

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.e-jds.com

Original Article

Predicting root fracture after root canal treatment and crown installation using deep learning

Wan-Ting Chang^a, Hsun-Yu Huang^a, Tzer-Min Lee^b,
Tsen-Yu Sung^c, Chun-Hung Yang^{d*}, Yung-Ming Kuo^{c**}

^a Department of Stomatology, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi, Taiwan

^b School of Dentistry, National Cheng Kung University, Tainan, Taiwan

^c Department of Electronic Engineering, National Formosa University, Yunlin, Taiwan

^d Department of Electronic Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan

Received 5 September 2023; Final revision received 15 October 2023

Available online 10 November 2023

KEYWORDS

Artificial intelligence;
Convolutional neural networks;
Decision-making;
Root canal treatment;
Vertical root fracture;
Treatment planning

Abstract *Background/purpose:* Vertical root fracture (VRF) is a prevalent reason for tooth extraction following root canal treatment and even after crown placement. Predicting fractures is challenging due to multifactorial nature. The current study aimed to predict the likelihood of fracture following root canal treatment and crown placement by developing a deep learning (DL) model.

Materials and methods: DL techniques were employed to analyze a dataset comprising 145 clinical cases consisting of 97 fractured teeth and 48 non-fractured teeth. This dataset spanned a five-year period and encompassed cases involving root canal therapy and crown installation. The analysis identified several root fracture-related parameters, which were incorporated into the DL system. The dataset consisted of 17 features presented in a mixed-type tabular format.

Results: The deep neural network (DNN) model surpassed the support vector machine (SVM) model with a higher accuracy (80.7 % vs. 71.7 %) and F1-score value (0.857 vs. 0.817) for predicting root fracture. Furthermore, in determining root fracture occurrence, it was observed that 17 significant characteristics in the DNN model outperformed the 7 features by 11.7 % in accuracy and 10 % in F1-score.

* Corresponding author. Department of Electronic Engineering, Southern Taiwan University of Science and Technology, No.1, Nantai St., Yungkuang Dist., Tainan City 710301, Taiwan.

** Corresponding author. Department of Electronic Engineering, National Formosa University, No.64, Wunhua Rd., Huwei Township, Yunlin County 632, Taiwan.

E-mail addresses: eliyang@stust.edu.tw (C.-H. Yang), ymkuo@gs.nfu.edu.tw (Y.-M. Kuo).

<https://doi.org/10.1016/j.jds.2023.10.019>

1991-7902/© 2023 Association for Dental Sciences of the Republic of China. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Conclusion: DL shows promise in predicting root fracture post root canal therapy and prosthesis, and it may have the potential to aid clinicians in assessing fracture risk and improving decision-making.

© 2023 Association for Dental Sciences of the Republic of China. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Vertical root fracture (VRF) is a frequent cause of tooth extraction following root canal therapy, with an incidence rate of 13.4%.¹ This fracture poses a significant threat to tooth prognosis during and after root canal therapy, and various factors influence it. Diagnosing a VRF is challenging, as it requires the fracture to progress to a particular stage before a visible radiographic image can be obtained.

Several factors influence VRFs, including tooth decay, which decreases cusp stiffness in teeth lacking marginal ridges.² Fractured teeth typically exhibit severe occlusal surface wear, a stiff cusp angle, and wear of the working side cusp tip.³ The incidence of root fracture can also be impacted by the techniques used during root canal therapy, such as the degree of canal taper. Pericervical dentin (PCD) is an important factor in determining tooth strength. PCD refers to the dentin that surrounds the crestal bone and extends approximately 4 mm from the alveolar crest.⁴ Preserving the strength of a tooth relies on maintaining the integrity of the PCD, as its deterioration can weaken the tooth. A conservative access opening during root canal therapy plays a crucial role in achieving this preservation.

Posterior teeth with coronal protection have a higher success rate than those without,⁵ but VRFs can still occur, indicating that several contributing factors exist. However, the current literature review lacks explicit information regarding the timing of tooth fracture (e.g., whether it occurs during root canal therapy or after prosthesis installation) and specific circumstances that could lead to tooth cracks. Despite numerous articles exploring the characteristics and potential causes of root fractures, accurately estimating tooth survival remains challenging due to the intricate interplay of various factors.

