3.2.2.3**, the upper and lower limits for R1/3, R2/3 and Rc by Nolla's staging technique were compiled in **Tables 45** and **46.**

	Sex	**R1/3 ↓ limit**	**R2/3 ↑ limit**	**R2/3 ↓ limit**	**Rc ↑ limit**
LL1	Male	1.29	1.28	0.71	0.70
	Female	1.34	1.33	0.74	0.73
LL2	Male	1.31	1.30	0.73	0.72
	Female	1.36	1.35	0.75	0.74
LL3	Male	1.24	1.23	0.69	0.68
	Female	1.29	1.28	0.72	0.71
LL4	Male	1.20	1.19	0.67	0.66
	Female	1.31	1.30	0.73	0.72
LL5	Male	1.11	1.10	0.62	0.61
	Female	1.16	1.15	0.65	0.64
LL6	Male	1.28	1.27	0.71	0.70
	Female	1.35	1.34	0.75	0.74
LL7	Male	1.42	1.41	0.79	0.78
	Female	1.52	1.51	0.85	0.84
LL8	Male	2.27	2.26	1.26	1.25
	Female	2.48	2.47	1.38	1.37
UL1	Male	1.31	1.30	0.73	0.72
	Female	1.38	1.37	0.77	0.76
UL2	Male	1.16	1.15	0.65	0.64
	Female	1.32	1.31	0.74	0.73
UL3	Male	1.14	1.13	0.64	0.63
	Female	1.19	1.18	0.66	0.65
UL4	Male	1.34	1.33	0.74	0.73
	Female	1.34	1.33	0.75	0.74
UL5	Male	1.25	1.24	0.69	0.68
	Female	1.31	1.30	0.73	0.72

UL6	Male	1.36	1.35	0.75	0.74
	Female	1.37	1.36	0.76	0.75
UL7	Male	1.39	1.38	0.78	0.77
	Female	1.42	1.41	0.79	0.78
UL8	Male	2.06	2.05	1.15	1.14
	Female	2.18	2.17	1.21	1.20

Table 46: Upper and lower limits of mandibular and maxillary root fraction according to Nolla's staging technique

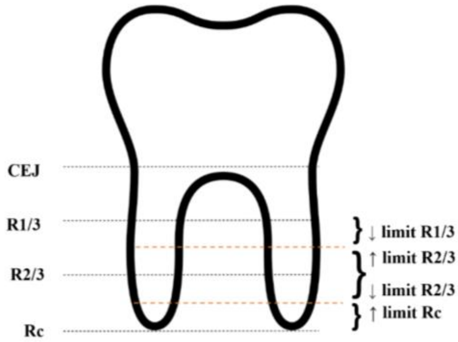


Figure 48: A mandibular molar with its root fraction landmarks as described by Nolla's staging technique, with the upper and lower limits

Landmarks defined by Nolla's staging technique such as CEJ, R1/3, R2/3, Rc, as well as the upper and lower limits were as illustrated in **Figure 48.**

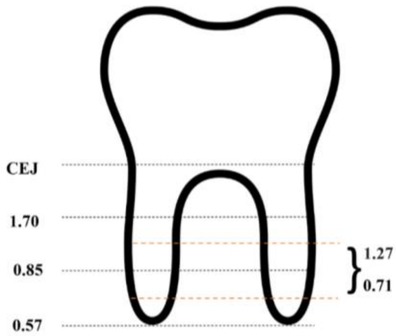


Figure 49: Example of a male's LL6 root fractions as described by Nolla's staging technique, and their crown-root ratios, as well as the upper and lower limits of R2/3

	Sex	R1/3	R1/3 ↓limit	R2/3 ↑limit	R2/3	R2/3 ↓limit	Rc ↑limit	Rc
LL6	Male	1.70	1.28	1.27	0.85	0.71	0.70	0.57

Table 47: Crown-root ratio for LL6 root fractions and their limits according to Nolla's staging technique specific to a male individual

Figure 49 provided an example of root fraction ratio assignment, as well as their limits based on the known sex of the subject. With reference to the figure and **Table 47**, the average ratios for a male individual's mandibular first molar (LL6) were 1.70 for R1/3, 0.85 for R2/3 and 0.57 for Rc. The upper limit for Rc was 0.70, while the limits for R2/3 were 0.71 (lower) and 1.27 (upper). R1/3 was confined only by its lower limit at 1.28. In reference to this example, should a male's LL6 crown-root ratio be measured within the range of 0.71-1.27, the appropriate stage to classify was R2/3. Alternatively, the tooth was considered as mature (or in stage Rc) if the crown-root ratio was found to be less than or equal to 0.70. If the ratio was more than or equal to 1.28, stage R1/3 was assigned. Only the intermediary stage (R2/3) was bounded by upper and lower limits. The most extreme stages on the other hand (R1/3 and Rc), were restricted to only one limit (either lower or upper).

4.2.12 Upper and lower limits for root fraction – Moorrees' staging technique

	Sex	R1/4	R1/4↓ limit	R1/2↑ limit	R1/2	R1/2↓ limit	R3/4↑ limit	R3/4	R3/4↓ limit	Rc ↑ limit	Rc
LL1	M	2.29	1.72	1.71	1.15	0.96	0.95	0.76	0.67	0.66	0.57
	F	2.37	1.78	1.77	1.19	0.99	0.98	0.79	0.69	0.68	0.59
LL2	M	2.32	1.74	1.73	1.16	0.97	0.96	0.77	0.68	0.67	0.58
	F	2.41	1.81	1.80	1.21	1.01	1.00	0.80	0.70	0.69	0.60
LL3	M	2.19	1.64	1.63	1.09	0.91	0.90	0.73	0.64	0.63	0.55
	F	2.29	1.72	1.71	1.15	0.96	0.95	0.76	0.67	0.66	0.57
LL4	M	2.13	1.60	1.59	1.07	0.89	0.88	0.71	0.62	0.61	0.53
	F	2.32	1.74	1.73	1.16	0.97	0.96	0.77	0.68	0.67	0.58
LL5	M	1.98	1.49	1.48	0.99	0.83	0.82	0.66	0.58	0.57	0.49
	F	2.06	1.55	1.54	1.03	0.86	0.85	0.69	0.61	0.60	0.52
LL6	M	2.27	1.70	1.69	1.13	0.95	0.94	0.76	0.67	0.66	0.57
	F	2.40	1.80	1.79	1.20	1.00	0.99	0.80	0.70	0.69	0.60
LL7	M	2.51	1.89	1.88	1.26	1.05	1.04	0.84	0.74	0.73	0.63
	F	2.71	2.03	2.02	1.35	1.13	1.12	0.90	0.79	0.78	0.68
LL8	M	4.04	3.03	3.02	2.02	1.69	1.68	1.35	1.18	1.17	1.01
	F	4.41	3.31	3.30	2.20	1.84	1.83	1.47	1.29	1.28	1.10
UL1	M	2.32	1.74	1.73	1.16	0.97	0.96	0.77	0.68	0.67	0.58
	F	2.45	1.84	1.83	1.23	1.03	1.02	0.82	0.72	0.71	0.61
UL2	M	2.06	1.55	1.54	1.03	0.86	0.85	0.69	0.61	0.60	0.52

	F	2.34	1.76	1.75	1.17	0.98	0.97	0.78	0.68	0.67	0.58
UL3	M	2.03	1.52	1.51	1.01	0.85	0.84	0.68	0.60	0.59	0.51
	F	2.20	1.63	1.62	1.06	0.89	0.88	0.71	0.62	0.61	0.53
UL4	M	2.38	1.79	1.78	1.19	0.99	0.98	0.79	0.69	0.68	0.59
	F	2.38	1.79	1.78	1.19	0.99	0.98	0.79	0.70	0.69	0.60
UL5	M	2.21	1.66	1.65	1.10	0.92	0.91	0.74	0.65	0.64	0.55
	F	2.32	1.74	1.73	1.16	0.97	0.96	0.77	0.68	0.67	0.58
UL6	M	2.41	1.81	1.80	1.21	1.01	1.00	0.80	0.70	0.69	0.60
	F	2.44	1.83	1.82	1.22	1.02	1.01	0.81	0.71	0.70	0.60
UL7	M	2.47	1.86	1.85	1.24	1.03	1.02	0.82	0.72	0.71	0.62
	F	2.52	1.89	1.88	1.26	1.05	1.04	0.84	0.74	0.73	0.63
UL8	M	3.66	2.75	2.74	1.83	1.53	1.52	1.22	1.07	1.06	0.92
	F	3.88	2.91	2.90	1.94	1.62	1.61	1.29	1.13	1.12	0.97

Table 48: Crown-root ratio of mandibular and maxillary root fraction according to Moorrees' staging technique, as well as their upper and lower limits. M=Male, F=Female

Through data obtained in **Table 42** and formulae stated in **3.2.2.3.**, the upper and lower limits for R1/4, R1/2, R3/4 and Rc by Moorrees' staging technique were compiled in **Tables 48** and **49**.

	Sex	**R1/4 ↓limit**	**R1/2 ↑limit**	**R1/2 ↓limit**	**R3/4 ↑limit**	**R3/4 ↑limit**	**Rc ↓limit**
LL1	M	1.72	1.71	0.96	0.95	0.67	0.66
	F	1.78	1.77	0.99	0.98	0.69	0.68
LL2	M	1.74	1.73	0.97	0.96	0.68	0.67
	F	1.81	1.80	1.01	1.00	0.70	0.69
LL3	M	1.64	1.63	0.91	0.90	0.64	0.63
	F	1.72	1.71	0.96	0.95	0.67	0.66
LL4	M	1.60	1.59	0.89	0.88	0.62	0.61
	F	1.74	1.73	0.97	0.96	0.68	0.67
LL5	M	1.49	1.48	0.83	0.82	0.58	0.57
	F	1.55	1.54	0.86	0.85	0.61	0.60
LL6	M	1.70	1.69	0.95	0.94	0.67	0.66
	F	1.80	1.79	1.00	0.99	0.70	0.69
LL7	M	1.89	1.88	1.05	1.04	0.74	0.73
	F	2.03	2.02	1.13	1.12	0.79	0.78
LL8	M	3.03	3.02	1.69	1.68	1.18	1.17
	F	3.31	3.30	1.84	1.83	1.29	1.28
UL1	M	1.74	1.73	0.97	0.96	0.68	0.67

	F	1.84	1.83	1.03	1.02	0.72	0.71
UL2	M	1.55	1.54	0.86	0.85	0.61	0.60
	F	1.76	1.75	0.98	0.97	0.68	0.67
UL3	M	1.52	1.51	0.85	0.84	0.60	0.59
	F	1.63	1.62	0.89	0.88	0.62	0.61
UL4	M	1.79	1.78	0.99	0.98	0.69	0.68
	F	1.79	1.78	0.99	0.98	0.70	0.69
UL5	M	1.66	1.65	0.92	0.91	0.65	0.64
	F	1.74	1.73	0.97	0.96	0.68	0.67
UL6	M	1.81	1.80	1.01	1.00	0.70	0.69
	F	1.83	1.82	1.02	1.01	0.71	0.70
UL7	M	1.86	1.85	1.03	1.02	0.72	0.71
	F	1.89	1.88	1.05	1.04	0.74	0.73
UL8	M	2.75	2.74	1.53	1.52	1.07	1.06
	F	2.91	2.90	1.62	1.61	1.13	1.12

Table 49: Upper and lower limits of mandibular and maxillary root fraction according to Moorrees' staging technique. M=Male, F=Female

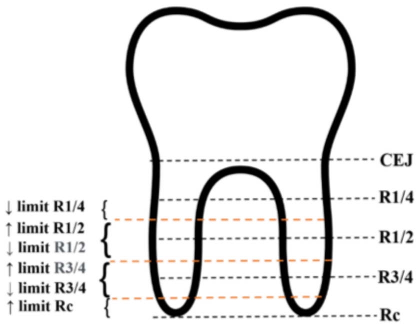


Figure 50: A mandibular molar with root fractions as described by Moorrees' staging technique as well as their upper and lower limits

Landmarks defined by Moorrees' staging technique such as CEJ, R1/4, R1/2, R3/4 and Rc were as represented in **Figure 50**. Their upper and lower limits were also defined.

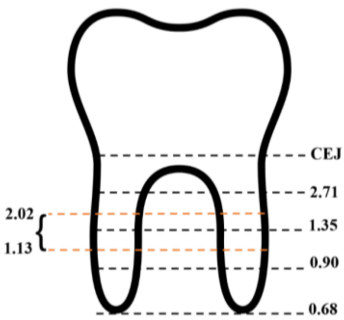


Figure 51: Example of a female's LL7 root fractions as described by Moorrees' staging technique, and their respective crown-root ratios, as well as the upper and lower limits for R1/2

	Sex	R1/4	R1/4↓ limit	R1/2↑ limit	R1/2	R1/2↓ limit	R3/4↑ limit	R3/4	R3/4↓ limit	Rc ↑ limit	Rc
LL7	F	2.71	2.03	2.02	1.35	1.13	1.12	0.90	0.79	0.78	0.68

Table 50: Crown-root ratio assigned for LL7 root fractions and their limits according to Moorrees' staging technique specific to a female individual. F=Female

Root fraction ratio assignment based on subject's known sex, with the upper and lower limits were as shown in **Figure 51**. With reference to the figure and **Table 50**, the average crown-root ratios for a mandibular second molar (LL7) in females were 2.71 for R1/4, 1.35 for R1/2, 0.90 for R3/4 and 0.68 for Rc. The limit for Rc was 0.78, while R3/4 was confined in the 0.79 (lower limit) and 1.12 (upper limit) range. R1/2 was limited within 1.13 (lower) and 2.02 (upper) while R1/4 was restricted only by its lower limit at 2.03. Therefore, the crown-root ratio of a female's mandibular second molar measured within the 1.13 – 2.02 range should be staged as stage R1/2, or R3/4 if the ratio was found in between 1.12 – 0.79. The tooth would be considered mature or in stage Rc if the crown-root ratio was less than or equivalent to 0.78 and R1/4 if it was more than to 2.02. Intermediary stages such as R1/2 and R3/4 were bounded by both upper and lower limits while the most extreme stages such as R1/4 and Rc were restricted to either a lower or an upper limit, respectively.

