Received: 28 November 2022 | Revised: 8 March 2023 | Accepted: 17 March 2023

DOI: 10.1111/jre.13122

ORIGINAL ARTICLE

Journal of PERIODONTAL RESEARCH WILEY

Deep learning improves prediction of periodontal therapy effectiveness in Chinese patients

Ruiyang Wang[1] | Ruixin Wang[2] | Tong Yang[2] | Jian Jiao[1] | Zhanqiang Cao[3] | Huanxin Meng[1] | Dong Shi[1]

[1]Department of Periodontology, Peking University School and Hospital of Stomatology & National Center for Stomatology & National Clinical Research Center for Oral Diseases & National Engineering Research Center of Oral Biomaterials and Digital Medical Devices, Beijing, China

[2]School of Computer Science, Peking University, Beijing, China

[3]Center for Information, Peking University School and Hospital of Stomatology & National Center for Stomatology & National Clinical Research Center for Oral Diseases & National Engineering Research Center of Oral Biomaterials and Digital Medical Devices, Beijing, China

Correspondence
Dong Shi, Department of Periodontology, Peking University School and Hospital of Stomatology & National Clinical Research Center for Oral Diseases & National Engineering Laboratory for Digital and Material Technology of Stomatology & Beijing Key Laboratory of Digital Stomatology, Beijing, China, 22 Zhongguancun South Avenue, Beijing 100081, China.
Email: dentstone@163.com

Abstract

The objective: This study aims to propose a new model to predict the specific treatment effectiveness at site level by analyzing massive amounts of periodontal clinical data with deep learning methods.

The background data discussing the present status of the field: In light of the low accuracy of current tools, the proposed models cannot fully meet the needs of clinical effectiveness prediction and cannot be applied to on site level prognosis development and formulation of specific treatment plan.

Materials and Methods: Periodontal examination data of 9273 Chinese patients were extracted and used to propose a Sequence-to-Sequence model after performing data management and reconstruction. The model was optimized by introducing the Attention mechanism.

Results: In the test set, the model obtained an average site-level probing depth (PD) accuracy (defined as the proportion of sites with <1 mm deviation of the predicted result from the true value) of 92.4% and high sensitivity (98.6%) for the pocket closure variable. For sites with baseline PD <5 mm, the model achieved a prediction accuracy of 94.6%, while it decreased to 79.9% at sites with PD ≥5 mm. In contrast, for teeth with initial mean PD ≥5 mm, the prediction accuracy significantly differed between molars and non-molars.

Conclusion: Our model is the first to predict the site-level effectiveness with high accuracy and sensitivity. Future prediction models should incorporate deep learning for improved clinical prediction.

KEYWORDS
big data, deep learning, modelling, periodontitis, prognosis, treatment prediction

1 | INTRODUCTION

Clinical outcomes in periodontal treatment have been confirmed to be influenced by several systemic and local factors such as age, gender, cigarette smoking, and patient compliance, among others.[1–3] The combined effects of various factors produce varying states and progression of periodontitis among different individuals which leads to great differences in treatment response.[4]

Since the last century, a number of prognostic models and risk assessment tools[2,5–8] based on the abovementioned factors have been developed to calculate the probability of periodontitis progression[9,10] and to predict tooth survival after therapy.[3,5,7,11] However,

the existing models have the following limitations: (1) most qualitative models[1,6,8,12] are empirically specified and grouped with no specific evidence of division[13]; (2) the models were developed based on small samples and traditional statistical methods such as the COX proportional-hazards model or logistic regression, which may lead to possible bias contributed from both sample selection and the multicollinearity between multiple factors at different levels[14]; (3) and no model can predict the treatment outcome at a given time point after periodontal therapy. Although new methods such as recursive partitioning, random forest, extreme gradient boosting, and other multivariate models have been recently applied,[15] the proposed models still cannot fully meet the needs of clinical effectiveness prediction and upon external validation across centers, presented low accuracy.[6,16] Furthermore, these newer models cannot be applied to onsite level prognosis development and formulation of specific treatment plan.

The electronic periodontal charting record system (EPCRS) in the Information Center of Peking University Hospital and School of Stomatology, Beijing, China, has collected clinical data on Chinese patients with periodontitis since 2007. The data accessed before 2015 has been applied to a preliminary analysis of periodontitis in Chinese patients, showing that parameters at the site, tooth, and individual levels contributed to non-surgical periodontal therapy effectiveness.[17,18] To date, the EPCRS has collected over a hundred thousand records of periodontal charts, including over a billion characters. For such massive data, more powerful computing tools are needed for retrospective analysis.[10]

Recently, deep learning (DL) has continued to evolve and has achieved excellent results in many dental practical scenarios.[19,20] Currently, DL is gradually used as a novel tool in periodontology, such as in the identification of microorganisms[21] and antibody,[22] and diagnosis of periodontal bone loss and periodontitis staging.[23] With the quantity of prior work laid out, it was natural to attempt the application of DL methods in processing large accumulations of clinical data.