Artificial intelligence (AI), specifically deep learning (DL) and machine learning (ML), has greatly advanced in recent years. ML is a broader field that includes various algorithms for learning from data, while DL is a specialized branch of ML that specifically deals with deep neural networks (DNNs). These technologies are increasingly being used in various fields, including dentistry, to traditionally automate human tasks and make accurate predictions. A review of 43 research articles in 2021 showed the growing prominence of AI applications in dentistry,⁶ and applications of AI in dentistry have gained prominence in recent years. To the best of our knowledge, this was the first study leveraging DL models with extensive data analysis to predict the

likelihood of tooth fractures post root canal and prosthetic treatment.

Materials and methods

Data acquisition

The clinical data used in this study were selected from January 2015 to June 2021 at the Department of Stomatology, Ditmanson Medical Foundation Chia-Yi Christian Hospital. The protocols were conducted under the approved guidelines of the Chia-Yi Christian Hospital Institutional Review Board (CYCH-IRB No. 2021115). To create the dataset for this study, data were extracted from the dental records of patients diagnosed with root fracture (ICD-10 code K0381) and from those without root fracture 5 years after root canal treatment and crown installation.

The diagnosis of fractured teeth was based on clinical symptoms, including a solitary deep and narrow pocket, the presence of lateral radiolucency or a combination of lateral and periapical (J-shaped) radiolucency identified through periapical radiography, or the presence of a gingival sinus tract.⁷ All the teeth were confirmed to be fractured upon extraction. The non-fractured teeth were diagnosed through stable clinical check-ups and periapical radiographs that showed no evidence or reduction of radiolucent lesions. In total, conforming clinical data were obtained from 145 teeth, comprising 97 fractured teeth and 48 non-fractured teeth. Each dataset contained 17 characteristics, as shown in [Table 1](#). These 17 characteristics could be divided into categorical and numerical features, which were ten categorical features and seven numerical features. All of the categorical features and some of the numerical features (the age at the time of treatment, quantity of remaining tooth walls, duration from completion of root canal treatment until the date of prosthetic installation, and tooth wear condition) were according to medical chart records and periapical radiographs. The periodontal condition was assessed based on the percentage of bone loss observed on periapical radiographs, which was divided into the coronal third, middle third, or apical third of the root.⁸ The remaining root canal thickness and PCD thickness were assessed using periapical radiographs. After length calibration using software (Planmeca Romexis dental software, Planmeca Oy, Helsinki, Finland), measurements of the remaining wall thickness of the middle

Table 1 Definition of categorical and numerical features for the study patients.

Categorical Features							
No.	Factors	Definition	Code	No.	Factors	Definition	Code
1	Sex	Male	1	7	Endodontical retreatment	Yes	17
		Female	2			No	18
2	Tooth position	Maxillary anterior teeth	3	8	Posts placement	Para post	19
		Maxillary premolar	4			Casting post	20
		Maxillary molar	5			Fiber post	21
		Mandibular anterior teeth	6			Screw post	22
		Mandibular premolar	7			No	23
		Mandibular molar	8			No	24
3	Previous dental fractures	Yes	9	9	Abutment of removable dentures or fixed partial dental prostheses	Abutment of fixed partial dental prostheses	25
		No	10			Abutment of removable dentures	26
4	Previous prostheses	Yes	11	10	Previous apicoectomy or root amputation	Both coincide	27
		No	12			No	28
5	Preoperative pain	Yes	13	Apicoectomy	29		
		No	14	Root amputation	30		
6	Percussion pain	Yes	15				
No	16						
Numerical Features							
No.	Definition			No.	Definition		
11	The age at the time of treatment			14	Tooth wear condition		
12	Quantity of remaining tooth walls			15	Periodontal condition		
13	Duration from completion of root canal treatment until the date of prosthetic installation			16	Remaining root canal wall thickness		
17	Pericervical dentin thickness						

root and PCD thickness at the cements enamel junction were recorded.