4.3 Part 3 – Crown-Root Ratio Accuracy

4.3.1 Descriptive analysis of subjects

	Male				Female			
Tooth	**N**	**Mean**	**SD**	**SE**	**N**	**Mean**	**SD**	**SE**
LL1	15	17.48	2.139	0.552	33	16.51	4.747	0.826
LL2	15	17.48	2.139	0.552	33	15.82	4.535	0.789
LL3	17	17.16	2.229	0.541	44	15.91	4.786	0.722
LL4	15	17.29	2.353	0.608	45	15.84	4.757	0.709
LL5	17	17.16	2.229	0.541	45	15.84	4.757	0.709
LL6	15	17.31	2.326	0.601	45	15.84	4.757	0.709
LL7	17	17.16	2.229	0.541	44	15.72	4.749	0.716
UL1	17	17.16	2.229	0.541	42	16.09	4.829	0.745
UL2	15	17.31	2.326	0.601	45	15.84	4.756	0.709
UL3	15	17.31	2.326	0.601	45	15.84	4.758	0.709
UL4	14	17.33	2.436	0.651	45	15.84	4.757	0.709
UL5	17	17.16	2.229	0.541	45	15.84	4.757	0.709
UL6	13	17.50	2.464	0.684	45	15.84	4.757	0.709
UL7	16	17.02	2.220	0.555	44	16.12	4.677	0.713

Table 51: Average age of subjects for each tooth measured

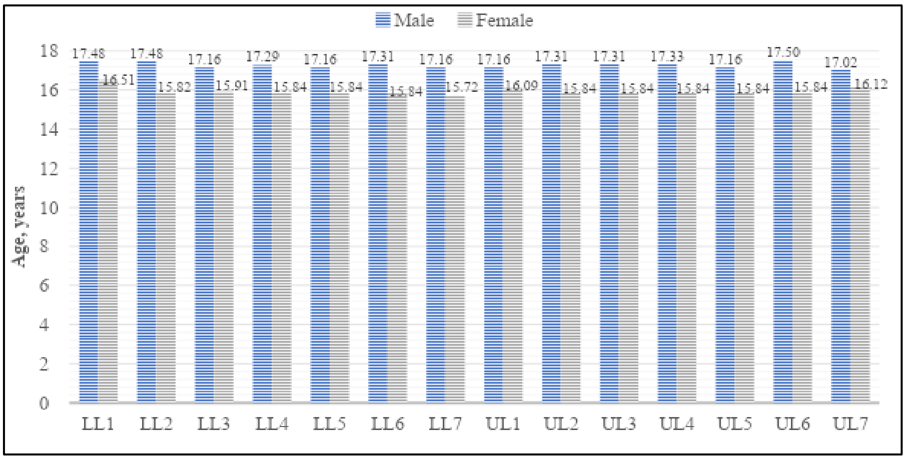


Figure 52: Average mean age per tooth measured

Mean age for males was 17.16 (SD=2.23/SE=0.54) and females was 15.84 (SD=4.76/SE=0.71) years. **Table 51** and **Figure 52** described the average age of subjects involved in crown-root ratio measurement and staging for each tooth separately. The average age range for males was between 17.02 – 17.50 years while a slightly younger and wider range of 15.72 – 16.51 years was observed in the other group. Females' LL7 recorded the youngest mean age while males' UL6 was the oldest.

4.3.2 Crown-root ratio

Tooth	Male				Female			
	N	Mean	SD	SE	N	Mean	SD	SE
LL1	15	0.62	0.070	0.018	33	0.58	0.068	0.012
LL2	15	0.65	0.075	0.019	33	0.60	0.081	0.014
LL3	17	0.61	0.074	0.018	44	0.61	0.105	0.016
LL4	15	0.60	0.054	0.014	45	0.67	0.181	0.027
LL5	17	0.53	0.041	0.010	45	0.74	0.381	0.057
LL6	15	0.58	0.063	0.016	45	0.60	0.060	0.009
LL7	17	0.67	0.096	0.023	44	0.93	0.409	0.062
UL1	17	0.63	0.055	0.013	42	0.62	0.051	0.008
UL2	15	0.57	0.061	0.016	45	0.58	0.072	0.011
UL3	15	0.56	0.062	0.016	45	0.62	0.160	0.024
UL4	14	0.67	0.111	0.030	45	0.73	0.268	0.040
UL5	17	0.63	0.071	0.017	45	0.80	0.387	0.058
UL6	13	0.63	0.107	0.030	45	0.64	0.090	0.013
UL7	16	0.67	0.103	0.026	44	0.93	0.512	0.077

Table 52: Mandibular and maxillary teeth crown-root ratio of test subjects

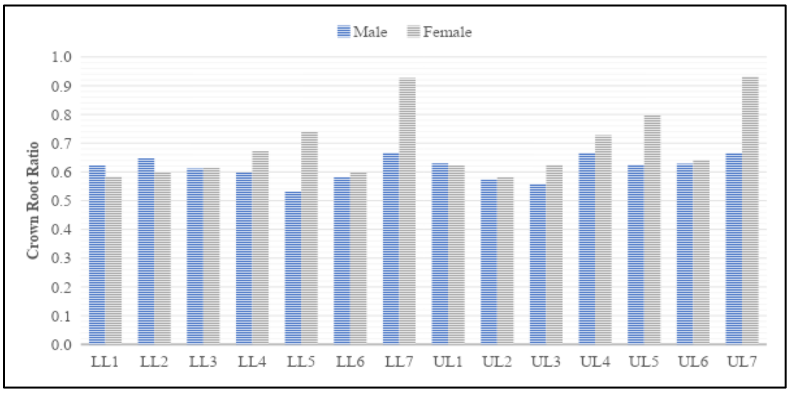


Figure 53: Crown-root ratio of mandibular and maxillary teeth by sex

A total of 818 mandibular and maxillary teeth were measured where the means of crown-root ratio ranged between 0.53 (male, LL5) to 0.93 (female, LL7 and UL7). UL5 of the female subjects exhibited the second highest crown-root ratio with 0.80 while the opposing mandibular tooth of the same sex group, LL5 followed with 0.74. Overall, male subjects produced a much smaller and consistent range of means and standard deviations while the opposite was observed in the other sex group (**Table 52**, **Figure 53**).

4.3.3 Scatterplot of crown-root ratio against Moorrees' stages

Scatterplots of individual crown-root ratio against their individual root stages according to Moorrees' staging technique were constructed for each mandibular and maxillary tooth to assess accuracy of the predicted ratios. Teeth staged within the limits were considered as 'accurate' and vice versa. A two-sample t test validated that no significant differences were observed between stages Rc, A1/2 and Ac (P>0.05) and therefore, stages A1/2 and Ac were combined into and defined by the same limits assigned for Rc. Thus, from here onwards, stages A1/2 and Ac will be referred to as Rc.

4.3.3.1 Mandibular central incisor

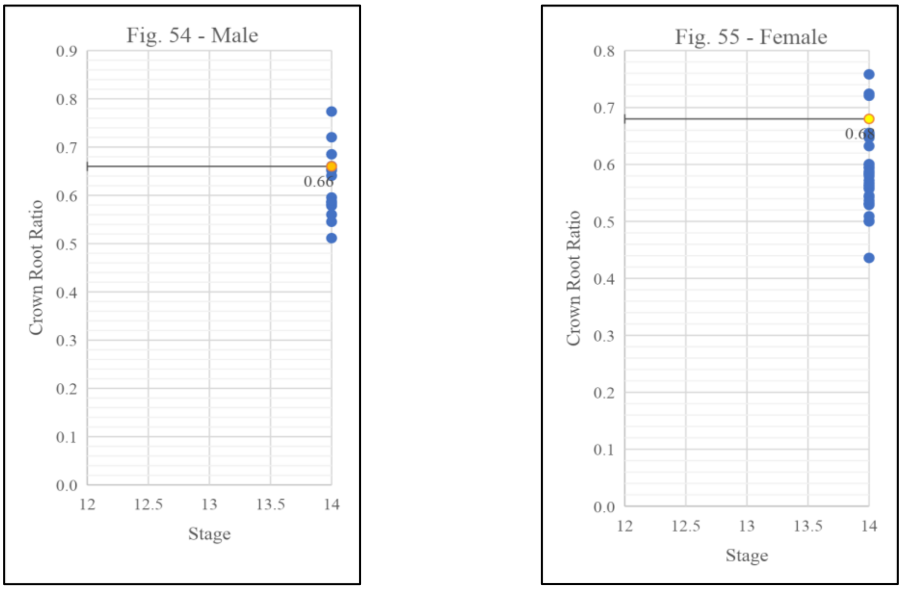


Figure 54 & **55**: Scatterplot of LL1 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stages 12 – 14=Rc

Male subjects were represented in **Figure 54.** A total of 15 teeth were scored where all 15 were in stage Rc. 12 teeth were scored within the upper limit of Rc (0.66) in contrast to 3. Among female subjects, 30 out of 33 were within their assigned upper Rc limit (0.68)

(**Figure 55**). Higher accuracy percentage was seen in the latter sex group (90.9%) compared to the former (80.0%), as summarised in **Table 53**:

	Male	Female
Stage	**Rc**	**Rc**
Teeth scored within limit(s)	12	30
Teeth scored outside limit(s)	3	3
Total teeth scored	15	33
Stage accuracy %	80.0	90.9
Total accuracy %	80.0	90.9

Table 53: Summary of LL1 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.2 Mandibular lateral incisor

Both sex groups concluded that all 48 of their LL2 were classified as Rc. However, for males, only 8 were appropriately scored below the upper limit (0.67) and the other 46.7% were scored otherwise (**Figure 56**). Female subjects on the other hand exhibited a much higher accuracy where 84.8% of the total staged appropriately (**Figure 57**). **Table 54** summarised these findings:

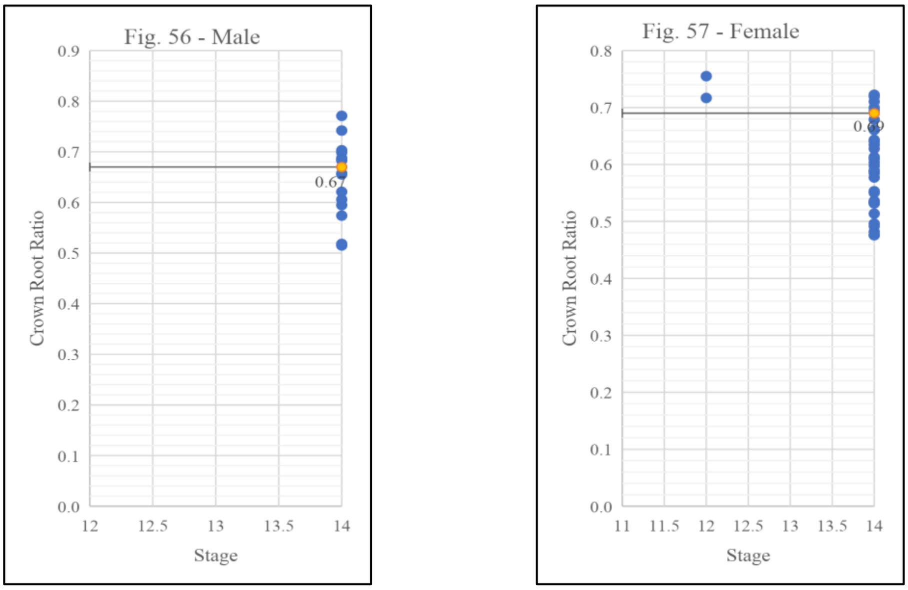
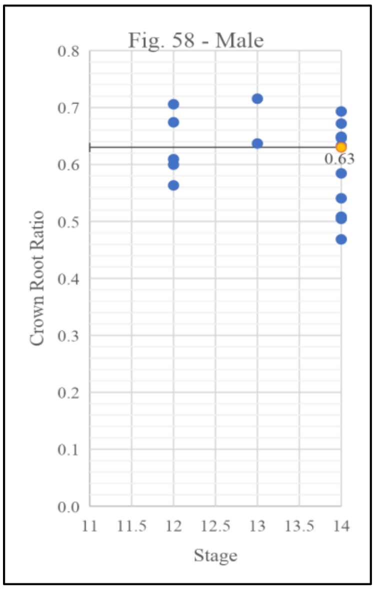


Figure 56 & **57**: Scatterplot of LL2 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stages 12 – 14=Rc

	Male	Female
Stage	**Rc**	**Rc**
Teeth scored within limits	8	28
Teeth scored outside limits	7	5
Total teeth scored	15	33
Stage accuracy %	53.3	84.8
Total accuracy %	53.3	84.8

Table 54: Summary of LL2 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.3 Mandibular canine