This study aimed to propose a novel predictive model for specific treatment effectiveness achieved after certain time at site level by analyzing massive periodontal clinical data using deep learning methods.

2 | MATERIALS AND METHODS

2.1 | Study population

The present model development and validation study employed data obtained from patients who had received periodontal therapy in the Department of Periodontology, Peking University School and Hospital of Stomatology, Beijing, China, between January 2015 and January 2021.

This study was approved by the human subjects ethics board of the Ethics Committee of the Peking University School and Hospital of Stomatology (approval number: PKUSSIRB-202167116) and was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2013. As this study was retrospective, informed consent was waived and ethics committee approval was obtained.

The inclusion criterions were as follows: (a) Patients in the context of clinical care diagnosed with periodontitis according to the classification scheme proposed at the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions[24] and were staged using the algorithm developed by Graetz et al[25]; (b) patients with complete records and periodontal chart recordings in EPCRS; (c) and patients with at least one record of re-evaluation after first-visit within 3 months.

The exclusions were as follows: (a) patients with medical records that cannot be verified or cannot be corresponded to a real individual; (b) patients who were not diagnosed with periodontitis; (c) and patients with contraindications to periodontal treatment, such as severe systemic disease (intolerant or not recommended for periodontal treatment)

The process of patient selection and screening is presented in Figure 1. A total of 65 389 patients were screened, of whom 9273 patients fulfilled the mentioned criteria and were included in the analysis.

2.2 | Data extraction and selected variables

On the basis of our previous study,[17,18] the following parameters were assessed and extracted at the initial visit (T0), first re-evaluation within 3 months after T0 (T1), and every re-evaluation (Tn) from the patient records stored in the EPCRS for analysis:

Patient level:

1. Age and remained teeth at T0
2. Frequency of periodontal maintenance (FPM): regular (receiving periodontal treatment at least once a year) or irregular
3. Sex (male or female)
4. Clinicians (periodontists or postgraduate students)

Tooth level:

1. Mazza bleeding index (BI) values recorded 30 s after probing (0–5)[26]
2. Tooth mobility (0–III)[27]
3. Furcation involvement (FI) for multirooted teeth measured by Glickman classification (0-IV)[28]

Site level

1. Probing depth (PD) measured at six sites (mesial, distal and middle sites of buccal and lingual surfaces) using a Williams periodontal probe;
2. Clinical attachment loss (CAL) measured by the distance from the cementoenamel junction to the bottom of the periodontal pocket

Tooth and site level data from the third molars were excluded; teeth lost and implants were marked specially.