Data preprocessing

Mixed-type tabular data are most commonly used when studying AI approaches in medical diagnosis based on patient history.⁹ As categorical features are usually high-dimensional, data preprocessing plays a vital role in improving the accuracy of AI models. The categorical features must be transformed into real number vectors as inputs for DL.¹⁰ In this study, ordinal encoding was utilized as the primary technique for encoding categorical features, which was the simplest approach. Each category was assigned an integer value, and the corresponding values can be found in Table 1.

After performing ordinal encoding, data normalization is necessary to enhance the performance of the DL algorithm¹¹ to prevent specific numerical features from dominating the training phase. This study employed a standardized min–max normalization technique for numerical and encoded categorical features. The normalization equation is described as follows:

$$X_{nom} = \frac{X - X_{min}}{X_{max} - X_{min}} \in [0, 1] \quad (1)$$

where X_{nom} is represented as a normalized feature, X is the input numerical or categorical feature, X_{max} is a maximum feature of a column, and X_{min} is a minimum feature of a column. Finally, the representation of features is changed into $[0, 1]$.

Implementation of classification models

The architecture of the system is shown in Fig. 1. For tabular data, ordinal encoding and min–max normalization have been mentioned. Here, a support vector machine (SVM) and a DNN were selected as the primary classification models.

Experimental environment and setting

The experimental environment included hardware devices and software configurations. The DL framework was TensorFlow 2.8.0, and the programming language was Python 3.7.13. The model was trained on an NVIDIA GeForce GTX 1050Ti GPU. Seventeen feature values were entered, with one output result in the output layer and three hidden layers, each with 425 nodes (17×25). The model

architecture adopted fully connected layers and selected ReLU as the activation function for the model. In the training phase, the learning rate was initially set as 0.001 with an Adam¹² optimizer to maximize the model performance. This model utilized batch normalization to enhance the stability during training.

Evaluation metrics

To effectively and objectively evaluate the methods proposed in this project, this study used accuracy and F1-score parameters as performance indicators and employed 5-fold cross-validation¹³ (Fig. 2) to objectively evaluate the classification performance of the model. The total number of data points was 145, and the batch size was set to 29 (145/5).

Before calculating the F1-score, it is necessary to first calculate the recall and precision through true positives (TPs), true negatives (TNs), false positives (FPs) and false negatives (FNs), as shown in Eqs. (2) and (3):

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

The occurrence of tooth fracture was defined as positive in this article, while the absence of tooth fracture was defined as negative. Then, the F1-score was calculated through Eq. (4).

$$F1 \text{ score} = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (4)$$

Results

Classification results

To understand the performance of the proposed method and prove that DNN was better than SVM, four parameters, including accuracy, recall, precision and F1-score, were adopted in the study. From Table 2, the accuracy and F1-score of the proposed method were 80.7 % and 0.857, respectively. With the SVM model, the accuracy and F1-score values were 71.7 % and 0.817, respectively. The results showed that the proposed method had a good performance and that the DNN model was better than SVM in predicting the occurrence of root fracture.

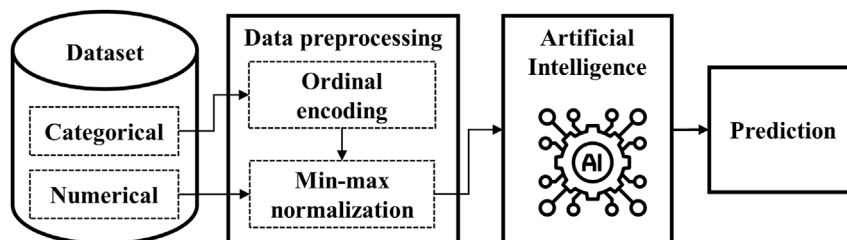


Fig. 1 Workflow of the deep learning framework for the occurrence of root fracture and survival rate.