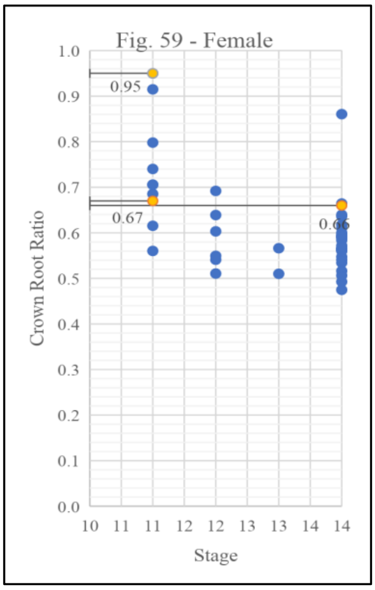


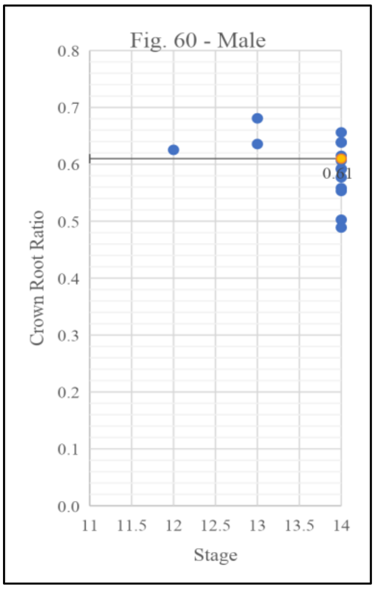
Figure 58 & **59**: Scatterplot of LL3 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 11=R3/4, 12 – 14=Rc

Figures 58 and **59** demonstrated 17 male subjects were staged in contrast to 44 females. Similar to previous finding, male subjects exhibited low accuracy in their Rc stage where only 10 properly scored below 0.63 (Rc upper limit). 35 of 44 female subjects were observed as mature, and only 2 were outside of the Rc limit, reflecting 94.3% accuracy. The other 9 female subjects were classified into stage R3/4 where 7 (77.8%) placed within the upper and lower limits (0.67-0.95). **Table 55** showed that a combination of both stages produced 90.9% accuracy in the group whereas males were only 58.8% accurate.

	Male	Female	
Stage	**Rc**	**R3/4**	**Rc**
Teeth scored within limits	10	7	33
Teeth scored outside limits	7	2	2
Total teeth scored	17	9	35
Stage accuracy %	58.8	77.8	94.3
Total accuracy %	58.8	90.9	

Table 55: Summary of LL3 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.4 Mandibular first premolar



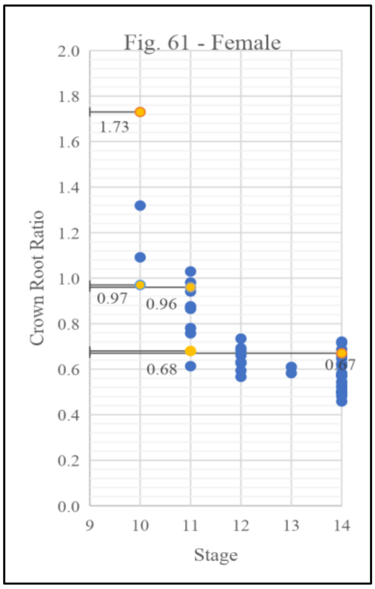


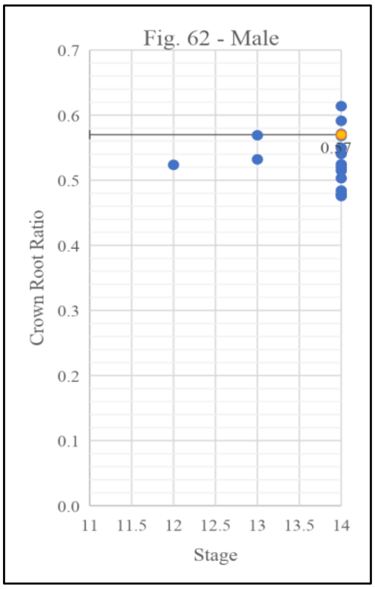
Figure 60 & **61**: Scatterplot of LL4 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 10=R1/2, 11=R3/4, 12 – 14=Rc

	Male	Female		
Stage	**Rc**	**R1/2**	**R3/4**	**Rc**
Teeth scored within limits	9	2	6	29
Teeth scored outside limits	6	0	3	5
Total teeth scored	15	2	9	34
Stage accuracy %	60.0	100.0	66.7	85.3
Total accuracy %	60.0	82.2		

Table 56: Summary of LL4 findings with stage and total accuracy (in percentage) for males and female subjects

40% (6 in 15) males' LL4 were staged outside of their upper Rc limit (>0.61) compared to their counterparts with 14.7% inaccuracy, who also presented with two other root fraction stages, R1/2 and R3/4 (**Figures 60, 61**). Remarkably, 100% accuracy was achieved in stage R1/2 while 2 out of 3 female participants were scored within the R3/4 range (0.68-0.96). Collectively, agreement in males was inferior than females by 22.2% (**Table 56**).

4.3.3.5 Mandibular second premolar



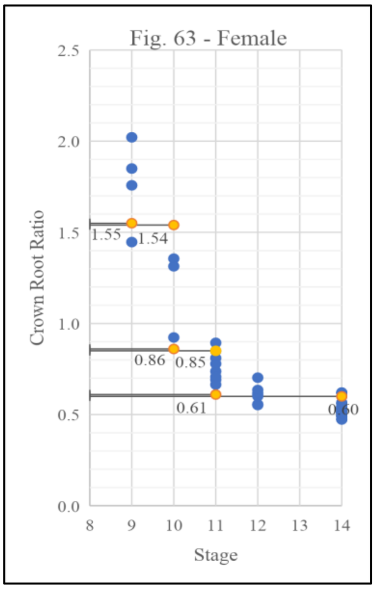


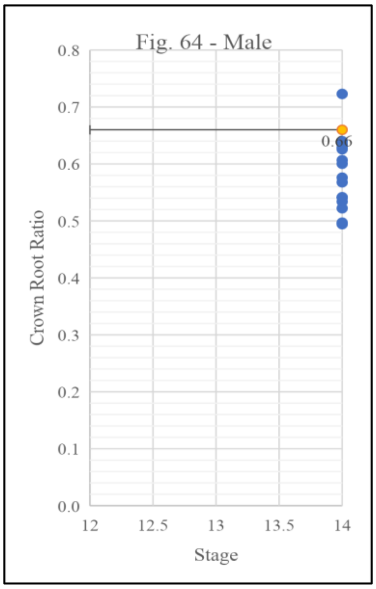
Figure 62 & **63**: Scatterplot of LL5 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 9=R1/4, 10=R1/2, 11=R3/4, 12 – 14=Rc

	Male	Female			
Stage	**Rc**	**R1/4**	**R1/2**	**R3/4**	**Rc**
Teeth scored within limits	15	3	3	7	25
Teeth scored outside limits	2	1	0	1	5
Total teeth scored	17	4	3	8	30
Stage accuracy %	88.2	75.0	100.0	87.5	83.3
Total accuracy %	88.2	84.4			

Table 57: Summary of LL5 findings with stage and total accuracy (in percentage) for males and female subjects

All root fraction stages described by Moorrees were present in LL5 of the female group (**Figure 63)**. Defined at 1.55 (≥1.55), 75% were staged correctly as their predicted means, while stage R1/2 displayed excellent agreement with 100%. The later root fraction stages also produced exceptional results with 87.5% and 83.3% for R3/4 and Rc, respectively. 100% male subjects achieved root maturity with 88.2% accuracy (**Figure 62**). **Table 57** provided summary of the results.

4.3.3.6 Mandibular first molar



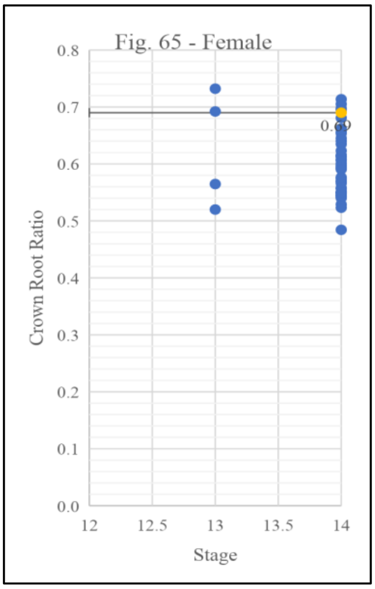


Figure 64 & **65**: Scatterplot of LL6 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stages 13 and 14=Rc

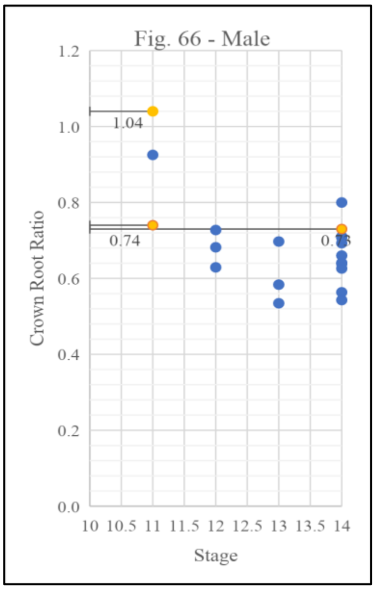
Both groups reached maturity in LL6 development. Accuracy was excellent where 93.8% of male subjects were within their predicted limit ≤0.66, as opposed to females with significantly lower result 84.4% (**Figure 64, 65** and **Table 58**).

	Male	Female
Stage	**Rc**	**Rc**
Teeth scored within limits	15	38
Teeth scored outside limits	1	7
Total teeth scored	16	45
Stage accuracy %	93.8	84.4
Total accuracy %	93.8	84.4

Table 58: Summary of LL6 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.7 Mandibular second molar

Merely two root fraction stages were observed in males as opposed to four in females (**Figures 66**, **67**). Total accuracy of the latter group dropped to 75% due to poor agreement in one of the stages. Both stages R1/2 and R3/4 recorded 75% agreement within the predicted upper and lower limits. 20 of 24 the female's LL7 were correctly classified as mature while only 25% of the R1/2 group had precise staging. The male subjects, however, exceeded the females' total accuracy by 19.1% where only 1 subject was misclassified in Rc stage and 100% accuracy in stage R3/4 (**Table 59**).



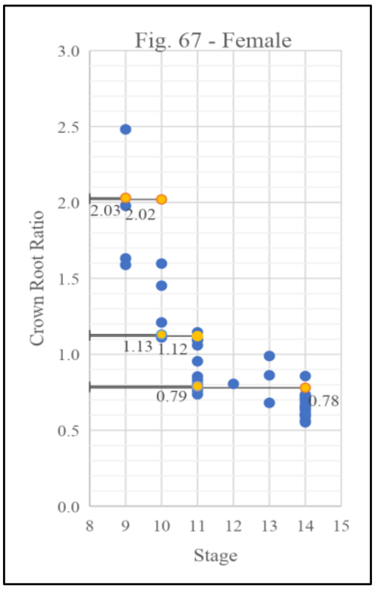


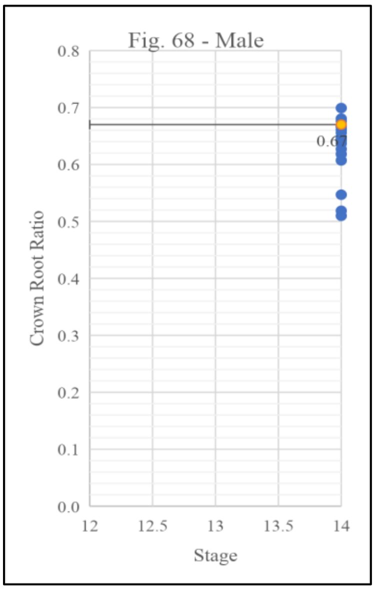
Figure 66 & **67**: Scatterplot of LL7 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 9=R1/4, 10=R1/2, 11=R3/4, 12 - 14=Rc

	Male		Female			
Stage	**R3/4**	**Rc**	**R1/4**	**R1/2**	**R3/4**	**Rc**
Teeth scored within limits	1	15	1	3	9	20
Teeth scored outside limits	0	1	3	1	3	4
Total teeth scored	1	16	4	4	12	24
Stage accuracy %	100.0	93.8	25.0	75.0	75.0	83.3
Total accuracy %	94.1		75.0			

Table 59: Summary of LL7 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.8 Maxillary central incisor

All 59 male and female subjects (100%) presented with mature roots. Restricted within their respective 0.67 and 0.71 Rc limits, more males were staged outside of their predicted ratios (11.8%) than females (3.4%). 58 in over 59 subjects were in stage 14, though all three stages were significantly identical and grouped as stage Rc (**Figures 68, 69** and **Table 60**).



