ABSTRACT

Background: Child marriage is recognised as a harmful practice. Recently, Indonesia enacted a new law to raise the minimum age of marriage from 16 to 19 years old for women, creating no minimum age difference between males and females. However, this improvement may be detrimental for individuals in remote areas with no legal documentation and the common practices of age falsification to reach the minimum age of marriage. Therefore, implementing an age estimation technique for juveniles is mandatory to reduce the risk of child marriage. Methods: this study used the third molar maturity index ($I_{3M}$) to distinguish an individual under or over 19 years old. $I_{3M}$ values from 222 digital OPGs aged between 15 to 23.99 years were calculated. The sample was randomly assigned as a training dataset ($n = 156$) and testing dataset ($n = 66$). The logistic regression model was created using a 5-fold cross-validation method, and the Youden's Index Value was used to establish the $I_{3M}$ cut-off value. Results: the logistic regression model showed significance in both sex and $I_{3M}$ value for predicting the probability of minimum age of marriage. $I_{3M}$ cut-off values of 0.08 and 0.09 for males and females, respectively, were taken. The accuracy of this test was 80% for both sexes in the testing dataset. Conclusions: the outcome of this pilot study showed a promising result of using $I_{3M}$ as a dental age estimation method to determine whether an individual is over or under 19 years old to comply with the newly enacted legal age of marriage in Indonesia. Future research should be carried out using a balanced age cohort for each sex and a more extensive training sample size to investigate the influence of sex in the cut-off value calculation.

INTRODUCTION

In forensic sciences, determining an individual’s age is a vital objective acquired through legal documents, including a medical chart or birth certificate. In the absence of these documents, an individual’s chronological age can be estimated through various age-related variables, including teeth. Dental development is highly correlated with chronological age, creating diverse methods to estimate an individual’s age. The practice of dental age estimation has been used in juveniles to determine whether an individual has passed a certain legal age threshold. At this age range, the third molar is the only tooth left to mineralise. Hence, third molar development was commonly used as the main parameter to differentiate between specific age threshold in juveniles.
However, most third molar development techniques have been divided into a finite number of stages, creating a higher error rate if the observed tooth falls between a particular stage. To overcome this limitation, Cameriere et al. researched the Third Molar Maturity Index (I_{3M}) technique, a ratio measurement on orthopantomogram (OPG) images of mandibular third molar and was proven to help distinguish an individual in different legal age thresholds.

In Indonesia, a new law has been enacted to prevent child marriage. It is commonly known that child marriage is a harmful practice. It promotes a higher incidence of sexually transmitted disease, intimate partner violence, and lower economic and educational quality. To prevent this, Indonesia’s new law raised the minimum age of marriage from 16 to 19 years old for women, creating no difference between male and female minimum age of marriage. However, applying the new legal age of marriage threshold is difficult for individuals with no legal documentation and prevent age falsification, a new study in Indonesian law to implement the applicability of the I_{3M} method is needed as there are no studies in determining I_{3M} cut-off value for Indonesian population who are younger or older than 19 years old for both male and female. Therefore, this study aimed to investigate the application of the I_{3M} technique to predict whether an individual is younger (<19 years old) or older than the minimum age of marriage (≥19 years old).

**MATERIAL AND METHODS**

**Sample**
In this retrospective observational study, we collected a total of 222 digital OPG images (M = 73, F = 149) from Indonesian children and juveniles between 15 and 23.99 years old from Pramita Laboratory, Semarang, Indonesia. The sample was selected based on the presence and clarity of the lower left mandibular third molar (LL3rdM) without any recorded developmental abnormalities or dental treatment. The anonymity of the sample was preserved while maintaining the information of patient number, sex, date of birth, and date of exposure. Sample age was obtained from the difference between the date of exposure and the date of birth. The required ethical approval was obtained from the institute’s ethics committee.

**Measurements**
Images were imported and enhanced for optimal visualization using Adobe Photoshop CC 2020 software built-in tools. Furthermore, LL3rdM measurement was performed by the first observer (RMB) in conjunction with Cameriere et al. method. The observer was blinded to the actual age of the image during the measurement. Tooth apical ends and length were analysed, and the I_{3M} was defined as follows: if the root development of LL3rdM was complete, the value of I_{3M} = 0. I_{3M} was calculated by the sum of total tooth apical ends inner margins (A_8 = A_8 + A_8) divided by tooth length (L) from apical ends to the highest point of the crown (I_{3M} = \frac{A_8}{L}).

