Creation of Modern Computational Prediction Model of Unerupted Canine and Premolars for Mixed Dentition Analysis in Thais

Theerada Mungthaweepongsa[*1], Somchai Manopatanakul[2] and Natchalee Srimaneekarn[3]

[1]Graduate Student, Master of Science Program in Advanced General Dentistry, Faculty of Dentistry, Mahidol University, Bangkok, Thailand
[2]Department of Advanced General Dentistry, Faculty of Dentistry, Mahidol University, Bangkok, Thailand
[3]Department of Anatomy, Faculty of Dentistry, Mahidol University, Bangkok, Thailand
[*]Corresponding author, E-mail: theerada.tm@gmail.com

Abstract

Artificial intelligence (AI) disrupts the world, and dentistry is not an exception. This AI disruption creates a modern computational tool that leads to a more accurate orthodontic diagnosis. The orthodontic prediction tool, mixed dentition analysis, which is essential for the prediction of unerupted permanent canine and premolars width, now becomes possible. This prediction model enables the dental crowding assessment; therefore, it facilitates the decision-making of whether an orthodontic treatment plan requires the removal of teeth or not. Hence, the purpose of this study was to formulate and test the novel prediction of the mixed dentition analysis for Thais using AI. The sample consisted of 180 Thai individuals (90 males and 90 females) aged 11 to 30 years old. The mesiodistal widths of all maxillary and mandibular fully erupted permanent teeth, except the second and third molar teeth, were measured with pointed end digital vernier calipers. Then, the simple linear regression analysis was performed to create prediction models based on the results from the corrected Akaike information criteria (AICc). Finally, leave-one-out cross-validation (LOOCV) was used to validate these predictions. As a result, the prediction models were developed in the form of prediction equations. For example, the male's maxillary prediction formula was Y= 0.78X + 27.40 where Y represented the predicted widths of canine and premolars and X represented the widths of mandibular central and lateral incisors. A validation as finished by LOOCV also showed that mean absolute errors (MAEs) were from 1.34 to 1.49 mm per arch. To summarize, the calculated mesiodistal width of the unerupted canine and premolars using the mandibular central and lateral incisors was proposed as another option. Within this margin of error, this novel model is considered easy and practical to predict the size of the unerupted teeth during mixed dentition for Thais.

Keywords: *mesiodistal tooth width, mixed dentition analysis, prediction model, Thais*

1. Introduction

In the 1950s, the concept of Machine intelligence or Artificial intelligence (AI) was conceived as the broadest term for machines capable of perception, reasoning, and learning. These AI capabilities enable computers to mimic the intelligence of the human brain. Machine learning is a subset of AI and employs algorithms that learn from data to make decisions or predictions, and its model performance improves when trained to big data over time (Figure 1; Tandon & Rajawat, 2020).