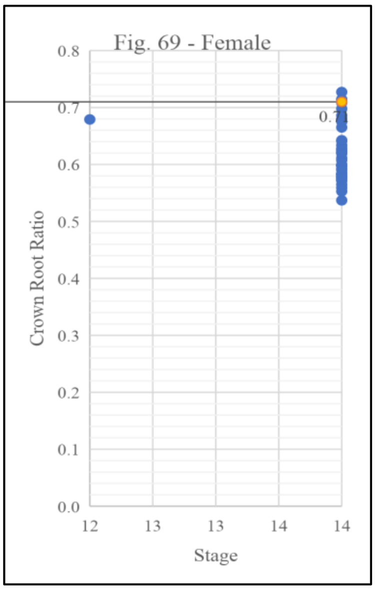


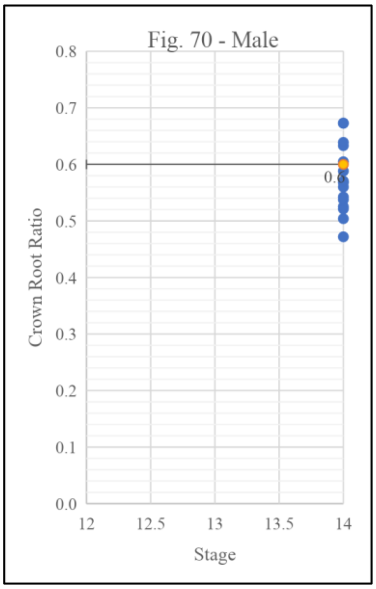
Figure 68 & **69**: Scatterplot of UL1 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stages 12 - 14=Rc

	Male	Female
Stage	**Rc**	**Rc**
Teeth scored within limits	15	41
Teeth scored outside limits	2	1
Total teeth scored	17	42
Stage accuracy %	88.2	97.6
Total accuracy %	88.2	97.6

Table 60: Summary of UL1 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.9 Maxillary lateral incisor

Both groups were distinguished by the number of root fraction stages present (**Figures 70** and **71**). All 15 male subjects were in stage 14 (Rc) and only 11 accurately staged within their predicted Rc limit (≤ 0.68). As a result, total and stage accuracy of the group declined to 73.3%. Female subjects on the other hand, projected with two root stages, stages R3/4 and Rc. One was precisely scored as stage R3/4 (100%) the other 41 of 44 as root mature (93.2%). In total, females displayed a much higher accuracy despite presenting with one additional root stage (**Table 61**).



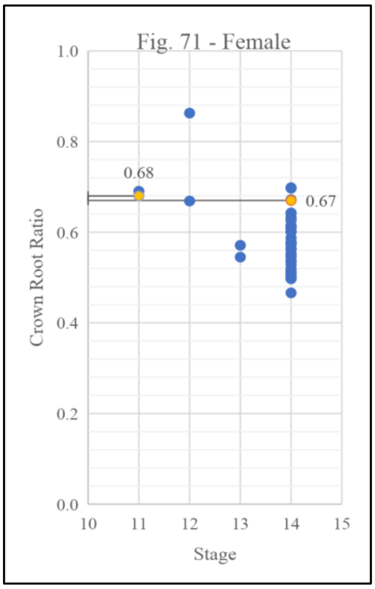


Figure 70 & **71**: Scatterplot of UL2 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 11=R3/4, 12 - 14=Rc

	Male	Female	
Stage	**Rc**	**R3/4**	**Rc**
Teeth scored within limits	11	1	41
Teeth scored outside limits	4	0	3
Total teeth scored	15	1	44
Stage accuracy %	73.3	100.0	93.2
Total accuracy %	73.3	93.3	

Table 61: Summary of UL2 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.10 Maxillary canine

Fifteen male and 34 female subjects were classified into stage Rc, as bounded by their limits (male=0.59, female=0.61) (**Figures 72**, **73**). Excellent agreement was achieved overall, where 12 males (80.0%) and 29 females (85.3%) were within their predicted ratios. Females presented with 2 other root fraction stages than males (R1/2 and R3/4). Both female candidates in stage R1/2 were within the limits (0.89-1.62) yielding 100%

accuracy whilst 33.3% of the R3/4 group were observed otherwise. **Table 62** verified that total accuracy of the female group exceeded males by 2.2%.

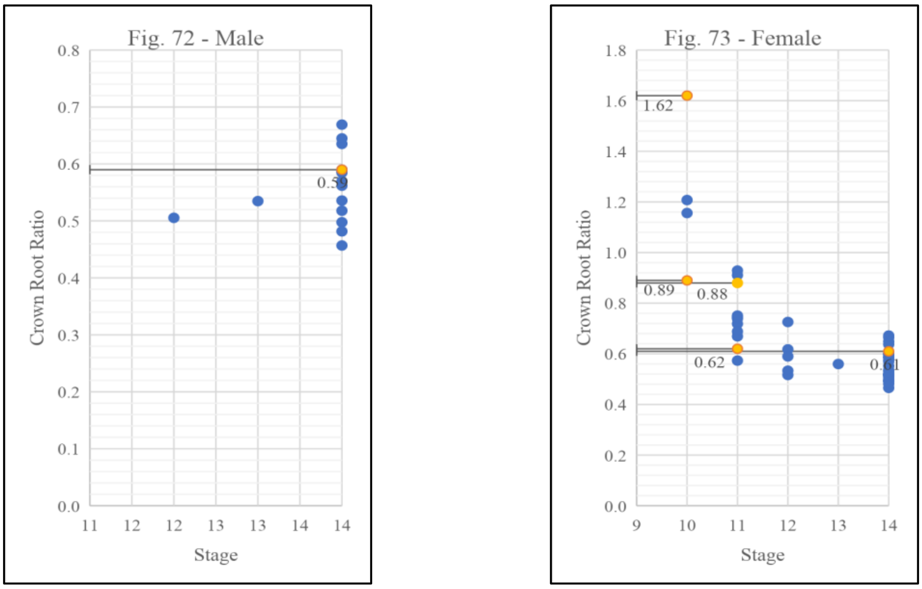


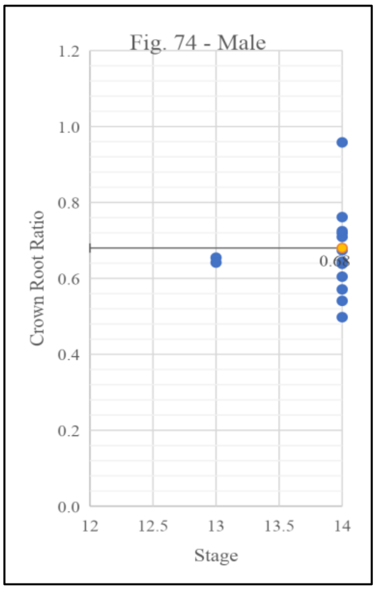
Figure 72 & **73**: Scatterplot of UL3 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 10=R1/2, 11=R3/4, 12 - 14=Rc

	Male	Female		
Stage	**Rc**	**R1/2**	**R3/4**	**Rc**
Teeth scored within limits	12	2	6	29
Teeth scored outside limits	3	0	3	5
Total teeth scored	15	2	9	34
Stage accuracy %	80.0	100	66.7	85.3
Total accuracy %	80.0	82.2		

Table 62: Summary of UL3 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.11 Maxillary first premolar

Higher total accuracy was observed in females than their male counterparts where 36 were accurately scored as their predicted crown-root ratios (80%) as opposed to 35.7% of the latter group who recorded beyond the Rc limit (>0.68). R3/4 had a modest agreement, where only 62.5% were as predicted. On the other hand, neither of the female subjects in stage R1/2 were accurate as both were staged outside of R1/4 lower limit (<1.79) (**Figures 74, 75** and **Table 63**).



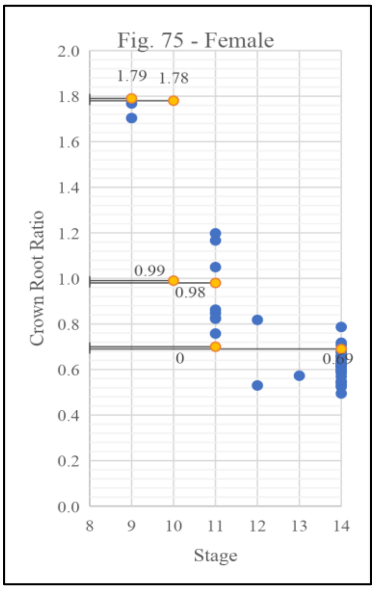


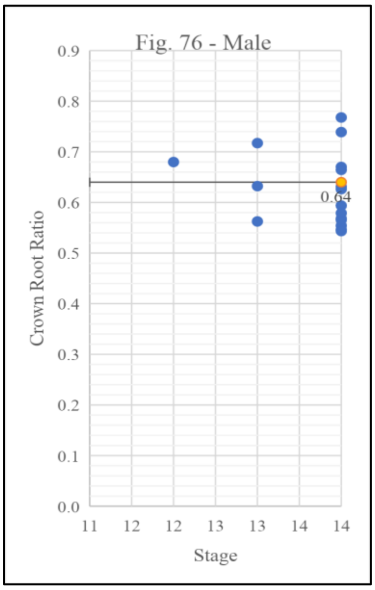
Figure 74 & **75**: Scatterplot of UL4 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 9=R1/4, 10=R1/2, 11=R3/4, 12 - 14=Rc

	Male	**Female**			
Stage	**Rc**	**R1/4**	**R1/2**	**R3/4**	**Rc**
Teeth scored within limits	9	0	0	5	31
Teeth scored outside limits	5	2	0	3	4
Total teeth scored	14	2	0	8	35
Stage accuracy %	64.3	0	0	62.5	88.6
Total accuracy %	64.3	80.0			

Table 63: Summary of UL4 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.12 Maxillary second premolar

Stage Rc was the only root stage displayed by the male group as compared to females who demonstrated all of Moorrees' four root fraction stages (**Figures 76, 77**). Females showed better agreement within the predicted Rc limit, ≤0.67 (27 agreements in 33 subjects) while males had a much lower agreement (11 in 17 subjects). On the other hand, 4 out of 6 female subjects (66.7%) showed accuracy within the R3/4 limits, while only 50% was recorded for both stages R1/4 and R1/2, although total number of subjects in both groups were significantly lower than the other two stages (**Table 64**).



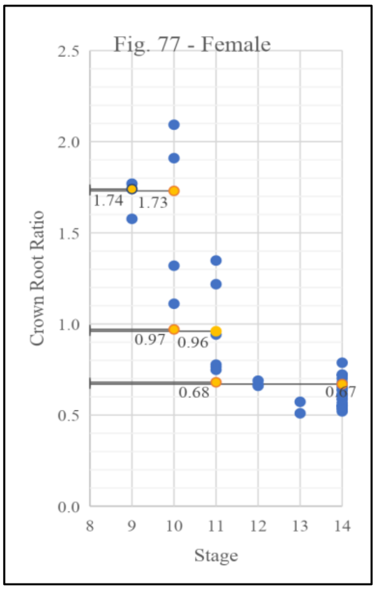


Figure 76 & **77**: Scatterplot of UL5 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 9=R1/4, 10=R1/2, 11=R3/4, 12 - 14=Rc

	Male	**Female**			
Stage	**Rc**	**R1/4**	**R1/2**	**R3/4**	**Rc**
Teeth scored within limits	11	1	2	4	27
Teeth scored outside limits	6	1	2	2	6
Total teeth scored	17	2	4	6	33
Stage accuracy %	64.7	50	50	66.7	81.8
Total accuracy %	64.7	75.5			

Table 64: Summary of UL5 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.13 Maxillary first molar

A total of 58 subjects were staged for the maxillary first molar (UL6) where all 13 males and 44 females were recorded in stage Rc. However, accuracy was moderate in males where only 9 crown-root ratios were within the limit (≤0.69) as opposed to excellent agreement in the female group where 36 were accurately projected. One lone female subject staged as R3/4 was also predicted within the stage's limits (0.71-1.01), reflecting 100% accuracy (**Figures 78, 79** and **Table 65**).

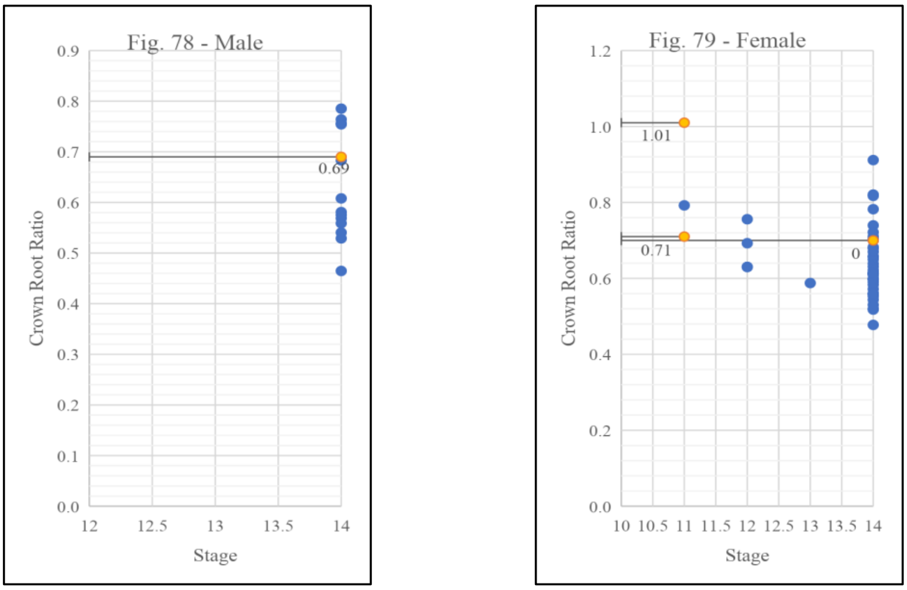


Figure 78 & **79**: Scatterplot of UL6 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 11=R3/4, 12 - 14=Rc

	Male	Female	
Stage	**Rc**	**R3/4**	**Rc**
Teeth scored within limits	9	1	36
Teeth scored outside limits	4	0	8
Total teeth scored	13	1	44
Stage accuracy %	69.2	100	81.8
Total accuracy %	69.2	82.2	

Table 65: Summary of UL6 findings with stage and total accuracy (in percentage) for males and female subjects