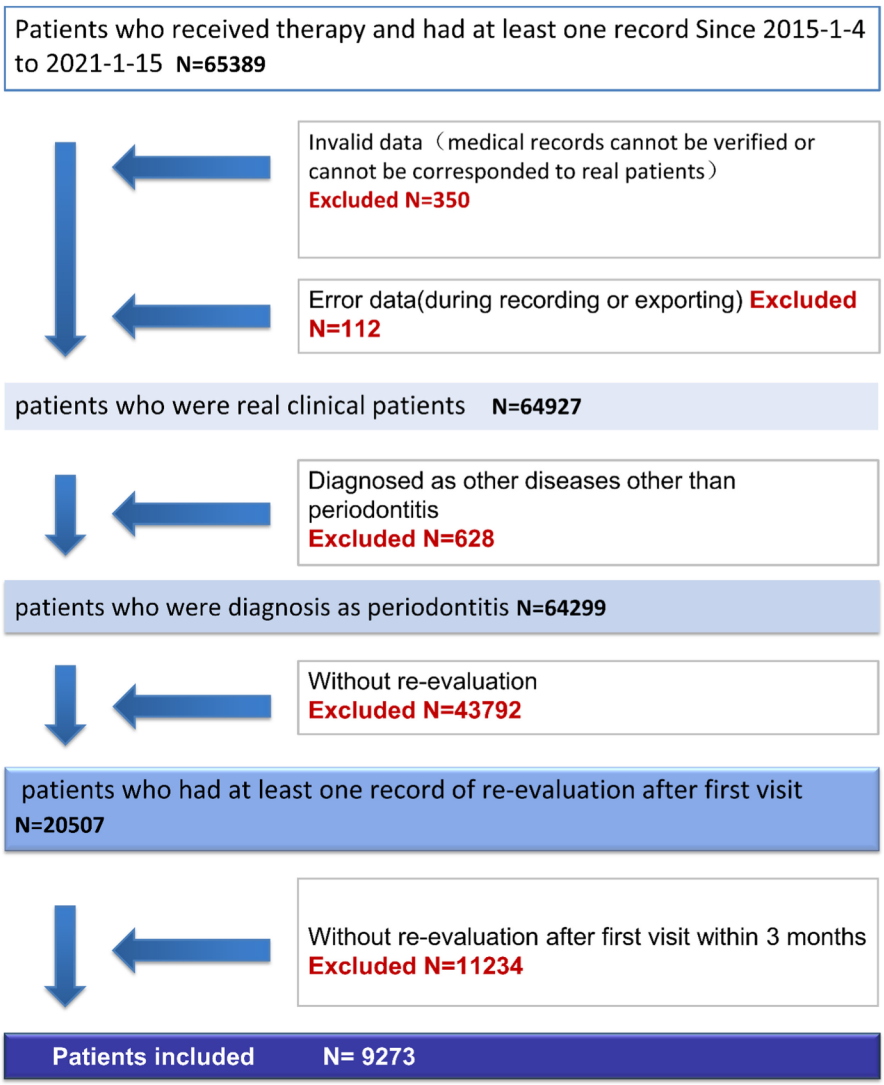

FIGURE 1 Flow chart of patients'selection and screening.

2.3 | Periodontal examination and treatment

2.3.1 | Periodontal examination

The periodontal clinical data of all subjects were examined and recorded simultaneously at the dental chair. Examination results by postgraduate students were double-checked by supervising periodontists. Each clinician had been calibrated with one senior periodontist, and the intra-class correlation of the examination results between the clinician and the senior periodontist before entering the clinic must reach 0.98, which showed high consistency.

2.3.2 | Standard treatment plan

Non-surgical periodontal treatment was performed after the initial examination as claimed in our previous article.[17,18,29] All patients were re-evaluated after 6–8 weeks of initial treatment. Pockets deeper than 5 mm with bleeding on probing or teeth with FI ≥2 were suggested for surgery; however, some patients refused. Furthermore, elective systemic and topical antimicrobial medication was provided at the indication of the periodontist responsible for the treatment. Full-mouth periodontal charting, oral hygiene index reinforcement, prophylaxis scaling, and SRP for residual pockets with PD≥4 mm sites were also performed for every re-visit during maintenance phase. Intervals of SPT were set at 3–6 months.

All periodontal examinations and treatments were performed by qualified clinical periodontists and postgraduate students who were systematically trained and calibrated in pre-clinical programs.

2.3.3 | Deep learning and statistical analysis

The study was conducted on the data of 9273 patients, 80% of whom were selected for training, 10% for verifying, and the remaining 10% for testing, k-fold cross-validation method was applied and k was set as 10.

The data were managed and reconstructed by *Python* software (Python Software Foundation. Python Language Reference,

version 3.7. Available at http://www.python.org). Statistical analysis and DL algorithm tests were performed using GPU server on FEATURIZE platform and IBM SPSS Statistics 24 software (IBM Corp. 2011).

2.4 | Data Management

2.4.1 | Processing of anomalous data

For data with obvious errors or anomalies the original periodontal chart was checked for correction.

2.4.2 | Filling the blanks

In DL, this leads to bias due to overfitting when the subgroups of the dependent variable to be predicted do not have equal sample sizes. To avoid this situation, it was necessary to create a balanced data set,[30] which was the reason the format of the exported data includes each variable of each tooth; however, this leads to a large number of null values generated by purpose (e.g., there cannot be FI in single-root tooth), and in our record, only positive results were recorded for clinical convenience. We used 0 to replace blanks generated by the negative results (e.g., Mobility and CAL), and we used the average to replace the lost data of a very small amount.

2.4.3 | Dataset reconstruction

To meet the data requirements for modeling, a reconstruction of the dataset was performed as summarized in Table 1.

2.4.4 | Model construction

The PD of each tooth site after specified treatment time was used to predict the effectiveness of periodontal treatment. In the initial stage, all 6 sites of 4 teeth (11, 26, 31 and 46) that is, a total of 24 sites were modeled separately using the random forest method. During model construction, inputs for this model were the personal information and all clinical data before treatment of the patients while the output was the clinical data after treatment, of which, both were sequential. As such, the model construction was considered using the Sequence-to-Sequence (Seq2Seq) method which was the main approach used to build the time series correlated model and map an input to output sequence through the deep neural network model (i.e. LSTMs or GRU).[19]

Seq2Seq can capture the semantic of the input sequence by the encoder module and predict the target sequence in the decoder.[31] This process consists of two links: encoding input (Encoder) and decoding output (Decoder). The periodontal data of each tooth before treatment is represented by a vector, and as the input of each time step on the Encoder side, and the PD values of each tooth after treatment are expressed as a vector, and as the output of each time step on the Decoder side.