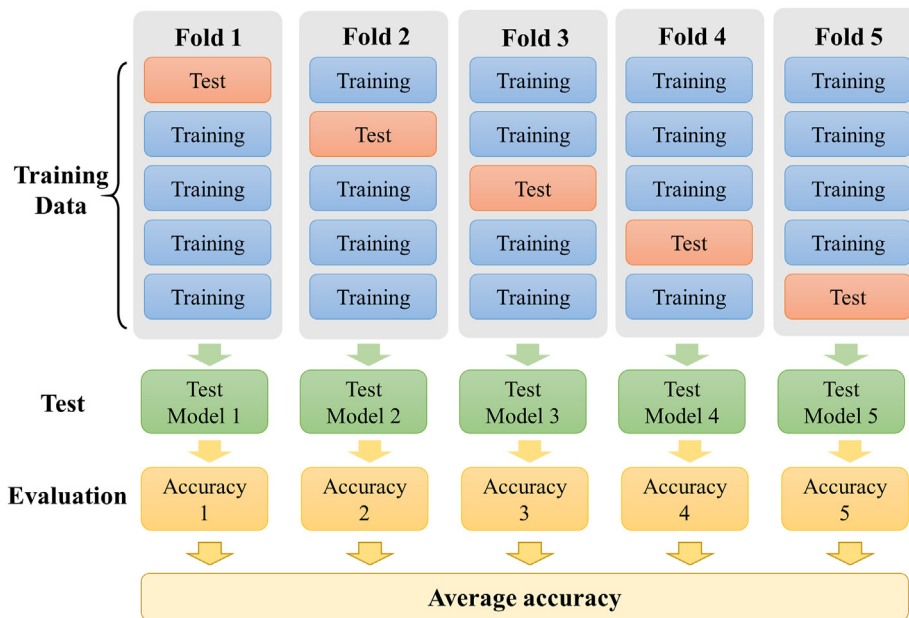


Fig. 2 5-fold cross-validation method.

Features influence analysis

Features play an important role in predicting the occurrence of root fracture. Additionally, feature reduction emerges as a critical concern in improving the efficiency of the model for training and testing. To assess the influence of each feature and obtain the significant features, this study employed statistical and analytical approaches to analyze 17 features individually for classifying root fractures. The F1 score was used as the evaluation metric because it is the most commonly used measure in boundary-based evaluation. Generally, the feature is a significant feature if the value of the F1-score is larger than 0.6. The ranking of features by F1-score is shown in Table 3. Thus, seven significant features were selected, including 1) previous dental fractures, 2) post placement, 3) tooth wear condition, 4) tooth position, 5) duration from completion of root canal treatment until the date of prosthetic installation, 6) age, and 7) PCD thickness. Subsequently, the 7 selected features were utilized as inputs for the DNN model to evaluate the predictive performance. Table 4 shows that the DNN with 17 features obtained a higher classification accuracy compared to the 7 significant features. Based on the DNN model with 17 features, the accuracy and F1-score values were 80.7 % and 0.867, respectively. At the same time, the classification results of the DNN model, using 17

features, showed an 11.7 % relative improvement in accuracy and a 10 % relative improvement in F1-score compared to the one using 7 features. The model, using 17 features, indeed obtained better results in predicting the occurrence of root fracture.

Discussion

DL is commonly used in applications such as image, audio, and text processing. Due to the different applications and tasks involved, various algorithms have been developed. Common examples include convolutional neural networks (CNNs) for image classification, generative adversarial networks (GANs) for image generation, long short-term memory (LSTM) for speech recognition and DL for general data classification.^{14,15} In this study, the utilized dataset consisted of 17 features presented in a tabular format. Therefore, DL was selected as the classification model for the present analysis.

Statistical methods have been commonly used in data analysis. However, such methods are unable to reason and learn like human logic. Therefore, ML technology has emerged as a rising trend. Well-known algorithms include decision tree, k-nearest neighbors (K-NN), SVM, and DL.

DNN, derived from artificial neural network (ANN), consists of the input, hidden, and output layers. The input layer receives and processes data, while the hidden layer is situated between the input and output layers. The number of layers and neurons in each hidden layer is typically determined experimentally. Increasing the number of hidden layers can handle more complex problems, but excessive layers may lead to convergence difficulties during the learning process, resulting in reduced performance or overfitting. Through experiments, this study found that using three hidden layers with 425 nodes (17×25) per layer demonstrated the best classification ability.

Table 2 Comparison of performance prediction between tooth fracture and non-tooth fracture with 17 features.

