Evaluation of Cameriere and Willems age estimation methods in panoramic radiographs of Brazilian children

ABSTRACT
The importance of age estimation in the forensic field is inherent to the process of establishing the biological profile of children, sub-adults and adults. The established profile might be useful for the identification of deceased victims or living individuals when it comes to age of legal interest. In parallel, age estimation is also investigated for clinical purposes, especially for the diagnosis of dental and bone maturation. Several studies were developed to provide accurate age estimation methods based on skeletal and dental development. This study aimed to apply and compare Cameriere’s and Willems’ methods for dental age estimation in a Brazilian sample. Two examiners performed image analysis and method application in 180 panoramic radiographs of Brazilian children aged 6-14 years old. The ages estimated with both methods revealed a good correlation with the chronological ages of Brazilian boys and girls. Cameriere’s method showed a slight underestimation of 0.05 years for girls and 0.03 for boys. Willems’ method, on the other hand, showed an overestimation of 0.47 years for girls and 0.39 for boys. Better age estimates were obtained combining the outcomes of both methods. In practice, Cameriere’s and Willems’ methods reached reliable outcomes and could be applied for dental age estimation purposes.

INTRODUCTION
Age estimation is fundamental in civil justice, especially when it is used in adoption cases1 or applied for asylum seekers,2 and even for any kind of questioned civil registration.3 In the criminal scenario, age estimation contributes to building anthropological profiles of victims4-7 and also supports investigations of alleged minor offenders.8 The process of estimating age in the living increased over time given the several issues inherent to the different countries worldwide, such as authoritarian policies, civil wars and extreme poverty. Countries that offer entry to those in need have the important role of providing personal documents and possibilities to work, and access to health and education. In this context, age estimation becomes a tool to assure human rights.9,10 In South America, Brazil is the country that shelters the highest number of refugees.11 Cameriere et al.12 designed a quantitative approach through a formula based on sex and the ratio between length and apex opening measurements of each lower left tooth. Based on the
classification of dental maturation stages), Willems et al. designed a method able to reach more accurate outcomes within a less time consuming process.

This study aimed to conduct an initial study of the application of Cameriere’s and Willems’ methods in a sample of digital panoramic radiographs from Brazilian children aged from 6 to 14 years. A secondary study focused on analyzing eventual differences in dental maturation between boys and girls.

MATERIAL AND METHODS
This project was approved by the Research Ethics Committee (Protocol: 06634919.7.0000.5419).

A total of 180 digital panoramic radiographs were selected: 90 girls and 90 boys from 6 to 14 years old (10 female and 10 male radiographs for each age, totaling 20 radiographs for each year) to conduct an initial study of the application of age estimation methods. The inclusion criteria were high image resolution and presence of the lower left teeth from central incisor to second molar. Digital panoramic radiographs that showed low image resolution, dental developmental changes, absence or fractures of the left lower teeth were excluded from the sample.

The radiographs were coded and tabulated in a Microsoft Excel® spreadsheet (Microsoft, Washington, USA) with respective code, gender, date of birth, date of X-ray taking and age in years. To apply the methods, the examiners had access only to the panoramic radiographs and their corresponding codes.

Both samples were analyzed by two previously calibrated examiners, who estimated the ages using Cameriere’s method (metric) and Willems’ method (non-metric). In addition, the estimated age by Cameriere’s method was added to the estimated age by Willems’ method and the final value divided by two, resulting in the third value used for statistical calculations: the mean age between the methods. The examiners repeated both methods in 30% of the sample after four weeks - radiographs from 30 girls and 30 boys, for intra-examiner analysis.

To get the necessary measurements to apply Cameriere’s method (distance between the internal sides of the open apex and the total length of the dental element - reduce the possible differences between angulations and distortions of the radiographs), the software ImageJ was used (Wayne Rasband; National Institutes of Health, USA).

For statistical analysis, the following software was used: R Core Team (R Foundation for Statistical Computing, Vienna, Austria) and SAS Statistical Software v9.3 (SAS Institute, Cary, North Carolina, USA). Agreement analysis was calculated via Intraclass Correlation Coefficient (ICC) with 95% confidence interval, in order to verify if the evaluators were calibrated and if the age estimation methods were reproducible. ICC was also used to verify the agreement between the outcomes of mean estimated age in Cameriere’s method and Willems’ method compared to the chronological age of the sampled children. To calculate the agreement and magnitude of the differences between the chronological and estimated ages, the Bland-Altman method and a Linear Mixed-Effects Regression were used.