Additional three kinds of models were also proposed to improve the accuracy during the experimental stage, which included (2) Seq2Seq + Attention, a mechanism that can encode different vectors according to each time step of the sequence; (3) Seq2Seq + Attention + simple self-training, constructed using model (1) to predict the sieved data and merge the predictions with the original incorporated dataset to generate a new dataset; (4) Seq2Seq + Attention + MixMatch, various semi-supervised learning methods were assembled, and the data were augmented to improve the generalization ability of the model by adding terms to the loss function.

All DL work was done in cooperation with the Computer Science Department of Peking University.

2.4.5 | Outcome

The accuracy at site level is defined as the predicted result deviating from the true value by less than 1 mm to be accurate. The proportion of sites with accurate predictions to the total sites was calculated and used as the accuracy of the model prediction.

Pockets with PD ≤4 mm were defined as pocket closure.[32] We calculated the proportion of the predicted pocket closure to compare with the reality outcome. If the predicted value was consistent with the actual value, the outcome was evaluated True; if it was inconsistent, it was deemed False. If the periodontal pocket was closed, it was evaluated as positive and negative if otherwise.

TABLE 1 The format of data set after reconstruction.

Name	Individual information and clinical data gained from T0	Time	Data of Tn
X	Individual and clinical data at T0	T1 (1st re-evaluation)-T0	Data of T1
X	Individual and clinical data at T0	T2 (2nd re-evaluation)-T0	Data of T2
…	…	…	…
X	Individual and clinical data at T0	Tx (last re-evaluation)-T0	Data of Tx

Note: Tx, the last visit.

To further evaluate the accuracy of the model, the precision, recall, *F*1 Score, and accuracy were calculated with the following formulas:

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}$$

$$F1\ \text{Score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$\text{Accuracy} = \frac{\text{true positive} + \text{true negative}}{\text{true postive} + \text{true negative} + \text{false positive} + \text{true negative}}$$

3 | RESULTS

3.1 | Study population

The demographic data of the participants are shown in Table 2. In total, data from 9273 patients were included in the analysis.

3.2 | The efficacy of treatment

Data from 1 557 864 sites were included into analysis, while the average PD of the baseline was 3.52 mm. After a mean observation period of 246.67 days, the mean PD was 2.72 mm and the overall PD decreased by approximately 0.79 mm ($p < .01$). At the first re-evaluation, approximately 90.26% of the sites with PD ≤4 mm, i.e., periodontal pocket closure was achieved, and at the last re-evaluation, approximately 8.40% of sites were with residual PD ≥5 mm.

3.3 | Accuracy of different models

In the initial stage, the accuracy of 24 sites using random forest was from 51.2% to 68.7%.

In order to find the optimal version, four kinds of models were proposed during the experimental stage, which include (1) Seq2Seq; (2) Seq2Seq + Attention; (3) Seq2Seq + Attention+simple self-training; (4) and Seq2Seq + Attention+MixMatch. Of note, since the the training process of the DL model has some randomness, the results of each experiment had small fluctuations; thus, the average and the optimal prediction accuracy were both calculated.

The results showed that for the four different models, the average prediction accuracies obtained were 89%, 92%, 92%, and 88%, respectively, with the highest prediction accuracy of 92.4% for model (2) Seq2seq + Attention, which became the final model.

With all the parameters determined, the final experiments were conducted (Figure 2), and three repetitions were carried out to eliminate any errors caused by randomness. The prediction accuracy of all three experiments reached more than 92.4%, indicating that the set of parameters can make the model performance more stable.

3.4 | Accuracy of different level

In real world setting, a PD ≥5 mm meant unpredictable prognosis and may be considered a better predictor of periodontal breakdown[33,34]; thus, we suggested that sites with PD ≥5 mm would show lower accuracy. Therefore, the prediction results were analyzed at tooth, site, and the combined (tooth & site) level to observe the prediction accuracy for sites with baseline PD ≥5 mm and <5 mm.

3.5 | At tooth level

Teeth with an average baseline of PD ≥5 mm were grouped together while the remaining were categorized under another group to calculate the prediction accuracy. The results showed that the prediction accuracy of the group with an initial average of PD ≥5 mm was only 83.5%, while the prediction accuracy of the other group with an initial average PD <5 mm was 94.2%. The amount of teeth with average baseline PD≥5 mm was calculated and the result was shown in Figure 3.

3.6 | At site level

Sites with baseline PD ≥5 mm were taken as one group and the others as the other group to calculate the prediction accuracy. The results showed that the prediction accuracy of those with sites of PD ≥5 mm was only 79.9%, while the prediction accuracy of those with sites with initial PD <5 mm was 94.6%, of which, the difference was statistically significant ($p < .01$). Out of comprehensive consideration, the percentage of prediction errors was counted by picking out the sites with baseline PD ≥5 mm in teeth of different positions. The detail was shown in Figure 4. The average error rate was 0.21 ± 0.04 (95% confidence interval [CI]: 0.211–0.214) in the lower teeth while in the upper teeth, average error rate was 0.198 ± 0.038 (95% CI: 0.197–0.199). Due to the large variation in the number of teeth included within each group, it was difficult to avoid a Class I error; thus, a between-group analysis was performed for molar and non-molar teeth, and the results showed a statistical difference ($p < .01$, 95% CI: 0.078–0.080).