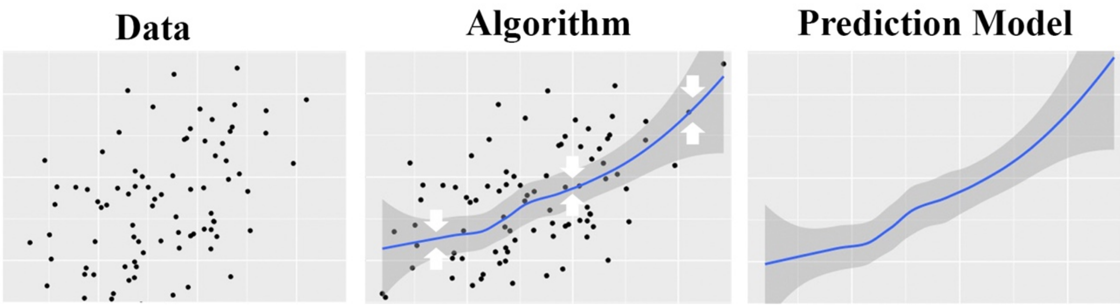

Figure 1 With AI, specifically machine learning, the data were turned into the algorithm of a prediction. Equipped with modern statistical freeware and computers, machine learning enables many precise prediction models. Specifically, with the use of loop calculations of correlation, regression, and goodness-of-fit parameters in RStudio software, the best prediction is at hand. Instantaneously, the goodness-of-fit parameter allows providers to select the most precise prediction model with a more practical and comprehensible error of the estimate. Additionally, by implementing more big data to this algorithm, it is capable of learning, which, as a result, provides a precise continuing loop calculation and improves its precision over time.

At this stage, AI is disrupting various fields including dentistry. The elegance of AI, or to be precise, machine learning, is that the machines can be trained to evaluate and memorize all various dental records (Chen, Stanley, & Att, 2020; Khanagar et al., 2021). As such, AI will soon become part of our daily lives and support an orthodontic diagnosis by analyzing large data set of dental models and radiographs (Shan, Tay, & Gu, 2021). Also, the capacity for analyzing large data sets of this modern computational system enables the formulation and verification of the prediction models. This process reduces the need for train and test sets of samples while providing more accurate prediction models (Donatelli & Lee, 2015).

Orthodontic mixed dentition analysis estimates the mesiodistal widths of unerupted permanent canine and premolars. This estimate will be compared to the amount of space available in the dental arch. Therefore, the prediction model acts as part of an orthodontic assessment of dental crowding or spacing. With modern technology, the final decision of whether to remove some teeth or not during orthodontic treatment is possible from this computational prediction model (Proffit, 2007). A novel computational system also offers a more accurate method for the selection and verification of the prediction models. Moreover, the modern statistical method with a modern powerful computer allows the use of data for both train and test sets of the sample. Furthermore, the selection of the best predictor variables is also possible, and it provides more accurate prediction models. A calculation of the massive databases used to take days to finish; but with the modern computer and smarter software, the calculation presently is finished within minutes. Hence, tooth width prediction calculated from a more massive database becomes feasible.

Several methods have been developed to predict the size of unerupted teeth. The most common methods among these are tables of Moyers (Moyers, 1988) and Tanaka and Johnston equations (Tanaka & Johnston, 1974). Both were formulated from the Northern European progenies. Since this tooth width prediction was reported to be accurate only when it was formulated and applied to the same racial group. This modern computational system also has made it possible to predict the tooth width for orthodontic diagnosis, especially from a more massive and specific database exclusively for Thais (Manopatanakul et al., 2018).

2. Objectives

The principal objective of this study was to formulate and verify the prediction equations of the mesiodistal widths of the unerupted permanent canine and premolars for Thais using AI. This study also aimed to calculate the estimation error of these equations in a more comprehensible millimeter unit.

3. Materials and Methods

In this cross-sectional study, the dental models of Thai patients were recruited from the Department of Orthodontics, Faculty of Dentistry, Mahidol University. Upon completion of the ethical approval from the

Institutional Review Board of the Faculty of Dentistry/Faculty of Pharmacy, Mahidol University (COE. No. MU-DT/PY-IRB 2020/064.0912), data collection was then initiated.

3.1 Sample and Data Collection

The dental casts were from 180 Thai patients (90 males and 90 females) at the age of 11 to 30 years. These dental casts showed fully erupted permanent teeth. These conditions were established as inclusion criteria. Exclusion criteria were subjects with caries, proximal restorations, attrition, or dental anomalies. The mesiodistal widths of all fully erupted permanent teeth, except the second and third molars teeth, were then measured with pointed end digital vernier calipers. To verify the reliability of the measurement of these mesiodistal tooth widths, two investigators were blinded to the subjects, and a high intraclass correlation coefficient (ICC = 0.90) was set to ensure the measurement agreement.