4.3.3.14 Maxillary second molar

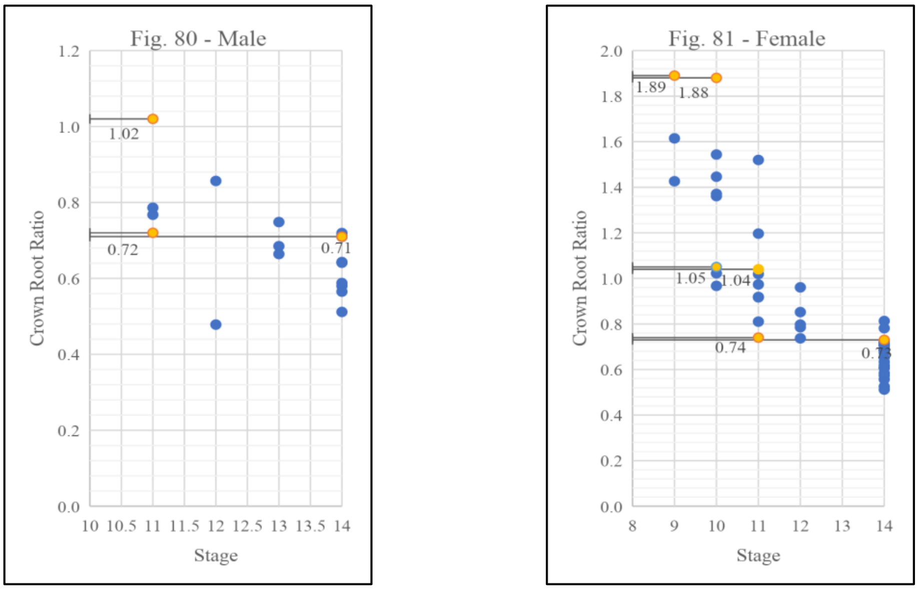


Figure 80 & **81**: Scatterplot of UL7 crown-root ratio vs stage as defined by Moorrees' in both sex groups. Stage 9=R1/4, 10=R1/2, 11=R3/4, 12 - 14=Rc

Comparatively, more root stages were observed in females (all four as defined by Moorrees) than that of males (**Figures 80** and **81**). Best stage accuracy was recorded by males' R3/4 where both subjects' ratios were measured within 0.72 – 1.02 limits. Least accuracy was seen in females' R1/4 where neither ratio was found within the stage's 1.89 limit. The remaining female root stages, on the other hand, displayed moderate agreement (R1/2=66.7%, R3/4=71.4%, Rc=78.6%). Majority of male subjects were staged as Rc and 85.7% of them were accurately predicted lower than 0.72 (≤0.71) (**Table 66**).

	Male		**Female**			
Stage	**R3/4**	**Rc**	**R1/4**	**R1/2**	**R3/4**	**Rc**
Teeth scored within limits	2	12	0	4	5	22
Teeth scored outside limits	0	2	2	2	2	6
Total teeth scored	2	14	2	6	7	28
Stage accuracy %	100	85.7	0	66.7	71.4	78.6
Total accuracy %	87.5		72.9			

Table 66: Summary of UL7 findings with stage and total accuracy (in percentage) for male and female subjects

4.3.4 Summary of Part 3 findings

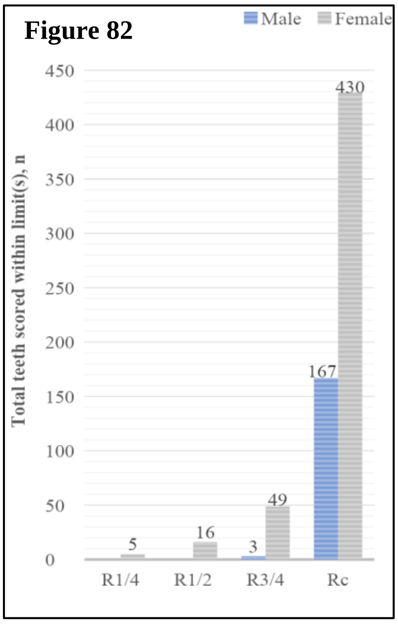
4.3.4.1 Summary of Part 3 findings by stage and sex

Stage	**R1/4**		**R1/2**		**R3/4**		**Rc**	
Sex	**M**	**F**	**M**	**F**	**M**	**F**	**M**	**F**
N within limit(s)	0	5	0	16	3	51	167	430
N outside limit(s)	0	9	0	5	0	19	49	64
Total teeth scored	0	14	0	21	3	70	216	494
Stage accuracy %	NA	35.7	NA	76.2	100.0	70.0	77.3	87.0

Table 67: Summary of stage and total accuracy (in percentage) for male and female subjects

Findings pertaining to subjects' total accuracy by stage and sex were summarised in **Table 67**. **Figures 82** and **83** illustrated the clear distinctions in the number of root stages and teeth scored between two sex groups. Male subjects presented with merely two Moorrees' root stages, R3/4 and Rc while the female group displayed all four. Among these 2 stages displayed by males, 216 (98.6%) were observed in stage Rc and only 3 (1.4%) were in R3/4. Females had a more diverse coverage with 82.5% in stage Rc, 11.7%

in R3/4, 3.5% in R1/2 and 2.3% in R1/4. Three of the later root fraction stages by females had excellent accuracy percentage (R1/2=76.2, R3/4=70.0, Rc=87.0) while stage R1/2 demonstrated poor accuracy with 35.7%. Excellent accuracy was demonstrated by both of males' root stages (R3/4=100%, Rc=77.3%).



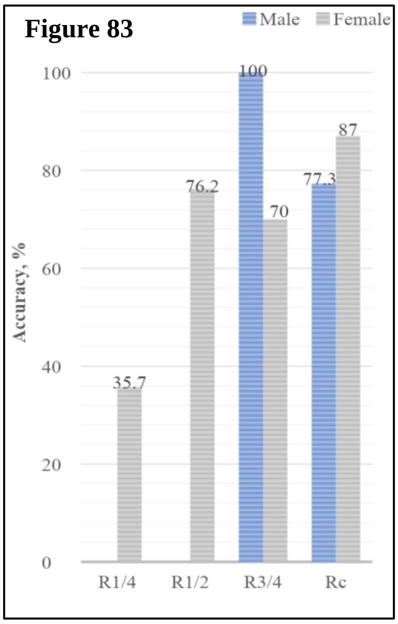


Figure 82 & **83**: Total teeth staged and accuracy (in percentage) by stage and sex

Side by side comparison between **Figures 82** and **83** showed that agreement might not be accurately represented due to uneven distribution of stages between the two sex groups. This was most clearly seen in males' R3/4 with 100%. When studied more closely, only 3 were observed in stage R3/4 whereas in the female group there was a significantly higher number, 49 and accuracy was 30.0% lower than males. Another clear contrast was in stage Rc where the disparity between sexes was small when expressed in percentage (9.7%), yet remained large when only number of teeth was considered (263 teeth).

4.3.4.2 Summary of Part 3 findings by stage

Data collected from male and female groups were combined to reflect total percentage agreement and total teeth scored per root fraction stage. Highest accuracy was proven by stage Rc (87%) while stage R1/4 proved to be the least accurate with 35.7%. Stages R1/2 and R3/4 also demonstrated excellent agreement with 76.2% and 73.2% each. The pattern was random and did not reflect the trend for total teeth scored, which was increasing from R1/4 through Rc (**Table 68, Figure 84)**.

Stage	R1/4	R1/2	R3/4	Rc
N within limit(s)	5	16	54	597
N outside limit(s)	9	5	19	113
Total teeth scored	14	21	73	710
Stage accuracy %	35.7	76.2	74	84.1
Total teeth %	1.7	2.6	8.9	86.8

Table 68: Summary of total stage accuracy

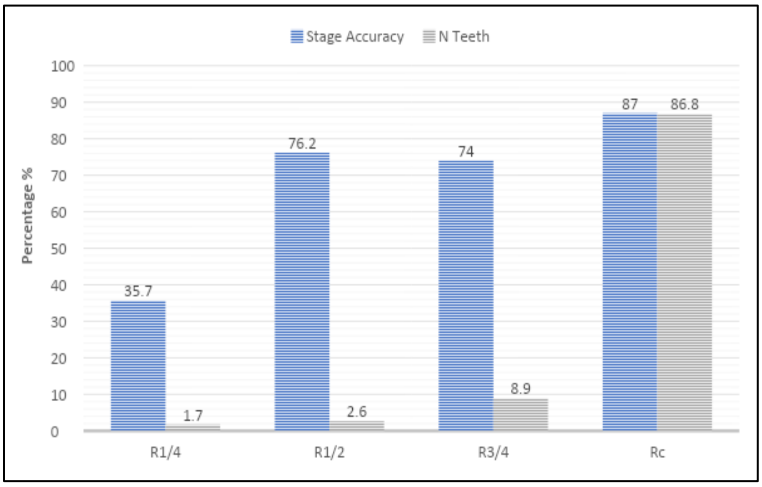


Figure 84: Total percentage agreement and teeth scored by stage

5. DISCUSSION

5.1 Reliability of tooth staging techniques

Most dental age estimation methods that employ staging techniques classify tooth development process in stages. This continuous process is divided into groups that may not represent their exact state of development. Unclear definition and differing interpretation of stages, uncertainties about the true length of a root in its mature form, inadequate training and experience, all contribute to bias. Although it is not always possible to achieve a faultless accuracy when predicting age, bias, represented by the estimated range should be reduced to the smallest value possible.

Kappa statistics is often used in medical science to measure observers' agreements and reliability of a method. When two observers perform an assessment, it is important that the ratings are given independently. It is also crucial that the assessment is done objectively and not based on guessing. However, guessing is inevitable at times, especially when it comes to a subjective assessment, such as the case in tooth staging. Cohen's kappa [59] addresses this issue by taking 'agreement by chance' into account and the value of kappa is corrected:

$$\kappa = \frac{\text{observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}} = \frac{\rho_o - \rho_e}{1 - \rho_e} = 1 - \frac{(1 - \rho_o)}{(1 - \rho_e)}$$

All three staging methods were proven to be highly reliable in this study. Moorrees' 14–stage technique was the most reliable (κ=0.938) as compared to the other two, even though the differences were not large. This contradicted findings by Martínez Gutiérrez et al. [68] when it was found that Moorrees' was the least precise as it demonstrated the biggest standard error as compared to Nolla's and Demirjian's in Venezuelan children. It also challenged the argument that increasing the number of stages might result in reduced precision of a staging technique [69,70]. However, when these techniques are analysed more critically, and stages are studied individually through crosstabulations, it is clear that Moorrees' in general had the lowest percentage agreement as compared to Nolla's and Demirjian's. Moreover, if the average percentage agreement is calculated individually for each technique, Moorrees' had the lowest raw agreement (61.1%), followed by Demirjian's (70.3%) and Nolla's (71.8%). The question arises – is kappa an accurate reflection of the true agreement for staging techniques? How much weight/penalty should be given when disagreements take place? Should agreement by chance be considered at all?

Another aspect of the agreement that should be analysed more closely is the high percentage in the final stages when root is mature and development is no longer observed (Nolla's stage 10, Moorrees' stage Ac, and Demirjian's stage H). By logic, it is more likely to achieve better agreement in this stage as 'root complete' is the final stage of a developing tooth, and one would presumably be more certain to score for this stage as less speculation is involved. It is also important to note that as compared to other stages, it only has a 'before' but no 'after' stage. As a result of this situation, there would be more disagreements in the intermediate stages, signifying more weights than the end stages. Maclure and Willett [71] acknowledged that "an intermediate category will often be subject to more misclassification than an extreme category because there are two directions in which to err away from the extremes". It was claimed that kappa is virtually meaningless for continuous data that are grouped into ordinal categories, much like tooth formation itself. While contribution of weighted kappa was acknowledged, the issue of the 'unfairness' when it comes to weighting the extreme categories was not properly discussed. Perhaps, the magnitude for weights should be given more to the extreme or end stages, relative to overall disagreements when discordance is present in tooth staging, as discussed by Sim and Wright [72]. By using linear weights when one disagreement is more (or less) serious than the other, such as the disagreement in final stages than intermediate ones, the kappa value will be different.

Despite all arguments against Moorrees' high reliability and lower average percentage agreement, credit must be given when it is due. As Dhanjal et al. [66] claimed that increasing the number of stages will not be helpful, it should be argued that in a continuous process such as dental development, giving more options for a tooth to be staged may prevent guessing from happening. For example, it is much easier to score a tooth as R1/2 when it appears to be halfway through maturity as compared to deciding whether it should be classified as R1/3 or R2/3. The number of samples in Dhanjal et al.'s [66] study may also play a part in their precision findings. Comparative to this study, Dhanjal et al [66] had a much smaller sample (*n*=127) and smaller samples have been known to yield a much bigger confidence interval (CI), thus indicating low precision as compared to studies with a much larger sample. The study also focused only on the third molars, and it would be more appropriate perhaps, to conclude that Moorrees' staging technique had a lower reliability when teeth are assessed individually (the third molar), rather than collectively.