3.7 | At individual level

We calculated the prediction accuracy between the different groups in the individual level of test set, and the results revealed a slightly higher accuracy in the male group (92.56%) than in the female group (92.22%), and decreases with increasing age (93.17%, 92.41%, and

TABLE 2 The demographic data of patient.

Category	N	%	Mean (SD)	Range
Age at T0			42.05 (13.03)	10,86
Gender				
Female	5023	54.17		
Male	4250	45.83		
FPM				
Regular	2271	24.49		
Irregular	7002	75.51		
Observation period (days)			246.67 (356.92)	302,126
Clinicians				
Periodontist	3678	39.66		
Postgraduate Student	5595	60.34		
Total	9273			

Note: T0, the initial visit.

Abbreviation: FPM, frequency of periodontal maintenance.

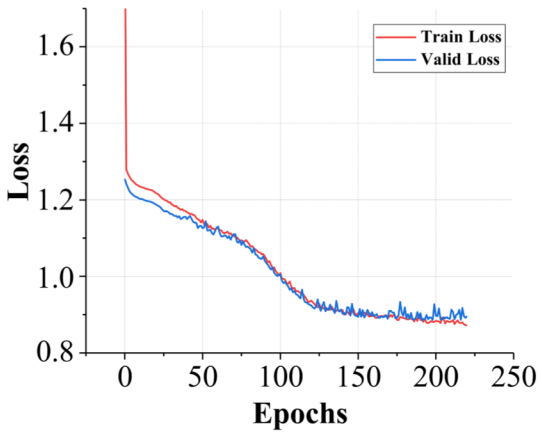

FIGURE 2 The train loss and valid loss with epochs adds of the model (2) Seq2seq + attention.

92.40% for ages <35, 35–50s, and >50 years old, respectively). The accuracy was slightly higher when the procedure was performed by a periodontist (92.8%) than when performed by students (91.8%). Since there was only 1 sample staged as stage I in the test set, stages I and II were combined for the statistics. As staging increases and the disease increased in complexity, a corresponding decrease in predictive accuracy from 96.9% to 91.1% occurred.

In order to improve the prediction accuracy, an optimization scheme of the model was proposed for the above findings. Adding a marker variable for each site to distinguish whether the PD of the site was ≥5 mm, and if it holds, the marker variable was equal to 1, otherwise it was equal to 0, so as to make the model more sensitive to the baseline PD. The experimental results showed that the average prediction accuracy of the scheme was about 92.4%, and the highest prediction accuracy reached 93.3%, which was an improvement compared with the previous optimal version.

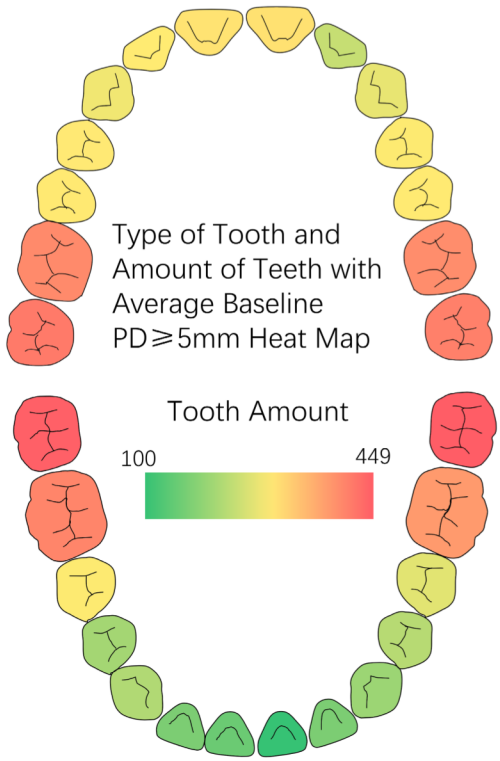

FIGURE 3 The amount of teeth with average baseline PD ≥5 mm in different kind of teeth.

By defining pockets ≤4 mm as the pocket closure, this model showed a precision (positive predictive value) of 93.1%, recall (sensitivity) of 98.6%, F1 Score of 95.7%, and accuracy of 92.0%.