Method	Accuracy	Recall	Precision	F1-score
SVM	71.7 %	0.947	0.719	0.817
DNN	80.7 %	0.937	0.806	0.867

SVM, support vector machine.

DNN, deep neural network.

Table 3 Features influence ranking by F1-score value.

Ranking	Features	Precision	Recall	F1-score
1	Previous dental fractures	0.674	0.938	0.784
2	Posts placement	0.722	0.856	0.783
3	Tooth wear condition	0.723	0.835	0.775
4	Tooth position	0.771	0.763	0.767
5	Duration from completion of root canal treatment until the date of prosthetic installation	0.753	0.691	0.720
6	The age at the time of treatment	0.738	0.639	0.685
7	Pericervical dentin thickness	0.794	0.557	0.655
8	Endodontical retreatment	0.681	0.505	0.580
9	Quantity of remaining tooth walls	0.712	0.485	0.577
10	Previous prostheses	0.712	0.485	0.577
11	Sex	0.729	0.443	0.551
12	Preoperative pain	0.774	0.423	0.547
13	Periodontal condition	0.714	0.412	0.523
14	Remaining root canal wall thickness	0.783	0.371	0.503
15	Percussion pain	0.756	0.320	0.449
16	Abutment of removable dentures or fixed partial dental prostheses	0.783	0.186	0.300
17	Previous apicoectomy or root amputation	0.889	0.082	0.151

Table 4 Feature influence analysis with a deep neural network (DNN).

Method	Accuracy	Recall	Precision	F1-score
7 features	69.0 %	0.774	0.768	0.767
17 features	80.7 %	0.937	0.806	0.867

To simplify clinical applications, this study also attempted to use fewer features for analysis. The results showed that the accuracy was slightly lower than that of the complete feature analysis. Using all features for classification performed better than using only the 7 most representative features, with accuracies of 80.7 % and 73.8 %, respectively. The classification of 17 features is better than that of 7 features. Based on the experimental results, three possible reasons are considered. 1) Although a subset of features with high F1 scores was selected as the model input, these features may lack dominant information and may not exhibit superior classification performance. 2) There may be no dominant feature for the 17 features, and each feature has a tiny contribution for prediction. 3) Some of the significant features are produced from the nodes of hidden layers in the DNN to compute the best classification results. Therefore, when developing a prediction model, all available features should be utilized as much as possible to improve the accuracy of the prediction.

However, greater F1 values of features can still be regarded as clinical references. The feature with the highest F1 score value was whether the tooth has had a crown fracture. This study demonstrated that teeth with previous fractures may be highly susceptible to root fracture. Other factors with F1 scores higher than 0.6 included the presence or absence of dental posts (0.783), the degree

of tooth wear (0.775), the tooth position (0.767), the duration between root canal completion and prosthesis installation (0.720), age (0.685), and PCD thickness (0.655). Mireku et al.¹⁶ found that the likelihood of VRFs occurring in teeth that have undergone root canal therapy with posts was higher among older patients and those with thin dentin.

VRFs commonly occur in the maxillary and mandibular premolars, maxillary molars with mesiobuccal roots, mandibular molars with mesial roots, and mandibular incisors. Among them, the highest incidence is found in mandibular first molars with mesial roots, followed by maxillary molars with mesiobuccal roots.⁷ The present study demonstrated that a higher F1 score was correlated with tooth position. The timing of prosthesis placement after root canal therapy was found to influence the F1 score, indicating that prosthesis placement should ideally be undertaken promptly following root canal therapy.

Taking preventive measures to minimize the occurrence of VRFs is essential for preserving the structural integrity and function of teeth. Identifying teeth susceptible to VRFs and using conservative endodontic procedures to minimize the risk is essential. Additionally, appropriate post-endodontic restorative procedures should be employed to reduce the incidence of terminal VRFs.¹⁷ The purpose of the DL model established in this study is to assist clinicians in accurately predicting the likelihood of VRFs occurrence, enabling them to provide relevant risk warnings and advice to patients early, thereby reducing the possibility of medical disputes.