The estimated differences were calculated from chronological age (chronological age less the age estimation for each method), in other words, when the estimated age was greater than the chronological age, the result was negative. The result was positive when there was an underestimate by the method, because the estimated age was subtracted from the chronological age, resulting in a positive value.

RESULTS
To verify the reproducibility of the methods and the calibration of the examiners, ICC was calculated with its respective 95% confidence interval. Considering that the ICC values approached 1: 0.99 and 0.98 for girls according to examiner 1; 0.98 and 0.96 according to examiner 2, using Cameriere’s and Willems’ methods respectively. For boys, the ICC resulted in a value of 1 and 0.99 according to examiner 1; 0.94 and 0.95 according to examiner 2, using Cameriere’s and Willems’ methods respectively.

A high correlation was detected between the ages estimated by each examiner in the main analysis and 4 weeks later. The methods were reproducible and the examiners were calibrated.

To assess whether or not the estimated age was close to the chronological age (using both methods), the ICC was calculated with its respective 95% confidence interval. The mean ages of Cameriere’s and Willems’ methods were considered during this analysis. The ICC values for girls and boys were close to 1: 0.94 for girls...
according to examiner 1 for both methods; 0.93 and 0.94 according to examiner 2, using Camereire’s and Willems’ methods respectively. For boys, the ICC resulted in a value of 0.92 and 0.94 according to examiner 1; 0.93 and 0.94 according to examiner 2, using Camereire’s and Willems’ methods respectively. In addition, the mean presented an ICC of 0.96 for girls for both examiners and a value of 0.94 for boys according to examiner 1 and 0.95 according to examiner 2. These results suggest that both methods were able to provide age estimates that were close to the chronological age. Interestingly, the ICC was higher when the mean ages estimated from both methods were combined.

To calculate the difference between chronological and estimated ages, Bland-Altman graphs were designed, which demonstrate a statistical graphical analysis for the comparison of the two methods (in this case, Camereire’s method or Willems’ method and the chronological age). The central axis is the mean of the differences between the two measurements, and thus, concordant measurements have difference values close to zero. Graphs were elaborated for girls (Figure 1) and boys (Figure 2).

**Figure 1.** Difference between the real ages and the estimated ages by both methods for girls, as well as their means. CM – Cameriere’s method; WM – Willems’ method; CMWM – Mean between Cameriere’s method and Willems’ method; 2SD – Two Standard Deviations.
Figure 2. Difference between the real ages and the estimated ages by both methods for boys, as well as their means. CM – Cameriere’s method; WM – Willems’ method; CMWM – Mean between Cameriere’s method and Willems’ method; 2SD – Two Standard Deviations.

It is possible to observe that the graphs that represent the difference between the chronological age and the age estimates with Willems’ method have the central axis towards higher negative values, a circumstance that demonstrates an overestimation of age by the method for both sexes – since the difference is calculated from the chronological age. This condition is not evidenced in Cameriere’s method, because the central axis is closer to zero, showing less difference with the chronological age.

To compare the chronological and the estimated age, in general, considering the random effect, a Linear Mixed-Effects Regression model was used to analyze the variables. The estimated differences (chronological age less the age estimation for each method) are presented with their respective $p$ values and 95% confidence intervals. Table 1 compares the chronological age with the estimated age of each method combining the analysis of examiners 1 and 2.
It is possible to observe that the estimated differences for Cameriere’s method are positive and close to zero, demonstrating a slight underestimation and small difference between the real and estimated ages. The estimated difference for Willems’ method is negative and more distant from the value zero, showing an overestimation and higher difference between the real and estimated ages by the method. Considering the confidence interval and the p-value, Cameriere’s method includes the zero value in both sexes, has a little distance between upper and lower limits and demonstrates a p-value higher than 0.05. It indicates lack of statistically significant differences between chronological and estimated ages. For the Willems’ method, the confidence interval is negative and does include zero values, providing evidence of statistically significant differences between chronological and estimated ages. The Willems’ method influences in the overestimation when the mean age is calculated using both methods. It is confirmed because the confidence interval was close to zero and the p-value maintained low, but the estimated differences between chronological and estimated ages were negative. In other words, ICC of the mean is close to 1, there is still an overestimation in relation to the chronological age.

