

## Dental age estimation using the pulp-to-tooth ratio in canines by neural networks

Maryam Farhadian<sup>1</sup>, Fatemeh Salemi<sup>2,\*</sup>, Samira Saati<sup>2</sup>, Nika Nafisi<sup>2</sup>

<sup>1</sup>Department of Biostatistics, School of Public Health and Research Center for Health Sciences, Hamadan University of Medical Sciences, Hamadan, Iran

<sup>2</sup>Department of Oral and Maxillofacial Radiology, Dental School, Hamadan University of Medical Sciences, Hamadan, Iran

### ABSTRACT

**Purpose:** It has been proposed that using new prediction methods, such as neural networks based on dental data, could improve age estimation. This study aimed to assess the possibility of exploiting neural networks for estimating age by means of the pulp-to-tooth ratio in canines as a non-destructive, non-expensive, and accurate method. In addition, the predictive performance of neural networks was compared with that of a linear regression model.

**Materials and Methods:** Three hundred subjects whose age ranged from 14 to 60 years and were well distributed among various age groups were included in the study. Two statistical software programs, SPSS 21 (IBM Corp., Armonk, NY, USA) and R, were used for statistical analyses.

**Results:** The results indicated that the neural network model generally performed better than the regression model for estimation of age with pulp-to-tooth ratio data. The prediction errors of the developed neural network model were acceptable, with a root mean square error (RMSE) of 4.40 years and a mean absolute error (MAE) of 4.12 years for the unseen dataset. The prediction errors of the regression model were higher than those of the neural network, with an RMSE of 10.26 years and a MAE of 8.17 years for the test dataset.

**Conclusion:** The neural network method showed relatively acceptable performance, with an MAE of 4.12 years. The application of neural networks creates new opportunities to obtain more accurate estimations of age in forensic research. (*Imaging Sci Dent 2019; 49: 19-26*)

**KEY WORDS:** Forensic Dentistry; Cone-Beam Computed Tomography; Regression Analysis, Neural Networks

### Introduction

In some clinical fields, such as the differential diagnosis of metabolic syndrome, pediatric dentistry, orthodontic therapy, and forensic radiology, it is highly important to know the chronological age of a living or dead person. Indeed, it is absolutely essential to have techniques for estimating the age of unknown corpses, victims of natural disasters, people who do not have specific identity documents, those who have lost their memory, people whose birthday is unknown, or people whose age is disputed, as in criminal cases. Age estimation is also applicable with

regard to people who migrate to other countries with an unknown age, as well as people with false identities.<sup>1-3</sup>

Several body parts can be used to estimate age, but in severe accidents, burns, or buried bodies, many parts of the body are damaged or lost, making them useless in practice. Previous studies have indicated that there is a solid relationship between dental age and physiological age, meaning that dental age can be used as a better estimate of physiological age than skeletal estimation or other common methods of age estimation.<sup>3</sup>

Teeth are good biologic markers for estimating the age of a person because they remain intact long after death. In children, the estimation of age based on teeth is relatively simple, as it is carried out based on the developmental stages of the teeth. Conversely, in adults, age estimation is a challenge in forensic dental science. The teeth, which are formed of enamel and dentin, are the hardest part of

Received August 6, 2018; Revised October 26, 2018; Accepted November 7, 2018

\*Correspondence to : Dr. Fatemeh Salemi

Department of Oral and Maxillofacial Radiology, Dental School, Hamadan University of Medical Sciences, Shahid Fahmideh Street, Hamadan 65178-38677, Iran

Tel) 98-81-38354250, E-mail) dr.salemi@yahoo.com

Copyright © 2019 by Korean Academy of Oral and Maxillofacial Radiology

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Imaging Science in Dentistry · pISSN 2233-7822 eISSN 2233-7830

the human body.<sup>4</sup> The dental pulp is mesenchymal tissue surrounded by a pulpal canal. Outside of the pulp, there are odontoblasts that form dentin throughout the lifetime, which can reduce the pulpal canal size. Dentin and tooth pulp experience age-related pathological and physiological changes.<sup>5-9</sup> Measuring these morphological changes requires cutting teeth, which is not feasible *in vivo*, so the methods used to estimate age depend on radiological imaging of the tooth.<sup>8-10</sup>

Each tooth can be utilized to determine age; however, the canines are particularly suitable candidates for age estimation since they are commonly found at older ages, are less prone to rot, and have a large root and pulp.<sup>5-8</sup>