If the tooth had not developed a bifurcation, the length between the inner crown margins was considered tooth apical end (A_8) (Figure 1).

**Statistical Analysis**
Measurements were collated in an Excel file (Microsoft Excel 365) and processed using R. Twenty five images were randomly selected after two weeks and recalculated by RMB and second observer (HE). RMB had over three years of experience, and HE had just been introduced to dental age estimation. Intraclass Correlation Coefficient (ICC) was used to estimate the Intra-Inter-Observer Reliability using the psych package.

The Caret package was used to calculate the k-fold cross-validation and the linear model. Consequently, 70 percent of the data were randomly selected (set.seed = 100) and assigned as a training dataset (n = 156). The remaining data were used as a testing dataset (n = 66). A logistic regression model was created on the dataset using 5-fold cross-validation with dependent variables of T = 1 and T = 0 for an individual over and under 19 years old, and predictive variables of I_{3M} and sex with s = 1 and s = 0 for male and female,
respectively. The model's predictive accuracy was determined using the receiver operating curve (ROC) and the area under the curve (AUC). Furthermore, the optimal cut-off value was established using the highest Youden's index (J) value with the cutpointr package.\textsuperscript{19}

The cut-off value performance was established in the testing dataset by the terms of accuracy (Eq. 1), Sensitivity (Se) or the percentage of the subjects ≥ 19 years old who had $I_{3M} < \text{cut-off value}$, and Specificity (Sp) or the percentage of the subjects < 19 years old who had $I_{3M} \geq \text{cut-off value}$ was calculated as follows:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

TP (True Positive) represented the results of those who were ≥ 19 years old and had $I_{3M} \leq \text{cut-off value}$. Alternatively, the results of those who were < 19 years old and had $I_{3M} > \text{cut-off value}$ were described as TN (True Negative). FP (False Positive) was the result of those who were < 19 years old and had $I_{3M} \leq \text{cut-off value}$. Finally, those who were ≥ 19 and had $I_{3M} > \text{cut-off value}$ listed as FN (False Negative).

Further evaluation of the model performance was done by calculating the positive predictive value (PPV), negative predictive value (NPV), and positive and negative likelihood ratios (LR +, LR -). PPV and NPV were calculated to examine how many of each positive (≥ 19 years old) and negative (< 19 years old) were correctly classified. LR+ indicated how many TP would be observed per FP. LR- indicated how many FN would be observed per TN.

Bayes post-test probability was calculated (Eq. 2) to help $I_{3M}$ cut-off value distinguished individuals < 19 years old from individuals ≥ 19 years old (i.e., the proportion of individuals with $I_{3M} \leq \text{cut-off value}$ which was ≥ 19 years old):

\[
p = \frac{Se \cdot P_0}{Se \cdot P_0 + (1 - Sp)(1 - P_0)}
\]

Where $p$ was post-test probability and $P_o$ was the proportion of individuals in the target population who were ≥ 19 years old, given that they were between 15-23.9 years old. $P_o$ was calculated from the data obtained from Statistics Indonesia (Badan Pusat Statistik).\textsuperscript{20}

**Figure 1.** Measurement example of third molar maturity index ($I_{3M}$). Root development was completed, $I_{3M} = 0$ (a). The distinct approach was applied when the tooth has developed a bifurcation ($I_{3M} = \frac{A_{81} + A_{82}}{L}$) (b), or not ($I_{3M} = \frac{A_{81} - A_{82}}{L}$) (c).
RESULTS

Table 1 shows the distribution of the training and testing dataset. Figure 2 shows the $I_{3M}$ value in each sex, with the first value of $I_{3M} = 0$ observed in males and females at the age of 21.5 and 20.3 years old, respectively. Pearson’s correlation coefficient between $I_{3M}$ and age was $-0.64$ ($p < 0.001$). The inter- and intra-rater agreement showed excellent results proving the repeatability of the measurement, with an ICC value of 0.98 and 0.96 for inter- and intra-rater agreement, respectively. The logistic regression model displayed the significance of sex ($p < 0.001$) and $I_{3M}$ ($p < 0.001$) as independent variables for predicting the minimum age of marriage. The model may be written as follows:

$$logit(p) = 2.5401 - 13.4052(I_{3M}) - 1.4954(s)$$

The ROC curve is presented in figure 3, with the AUC value of 0.91. As sex became significant as an independent variable, the $I_{3M}$ cut-off value was derived differently for each sex to achieve better accuracy. The $I_{3M}$ for male and female were 0.08 ($J = 0.76$) and 0.09 ($J = 0.76$), respectively. The performance of each cut-off value was analysed in the testing dataset with 80% accuracy in both sexes. The female testing dataset achieved better overall performances ($Se$, $Sp$, PPV, NPV, LR+, and LR-) (Table 2-4). Bayes post-test probability showed that the probability of male and female subjects, with $I_{3M} \leq$ the indicated cut-off value for each sex was 19 years old or older, were 0.87 and 0.92, respectively.

**Table 1.** Age and sex distribution on training and testing dataset

<table>
<thead>
<tr>
<th>Age (Years)</th>
<th>Male Training Dataset</th>
<th>Female Training Dataset</th>
<th>Total Training Dataset</th>
<th>Male Testing Dataset</th>
<th>Female Testing Dataset</th>
<th>Total Testing Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-15.99</td>
<td>10</td>
<td>13</td>
<td>23</td>
<td>2</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>16-16.99</td>
<td>6</td>
<td>12</td>
<td>18</td>
<td>7</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>17-17.99</td>
<td>10</td>
<td>8</td>
<td>18</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>18-18.99</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>19-19.99</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>1</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>20-20.99</td>
<td>2</td>
<td>18</td>
<td>20</td>
<td>1</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>21-21.99</td>
<td>4</td>
<td>14</td>
<td>18</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>22-22.99</td>
<td>4</td>
<td>14</td>
<td>18</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>23-23.99</td>
<td>4</td>
<td>7</td>
<td>11</td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>104</td>
<td>156</td>
<td>21</td>
<td>45</td>
<td>66</td>
</tr>
</tbody>
</table>
**Figure 2.** Box plots showing the relationship between $I_{3M}$ value and chronological age between males and females.

**Figure 3.** Receiver operating characteristic curve for "19 Years or Older" with an Area Under Curve of 0.91.
**DISCUSSION**

The use of $I_3M$ to determine the legal age of marriage in the Indonesian population showed an acceptable result in this initial study, where the age threshold is one year old higher than the original study. The original study by Cameriere et al. was conducted to determine the probability of an individual being older or younger than 18 years old with a cut-off value of $0.08$. This cut-off value has been tested and validated in many countries with high accuracy. Furthermore, the $I_3M$ method also has its versatility in the different legal age thresholds. Balla et al. (2019) applied the $I_3M$ method to derive a cut-off value to predict if an individual has reached the age of 16 in the Indian population, resulting in a cut-off value of 0.293 with an accuracy of 88 and 88.7 percent for both males and females. Another cut-off value was also calculated in the Indian population to determine whether an individual has reached the age of 14 in compliance with the child labour laws. Hence, a new cut-off value needs to be calculated for each legal age implementation applied in the respective population.

In this study, our finding suggests that the third molar development in Indonesian juveniles was slower than in other countries. This might be explained by the optimal cut-off value achieved by our study sample to reach 19 years old. $I_3M$ has a reverse correlation coefficient with chronological age, meaning that the $I_3M$ value will decrease as the individual gets older. Furthermore, Santiago et al. (2018) reported that most of the $I_3M$ studies have a high accuracy in using the $I_3M$ cut-off value of 0.08 to determine the age of the majority, which is 18 years old. Compared to our study, the $I_3M$ values only reached 0.08 and 0.09 at the age of 19 years old, which makes the state of $I_3M$ value in Indonesian male juveniles in the age of 19 is equal to 18 years old juveniles in other countries, and even less developed in female. This result is in line with Tangmose et al. (2015), who recommended that genetic differences, including ethnic origin, play a vital role in the third molar development rate. However, a direct comparison with other Indonesian populations or other cut-off values with 19 years old age threshold is not available due to the non-existent data.