4 | DISCUSSION

Emerging scientific evidence has inspired the development of various periodontal risk assessment tools in calculating the probability of periodontitis progression[9,10] and tooth survival after therapy.[3,5,7,11] Most of these tools calculate the risk of tooth loss as either a continuous or an ordinal (categorical) variable.[16] The use of descriptive words like good, fair, or questionable can be perceived subjectively by clinicians due to heterogeneity in how definitions are interpreted and applied.[35] Most of these tools are based on samples of less than a thousand, while some are based on less than a hundred samples which has been considered to bring huge bias and cannot be applied in practice.[19] Existing models are not able to predict efficacy after a specified arbitrary period of time due to mathematical and statistical limitations, and are not able to focus on continuous changes over the entire treatment time series, which is the reason for the use of the keywords "long-term" and "short-term." However, there cannot be certain clarification to define these two definitions.[4] Additionally, although many of these models have showed promising results in the populations they were originally developed in, external validity of prior

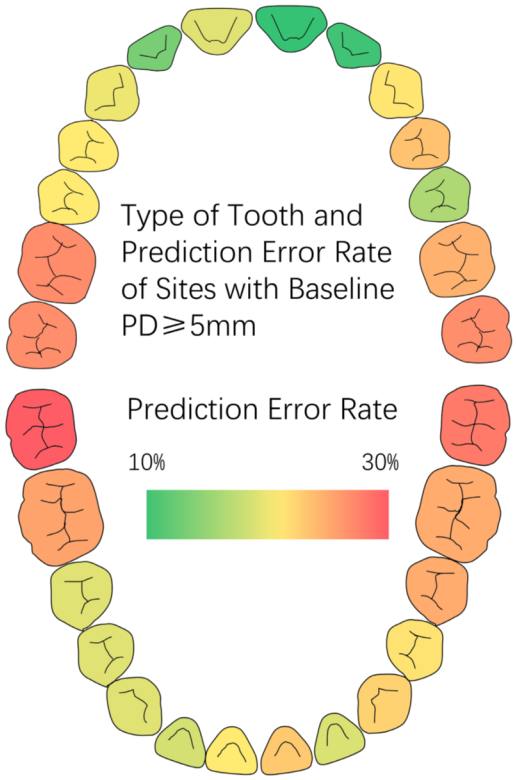

FIGURE 4 The percentage of prediction errors of sites with baseline ≥5 mm in different teeth.

findings remains to be tested on a new cohort.[8,16] Some studies have published validation results[16,35]; however, most models showed low accuracy (AUC ranged between 0.52 and 0.67) in externally validation across centers.[15]

Due to the inadequacy of the existing models, a new model is proposed to eliminate the artificial division of time points or classification criteria, and instead, clinical continuous variable is to be used as the description of the prediction while at the same time predicting the progression and change in treatment outcomes over time. The emergence of DL methods makes this idea possible.

At the stage of exploration, first we used random forest on 24 sites, the accuracy of these models was relatively low (<70%) and tedious. Then four models have been proposed. Initially, the Seq2Seq model was proposed; however, due to the limitations of the model itself, the large information lost during the modeling process would be difficult to recover. By introducing the Attention mechanism, data loss in the model is reduced and the prediction accuracy significantly improved.

In this study, there were approximately 100 000 pieces of real data, of which more than half were sieved out based on the inclusions and exclusions criteria, leaving the data underutilized. To effectively utilize these unlabeled data, semi-supervised learning was introduced in this project during the pilot phase, which was model (3) Seq2Seq + Attention+MixMatch and model (4) Seq2Seq + Attention+SST. However, the accuracy of these two methods was not as good as method (2), and were, therefore, discarded. Combining the above reasons with the comparison of accuracy, we choose Seq2Seq + Attention mechanism as the final model.

In the final model, the PD of each tooth site of the patient after treatment was used as the prediction of the efficacy of periodontal treatment, and the average accuracy achieved was 92.4%. The average percentage of error prediction between molars and non-molars showed a statistically significant difference ($p < .01$). Further, upon statistical analysis, three possible reasons were revealed: (1) the mean amount of teeth at each position with an average baseline of PD ≥5 mm differed between the molars (387.5) and non-molars (147.2) in test dataset; (2) molars were considered inferior treatment outcomes and worse prognosis due to anatomic factors such as presence of furcation, concavities on the root surfaces and cervical enamel projections[36,37]; (3) and the unique anatomical characteristics of Chinese descendants can be another reason.[38,39]

This study was the first site-level precise prediction model which was built on mass followed cohorts of Chinese patients treated for periodontitis. Results of this study reflect the real-world situation that the frequency of subjects with severe periodontitis (stage III or IV) was about 92%, much higher than previous finding in data from the Fourth National Oral Health Survey (2015–2016).[29] After treatment, periodontal pockets closed in approximately 91.6% of sites, with a PD change of approximately 0.79 mm, which is consistent with our previous findings.[17]

However, there are limitations to this study, one being the data of the model were gathered from one center. This may result in overfitting and lead to high apparent performance. When models were externally validated across centers, model performance decreased.[40] Another drawback is that the systematical status, smoking status, and radiographic information of the patients were not included, due to the need to manually consult the medical record system to supplement these pieces of information, a heavy task that has not yet been completed.