3.2 Data analysis and Statistical Analysis

To investigate the symmetry and sexual dimorphism of the tooth width, the tooth widths from the left were compared to the right side using the paired t-test, and female tooth widths were also compared with male tooth widths using the independent samples t-test. A simple linear regression analysis was performed to create and validate these prediction models. To search for the best predictors, the corrected Akaike information criteria (AICc) was finished. Bland Altman (BA) agreement analysis was also included to test the measurement error of the predictors. Then, the leave-one-out cross-validation (LOOCV) was conducted to validate these prediction models. Mean absolute error and root mean square error (MAE and RMSE) were also finished accordingly. All data and statistical analysis were analyzed using RStudio software (version 1.2.1335).

4. Results and Discussion

4.1 Result

Individual tooth widths as separated and combined genders are shown in Table 1. The ICC was 0.90, indicating high interrater reliability. Further, the normality test of the sum of the measured mesiodistal width of all permanent teeth was completed using the Shapiro-Wilk test, and it showed a normal distribution (p>0.05). As a result, these normally distributed variables were statistically analyzed using the parametric methods.

Table 1 Descriptive statistics for the maxillary and mandibular mesiodistal widths of individual teeth in millimeter units are shown. These tooth widths are shown as separated and combined genders.

Maxillary tooth	Male		Female		Combined genders	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Central incisor	8.78	0.54	8.45	0.54	8.62	0.56
Lateral incisor	7.27	0.53	7.06	0.57	7.16	0.56
Canine	8.20	0.43	7.81	0.47	8.01	0.49
First premolar	7.54	0.38	7.36	0.43	7.45	0.42
Second premolar	7.04	0.43	6.91	0.44	6.98	0.44
First molar	10.40	0.47	10.13	0.44	10.27	0.47
Mandibular tooth	**Male**		**Female**		**Combined genders**	
	Mean	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
Central incisor	5.46	0.36	5.37	0.39	5.42	0.38
Lateral incisor	6.13	0.38	5.93	0.42	6.03	0.41
Canine	7.18	0.44	6.72	0.43	6.95	0.49
First premolar	7.50	0.40	7.29	0.40	7.39	0.41
Second premolar	7.35	0.39	7.20	0.44	7.27	0.42
First molar	11.35	0.47	10.97	0.50	11.16	0.52

S.D.: Standard deviation

Determination of new formulas

When the sum of the mesiodistal widths of canines and premolars of the right and left sides were compared, a paired t-test showed no significant difference in the mandibular tooth (p>0.05). Therefore, these equations were formulated using the left and right sides of the teeth combined. However, a significant difference was found between the male and female samples ($p<0.05$), which indicated that the prediction models should be analyzed separately regarding gender. Then, to find the best predictor from the erupted tooth widths, AICc was used. The AICc indicated that the sum of four lower incisors was the best predictor. BA analysis showed that the mean error of measurements of these predictors combined was only -0.27 mm, with the level of agreement (LoA) estimated at 95% confidence intervals of 1.08 to -1.62 mm (Figure 2).

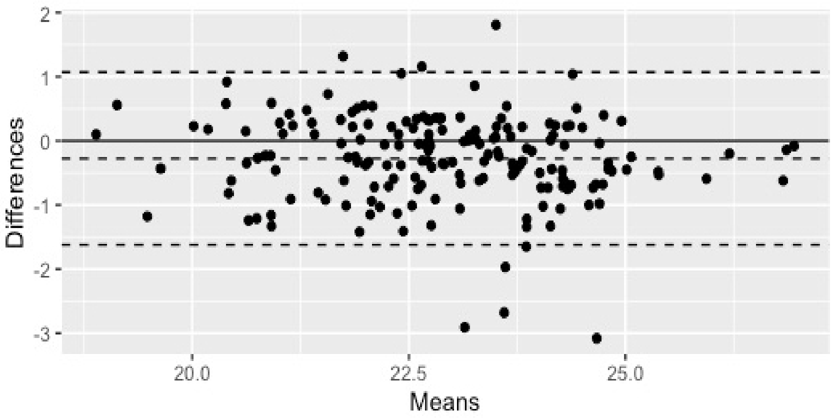

Figure 2 Bland Altman agreement analysis showed the mean differences between two observers' measurements of the predictors. The results showed the average difference and the level of agreement (LoA) estimated at 95% confidence intervals. The mean difference was -0.27 mm. Ninety-five percent LoA were from 1.08 to -1.62 mm.