Overall, mandible had better precision than maxilla except for Moorrees' technique although the difference is almost negligible (0.004). This is due to the fact that on panoramic radiographs, mandibular teeth were easier to assess as there were fewer anatomical superimpositions as compared to the maxilla's [73] as suggested by Demirjian [40] in his technique when only the permanent mandibular left teeth were to be scored. It is also interesting to note, and perhaps unsurprisingly, that for all three techniques the maxillary first molar was the least reliable when teeth were studied individually. This is most possibly due to the anatomy and morphology of the tooth itself. Maxillary first molars are typically presented with three roots (palatal, mesio-buccal, and disto-buccal) in most world populations [74]. Although they may be presented with four, two and single roots, the prevalence is very low [75]. The proximity of the maxillary first molars to the maxillary sinus floor has also been extensively discussed in literature [76-78]. It seems apparent that because of this proximity, the tendency for misclassifying the maxillary first molars would be high due to radiographic superimpositions. Also, due to the presence of the palatal root, which appears less visible on majority of radiographs as compared to the buccal roots, made it more difficult to stage. High precision in maxillary lateral incisor and mandibular second molar may also be attributed to their location and lack of superimposition with other anatomical structures.

McHugh [79] supported Cohen's [63] argument that kappa values may be too lenient for medical research. For instance, kappa of 0.60 is considered good in general, but in cancer treatment for example, it may be considered unacceptable. Such is the case in staging technique. It may be worth revisiting and re-evaluating kappa values of dental age estimation methods in past and future publications, and whether they should be more stringent or relaxed.

Visual illusions are not an uncommon occurrence in dentistry. Mach band effect [80] refers to an optical illusion phenomenon where the retina perceives contrasting shades in radiographs to be greater than they actually are. The extreme difference of contrast in enamel (radiopaque or white) and air (radiolucent or black), results in illusion where caries is perceived when in fact, is not present. Another example of visual illusion is the parallax effect, which is the displacement of an object due to the position of the x-ray beam in relation to the patient or an object. Buckle et al [81] categorised these visual illusions into three levels – image formation (parallax), sensation (Mach band effect), and perception. The basic building block of perception consists of a series of primitive shape, which is then combined to produce an object to match a database from visual memory.

Perception is a result of a positive match with this database. Within the perception level as described by Buckle et al [81], it is further categorised into ambiguous figures, fictional illusion and perceptual illusion. Visual memory greatly influences quality of root staging when radiographs are assessed in large volumes. It may not be possible to erase visual memory, but taking several rests *could* make a difference. Perceptual illusion on the other hand, is a phenomenon when two objects, despite being similar in size, are perceived differently. Subjective root staging can be categorised as perceptual illusion too, which could be improved by using a ruler to measure the dimensions as suggested by Yeung [82].

5.2 Crown-root ratio

Objective measurements have been suggested to improve reproducibility in age estimation techniques by Kullman et al. [83]. The study of 56 panoramic radiographs focused on defining anatomical landmarks and digitising plain radiographs. It was found that reproducibility was generally better when observers were calibrated, and knowledge of landmarks were briefed beforehand. It was also suggested that measurement of tooth length was to be adopted in order to improve subjectivity of tooth staging techniques. However, the issues of distortion and vertical magnification in panoramic radiographs were not addressed [84,85]. Thus, raw measurements should be avoided, and averaged into ratios to overcome this limitation.

In prosthodontics, crown-root ratio analysis is useful in designing dental prosthetics [86 – 88] while orthodontists use the method to assess short roots prior to treatment planning and predicting treatment prognosis [55]. Although calculation in most published studies was done in reverse, root-crown instead crown-root, data conversion for the purpose of comparison with this study was simplistic by applying simple arithmetic. For example:

Root-crown ratio = *root length/crown height*, and let *root-crown ratio* be 2.00

Root length/crown height = 2.00/1.00, therefore *crown height/root length* = 1.00/2.00

Crown-root ratio = 1.00/2.00 = 0.50

Crown-root ratios in other published studies ranged relatively similarly with present study (without the third molars), although Hölttä et al. [55] indicated longer roots i.e., smaller crown-root ratio compared to other groups. Generally, third molars provide minimal value for prosthodontics and orthodontics treatment. A retrospective cohort study of Belgian patients reported very little interest as only 4.3% would justify the retainment of the third molars for these two reasons [89]. This is reflected across published studies where data for the third molars' crown-root ratio were not made available for comparison with present study. Nonetheless, this study highlighted the fact that the third molars had the highest mean ratios and standard deviations compared to other permanent teeth. This indicated that the third molars had the shortest roots, and that the root length was the most variable. Such findings also suggested that crown-root ratio may not be a suitable method for the third molars. However, this may be partially corrected by increasing the sample size. Additionally, it would also be advantageous if the third molars are not disregarded in future studies such as Sindi A.S et al. [86] where the sample was composed of adult subjects where the third molars are mature and no longer developing.

Another significant finding was that among sixteen tooth types that were measured, sex differences were found in 50% of them. This was also found in other ethnic groups by Hölttä et al. (Finnish) [55], Sindi A.S et al. (Saudi) [86] and Yun et. Al (South Korean) [87] as well as Swedish children by Jakobsson and Lind [90], although the latter mainly investigated differences in permanent maxillary incisors. Because of these significant differences, it would be necessary to discriminate the data for each sex in future studies and pooling should be avoided.

Stages Rc, A1/2, and Ac in this study were assumed as comparable and therefore combined into one stage, Rc. This is because the differences between stages R1/4, R1/2 and Rc were in root lengths (metric) while A1/2 and Ac were different in their form, shape and structure (morphology). Moreover, t–tests between stage Rc and stages A1/2 and Ac showed that root length does not increase beyond stage Rc. Generally, data validation in the final part of the study indicated that crown-root ratio had a potential as an objective method when subjectivity is present during staging. Stage Rc's accuracy was the highest (87%) and this may be influenced by the total teeth tested (n=710) as the result of stages Rc, A1/2 and Ac grouped as one. In addition, as clarified in **5.1**, the higher accuracy for stage Rc may also be due to the fact that stage Rc was ranked last and therefore room for guessing was reduced. In contrast to stage Rc finding, R1/4 was the least accurate. Similar argument regarding total teeth tested could be made as it was the lowest (n=14) among other stages. Although stage R1/4 was also at the extreme side similar to Rc, guessing was more likely in this stage as one had to choose whether the root is either 1/4 or 1/2 (metric), compared to stage Rc, either root is complete or incomplete (morphologic). Interestingly, although both

stages R1/2 and R3/4 were subjected to greater chance for guessing, they both demonstrated better accuracy as more teeth were staged within their limits than R1/4. Perhaps the chance or risk for guessing was not as high as it was assumed, especially for stage R3/4. This is because the possible guessing options for this stage were three - R1/2, R3/4, or Rc. On one hand, when deciding whether a tooth should be classified as either stage R3/4 or Rc, it was presumably easier to eliminate Rc as the option as this stage showed clear and definite signs of root development having reached completion. Thus, the chance for guessing have reduced from three options (stages R1/2, R3/4, or Rc) to only two (stages R1/2 or R3/4).

Staging for R1/2 might be made less complicated by morphological observations. During this stage, indicators such as the shape of developing root canal walls (isosceles triangle), and crown height – root length comparison (which should be more or less similar) were used as a determining factor. However, total teeth scored (or lack thereof) in these fraction stages might have played a huge and significant role, especially for stage R3/4. Research involving more teeth in root fraction stages i.e., younger subjects may be necessary to confirm this hypothesis.

5.3 Limitations of material and difficulty of analysis

As panoramic radiographs were the primary material for this study, visual for tooth staging and measurement of the anterior teeth was affected by superimposition of cervical spine ghost image that caused exaggeration in radiopacity in the anterior region. Conversely, areas where radiolucency was increased such as the glossopalatal space, visual was also obscured, and this was a result of inadequate instruction during the x-ray exposure. This could be easily corrected by instructing the patient/subject to place the tongue on the hard palate for the whole duration of exposure. Another reason where radiograph quality was compromised is when subjects were placed outside of the focal trough resulting in image distortion and blurring of image, possibly caused by positioning error by the operator or patients with skeletal class II and III profiles. As discussed in **5.1**, magnification in panoramic radiographs can be corrected by converting raw measurements into ratios. Blurry images are more difficult to rectify, but can be avoided in the first place by making sure the patient/jaw is in the focal trough, even for patients with severe skeletal patterns.

Root length measurement was complicated by root length variations and morphology especially in multirooted teeth. Primary justification for selecting the longest root in mature teeth was to standardise all measurements. Another option was to consider the shorter roots or to measure and average the lengths of all roots. As discussed in **5.1**, the palatal root, which usually was the longest of all maxillary molar roots, was the least visible root on the panoramic radiographs and this might compromise precision. Similar limitation could be observed for immature teeth where the most apical point was selected as the root length. This was the most direct and clearest definition taken to maintain consistency in other cases. Clearly, choosing one option over the other could be disadvantageous, or vice versa, towards this study. For example, choosing the longest root instead of the shorter ones would result in differing set of mean ratios, as well as upper and lower limits. Subsequently, root fraction stage assignations would also be affected. These were the major limitations and differences between possible options, their influence towards results would require careful investigations in the future.

The number of subjects tested for validity was another drawback in this study. Initially, the study was designed to be tested with a longitudinal sample for a different objective, however, due to insufficient number of subjects recruited, the objective was subsequently amended. As discussed previously in **5.2**, stages with the fewest number of teeth were R1/4 and R1/2 and it was suggested that a larger sample may partially rectify the skewed outcome. Other than increasing the sample size, lowering the subjects age group would likely increase the chance for teeth to be staged in these two root fraction stages and indirectly expand the sample size. Results could also be improved by having a more uniformed sample size for each stage as this would allow better representation.

Another limitation with regard to the sample is that the subjects were recruited of mixed ethnicities - Malay, Chinese and Indian and the sample was pooled together. Although this presented no influence in the tooth staging reliability part of the study, as reliability of the techniques depended on the assessor’s ability to produce similar results, it is unknown whether differing ethnicities would yield significant differences for the crown-root measurement. Further, since there were significant differences between male and female crown-root ratios, having a balanced number of male and female subjects in Part 3 may possibly have a more positive impact on the study. Also, accuracy of crown-root measurement could be optimised by having a pre-assessment exercise on a small sample prior to the study to increase familiarity and improve technique efficiency. This would also allow clear definitions of landmarks for ratio measurement prior to Part 2 taking place.

5.4 Future research and applications

Cone beam computed tomography (CBCT) is an excellent imaging alternative for panoramic radiographs as it provides a three-dimensional (3D) view of the maxillofacial region, and more importantly, the resultant image is not affected by distortion or ghost image superimpositions. CBCT imaging software would provide various tools to improve visual assessment such as image magnification with minimal pixelation, and anatomical landmark measurement. As visual is enhanced, errors in the crown height and root length measurement could be minimised for better accuracy. Data comparison with present study can be made to ascertain significant differences between the imaging modalities. If no or minimal differences are observed, the use of panoramic radiographs in further studies may be acceptable and is also advantageous as panoramic radiographs are much easier to access and radiation dose is also significantly lower than the CBCT.

Upon collection of sufficient data, a semi-automated software may be developed to assist with tooth staging for dental age estimation. Important anatomical landmarks (such as cusp tip, CEJ and root apices) are detected for the crown height and root length, and the collective crown-root ratio data (preferably that are sex and ethnic-specific) for each root fraction stages can be used to predict the stage a tooth is in. The nature of the semi-automated software would allow the operator to correct the errors made in the landmark detection. The stages of the permanent teeth are then translated into a dental age based on published data. Further advancement into artificial intelligence (AI) software would allow accurate and reliable landmark detection, thus eliminating the need to correct errors and expediting the age estimation procedure. Another option is to study the relationship between crown-root ratio and dental age, and if correlations between the variables are significant, crown height and root length measurements can be directly translated into dental age without possibly going through the staging process.

The intended outcome of this research, especially in the field of forensic odontology was to improve reliability by removing the arbitrary aspect of staging methods. The subsequent by-product of this improvement was to maximise accuracy by minimising bias, which is defined by statistical error. Bunyarit et al. [91–93] applied the 8-tooth framework by Chaillet and Demirjian [94] and tabulated conversion of maturity score to dental age for Malaysian children by sex and ethnicity, Malay, Chinese and Indian. However, findings demonstrated that dental age was largely underestimated in comparison to chronological age for all three ethnic groups. Liversidge [95] undertook a similar study using Moorrees' 14–stage technique on Caucasian and Bangladeshi children in London. Interestingly, however, this has not yet been adapted for the Malaysian population. Further research is recommended to fill this knowledge gap, and crown-root ratio should be considered as an aid in the staging process. Similarly, past adaptations like Bunyarit's [91–93] could also benefit by adopting and applying the present quantitative method on the same sample used previously. Comparisons in accuracy, reliability and other relevant findings between the published and present studies could be made. At present, crown-root ratio specific to other populations may be limited, but the data obtained in this study could be used as a standard until more ethnicity and sex-specific data is available.

Lastly, the reduction in the number of late root stages by combining stages Rc, A1/2 and Ac into one singular stage could play a significant improvement for forensic odontologists. Since the finding of this study showed no significant differences in the late root stages crown-root ratio, it would allow for better productivity as time efficiency is increased. However, the relationships between these morphological changes (apex maturation process from stage Rc to Ac), crown-root ratio, and dental age is another research gap remains to be explored.

5.5 Conclusion

In conclusion. this study demonstrated that reliability of permanent teeth using Demirjian scoring is marginally more reliable than Nolla or Moorrees. Apart from kappa values, crosstabulation and percentage agreement may be as important to represent raw agreement between observers/observations. Crown-root ratio as a more objective approach has the potential to predict root fractions stages accurately in dental age estimation.