---

**Table 2.** Contingency table describing discrimination performance of $I_3M$ cut-off value ($0.09$) on Female testing dataset

<table>
<thead>
<tr>
<th>Females</th>
<th>Age (years)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 19</td>
<td>≥ 19</td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_3M &gt; 0.09$</td>
<td>16 TN</td>
<td>8 FN</td>
<td></td>
</tr>
<tr>
<td>$I_3M ≤ 0.09$</td>
<td>1 FP</td>
<td>20 TP</td>
<td></td>
</tr>
</tbody>
</table>

TN = True Negative, FN = False Negative, TP = True Positive, FP = False Positive

**Table 3.** Contingency table describing discrimination performance of $I_3M$ cut-off value ($0.08$) on Male testing dataset

<table>
<thead>
<tr>
<th>Males</th>
<th>Age (years)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; 19</td>
<td>≥ 19</td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_3M &gt; 0.08$</td>
<td>13 TN</td>
<td>3 FN</td>
<td></td>
</tr>
<tr>
<td>$I_3M ≤ 0.08$</td>
<td>1 FP</td>
<td>4 TP</td>
<td></td>
</tr>
</tbody>
</table>

TN = True Negative, FN = False Negative, TP = True Positive, FP = False Positive

**Table 4.** Performance description of each $I_3M$ cut-off value in each sex

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Se</td>
<td>0.57</td>
<td>0.71</td>
</tr>
<tr>
<td>Sp</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>PPV</td>
<td>0.57</td>
<td>0.71</td>
</tr>
<tr>
<td>NPV</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>LR+</td>
<td>7.125</td>
<td>11.83</td>
</tr>
<tr>
<td>LR-</td>
<td>0.46</td>
<td>0.3</td>
</tr>
<tr>
<td>PTP</td>
<td>0.87</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Se = Sensitivity, Sp = Specificity, PPV = Positive Predictive Value, NPV = Negative Predictive Value, LR+ = Positive Likelihood Ratio, LR- = Negative Likelihood Ratio, PTP = Bayes’ Post-Test Probability
The $I_{3M}$ cut-off values were taken differently on each sex. This cut-off value separation was done because sex was a significant independent variable in our model ($p < 0.001$). Furthermore, we observed that deriving the cut-off value equally for both males and females ($I_{3M} = 0.094$) gave an overall lower accuracy (0.79 for both males and females). However, the difference in cut-off value performance between males and females should be used carefully since the training sample was not balanced between the sexes. It is essential to note that multiple studies with balanced samples found that sex was not a significant predictor in their logistic regression model. Consequently, the choice of differentiating $I_{3M}$ cut-off value between a specific group (i.e., sex) should depend on the significant independent variable in the model or achieving a better particular value.

We present the study result with various performance descriptions in Table 4, most notably the $Se$ and $Sp$ values. In this context of the study, the $Se$ value represents the percentage of subjects over 19 years old and have the $I_{3M}$ below the cut-off value and are therefore correctly specified as an individual who can be married (TP). Thus, a high $Se$ value represents a low FP, which classifies an individual under 19 years old and classified as being able to marry. However, the $Se$ value in the male samples was lower than the female, which could also be explained by the sex imbalance in our training dataset. On the contrary, the $Sp$ value — which represents the percentage of subjects under 19 years old and having $I_{3M}$ above or equal to the cut-off value — was found to be high in both sexes. Moreover, avoiding a higher FP value is commonly done in the field of age estimation. However, a recent study by Rumble et al. (2018) reported that wealth and education had a significant impact on child marriage ($p < 0.05$). Furthermore, Grijns et al. (2018) reported that the state legal system was creatively interpreted in rural areas and was commonly influenced by religious beliefs. Further studies should address the major drawback of this study by using balanced age cohorts for each sex to calculate the significance of sex in the logistic model and its cut-off value threshold. After all, these findings showed that eliminating child marriage in Indonesia is a complicated matter and using the $I_{3M}$ value to assess dental age can help to assess the individual’s biological maturity.

**CONCLUSIONS**

The outcome of this pilot study showed a promising result of using $I_{3M}$ as a dental age estimation method to determine whether an individual was over or under 19 years old to comply with the newly enacted legal age of marriage in Indonesia. The results indicate that future research should be carried out using balanced age cohorts for each sex and a more extensive training sample size to investigate the influence of sex in the cut-off value calculation. In addition, incorporating other testing systems, such as psychological evaluation, can be used further to improve the quality of minimum age of marriage assessment.

**ACKNOWLEDGEMENTS**

The authors would like to thank Gunawan Wibisono., DDS, M.Sc, Djoko Priyanto., DDS, M.Sc and Pramita Laboratory Semarang for the assistance in obtaining the OPGs.

**FUNDING**

This study is fully funded by the Faculty of Medicine, Universitas Diponegoro internal funds No.85/UN.7.6.4.2/HK/2020.