Therefore, the most practicable equations were formulated from the simple linear regression analysis using four incisors as the predictors. The prediction equations were presented as follows.

Male:	Maxillary arch:	$Y = 0.78X + 27.40$
	Mandibular arch:	$Y = 0.80X + 25.58$
Female:	Maxillary arch:	$Y = 0.82X + 25.61$
	Mandibular arch:	$Y = 0.88X + 22.43$

From these prediction formulas, Y represented the predicted widths in the millimeter unit of the canine and premolars of one arch while X represented the widths of the mandibular central and lateral incisors also in the millimeter unit.

Validation of prediction models

The prediction model performance of the new equations was completed via LOOCV. When assessing the performance of models, the correlation coefficient (r), determination coefficient (r^2), RMSE, and MAE for both male and female prediction models were determined. The correlation coefficients between the mesiodistal width of the sum of the canine and premolars per arch and the sum of mandibular central and lateral incisors ranged from 0.52 to 0.64. Also, the MAEs of the maxillary formulas were 1.34 mm (RMSE = 1.77) and 1.51 mm (RMSE = 1.84) for male and female equations, respectively (Table 2). Likewise, for the mandibular formulas, the MAEs were 1.49 mm (RMSE = 1.77) and 1.31 mm (RMSE = 1.65). While r and r^2 of this study are low, the MAEs were comparable to other similar studies.

Table 2 The goodness-of-fit parameters, i.e., RMSEs and MAEs, using LOOCV for both maxillary and mandibular arches were showed. The correlation coefficient (r) and determination coefficient (r^2) also included a relative comparison of regression parameters among various populations.

Study, year Racial group (N)	Sex	Arch	r	r^2	RMSE	MSE	MAE/Arch (mm)	MAE/Quadrant (mm)
Lee-Chan et al., 1998	M+F	Mx	0.64					
Asian-American (201)		Md	0.66					
Yuen et al., 1998	M	Mx	0.77	0.63				0.67
Hong Kong Chinese (112)		Md	0.79	0.60				0.79
	F	Mx	0.69	0.43				0.60
		Md	0.65	0.48				0.80
Moghimi et al., 2012	M+F	Mx	0.74			1.95		
Iranian (106)		Md	0.70			1.86		
Manopatanakul et al., 2018	M+F	Mx	0.70		0.70			0.55
Vietnamese (8)		Md	0.75		0.86			0.72
The present study (180)	M	Mx	0.52	0.24	1.77		1.34	0.67
		Md	0.53	0.25	1.77		1.49	0.75
	F	Mx	0.57	0.29	1.84		1.51	0.76
		Md	0.64	0.38	1.65		1.31	0.66

M: Male, F: Female, Mx: Maxilla, Md: Mandible

RMSE: Root mean square error, MSE: Mean square error, MAE: Mean absolute error

4.2 Discussion

The smartness of AI or machine learning, specifically for this study, is that the machines can be trained to evaluate and memorize large data sets. With this learning process of the massive data, AI delivers the most pragmatic prediction. Besides, AI facilitates orthodontic diagnosis by completing time-consuming and boring tasks in a very short time, thus, reducing human error from this big data calculation. The diagnosis will also become more practical and accurate. Shortly, AI will carry out the orthodontic diagnosis by analyzing the models, radiographs, and photographs. It may further help to analyze the progression of treatment and alerting clinicians for early detection of unwanted tooth movement. Eventually, AI will lead to a much better final result of orthodontic treatment (Khanagar et al., 2021; Shan et al., 2021).