As a noninvasive method, dental radiography provides a valuable toolkit for clinicians to use in determining dental age. The ability to perform this technique in living subjects, along with other benefits such as low cost, simplicity, and reliability, have led many researchers to investigate this method in further detail.<sup>11-13</sup>

Previous studies have primarily used regression models to estimate age using different dental parameters.<sup>11,12</sup> A major disadvantage of linear regression models is that they only work well to model a linear function. If the relationship between the output and input is nonlinear, linear regression can only establish a local approximation, and when used at a global level, the approximation can become highly inaccurate. If the output is a nonlinear function of the input, then nonlinear regression models often provide a better fit. However, in such cases, a particular type of nonlinear function must be specified in advance. For example, it must be determined whether the relationship between the output variable and a particular input is exponential, quadratic, logarithmic, or another form. Thus, a particular type of nonlinearity is forced to be implemented a priori. In many applied problems, this may not be possible, because the kind of nonlinear behavior may not be known in advance or the nonlinear behavior may vary across different situations. If an improper nonlinear regression function is selected, it may represent a nonlinear relationship with less precision than its linear counterpart. In the situation described above, where the relationship between the output and input is nonlinear and the form of nonlinearity is not specified, an self-adjusting approach with more flexibility that can accommodate various types of nonlinear behavior, including a wide class of nonlinear mappings, is required.<sup>14-16</sup> Neural networks are well-suited for a very broad class of nonlinear approximations and mappings. As important standard machine learning procedures for classification and regression, neural networks have recently

become widespread in many disciplines, including biology and medicine.<sup>14-16</sup> Of particular note, feed-forward neural networks are nonparametric statistical models for deriving nonlinear relationships among data.<sup>14</sup>

In this regard, there is still a great need for a more accurate method capable of providing a reliable and precise estimation of age. Additionally, the need for advanced methods of data analysis in the field of dentistry is on the rise. As a result, neural networks could represent another option for solving complex prediction problems. They are increasingly become popular in medicine, and are especially successful in modeling nonlinear relationships between the predicted variable and the input data.<sup>15,16</sup> Indeed, applications of neural networks in other domains have shown this approach to perform better than regression models.

However, few studies have explored neural networks in the domain of dentistry. For example, Devito et al. (2008) used an artificial neural network to classify proximal dental caries with the goal of predicting whether an orthodontic treatment required extraction.<sup>17</sup> Moghimi et al.<sup>18</sup> used artificial neural networks to predict the size of unerupted canines and premolars.

This study aimed to assess the possibility of exploiting neural networks for estimating age using the pulp-to-tooth ratio in canines as a non-destructive method. In addition, the predictive performance of neural networks was compared with that of a linear regression model.

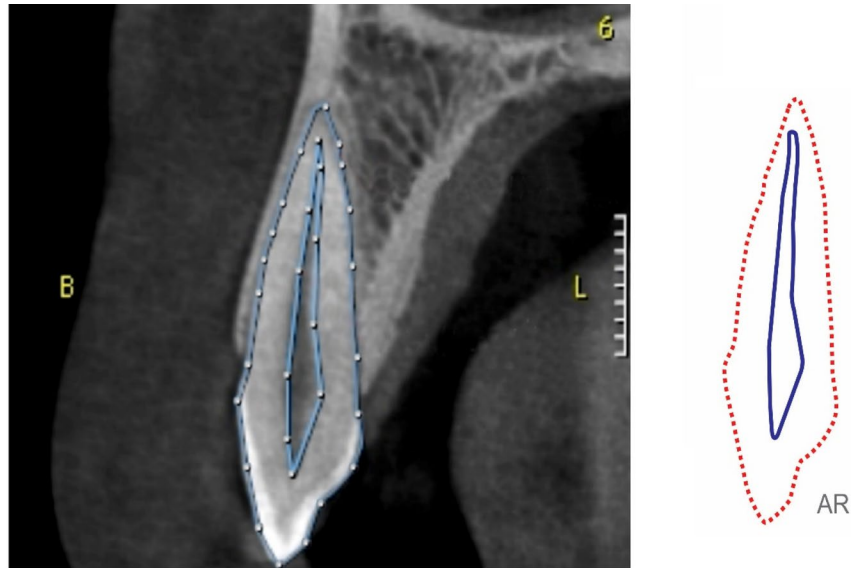
## Materials and Methods

### Materials

In this study, archived cone-beam computed tomographic (CBCT) scans from 300 patients who had been referred to 2 private jaw radiology centers in Hamadan, Iran, including 142 women and 158 men aged between 14 and 60 years, were studied. The average age of the participants was 36 years.