**DISCUSSION**

Before estimating age in large samples, it is necessary to calibrate the examiners in order to analyze if the method is reproducible. This methodological set up might improve the quality of the age estimation procedure and might standardize the analysis between and within examiners. In the present study, it is possible to observe a significant correlation between the first and the second examiner analyses performed for age estimation (ICC between 0.94 and 1). Those values confirm the reproducibility of Cameriere’s and Willems’ methods, as well as the calibration of examiners. Fernandes et al. analyzed exclusively Cameriere’s method, re-evaluating 20 orthopantomographs for intra-examiner analysis, which resulted in a p-value of 0.315 for the first examiner and 0.193 for the second examiner. These outcomes demonstrated no evidence of difference for intra-examiner reproducibility. Similarly, Galic et al. analyzed the reproducibility of the Willems’ method, re-examining 10% of the total sample. In their study, the intra-examiner analysis resulted in a Kappa coefficient of 0.811 – also confirming reproducibility. It must be noted that Willems’ method is based on the developmental stages of Dermijian et al.
Designed as a qualitative technique, this approach is susceptible to subjectivity, interpretation and description. However, Dermijian stages are clearly defined, facilitating the classification of the development of each tooth.\(^{15}\)

Another fact that justifies the high reproducibility is the quantitative characteristic of Cameriere’s method, which is based on mathematical measurements and formulae with objective, statistical and numerical approach that minimize possible errors of interpretation – which might result in higher reliability.\(^{18}\)

After the calibration and application of both methods in the present study, it was possible to observe high correlation values between the chronological and the estimated ages in the total sample. This outcome confirms that both age estimation methods demonstrate significant applicability in the studied sample. In addition, the mean of ages estimated by combining Cameriere’s and Willems’ methods revealed a maximum increase of 0.03 in the ICC.

Regarding the different correlation values between the methods, it is noted that Cameriere's method was slightly lower than the values corresponding to Willems’ method. This difference is similar to the study by El-Bakary et al.,\(^{18}\) which estimated age in 286 panoramic radiographs of Egyptian children from 5 to 16 years old, obtaining 98.02% accuracy with Cameriere’s method and 98.62% with Willems’ method. However, in this case, the difference between accuracies can be strictly correlated with the sample with subjects over 15 years of age, since the sample of Cameriere’s study corresponding individuals from 5 to 15 years old.

Similarly, Rai and Anand,\(^{19}\) in a study with 75 panoramic radiographs from individuals aged 5 to 14 years, in India, also concluded that the Cameriere’s method has a lower accuracy compared to the Willems’ method. This slight difference can be explained by the sample number used by Cameriere et al.\(^{12}\) in their study, which used 455 radiographs of Italian individuals from 5 to 15 years old. This sample number differs from the research by Willems et al.,\(^{14}\) performed with 2,523 panoramic radiographs of a Belgian Caucasian population. Moreover, the method designed by Cameriere et al.\(^{12}\) involves more measurement and calculation steps, requiring training and experience because it is associated with a longer learning curve.\(^{18}\)

Besides that, there is a better distribution of estimated ages around of chronological ages for Cameriere’s method. Although the values are further from the central oblique line when compared to the Willems’ method graphs in the ICC graphs, they are more evenly arranged around of the central oblique line. This result is also demonstrated by the low value of the differences between the chronological and estimated ages (0.05 years for girls and 0.03 years for boys). The \(p\)-values and confidence intervals do not provide evidence of difference between the chronological and estimated ages, since the \(p\)-values are higher than 0.05 and the confidence intervals include the zero value.

The positive value of the difference between the chronological and estimated ages demonstrates that there was a slight underestimation by Cameriere’s method. Although the minimal underestimation present in this study, there are reports in the literature of higher underestimations, as Luz et al.\(^{20}\) who applied Cameriere’s method to a Croatian and a Brazilian sample. Similarly, Fernandes et al.\(^{16}\) reported that Cameriere’s method underestimated the ages of 54.4% of Brazilian individual's sample, a trend that was verified mainly from 11 years of age.