However, the current study had several limitations. First, the data sample used was relatively small and unevenly distributed, which could introduce bias and potentially impact the performance of the DL model. A larger and more diverse dataset would enhance the generalizability

and reliability of the model's predictions. Second, the model's performance could be enhanced by including more features and clinical variables. In addition to the 17 tabular features used in the study, other factors, such as occlusal forces, tooth mobility, and root canal treatment quality, may play a role in VRFs. Integrating these variables into the DL model could improve its accuracy and predictive abilities. Third, the study utilized a learning model to predict the result, and the experimental result was acceptable. However, relying on one model may sometimes be unable to obtain the high-level relationship between features, which could reduce the prediction performance. Future studies could consider a hybrid DL model or ensemble learning for improved accuracy.

In conclusion, the use of a DL model in this study demonstrated the potential of AI in predicting tooth fracture after root canal therapy and prosthesis placement. By assisting clinicians in accurately assessing fracture risk and providing timely risk warnings and advice to patients, this model has the potential to improve clinical decision-making and reduce avoidable complications. Overall, this study provided important insights into the potential of DL models for predicting tooth fracture and highlighted the need for further research to advance our understanding and management of this clinical condition.

Declaration of competing interest

The authors have no conflicts of interest relevant to this article.

Acknowledgments

This study was supported by the Ditmanson Medical Foundation Chia-Yi Christian Hospital, Taiwan (Grant R111-27). The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

References

1. Touré B, Faye B, Kane AW, Lo CM, Niang B, Boucher Y. Analysis of reasons for extraction of endodontically treated teeth: a prospective study. *J Endod* 2011;37:1512–5.

2. Reeh ES, Messer HH, Douglas WH. Reduction in tooth stiffness as a result of endodontic and restorative procedures. *J Endod* 1989;15:512–6.
3. Dai L xia, Tong W liang, Guo J ling. The role of occlusal factors in the occurrence of vertical root fracture. *Shang Hai Kou Qiang Yi Xue* 2013;22:68–71.
4. Clark D, Khademi J. Modern molar endodontic access and directed dentin conservation. *Dent Clin North Am* 2010;54:249–73.
5. Sorensen JA, Martinoff JT. Intracoronal reinforcement and coronal coverage: a study of endodontically treated teeth. *J Prosthet Dent* 1984;51:780–4.
6. Khanagar SB, Al-ehaideb A, Maganur PC, et al. Developments, application, and performance of artificial intelligence in dentistry – a systematic review. *J Dent Sci* 2021;16:508–22.
7. Tamse A, Fuss Z, Lustig J, Kaplavi J. An evaluation of endodontically treated vertically fractured teeth. *J Endod* 1999;25:506–8.
8. Papapanou PN, Sanz M, Buduneli N, et al. Periodontitis: consensus report of workgroup 2 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. *J Clin Periodontol* 2018;45(Suppl 20):S162–70.
9. Ulmer D, Meijerink L, Cinà G. Trust issues: uncertainty estimation does not enable reliable OOD detection on medical tabular data. *Proc Machine Learn Health* 2020;136:341–54.
10. Hancock JT, Khoshgoftaar TM. Survey on categorical data for neural networks. *J Big Data* 2020;7:1–41.
11. Mahmoud NM, Fouad H, Alsadon O, Soliman AM. Detecting dental problem related brain disease using intelligent bacterial optimized associative deep neural network. *Cluster Comput* 2020;23:1647–57.
12. Kingma DP, Ba J. *Adam: A Method for Stochastic Optimization*. 2014. arXiv:1412.6980.
13. Stone M. Cross-validated choice and assessment of statistical predictions. *J R Stat Soc Series B Methodol* 1974;36:111–33.
14. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017;60:84–90.
15. Li Z, Liu F, Yang W, Peng S, Zhou J. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transact Neural Networks Learn Syst* 2022;33:6999–7019.
16. Mireku AS, Romberg E, Fouad AF, Arola D. Vertical fracture of root filled teeth restored with posts: the effects of patient age and dentine thickness. *Int Endod J* 2010;43:218–25.
17. Patel S, Bhuva B, Bose R. Present status and future directions: vertical root fractures in root filled teeth. *Int Endod J* 2022;55:804–26.