The inclusion criteria for the study were as follows: age of 14-60, complete development of the maxillary canine tooth, and complete formation of the maxillary root canal. Scans were excluded if they showed root infiltration, extensive repair, decay, or a congenital anomaly in the canine canal. Additionally, CBCT scans with insufficient resolution were excluded from the study.

In order to measure morphological variables, cross-sectional CBCT images of the maxillary canine were reconstructed with 1-mm slice thickness and a 1-mm interval. The pulp-to-tooth-area ratio (AR) (Fig. 1), the pulp-to-tooth-length ratio (P) (Fig. 2), the buccolingual pulp-to-



**Fig. 1.** Pulp-to-tooth-area ratio (AR).



**Fig. 2.** Pulp-to-tooth-length ratio (P).

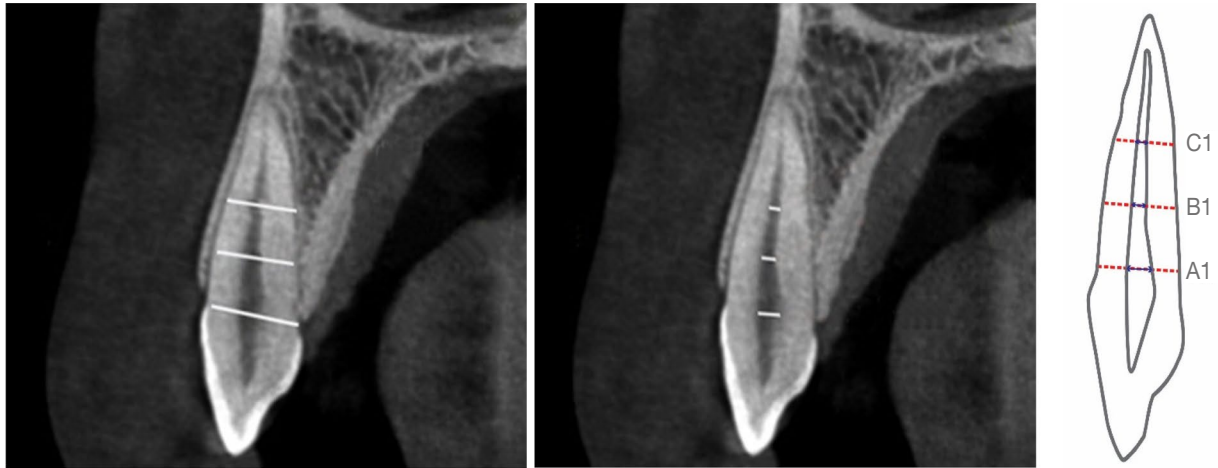
tooth-width ratio at the cementoenamel junction (CEJ) (A1), the mesiodistal pulp-to-tooth-width ratio at the CEJ (A2), the buccolingual pulp-to-tooth-width ratio at the mid-root (C1), the mesiodistal pulp-to-tooth-width ratio at the mid-root (C2), the buccolingual pulp-to-tooth-width ratio at the middle of A and C (B1), and the mesiodistal pulp-to-tooth-width ratio at the middle of A and C (B2) (Figs. 3 and 4) were measured.

All teeth were imaged using a Cranex 3D system (Soredex, Helsinki, Finland), with the exposure settings of 90 kVp, 8 mA, and 6.12 s, and saved in OnDemand software (CyberMed Inc., Seoul, Korea).