9. REFERENCES

1. Catherine Adams. Forensic Odontology: An Essential Guide. First Edition. John Wiley & Sons, Ltd; 2014. Disaster victim identification.
2. Refugees, U. (2020). UNHCR Global Trends 2018. [online] UNHCR. Available at: https://www.unhcr.org/uk/statistics/unhcrstats/5d08d7ee7/unhcr-global-trends-2018.html [Accessed 23 Feb. 2020].
3. Eurostat Press Office. (2020, February 23). Asylum in the EU Member States 580 800 first-time asylum seekers registered in 2018, down by 11% compared with 2017 Syrians, Afghans and Iraqis continued to be the top citizenships. European Commission. https://ec.europa.eu/eurostat/documents/2995521/9665546/3-14032019-AP-EN.pdf/eca81dc5-89c7-4a9d-97ad-444b6bd32790
4. Refugees, U. (2020). UNHCR Global Trends 2017. [online] UNHCR. Available at: https://www.unhcr.org/uk/statistics/unhcrstats/5b27be547/unhcr-global-trends-2017.html [Accessed 23 Feb. 2020].
5. UNHCR The UN Refugee Agency. UNHCR Statistical Yearbook 2009 Chapter IV—Asylum and refugee status determination: UNHCR; 2009.
6. Nuzzolese, E. and Di Vella, G., 2008. Forensic dental investigations and age assessment of asylum seekers. *International dental journal*, *58*(3), pp.122-126.
7. Trines, S. (2020). *The State of Refugee Integration in Germany in 2019*. [online] WENR. Available at: https://wenr.wes.org/2019/08/the-state-of-refugee-integration-in-germany-in-2019 [Accessed 23 Feb. 2020].
8. Nuzzolese, E., Solarino, B., Liuzzi, C. and Di Vella, G., 2011. Assessing chronological age of unaccompanied minors in Southern Italy. *The American journal of forensic medicine and pathology, 32*(3), pp.202-207.
9. Krishan, K., Kanchan, T. and Garg, A.K., 2015. Dental evidence in forensic identification–An overview, methodology and present status. *The open dentistry journal*, *9*, p.250.
10. Timmins, K., Liversidge, H., Farella, M., Herbison, P. and Kieser, J., 2012. The usefulness of dental and cervical maturation stages in New Zealand children for disaster victim identification. *Forensic science, medicine, and pathology*, *8*(2), pp.101-108.
11. Valenzuela, A., Martin-De Las Heras, S., Marques, T., Expósito, N. and Bohoyo, J.M., 2000. The application of dental methods of identification to human burn victims in a mass disaster. *International journal of legal medicine*, *113*(4), pp.236-239.
12. Mughal, A.M., Hassan, N. and Ahmed, A., 2014. Bone age assessment methods: A critical review. *Pakistan journal of medical sciences, 30*(1), p.211.
13. Gulati, A., Taneja, J.R., Chopra, S.L. and Madan, S., 1991. Inter-relationship between dental, skeletal and chronological ages in well-nourished and mal-nourished children. *Journal of the Indian Society of Pedodontics and Preventive Dentistry*, *8*(1), pp.19-23.
14. Dahlberg, P.S., Mosdøl, A., Ding, Y., Bleka, Ø., Rolseth, V., Straumann, G.H., Skjerven-Martinsen, M., Delaveris, G.J.M. and Vist, G.E., 2019. A systematic review of the agreement between chronological age and skeletal age based on the Greulich and Pyle atlas. *European radiology*, *29*(6), pp.2936-2948.
15. Roche, A.F., Davila, G.H. and Eyman, S.L., 1971. A comparison between Greulich-Pyle and Tanner-Whitehouse assessments of skeletal maturity. *Radiology*, *98*(2), pp.273-280.

16. Greulich, W. and Pyle, S. (1964). *Selected standard plates from the Radiographic atlas of skeletal development of the hand and wrist, second edition, 1959*. Stanford, Calif.: Stanford University Press.
17. Tanner, J.M., Whitehouse, R.H., Cameron, N., Marshall, W.A., Healy, M.J.R. and Goldstein, H., 2001. *Assessment of skeletal maturity and prediction of adult height (TW2 method)* (pp. 1-110). London: Saunders.
18. Khan, K. and Elayappen, A.S., 2012. Bone growth estimation using radiology (Greulich–Pyle and Tanner–Whitehouse methods). In *Handbook of Growth and Growth Monitoring in Health and Disease* (pp. 2937-2953). Springer, New York, NY.
19. Lamparski, D.G., 1972. Skeletal assessment utilizing cervical vertebrae. *Master of dental thesis*, *Univ of Pittsburgh*, *School of Dental Medicine*.
20. Chapman, S.M., 1972. Ossification of the adductor sesamoid and the adolescent growth spurt. *The Angle Orthodontist*, *42*(3), pp.236-244.
21. Grave, K.C. and Brown, T., 1976. Skeletal ossification and the adolescent growth spurt. *American journal of orthodontics*, *69*(6), pp.611-619.
22. Román, P.S., Palma, J.C., Oteo, M.D. and Nevado, E., 2002. Skeletal maturation determined by cervical vertebrae development. *The European Journal of Orthodontics*, *24*(3), pp.303-311.
23. Kraus, B.S., 1959. Calcification of the human deciduous teeth. *The Journal of the American Dental Association*, *59*(6), pp.1128-1136.
24. Moorrees, C.F., Fanning, E.A. and Hunt Jr, E.E., 1963. Age variation of formation stages for ten permanent teeth. *Journal of dental research*, *42*(6), pp.1490-1502.
25. Liversidge, H.M. and Marsden, P.H., 2010. Estimating age and the likelihood of having attained 18 years of age using mandibular third molars. *British dental journal*, *209*(8), p.E13.
26. Kreiborg, S. and Jensen, B.L., 2018. Tooth formation and eruption–lessons learnt from cleidocranial dysplasia. *European journal of oral sciences*, *126*, pp.72-80.
27. Ranta, R., 1986. A review of tooth formation in children with cleft lip/palate. *American journal of orthodontics and dentofacial orthopedics*, *90*(1), pp.11-18.
28. Esan, T.A. and Schepartz, L.A., 2020. Does nutrition have an effect on the timing of tooth formation?. *American Journal of Physical Anthropology*, *171*(3), pp.470-480.
29. Elamin, F. and Liversidge, H.M., 2013. Malnutrition has no effect on the timing of human tooth formation. *PloS one, 8*(8).
30. Damm, D.D. and Fantasia, J.E., 2000. Failure of eruption. Adenomatoid odontogenic tumor. *General dentistry*, *48*(6), pp.650-722.
31. Rad, A.S. and Reid, J., 1996. Delayed eruption of a permanent molar associated with a complex odontoma: report of case. *ASDC journal of dentistry for children*, *63*(4), p.299.
32. Wg, S., 1983. Hine MK, Levy BM. A textbook of oral pathology. *WB Saunders Co.: Philadelphia*, *26*, p.5.
33. Neville, B.W., Damm, D.D., Allen, C.M. and Chi, A.C., 2015. *Oral and maxillofacial pathology*. Elsevier Health Sciences.
34. Barbería, E.L., Marañes, J.P., Mourelle, M.M. and Moreno, J.G., 1988. Tooth eruption in children with growth deficit. *Journal of the International Association of Dentistry for Children*, *19*(2), pp.29-35.
35. Suri, L., Gagari, E. and Vastardis, H., 2004. Delayed tooth eruption: pathogenesis, diagnosis, and treatment. A literature review. *American Journal of Orthodontics and Dentofacial Orthopedics*, *126*(4), pp.432-445.

36. Chaudhry, K., Agarwal, A. and Rehani, U., 2010. Interrelationship among dental, skeletal and chronological ages in urban and rural female children. *International journal of clinical pediatric dentistry*, *3*(2), p.79.
37. Nolla, C.M., 1952. *The development of permanent teeth* (Doctoral dissertation, University of Michigan).
38. Gleiser, I. and Hunt Jr, E.E., 1955. The permanent mandibular first molar: its calcification, eruption and decay. *American journal of physical anthropology*, *13*(2), pp.253-283.
39. Anderson, D.L., Thompson, G.W. and Popovich, F., 1976. Age of attainment of mineralization stages of the permanent dentition. *Journal of Forensic Science*, *21*(1), pp.191-200.
40. Demirjian, A., Goldstein, H. and Tanner, J.M., 1973. A new system of dental age assessment. *Human biology*, pp.211-227.
41. AlQahtani, S.J., Hector, M.P. and Liversidge, H.M., 2010. Brief communication: the London atlas of human tooth development and eruption. *American Journal of physical anthropology*, *142*(3), pp.481-490.
42. Olze, A., Van Niekerk, P., Ishikawa, T., Zhu, B.L., Schulz, R., Maeda, H. and Schmeling, A., 2007. Comparative study on the effect of ethnicity on wisdom tooth eruption. *International journal of legal medicine*, *121*(6), pp.445-448.
43. Ubelaker, D.H., 1978. Estimating sex, stature and age. *Human skeletal Remains, excavation, analysis, interpretation. Chicago*.
44. Schour, I., 1941. The development of the human dentition. *Journal of the American Dental Association*, *28*, pp.1153-1160.
45. Schour I, Massler M. 1941b. Development of human dentition chart, 2nd ed. Chicago: American Dental Association.
46. Garn, S.M., Lewis, A.B. and Polacheck, D.L., 1959. Variability of tooth formation. *Journal of dental research*, *38*(1), pp.135-148.
47. Smith, B.H., 1991. Standards of human tooth formation and dental age assessment. Wiley-Liss Inc..
48. Ghougassian, S.S. and Ghafari, J.G., 2014. Association between mandibular third molar formation and retromolar space. *The Angle Orthodontist*, *84*(6), pp.946-950.
49. Demirjian, A. and Goldstein, H., 1976. New systems for dental maturity based on seven and four teeth. *Annals of human biology*, *3*(5), pp.411-421.
50. Liversidge, H.M., 2012. The assessment and interpretation of Demirjian, Goldstein and Tanner's dental maturity. *Annals of human biology*, *39*(5), pp.412-431.
51. Cameriere, R., Ferrante, L. and Cingolani, M., 2006. Age estimation in children by measurement of open apices in teeth. *International Journal of Legal Medicine*, *120*(1), pp.49-52.
52. Thevissen, P.W., Fieuws, S. and Willems, G., 2011. Third molar development: measurements versus scores as age predictor. *Archives of oral biology*, *56*(10), pp.1035-1040.
53. Garn, S.M., Lewis, A.B. and Bonné, B., 1962. Third molar formation and its development course. *The Angle Orthodontist*, *32*(4), pp.270-279.
54. Rowlands, P., Poling, R., Slater, D., Hobson, R. and Steen, N., 2006. Can crown-root ratio predict premolar eruption?. *American journal of orthodontics and dentofacial orthopedics*, *129*(3), pp.331-336.
55. Hölttä, P., Nyström, M., Evälahti, M. and Alaluusua, S., 2004. Root–crown ratios of permanent teeth in a healthy Finnish population assessed from panoramic radiographs. *The European Journal of Orthodontics*, *26*(5), pp.491-497.
56. Othman, N., Taib, H. and Mokhtar, N., 2011. Root-crown ratios of permanent teeth in Malay patients attending HUSM Dental Clinic. *Archives of Orofacial Sciences*, *6*(1), pp.21-26.

57. Ferrante, L. and Cameriere, R., 2009. Statistical methods to assess the reliability of measurements in the procedures for forensic age estimation. *International journal of legal medicine*, *123*(4), pp.277-283.
58. Sun, S., 2011. Meta-analysis of Cohen's kappa. *Health Services and Outcomes Research Methodology*, *11*(3-4), pp.145-163.
59. Shrout, P.E., 1998. Measurement reliability and agreement in psychiatry. *Statistical methods in medical research*, *7*(3), pp.301-317.
60. Staaf, V., Mörnstad, H. and Welander, U., 1991. Age estimation based on tooth development: a test of reliability and validity. *European Journal of Oral Sciences*, *99*(4), pp.281-286.
61. Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, *20*(1), pp.37-46.
62. Zaiontz, C., 2014. Real statistics using Excel: Cohen's kappa.
63. Cohen, J., 1968. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, *70*(4), p.213.
64. Feinstein, A.R. and Cicchetti, D.V., 1990. High agreement but low kappa: I. The problems of two paradoxes. *Journal of clinical epidemiology*, *43*(6), pp.543-549.
65. Wan, T.A.N.G., Jun, H.U., Hui ZHANG, P.W. and Hua, H.E., 2015. Kappa coefficient: a popular measure of rater agreement. *Shanghai archives of psychiatry*, *27*(1), p.62.
66. Dhanjal, K.S., Bhardwaj, M.K. and Liversidge, H.M., 2006. Reproducibility of radiographic stage assessment of third molars. *Forensic science international*, *159*, pp.S74-S77.
67. National Radiological Protection Board, 1994. Guidelines on Radiology Standards for Primary Dental Care: Report by the Royal College of Radiologists and the National Radiological Protection Board.
68. Martínez Gutiérrez, V.M. and Ortega-Pertuz, A.I., 2017. Comparison of Nolla, Demirjian and Moorrees methods for dental age calculation for forensic purposes. *Revista Odontológica Mexicana*, *21*(3), pp.155-164.
69. Fanning, E.A., 1961. A longitudinal study of tooth formation and root resorption. *New Zealand Dent. J.*, *57*, pp.202-217.
70. Thorson, J. and Hägg, U., 1991. The accuracy and precision of the third mandibular molar as an indicator of chronological age. *Swedish dental journal*, *15*(1), pp.15-22.
71. Maclure, M. and Willett, W.C., 1987. Misinterpretation and misuse of the kappa statistic. *American journal of epidemiology*, *126*(2), pp.161-169
72. Sim, J. and Wright, C.C., 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical therapy*, *85*(3), pp.257-268.
73. White, S.C. and Pharoah, M.J., 2014. *Oral radiology-E-Book: Principles and interpretation*. Elsevier Health Sciences.
74. Mufadhal, A.A., Aldawla, M.A. and Madfa, A.A., 2019. External and Internal Anatomy of Maxillary Permanent First Molars. In *Basics of Dental Morphology and Anatomy*. IntechOpen.
75. Versiani, M.A., Basrani, B. and Sousa-Neto, M.D. eds., 2018. *The root canal anatomy in permanent dentition*. Springer.
76. Fry, R.R., Patidar, D.C., Goyal, S. and Malhotra, A., 2016. Proximity of maxillary posterior teeth roots to maxillary sinus and adjacent structures using Denta scan®. *Indian journal of dentistry*, *7*(3), p.126.
77. Tian, X.M., Qian, L., Xin, X.Z., Wei, B. and Gong, Y., 2016. An analysis of the proximity of maxillary posterior teeth to the maxillary sinus using cone-beam computed tomography. *Journal of endodontics*, *42*(3), pp.371-377.
78. Pagin, O., Centurion, B.S., Rubira-Bullen, I.R.F. and Capelozza, A.L.A., 2013. Maxillary sinus and posterior teeth: accessing close relationship by cone-beam