In the future, specific treatment procedures will be included in the analysis by sorting out the medical records, intending to realize further model accuracy improvements.

5 | CONCLUSION

In this study, we propose a prediction model that is the first to predict specific effectiveness at the site level by applying deep learning methods to the massive amounts of periodontal clinical data. The novel prediction model achieved high accuracy and sensitivity which confirms the feasibility and applicability of deep learning in prediction.

AUTHOR CONTRIBUTIONS

Ruiyang, Wang: Contributed to conception, design, data acquisition, analysis, and interpretation, drafted the manuscript; Ruixin, Wang: Contributed to design, data analysis, and drafted

the manuscript; Tong, Yang: Contributed to design, data analysis, and critically revised manuscript; Jian, Jiao: Contributed to data acquisition, and critically revised manuscript; Zhanqiang, Cao: Contributed to data acquisition, and critically revised manuscript; Dong, Shi: Contributed to conception, design, data acquisition, and critically revised manuscript; Huanxin, Meng: Contributed to conception and critically revised manuscript. All authors gave their final approval and agree to be accountable for all aspects of the work.

FUNDING INFORMATION

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by grants from National Key R&D of Program of China (Grant NO. 2022YFC2504200) and Peking University School and Hospital of Stomatology (Grant NO.PKUSSNCT-21A03).

CONFLICT OF INTEREST STATEMENT

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

DATA AVAILABILITY STATEMENT

Research data are not shared.

REFERENCES

1. Checchi L, Montevecchi M, Gatto MRA, Trombelli L. Retrospective study of tooth loss in 92 treated periodontal patients. *J Clin Periodontol*. 2002;29(7):651-656.
2. Leininger M, Tenenbaum H, Davideau J-L. Modified periodontal risk assessment score: long-term predictive value of treatment outcomes. A retrospective study. *J Clin Periodontol*. 2010;37(5):427-435.
3. Eickholz P, Kaltschmitt J, Berbig J, Reitmeir P, Pretzl B. Tooth loss after active periodontal therapy. 1: patient-related factors for risk, prognosis, and quality of outcome. *J Clin Periodontol*. 2008;35(2):165-174.
4. Kornman KS. Contemporary approaches for identifying individual risk for periodontitis. *Periodontol 2000*. 2018;78(1):12-29.
5. Santos F, Beato F, Machado V, Proença L, Mendes JJ, Botelho J. Early tooth loss after periodontal diagnosis: development and validation of a clinical decision model. *Int J Environ Res Public Health*. 2021;18(3):1363.
6. McGuire MK, Nunn ME. Prognosis versus actual outcome. III. The effectiveness of clinical parameters in accurately predicting tooth survival. *J Periodontol*. 1996;67(7):666-674.
7. Martinez-Canut P, Alcaraz J, Alcaraz J Jr, et al. Introduction of a prediction model to assigning periodontal prognosis based on survival time. *J Clin Periodontol*. 2018;45(1):46-55.
8. Kwok V, Caton JG. Commentary: prognosis revisited: a system for assigning periodontal prognosis. *J Periodontol*. 2007;78(11):2063-2071.
9. Lang NP, Tonetti MS. Periodontal diagnosis in treated periodontitis. *J Clin Periodontol*. 1996;23(3):240-250.
10. Lang NP, Suvan JE, Tonetti MS. Risk factor assessment tools for the prevention of periodontitis progression a systematic review. *J Clin Periodontol*. 2015;42(S16):S59-S70.
11. Becker W, Becker BE, Berg LE. Periodontal treatment without maintenance: a retrospective study in 44 patients. *J Periodontol*. 1984;55(9):505-509.
12. Hirschfeld L, Wasserman B. A long-term survey of tooth loss in 600 treated periodontal patients. *J Periodontol*. 1978;49(5):225-237.
13. Beck JD, Philips K, Moss K, Divaris K, Morelli T, Offenbacher S. Advances in precision oral health. *Periodontol 2000*. 2020;82(1):268-285.
14. Tu YK. Linear mixed model approach to network meta-analysis for continuous outcomes in periodontal research. *J Clin Periodontol*. 2015;42(2):204-212.
15. Krois J, Graetz C, Holtfreter B, Brinkmann P, Kocher T, Schwendicke F. Evaluating modeling and validation strategies for tooth loss. *J Dent Res*. 2019;98(10):1088-1095.
16. Schwendicke F, Schmietendorf E, Plaumann A, Sälzer S, Dörfer CE, Graetz C. Validation of multivariable models for predicting tooth loss in periodontitis patients. *J Clin Periodontol*. 2018;45(6):701-710.
17. Jiao J, Shi D, Cao Z-q, et al. Effectiveness of non-surgical periodontal therapy in a large Chinese population with chronic periodontitis. *J Clin Periodontol*. 2017;44(1):42-50.
18. Jiao J, Zhang L, Meng H-X, et al. Clinical performance of non-surgical periodontal therapy in a large Chinese population with generalized aggressive periodontitis. *J Clin Periodontol*. 2018;45(10):1184-1197.
19. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface*. 2018;15(141):20170387.
20. Kühnisch J, Meyer O, Hesenius M, Hickel R, Gruhn V. Caries detection on intraoral images using artificial intelligence. *J Dent Res*. 2022;101(2):158-165.
21. Wang CW, Hao Y, Di Gianfilippo R, et al. Machine learning-assisted immune profiling stratifies peri-implantitis patients with unique microbial colonization and clinical outcomes. *Theranostics*. 2021;11(14):6703-6716.
22. Huang W, Wu J, Mao Y, et al. Developing a periodontal disease antibody array for the prediction of severe periodontal disease using machine learning classifiers. *J Periodontol*. 2020;91(2):232-243.
23. Chang H-J, Lee S-J, Yong T-H, et al. Deep learning hybrid method to automatically diagnose periodontal bone loss and stage periodontitis. *Sci Rep*. 2020;10(1):7531.
24. Tonetti MS, Greenwell H, Kornman KS. Staging and grading of periodontitis: framework and proposal of a new classification and case definition. *J Clin Periodontol*. 2018;45(S20):S149-S161.
25. Graetz C, Mann L, Krois J, et al. Comparison of periodontitis patients' classification in the 2018 versus 1999 classification. *J Clin Periodontol*. 2019;46(9):908-917.
26. Mazza JE, Newman MG, Sims TN. Clinical and antimicrobial effect of stannous fluoride on periodontitis. *J Clin Periodontol*. 1981;8(3):203-212.
27. Lindhe J, Lang NP, Karring T, eds. *Clinical Periodontology and Implant Dentistry*. John Wiley & Sons Ltd.; 2015.
28. Glickman I, ed. *Clinical Periodontology: Prevention, Diagnosis, and Treatment of Periodontal Disease in the Practice of General Dentistry*. 4th ed. Saunders; 1972.
29. Jiao J, Jing W, Si Y, et al. The prevalence and severity of periodontal disease in mainland China: data from the fourth National Oral Health Survey (2015–2016). *J Clin Periodontol*. 2021;48(2):168-179.
30. Yakar M, Etiz D, Metintas M, Ak G, Celik O. Prediction of radiation pneumonitis with machine learning in stage III lung cancer: a pilot study. *Technol Cancer Res Treat*. 2021;20:15330338211016373.
31. Tang YJ, Pang YH, Liu B. IDP-Seq2Seq: identification of intrinsically disordered regions based on sequence to sequence learning. *Bioinformatics*. 2021;36(21):5177-5186.
32. Tomasi C, Wennström JL. Is the use of differences in the magnitude of CAL gain appropriate for making conclusions on the efficacy of non-surgical therapeutic means? *J Clin Periodontol*. 2017;44(6):601-602.
33. Citterio F, Gualini G, Chang M, et al. Pocket closure and residual pockets after non-surgical periodontal therapy: a systematic review and meta-analysis. *J Clin Periodontol*. 2022;49(1):2-14.