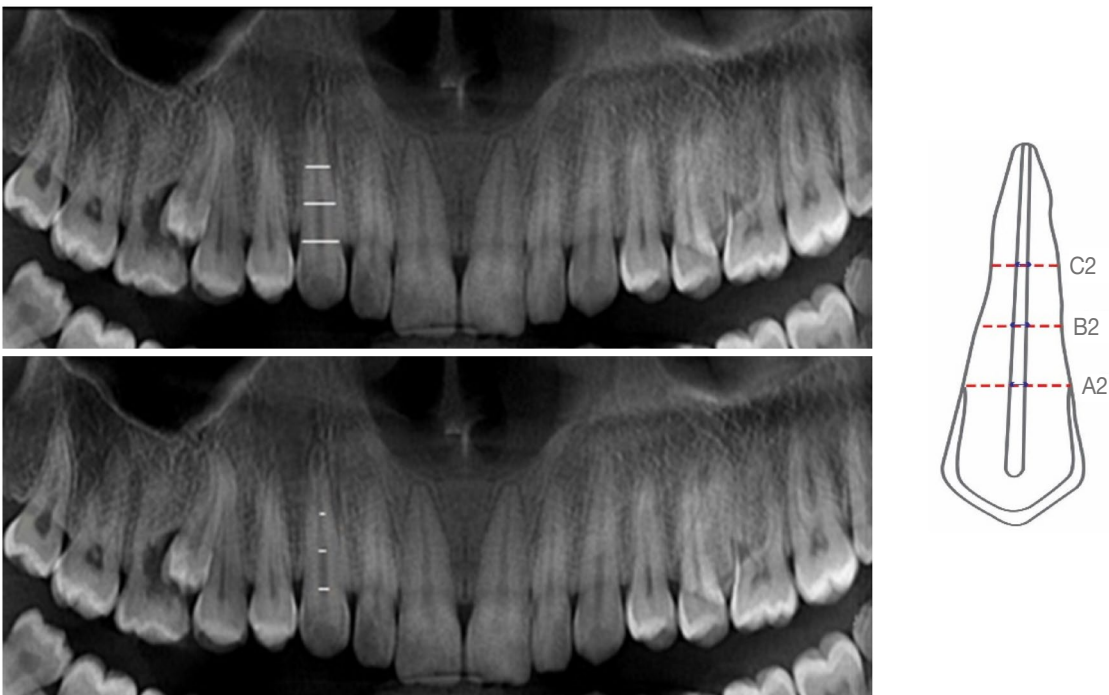
The scans were selected using cross-sections that passed

precisely through the center of the canine tooth. Twenty points around the edges of the tooth and 10 points around the edges of the pulp were determined, and the points were connected to calculate the pulp surface and tooth surface area in millimeters squared. Dental length, pulp length, buccolingual tooth, and pulp width in the 3 areas were measured through length measurements in OnDemand software using the cross-sectional area. In addition, the mesiodistal width of the tooth and pulp in these 3 areas was also measured on the panoramic reconstruction obtained through the software. The age and sex of each patient were also obtained from the form completed before the patient was imaged.

Inter-observer agreement (reproducibility) was deter-



**Fig. 3.** Buccolingual width measurements of the tooth and pulp at 3 levels in a cross-sectional image.



**Fig. 4.** Mesiodistal width measurements of the tooth and pulp at 3 levels in a panoramic image.

mined using data from 2 independent examiners. Intra-observer agreement (repeatability) was also evaluated by having both the examiners assess one-half of the CBCT images in 2 separate sessions. The intra-class correlation coefficient was computed to assess the reliability of the measurements recorded by the 2 examiners. The values of the intra- and inter-class correlation coefficients were very high ( $\rho=0.99$ ).

In order to evaluate the regression and neural network models, in each fold, the data set was randomly divided

into training (9/10 of the data, including 270 samples) and test (1/10 of the data, including 30 samples) sets, for a total of 100 times. The methods (regression and neural networks) were applied to the training set, and the test set was used to calculate the evaluation measures. These methods were compared in terms of the mean of the evaluation criteria values. The predictive performance of the regression and neural network models was assessed using the  $R^2$  statistic, mean absolute error (MAE), and root mean square error (RMSE).

To select the best networks in terms of the number of hidden neurons, different structures of neural networks with 1-10 hidden layers were developed. In more detail, to establishment the neural network model, feed-forward multilayer perceptron networks with a sigmoid activation function in the hidden layer were used. The number of neurons was 8 in the input layer, 7 in the hidden layer, and 1 in the output layer. The learning rate was set to 0.01.

In order to fit a multiple linear regression model, variables relating to the pulp-to-tooth ratio (including A, P, A1, A2, B1, B2, C1 and C2), which were defined above, were considered as independent variables and age was regarded as the dependent variable. The neural networks were developed in R 3.2.2 statistical software (R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org>), using the NeuralNetTools and nnet packages.

### Artificial neural networks

Artificial neural networks are derived from the human nervous system and follow the same learning process; that is, they learn through examples. A neural network comprises different layers of units that are interconnected. For classification and regression problems, feed-forward artificial neural networks are the most commonly utilized network design. They are the equivalent of non-linear multivariate regression methods.<sup>14,15</sup>

As a feed-forward network, the multi-layer perceptron comprises an input layer, H hidden layers, and an output layer (Fig. 5).