Space analysis is considered crucial in orthodontic diagnosis and treatment planning. The difference between available and required spaces in the arch can be predicted during mixed dentition. As mentioned, the result of this prediction model facilitates the calculation of space for teeth alignment, which leads to the decision whether to expand the dental arch, distalization of posterior teeth, stripping, or aggressive extraction of some teeth to provide space for teeth alignment (Proffit, 2007). The prediction models based on the mesiodistal width of the erupted permanent teeth are used to predict the width of unerupted canine and premolars. With this information, the aforementioned calculation and prediction results ensue (Manopatanakul et al., 2018).

Traditional mixed dentition analyses were performed by regression equations utilizing the conventional statistical analysis and/or radiographic evaluation of the dentition performed on only descendants of the northern European origin. (Hixon, 1958; Moyers, 1988; Staley & Kerber, 1980; Tanaka & Johnston, 1974). Thus, the relative comparison of regression parameters with other populations was suggested and performed by studies conducted on American born Asian (Lee-Chan, Jacobson, Chwa, & Jacobson, 1998), Hong Kong Chinese (Yuen, Tang, & So, 1998), Iranian (Moghimi, Talebi, & Parisay, 2012), and Vietnamese populations (Manopatanakul et al., 2018). Accordingly, several equations have been proposed for populations of different ethnic groups. At present, modern computational statistical analyses allow researchers to select the best prediction models from big data, develop a more reliable equation, and also permit the verification of the models. Besides, the best prediction models were chosen using AICc that is a fined technique for evaluating how well a model fits the data it was generated from. In statistics, the AICc is used to compare different possible models and determine which one is the best fit for the data (Akaike, 1974). To achieve the best possible prediction accuracy, this prediction model must be carefully validated.

The traditional simple validation methods are simple split or holdout methods. The data were used as the training set from which the prediction models were computed, and the other set was for validating the model. Since this study aimed to formulate and validate the prediction models for Thais, the AICc and LOOCV were completed. In detail, the full training data set was divided by random sampling, and one of these resulting sets was used as a training set (Donatelli & Lee, 2015; James G, 2013).

After fitting the model to these data, the model was then tested on the left-out data set. This process was repeated until all folds had been used to test the data sets. Then, the MAE, with the millimeter unit, was calculated as a metric appropriate for the prediction model. The MAE also allows the clinicians to easily understand the clinical insignificance. These errors were then compared with other studies on the degree of crowding against the limit of space available (Table 2). With the severe crowding set at 6-8 mm for extraction decision, this MAE allows the clinicians to visualize that this prediction error is justifiable. Unlike r, the MAE demonstrates the clinical insignificance of the error of the prediction model. In summary, these modern computational methods promote the strength of this study. On the other hand, the most significant weakness of this study was that the capacity of this modern calculation was not fully utilized. Its full capacity of calculation could be extended to perform on the sample size of way more than one thousand. Hence, it should be suggested here that further studies should be performed on a much larger sample size. Another weakness is that AI is still new and unfamiliar to dentistry. It took grit to promote the constant learning curve to conduct this study. Explaining these modern issues in a very concise way also limited the elaboration of the context. However, as mentioned, AI disruption is inevitable, and eventually, will be part of our daily lifestyle. Therefore, it would be smart to learn this AI technology fast and hard today. Besides, since tooth size variability exists among different ethnic groups, more studies can be carried out in the neighboring Asian countries. Further, the application of different and more advanced prediction models should be supported. The easy access to the clinicians, for example, mobile phone platforms should promote the popularity of these prediction models, which will, in turn, promote the AI application to dentistry.