34. Matuliene G, Pjetursson BE, Salvi GE, et al. Influence of residual pockets on progression of periodontitis and tooth loss: results after 11 years of maintenance. *J Clin Periodontol*. 2008;35(8):685-695.
35. Saleh MHA, Dukka H, Troiano G, et al. External validation and comparison of the predictive performance of 10 different tooth-level prognostic systems. *J Clin Periodontol*. 2021;48(11):1421-1429.
36. Heitz-Mayfield LJA, Trombelli L, Heitz F, Needleman I, Moles D. A systematic review of the effect of surgical debridement vs. non-surgical debridement for the treatment of chronic periodontitis. *J Clin Periodontol*. 2002;29(S3):92-102.
37. Matthews DC, Tabesh M. Detection of localized tooth-related factors that predispose to periodontal infections. *Periodontol 2000*. 2004;34(1):136-150.
38. Goh EXJ, Ong MMA. Anatomical, microbiological, and genetic considerations in treatment of Chinese periodontal patients. *J Invest Clin Dent*. 2019;10(1):e12381.
39. Zee KY, Chiu MLB, Holmgren CJ, Walker RT, Corbet EF. Cervical enamel projections in Chinese first permanent molars*. *Aust Dent J*. 1991;36(5):356-360.
40. Zhan Y, Holtfreter B, Meisel P, et al. Prediction of periodontal disease: modelling and validation in different general German populations. *J Clin Periodontol*. 2014;41(3):224-231.

How to cite this article: Wang R, Wang R, Yang T, et al. Deep learning improves prediction of periodontal therapy effectiveness in Chinese patients. *J Periodont Res*. 2023;00:1-9. doi:10.1111/jre.13122