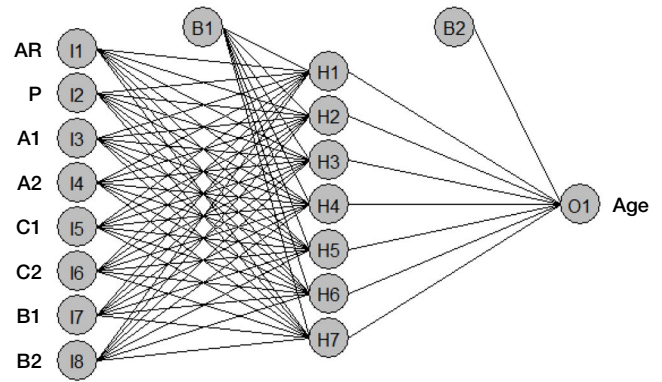
A neural network entails a set of examples (i.e., inputs and outputs) to detect the best approximated association between input and output values. This kind of learning is known as supervised learning.

During network training, a back-propagation algorithm should be used to estimate the weights that minimize the error function. This kind of algorithm is known as back-propagation, since it is based on the backward propagation of errors from output neurons to neurons of the first layer.<sup>14</sup> In fact, the aim of back-propagation is to update each of the weights of the network so they cause the actual output to be closer to the target output. Doing so minimizes the error for each output neuron and eventually reduces the overall network error.

Equation (1) represents the basic expression of the structure of the artificial neural network model:

$$y_i = F(z_i)$$

$$z_i = \sum_{i=1}^n w_i x_i + b_i \quad (1)$$



**Fig. 5.** Structure of the developed feed-forward neural networks for age estimation.

Where  $y_i$  represents the output variable (age),  $Z_i$  refers the weighted sum of the inputs,  $x_i$  denotes the calculated value for the input variables (related to the pulp-to-tooth ratio), and  $w_i$  and  $b_i$  are the weight and the bias, respectively.

In neural networks, complex nonlinear mappings between the input and output variables are learned through activation functions. The most typical activation function is the sigmoid function (equation 2) which is used as an activating function:

$$F(z_i) = \frac{1}{1 + e^{-\mu z_i}} \in (0,1) \text{ for } \mu > 0 \quad (2)$$

### Training and test sets

Network weights are derived using the training set, and the model performance should be tested on the test set. In order to obtain a more realistic estimate of the way the model performs with unseen data, part of the original data must be set aside and excluded from the training process. This data set is regarded as the test set.<sup>14</sup>

In the present study, the same training and test sets were exploited for the 2 methods, in order to obtain comparable results. The data in the test sets had the actual age recorded, making it possible to compare the predicted and actual age. This helped to assess the predictive performance of the model.

The discrepancy between the predicted and observed age was assessed in each test and training dataset. To quantify predictive performance, the mean absolute error (MAE) and root mean squared error (RMSE) were utilized. The determination coefficient ( $R^2$ ) was calculated as well. The MAE and RMSE were defined as follows:

$$RMSE = \sqrt{\frac{\sum (Age - Predicted\ Age)^2}{n}}$$

$$MAE = \frac{\sum |Age - Predicted Age|}{n}$$

## Results

Descriptive statistics of the input (pulp-to-tooth ratio) and output (age) variables used for the development of the prediction models are displayed in Table 1. The Pearson correlation coefficients between age and all the measured variables for all participants are presented in Table 2. As shown, a negative correlation coefficient for all variables was found, indicating that with increasing age, the magnitude of these variables decreased.

Neural networks with 7 neurons in the hidden layer were found to be the best structure for the prediction model of age, as shown in Figure 1. The prediction performance of the developed models in the test and training sets is presented in Table 3.

The results indicated that, for both models, the evaluation criteria showed lower values for the training set than for the test set. Additionally, for all criteria in both sets, the neural network performed better than the regression model.

The results showed that the prediction errors of the developed neural network model were acceptable, with an

RMSE of 4.40 years and an MAE of 4.12 years for the unseen dataset (test sets) (Table 3). However, the prediction errors of the regression model were greater than those of the neural network, with an RMSE of 10.26 years and an MAE of 8.17 years for the test dataset.

When assessed using the paired t-test, all evaluation criteria showed significant differences between the neural network and regression models in terms of performance in the test and training sets ( $P < 0.05$ ).