5. Conclusion

This study created the prediction models to estimate the width of the unerupted permanent canine and premolars for Thais using AI. The sum of the mandibular central and lateral incisors was used as the best predictors to develop new prediction models. A modern computational method was applied to select, develop, and verify these formulas. This modern method also evaluated the applicability, consistency, and accuracy of these predictions, especially in terms of clinical practicality. It also provided the result of an error in a more comprehensible millimeter unit. As a result, clinical significance was easily evaluated. Thus, these prediction models would be highly useful for patients and providers. More importantly, this study also extends the wider AI usage in dentistry.

6. References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723. doi:10.1109/TAC.1974.1100705

Chen, Y. W., Stanley, K., & Att, W. (2020). Artificial intelligence in dentistry: current applications and future perspectives. *Quintessence Int, 51*(3), 248-257. doi:10.3290/j.qi.a43952

Donatelli, R., & Lee, S. (2015). How to test validity in orthodontic research: a mixed dentition analysis example. *Am J Orthod Dentofacial Orthop, 147*(2), 272-279. doi:10.1016/j.ajodo.2014.09.021

Hixon, E. H., Oldfather, R. E. (1958). Estimation of the sizes of unerupted cusp and bicuspid teeth. *Angle Orthod, 28*, 236-240.

James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). *An Introduction to Statistical Learning*. New York: Springer.

Khanagar, S. B., Al-Ehaideb, A., Maganur, P. C., Vishwanathaiah, S., Patil, S., Baeshen, H. A., Sarode, S. C., Bhandi, S. (2021). Developments, application, and performance of artificial intelligence in dentistry – A systematic review. *Journal of Dental Sciences, 16*(1), 508-522. doi:https://doi.org/10.1016/j.jds.2020.06.019

Lee-Chan, S., Jacobson, B. N., Chwa, K. H., & Jacobson, R. S. (1998). Mixed dentition analysis for Asian-Americans. *Am J Orthod Dentofacial Orthop, 113*(3), 293-299. doi:10.1016/s0889-5406(98)70300-2

Manopatanakul, S., Huong, V. T., Thiradilok, S. (2018). The prediction formula of mesiodistal width of unerupted permanent canine and premolars from a group of Vietnamese, a preliminary study. *M Dent J, 38*(2), 113-123. doi:https://he02.tci-thaijo.org/index.php/mdentjournal/article/view/179309

Moghimi, S., Talebi, M., & Parisay, I. (2012). Design and implementation of a hybrid genetic algorithm and artificial neural network system for predicting the sizes of unerupted canines and premolars. *Eur J Orthod, 34*(4), 480-486. doi:10.1093/ejo/cjr042

Moyers, R. E. (1988). *Handbook of Orthodontics* Chicago: Year Book Medical Publishers.

Proffit, W. Fields., H., Sarver, D. (2007). *Contemporary Orthodontics* (7 ed.). St. Louis, MO: Mosby Elsevier.

Shan, T., Tay, F. R., & Gu, L. (2021). Application of Artificial Intelligence in Dentistry. *Journal of Dental Research, 100*(3), 232-244. doi:10.1177/0022034520969115

Staley, R. N., & Kerber, P. E. (1980). A revision of the Hixon and Oldfather mixed-dentition prediction method. *Am J Orthod, 78*(3), 296-302. doi:10.1016/0002-9416(80)90274-2

Tanaka, M. M., & Johnston, L. E. (1974). The prediction of the size of unerupted canines and premolars in a contemporary orthodontic population. *J Am Dent Assoc, 88*(4), 798-801.

Tandon, D., & Rajawat, J. (2020). Present and future of artificial intelligence in dentistry. *J Oral Biol Craniofac Res, 10*(4), 391-396. doi:10.1016/j.jobcr.2020.07.015

Yuen, K. K., Tang, E. L., & So, L. L. (1998). Mixed dentition analysis for Hong Kong Chinese. *Angle Orthod, 68*(1), 21-28. doi:10.1043/0003-3219(1998)068<0021:Mdafhk>2.3.Co;2