computed tomographic scanning in a Brazilian population. *Journal of endodontics*, *39*(6), pp.748-751.
79. McHugh, M.L., 2012. Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, *22*(3), pp.276-282.
80. Mach, E. (1865). Über die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut. *Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademic der Wissenschaftlen*, *52*, 303-322.
81. Buckle, C. E., Udawatta, V., & Straus, C. M. (2013). Now you see it, now you don’t: visual illusions in radiology. *Radiographics*, *33*(7), 2087-2102.
82. Yeung, A. W. K. (2022, March). The Diagnostic Relevance and Interfaces Covered by Mach Band Effect in Dentistry: An Analysis of the Literature. In *Healthcare* (Vol. 10, No. 4, p. 632). MDPI.
83. Kullman, L., Tronje, G., Teivens, A. and Lundholm, A., 1996. Methods of reducing observer variation in age estimation from panoramic radiographs. *Dentomaxillofacial Radiology*, *25*(4), pp.173-178.
84. Ladeira, D.B.S., Cruz, A.D., Almeida, S.M. and Bóscolo, F.N., 2012. Influence of the intergonial distance on image distortion in panoramic radiographs. *Dentomaxillofacial Radiology*, *41*(5), pp.417-421.
85. Vazquez, L., Nizamaldin, Y., Combescure, C., Nedir, R., Bischof, M., Dohan Ehrenfest, D.M., Carrel, J.P. and Belser, U.C., 2013. Accuracy of vertical height measurements on direct digital panoramic radiographs using posterior mandibular implants and metal balls as reference objects. *Dentomaxillofacial Radiology*, *42*(2), p.20110429.
86. Sindi, A. S., Al Sanabani, F., Al-Makramani, B. M., Mattoo, K., Adawi, H., Al-Mansour, H., ... & Jain, S. (2022). A Radiographic Study of the Root-to-Crown Ratio of Natural Permanent Teeth in 81 Saudi Adults. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, *28*, e936085-1.
87. Yun, H. J., Jeong, J. S., Pang, N. S., Kwon, I. K., & Jung, B. Y. (2014). Radiographic assessment of clinical root-crown ratios of permanent teeth in a healthy Korean population. *The journal of advanced prosthodontics*, *6*(3), 171-176.
88. Tada, S., Allen, P. F., Ikebe, K., Zheng, H., Shintani, A., & Maeda, Y. (2015). The impact of the crown-root ratio on survival of abutment teeth for dentures. *Journal of dental research*, *94*(9_suppl), 220S-225S.
89. De Bruyn, L., Vranckx, M., Jacobs, R., & Politis, C. (2020). A retrospective cohort study on reasons to retain third molars. *International Journal of Oral and Maxillofacial Surgery*, *49*(6), 816-821.
90. Jakobsson, R. and Lind, V., 1973. Variation in root length of the permanent maxillary central incisor. *European Journal of Oral Sciences*, *81*(4), pp.335-338.
91. Bunyarit, S. S., Nambiar, P., Naidu, M. K., Ying, R. P. Y., & Asif, M. K. (2021). Dental age estimation of Malay children and adolescents: Chaillet and Demirjian's data improved using artificial multilayer perceptron neural network. Pediatric Dental Journal, 31(2), 176-185.
92. Bunyarit, S. S., Jayaraman, J., Naidu, M. K., Yuen Ying, R. P., Nambiar, P., & Asif, M. K. (2020). Dental age estimation of Malaysian Chinese children and adolescents: Chaillet and Demirjian’s method revisited using artificial multilayer perceptron neural network. Australian Journal of Forensic Sciences, 52(6), 681-698.
93. Bunyarit, S. S., Nambiar, P., Naidu, M., Asif, M. K., & Poh, R. Y. Y. (2022). Dental age estimation of Malaysian Indian children and adolescents: applicability of Chaillet and Demirjian’s modified method using artificial neural network. Annals of Human Biology, 49(3-4), 192-199.

94. Chaillet, N., & Demirjian, A. (2004). Dental maturity in South France: A comparison between Demirjian's method and polynomial functions. *Journal of forensic sciences*, *49*(5), JFS2004037.
95. Liversidge, H. M. (2009). Permanent tooth formation as a method of estimating age. In *Comparative dental morphology* (Vol. 13, pp. 153-157). Karger Publishers.

LIST OF APPENDICES

Appendix 01 – Pilot study – Staging methods reliability (kappa)

N	Male=3, Female=7		
n	277		
Age	6 – 15 years		
Method	**Demirjian**	**Moorrees**	**Nolla**
Kappa	0.939	0.938	0.940

Appendix 02 – Pilot study – Nolla's staging method crosstabulation

		Obs2											
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
Obs 1	**0**	11	0	0	0	0	0	0	0	0	0	0	**11**
	1	0	2	0	0	0	0	0	0	0	0	0	**2**
	2	0	0	7	2	0	0	0	0	0	0	0	**9**
	3	0	0	0	4	0	0	0	0	0	0	0	**4**
	4	0	0	0	0	9	0	0	0	0	0	0	**9**
	5	0	0	0	0	0	8	0	0	0	0	0	**8**
	6	0	0	0	0	0	0	20	10	0	0	0	**30**
	7	0	0	0	0	0	0	1	36	9	0	0	**46**
	8	0	0	0	0	0	0	0	5	44	1	0	**50**
	9	0	0	0	0	0	0	0	0	13	13	8	**34**
	10	0	0	0	0	0	0	0	0	0	0	114	**114**
	Total	**11**	**2**	**7**	**6**	**9**	**8**	**21**	**51**	**66**	**14**	**122**	**317**

Appendix 03 – Pilot study – Moorrees' staging method crosstabulation

		Obs2												
		C_{oc}	**$Cr_{1/2}$**	**$Cr_{3/4}$**	**Cr_c**	**R_i**	**Cl_i**	**$R_{1/4}$**	**$R_{1/2}$**	**$R_{3/4}$**	**R_c**	**$A_{1/2}$**	**Ac**	**Total**
Obs1	**C_{oc}**	3	0	0	0	0	0	0	0	0	0	0	0	**3**
	$Cr_{1/2}$	0	1	1	0	0	0	0	0	0	0	0	0	**2**
	$Cr_{3/4}$	0	0	2	3	0	0	0	0	0	0	0	0	**5**
	Cr_c	0	0	0	12	0	0	0	0	0	0	0	0	**12**
	R_i	0	0	0	0	3	1	0	0	0	0	0	0	**4**
	Cl_i	0	0	0	1	4	13	0	0	0	0	0	0	**18**
	$R_{1/4}$	0	0	0	0	0	5	23	8	0	0	0	0	**36**
	$R_{1/2}$	0	0	0	0	0	0	5	15	14	0	0	0	**34**
	$R_{3/4}$	0	0	0	0	0	0	0	2	30	1	0	0	**33**
	R_c	0	0	0	0	0	0	0	0	2	14	9	0	**25**
	$A_{1/2}$	0	0	0	0	0	0	0	0	0	0	7	13	**20**
	Ac	0	0	0	0	0	0	0	0	0	0	2	107	**109**
	Total	**4**	**1**	**3**	**16**	**7**	**19**	**28**	**25**	**46**	**15**	**18**	**120**	**301**

Appendix 04 – Pilot study – Demirjian's staging method crosstabulation

		Obs2								
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**Total**
Obs1	**A**	5	3	0	0	0	0	0	0	**8**
	B	0	2	0	0	0	0	0	0	**2**
	C	0	0	9	2	0	0	0	0	**11**
	D	0	0	5	5	0	0	0	0	**10**
	E	0	0	0	1	63	8	0	0	**72**
	F	0	0	0	0	3	35	0	0	**38**
	G	0	0	0	0	0	9	21	0	**30**
	H	0	0	0	0	0	0	11	122	**133**
	Total	**6**	**5**	**14**	**8**	**66**	**52**	**32**	**122**	**304**

Appendix 05 – Pilot study – Crown-root ratio intra class correlation (ICC)

N	Male=3, Female=7
n	318
Age	6 – 15 years
ICC	0.959

Appendix 06 – First page of publication

Forensic Science International 346 (2023) 111654

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Reliability and limitations of permanent tooth staging techniques

A.H. Abdul Rahim [a,b,1], J.A. Davies [b], H.M. Liversidge [b,*]

[a] Oral & Maxillofacial Clinical Sciences, Faculty of Dentistry, Universiti Malaya, Kuala Lumpur, Malaysia
[b] Queen Mary University of London, Dental Institute, London, United Kingdom

ARTICLE INFO

Article history:
Received 12 December 2022
Received in revised form 21 March 2023
Accepted 23 March 2023
Available online 24 March 2023

Keywords:
Dental age estimation
Reliability
Cohen's kappa
Percentage agreement

ABSTRACT

Reliability, or repeatability, of permanent tooth staging techniques is usually expressed as Cohen's Kappa. This single value obscures information about the quantity and allocation of disagreements. In this study we assess and compare intra-observer reliability of permanent tooth staging techniques described by Nolla, Moorrees et al. and Demirjian et al. The sample was panoramic radiographs of healthy dental patients made up of 100 males and 100 females aged 6–15 years. All permanent teeth on the left side (excluding third molars) were scored twice. Weighted Kappa and percentage agreement were calculated. Results show Kappa values for all teeth combined as 0.918, 0.922 and 0.938 for Demirjian (number of teeth N = 2682), Nolla (N = 2698) and Moorrees (N = 2674) respectively. A comparison of Kappa values between upper and lower teeth showed marginally higher values for upper incisors and lower molar for all three scoring methods. Small differences in Kappa values were noted between tooth types with the upper first molar having smaller values than other teeth. Percentage agreement ranged from 81 % (Moorrees), 86 % (Nolla) to 87 % (Demirjian). Tooth stage differences between first and second assessments were not more than one stage. Our findings show that Demirjian scoring is marginally more reliable than Nolla or Moorrees. We suggest that (1) data for reliability are tabulated in full to show the quantity and allocation of disagreement between first and second readings, and (2) that the reliability sample is sufficiently large with a wide age range to include multiple different tooth stages.

1. Introduction

Dental age estimation is a method where age of an individual is predicted by observing age-related dental variables in relation to the chronological age, such as the crown and root development in children [1–6]. Tooth staging methods are commonly used to estimate dental age in children and young adults prior to full maturity of the permanent dentition. These qualitative, categorical methods involve observation of tooth formation usually from dental radiographs, where the tooth formation process is divided into consecutive stages, each of which can estimate age from published reference data.

Several staging methods have been described each having a different number of stages. Nolla's staging method [1] divides tooth formation into 11 main stages, from absence of a crypt to the completed apical root end. Crown and root formation are expressed in fractions (one or two thirds). Similarly, Moorrees et al.'s method [2] (hereafter cited as Moorrees) also expressed developing crowns and roots in fractions but using quarters (1/4, 1/2 and 3/4) with a total of fourteen developmental stages. Demirjian et al.'s staging method [5] (hereafter cited as Demirjian) categorised tooth development process into eight stages from A (initial mineralisation) to H (mature apex). The criteria for tooth staging for the Demirjian method [5] focus on morphological and structural changes seen in tooth development, rather than the estimated fractions of heights and lengths of the crown or root. One of the criteria for Demirjian stage E and F states that root length must be less or more than the crown height for a tooth to be classified into either stage, respectively. This shows that relative morphological comparison (crown height/root length) is inevitable in most staging techniques.

Reliability is an important statistical concept in dental age estimation. The terms 'reliability' and 'validity' are often used in literature interchangably to address the accuracy of an age estimation method although they do not carry the same meaning nor function. Reliability, or precision is "the degree to which further measurements or calculations give the same or similar results" [7]. It can be classified into two categories, *repeatability* and *reproducibility*. Repeatability is the ability of an observer (intra-examiner) to produce the same results after two or more observations. Reproducibility, on

* Corresponding author.
E-mail address: h.m.liversidge@qmul.ac.uk (H.M. Liversidge).

[1] Sponsorship: Ministry of Higher Education, Malaysia.

https://doi.org/10.1016/j.forsciint.2023.111654