A comparison of prediction performance of the developed models in different age groups in the test sets is shown in Table 4. After dividing participants into 5 age groups, in all subgroups of age (except for the oldest age group), the neural network model showed better performance than the regression model ( $P < 0.05$ ). However, the difference in performance between the 2 models in the oldest age group was negligible ( $P > 0.05$ ).

The best performance based on the MAE (2.53 years for the neural network model and 4.78 years for the regression model) and RMSE (2.66 years for the neural network model and 6.16 years for the regression models) for both models was in the 40-to 50-year age group. For this age group, the neural networks performed better, with less error, than regression.

**Table 1.** Descriptive statistics of inputs (pulp/tooth ratio) and output (age) variables

| Variables   | Mean $\pm$ SD   | Range     |
|---|-----------------|-----------|
| Age   | 36.0 $\pm$ 12.6 | 14-60     |
| Pulp to tooth area ratio (AR)   | 0.16 $\pm$ 0.04 | 0.08-0.25 |
| Pulp to tooth length ratio (P)  | 0.77 $\pm$ 0.06 | 0.56-0.88 |
| Buccolingual pulp to tooth width ratio at CEJ (A1)                      | 0.20 $\pm$ 0.04 | 0.08-0.31 |
| Mesiodistal pulp to tooth width ratio at the CEJ (A2)                   | 0.27 $\pm$ 0.05 | 0.14-0.40 |
| Buccolingual pulp to tooth width ratio at the mid-root (C1)             | 0.17 $\pm$ 0.05 | 0.06-0.30 |
| Mesiodistal pulp to tooth width ratio at the mid-root (C2)              | 0.16 $\pm$ 0.05 | 0.08-0.41 |
| Buccolingual pulp to tooth width ratio at in the middle of A and C (B1) | 0.20 $\pm$ 0.05 | 0.05-0.32 |
| Mesiodistal pulp to tooth width ratio at in the middle of A and C (B2)  | 0.25 $\pm$ 0.06 | 0.10-0.41 |

**Table 2.** Pearson's correlation coefficient between age and the measured variables (n = 300)

| Variables               | B2     | B1     | C2     | C1     | A2     | A1     | P      | AR     |
|-------------------------|--------|--------|--------|--------|--------|--------|--------|--------|
| Correlation coefficient | -0.579 | -0.428 | -0.246 | -0.401 | -0.600 | -0.406 | -0.160 | -0.779 |
| P-value                 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |

**Table 3.** The prediction performance of the developed models in the train and the test sets

| Multiple linear regression |        |       |                |       |       | Neural networks |       |       |                |       |       |
|----------------------------|--------|-------|----------------|-------|-------|-----------------|-------|-------|----------------|-------|-------|
| Test                       |        |       | Train          |       |       | Test            |       |       | Train          |       |       |
| R <sup>2</sup>             | RMSE   | MAE   | R <sup>2</sup> | RMSE  | MAE   | R <sup>2</sup>  | RMSE  | MAE   | R <sup>2</sup> | RMSE  | MAE   |
| 0.506                      | 10.261 | 8.176 | 0.664          | 7.240 | 7.175 | 0.857           | 4.403 | 4.121 | 0.901          | 3.573 | 3.057 |

RMSE: root mean square error, MAE: mean absolute errors, R<sup>2</sup>: coefficient of determination

**Table 4.** Comparison of prediction performance of the developed models in age group (in the test sets)

| Age group (year) | N* | RMSE  |        | MAE   |        | R <sup>2</sup> |        | MEAN (Error) |         |
|------------------|----|-------|--------|-------|--------|----------------|--------|--------------|---------|
|                  |    | NN    | REG    | NN    | REG    | NN             | REG    | NN           | REG     |
| < 20             | 40 | 4.831 | 13.346 | 4.123 | 11.927 | 0.492          | 0.016  | -2.153       | -10.661 |
| 20-30            | 75 | 3.791 | 11.969 | 3.080 | 10.638 | 0.667          | 0.282  | -1.004       | -9.717  |
| 30-40            | 68 | 3.481 | 9.162  | 3.301 | 7.278  | 0.584          | 0.085  | -0.994       | -5.249  |
| 40-50            | 70 | 2.666 | 6.166  | 2.528 | 4.785  | 0.675          | 0.099  | 0.245        | -3.388  |
| > 50             | 47 | 5.796 | 11.477 | 4.549 | 7.991  | 0.336          | -0.009 | 4.247        | 5.803   |

\*: Number, RMSE: root mean square error, MAE: mean absolute errors, R<sup>2</sup>: coefficient of determination

The worst performance of the neural networks in the test sets was registered in the oldest age group (> 50 years), with an MAE of 4.55 years. In contrast, the worst performance of the regression model in the test sets was observed for the youngest age group (< 20 year).