Despite the higher ICC presented in the Willems’ method, it pointed to an overestimation of age when compared to the chronological age, a circumstance that can be observed by the negative results of the estimated differences. The difference for the Willems’ method resulted in a value of -0.47 years for girls and -0.39 for boys. The \(p\)-value is less than 0.05 and the confidence interval does not include the zero value, providing evidence of differences between the chronological and estimated ages.

The overestimation of age presented by Willems’ method might be justified by the original study of Willems’ et al.\(^{14}\) that simplified the two score tables originally created by Dermijian et al.\(^{13}\) In the literature, both methods have several reports of overestimation in different populations.\(^{17,18,21}\)

In the Bosnia and Herzegovina community, Galic et al.\(^{17}\) demonstrated an overestimation of age of 0.25 years for girls and 0.42 for boys. Even in the Brazilian population, there are studies that present that characteristic, such as Franco et al.\(^{22}\) who indicated a slight difference of -0.17 years for girls and -0.38 for boys.

On the other hand, in the study by Apaydin and Yasar,\(^{21}\) the estimated ages by Willems’ method
had an underestimation of 0.06 years. Through the studies by Angelakopoulos et al. and Halilah et al., it was possible to observe an overestimation of Cameriere’s method in younger children in a sample from 6 to 14 years old from South Africa and from 5 to 16 years old from Germany.

In this context, it is essential to distinguish the possible consequences of overestimating or underestimating an individual’s chronological age. In the civil sphere, false positives or false negatives tend not to have different weights because one does not have worse consequences than the other, since civil law covers numerous circumstances and each case must be assessed separately. However, in the criminal context, overestimation can have undesirable implications, because the individual can be considered imputable and legally punished for a false positive result, since it could only be targeted by educational protective measures.

The different sexes presented high correlation coefficients. However, when the differences between the chronological and estimated ages of the female sample are compared with the male sample, Cameriere’s method shows a higher tendency to underestimate and Willems’ method a higher tendency to overestimate the age of girls. This tendency is evident when the values of the differences for each sex are observed: when Cameriere’s method was used, the difference value for girls was 0.05 and for boys 0.03; when age was estimated by Willems’ method, the difference for girls was ~0.47 years and for boys was ~0.39 years. Concluding that the boys’ values were better distributed around the chronological age, decreasing the difference with the estimated age.

Wolf et al. also found a higher difference for girls (0.08 years for girls and 0.07 for boys). However, Apaydin and Yasar found a smaller error for girls when using Cameriere’s method, with an underestimation of 0.603 years for boys and 0.550 for girls.

The higher overestimation for girls with Willems’ method was also described by the author during his study with a sample of Belgian Caucasian children, with an overestimation for girls of 0.2 years and 0.1 for boys. Apaydin and Yasar also found a higher error for girls when using Willems’ method, with an underestimation of 0.062 years for girls and 0.056 for boys. However, in the study by El-Bakary, Hammad and Mohammed, girls had lower errors level compared to boys when using Willems’ method (overestimation of 0.14 years for girls and 0.29 for boys).

Regarding the different results of the studies with higher overestimation for girls or boys, Rai et al. propose that there are several genetic, environmental and geographical causes, such as the nutritional and socioeconomic status of a given population. The difference between the sexes is mainly due to advanced girls’ development when compared to boys’ development, especially when it comes to dental formation, since girls reach almost all stages of dental development before boys.

It is evident that both methods are applicable to girls and boys in the studied sample, since they showed high agreement with chronological age. The mean estimated age obtained by combining both methods had a slight increase in agreement with the chronological age in relation to the mean estimated age obtained with each method separately.

ACKNOWLEDGMENTS
This work was conducted with financial support from CAPES/BRAZIL – Coordination for the Improvement of Higher Education Personnel, and Federal Police of Brazil (Prof-Forenses 001, Edital CAPES 25/2014). This study was financed in part by CAPES/BRAZIL - finance code: 001.

REFERENCES