In the first 3 subgroups of age, the average age predicted by the neural network was lower than the actual age, while in the older 2 age groups, the average age estimated by the neural networks was higher than the actual age. The performance of the regression model in this regard was similar to that of the neural networks.

## Discussion

In the present study, neural networks were applied as a proposed method for age prediction based on dental parameters. The predictive performance of the regression and neural network models was tested by cross-validation, consisting of 10-times 10-fold experiments. The prediction errors for predicting age based on the developed neural network model were in the acceptable level, with an MAE of 4.12 years for unseen subjects in the test sets, and 3.05 years for the training data.

The developed prediction model can be considered as a useful tool for age estimation in the field of forensic odontology. The advantage of neural networks over classical statistical methods, such as simple or multiple linear regression models, is that neural networks are able to model more complex, non-linear relationships between input variables and outputs.<sup>13</sup>

Both panoramic radiography and periapical radiography can be used to measure the pulpal tooth ratio. However, these methods can fail due to the limitations of 2-dimensional images, magnification, and distortion.<sup>19</sup> Computed tomography (CT) is the most ideal and accurate method for measuring the pulp-to-tooth ratio. However, due to the lower radiation dose and the higher image resolution of CBCT than CT, recent research on age estimation has been carried

out using CBCT images. Considering the limited studies on the use of CBCT for age estimation, we conducted a study on age estimation based on CBCT images from the canine teeth and calculation of the pulp-to-tooth ratio.<sup>20</sup>

All teeth can be used for age estimation. However, canine teeth are especially appropriate for this purpose, because they are less likely to be rotten, have a large root and pulp, and are commonly present in older individuals. Furthermore, they are less susceptible to wear due to diet than are the posterior teeth, and are less probable than other anterior teeth to suffer wear; therefore, these single-root teeth with the largest pulp area are the easiest to analyze.<sup>7,21</sup>

Although fitted models in some previous studies yielded good predictive performance within the sample, these predictive models were vulnerable to overfitting.<sup>6,22</sup> The reason for this is that the same observations are utilized for both model development and testing, which can lead to overestimation of the model's predictive performance. A more accurate estimation of a model's performance requires various observations sets to train and test the predictive model.<sup>14,23</sup> Thus, in this study, instead of working with all of the data, which is a major mistake that has been made in some studies, the data were randomly partitioned into separate training and test sets. Moreover, the evaluation of the predictive model based on a single test set is profoundly influenced by the data splitting process. Therefore, the process of data splitting was repeated 100 times. Furthermore, for the comparison of the performance of the 2 methods, this division was considered to be the same for both methods. This process enhances the generalizability of the findings.<sup>23</sup>

Different studies have reported the standard error of estimation, the MAE, or the standard deviation as measures of accuracy. The variety of approaches that have been used both to estimate and to report the obtained accuracy makes it impossible to confirm whether a given method is superior to another.<sup>24</sup>

Further investigations with different CBCT scanners, dif-

ferent resolutions of the image sections, and different types of teeth are required to draw conclusions regarding the use of CBCT for age estimation. Nonetheless, it should be noted that the results of a study on a particular population cannot be generalized to other populations. Consequently, this study should be replicated using a wider statistical population in which race and other relevant factors are taken into account, in addition to age and sex.

The results of this study suggest the possibility of developing a new tool using neural networks to predict age based on dental findings. It is thus recommended that this method should be used in similar prediction contexts. It is concluded that neural network models are useful for age prediction and should be explored further as more data become available. Additionally, the use of this method and other methods of machine learning is recommended in the field of dentistry. In this regard, complementary research can be performed to achieve further improvements.

### Acknowledgements

This work was part of a doctoral degree of dentistry thesis that was supported by the Dental Research Center, Vice Chancellor of Research, Hamadan University of Medical Sciences.

### References

1. Cameriere R, Cunha E, Sassaroli E, Nuzzolese E, Ferrante L. Age estimation by pulp/tooth area ratio in canines: study of a Portuguese sample to test Cameriere's method. *Forensic Sci Int* 2009; 193: 128.e1-6.
2. Rai A, Acharya AB, Naikmasur VG. Age estimation by pulp-to-tooth area ratio using cone-beam computed tomography: a preliminary analysis. *J Forensic Dent Sci* 2016; 8: 150-4.
3. Bolanos MV, Manrique MC, Bolanos MJ, Briones MT. Approaches to chronological age assessment based on dental calcification. *Forensic Sci Int* 2000; 110: 97-106.
4. Cameriere R, Ferrante L, Cingolani M. Precision and reliability of pulp/tooth area ratio (RA) of second molar as indicator of adult age. *J Forensic Sci* 2004; 49: 1319-23.
5. Biuki N, Razi T, Faramarzi M. Relationship between pulp-tooth volume ratios and chronological age in different anterior teeth on CBCT. *J Clin Exp Dent* 2017; 9: e688-93.
6. Jagannathan N, Neelakantan P, Thiruvengadam C, Ramani P, Premkumar P, Natesan A, et al. Age estimation in an Indian population using pulp/tooth volume ratio of mandibular canines obtained from cone beam computed tomography. *J Forensic Odontostomatol* 2011; 29: 1-6.
7. Babshet M, Acharya AB, Naikmasur VG. Age estimation in Indians from pulp/tooth area ratio of mandibular canines. *Forensic Sci Int* 2010; 197: 125.e1-4.
8. Cameriere R, Ferrante L, Belcastro MG, Bonfiglioli B, Rastelli E, Cingolani M. Age estimation by pulp/tooth ratio in canines by peri-apical X-rays. *J Forensic Sci* 2007; 52: 166-70.
9. Kvaal SI, Kolltveit KM, Thomsen IO, Solheim T. Age estimation of adults from dental radiographs. *Forensic Sci Int* 1995; 74: 175-85.
10. Cameriere R, Ferrante L, Cingolani M. Variations in pulp/tooth area ratio as an indicator of age: a preliminary study. *J Forensic Sci* 2004; 49: 317-9.
11. Juneja M, Devi YB, Rakesh N, Juneja S. Age estimation using pulp/tooth area ratio in maxillary canines - a digital image analysis. *J Forensic Dent Sci* 2014; 6: 160-5.
12. Graham JP, O'Donnell CJ, Craig PJ, Walker GL, Hill AJ, Cirillo GN, et al. The application of computerized tomography (CT) to the dental ageing of children and adolescents. *Forensic Sci Int* 2010; 195: 58-62.
13. Maret D, Molinier F, Braga J, Peters OA, Telmon N, Treil J, et al. Accuracy of 3D reconstructions based on cone beam computed tomography. *J Dent Res* 2010; 89: 1465-9.
14. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference and prediction. Springer series in statistics. 2nd ed. New York, NY: Springer; 2009.
15. Lisboa PJ. A review of evidence of health benefit from artificial neural networks in medical intervention. *Neural Netw* 2002; 15: 11-39.
16. Farhadian M, Aliabadi M, Darvishi E. Empirical estimation of the grades of hearing impairment among industrial workers based on new artificial neural networks and classical regression methods. *Indian J Occup Environ Med* 2015; 19: 84-9.
17. Devito KL, de Souza Barbosa F, Felipe Filho WN. An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2008; 106: 879-84.
18. Moghimi S, Talebi M, Parisay I. 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* 2012; 34: 480-6.
19. Eskandarloo A, Mirshekari A, Poorolajal J, Mohammadi Z, Shokri A. Comparison of cone-beam computed tomography with intraoral photostimulable phosphor imaging plate for diagnosis of endodontic complications: a simulation study. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2012; 114: e54-61.
20. Singaraju S, Sharda P. Age estimation using pulp-tooth area ratio: a digital image analysis. *J Forensic Dent Sci* 2009; 1: 37-41.
21. De Angelis D, Gaudio D, Guercini N, Cipriani F, Gibelli D, Caputi S, et al. Age estimation from canine volumes. *Radiol Med* 2015; 120: 731-6.
22. Bagherpour A, Anbiaee N, Partovi P, Golestani S, Afzalinasab S. Dental age assessment of young Iranian adults using third molars: a multivariate regression study. *J Forensic Leg Med* 2012; 19: 407-12.
23. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med* 2004; 66: 411-21.
24. Marroquin TY, Karkhanis S, Kvaal SI, Vasudavan S, Kruger E, Tennant M. Age estimation in adults by dental imaging assessment systematic review. *Forensic Sci Int* 2017; 275: 203-11.