

# SCIENTIFIC REPORTS



OPEN

## Predicting postoperative facial swelling following impacted mandibular third molars extraction by using artificial neural networks evaluation

Wei Zhang<sup>1,2</sup>, Jun Li<sup>3</sup>, Zu-Bing Li<sup>1,4</sup> & Zhi Li<sup>1,4</sup>

Patients' postoperative facial swelling following third molars extraction may have both biological impacts and social impacts. The purpose of this study was to evaluate the accuracy of artificial neural networks in the prediction of the postoperative facial swelling following the impacted mandibular third molars extraction. The improved conjugate grads BP algorithm combining with adaptive BP algorithm and conjugate gradient BP algorithm together was used. In this neural networks model, the functional projective relationship was established among patient's personal factors, anatomy factors of third molars and factors of surgical procedure to facial swelling following impacted mandibular third molars extraction. This neural networks model was trained and tested based on the data from 400 patients, in which 300 patients were made as the training samples, and another 100 patients were assigned as the test samples. The improved conjugate grads BP algorithm was able to not only avoid the problem of local minimum effectively, but also improve the networks training speed greatly. 5-fold cross-validation was used to get a better sense of the predictive accuracy of the neural network and early stopping was used to improve generalization. The accuracy of this model was 98.00% for the prediction of facial swelling following impacted mandibular third molars extraction. This artificial intelligence model is approved as an accurate method for prediction of the facial swelling following impacted mandibular third molars extraction.

The extraction of impacted mandibular third molars is one of the most common surgical events. Patients' postoperative facial swelling following third molars extraction may have both biological impacts and social impacts. The severity of the facial swelling is depending on the case and the facial swelling originated in an inflammatory process that initiated by surgical trauma. The facial swelling is more likely occurring in those patients who received impacted mandibular third molars extraction with utilizing flap operation and bone removal. The factors thought to influence the incidence of facial swelling after third molar removal include patients' age, gender, physique and oral hygiene<sup>1-5</sup>. In addition, the incidence of facial swelling is related to the type of third molar, the degree of impaction, and ease of extraction operations<sup>1-8</sup>. As far as a specific patient is concerned, most of oral surgeons could predict the incidence of the facial swelling solely based on their personal experience. The first attempt to create a model of prediction was published by MacGregor<sup>7</sup>, who established a multivariate model based on radiographic findings. Berge and Bøe<sup>8</sup> attempted to predict the extent of postoperative morbidity following third molar surgery through multiple regression analysis. Although a number of efforts at determining a model for this assessment have been made, none of them could be praised for universal reliability<sup>1</sup>. Obviously, the main reason was the difficulty of quantification of the magnitude of the contributions of the different categories of variables.

<sup>1</sup>The State Key Laboratory Breeding Base of basic science of Stomatology (Hubei-MOST) and Key Laboratory for Oral Biomedical Engineering of the Ministry of Education (Wuhan University), Wuhan, China. <sup>2</sup>Department of Endodontics, School and Hospital of Stomatology, Wuhan University, Wuhan, China. <sup>3</sup>School of Mechanical and Materials Engineering, Jiujiang University, Jiujiang, China. <sup>4</sup>Department of Oral and Maxillofacial Surgery, School and Hospital of Stomatology, Wuhan University, Wuhan, China. Correspondence and requests for materials should be addressed to Z.L. (email: zhili@whu.edu.cn)

Received: 5 December 2017

Accepted: 16 July 2018

Published online: 16 August 2018

Artificial neural networks principle can be used to analyze the multiple factors and the relationships between the factors which have unclear relationships, and it is a system based on the imitation of human brain structure and function. Neural networks method is a branch of biomedical engineering known as artificial intelligence (AI). Artificial neural networks, due to their excellent ability of non-linear mapping, generalization, self-organization and self-learning, have been proved to be of widespread utility in digital signal processing, system modeling, automatic control, and others<sup>9–12</sup>. The applications of neural networks have received a great deal of attentions in clinical medicine. By properly choosing neural networks structures and training the weights, researchers could use neural networks to perform medical outcome prediction<sup>13–18</sup>. Neural networks have the capacity to “learn” how to make a diagnosis through the information presented to them<sup>19–21</sup>.

In this present study, an AI model was established with improved conjugate gradient backpropagation (BP) algorithm to predict facial swelling following impacted mandibular third molars extraction. The software we used for the experiments is MATLAB's Neural Network Toolbox<sup>22</sup>. 5-fold cross-validation is used to get a better sense of the predictive accuracy of the neural network and early stopping is used to improve generalization. The purpose of this study was to evaluate the accuracy of this improved neural networks model in predicting.

## Methods and Materials

**Study population.** The study protocol was approved by the Ethics Committee of the School and Hospital of Stomatology, Wuhan University, Wuhan, China. All patients were informed of the clinical study, and they provided signed informed consent. The study was conducted in full accordance with ethical principles, including the Declaration of Helsinki. All methods were performed in accordance with the relevant guidelines and regulations. The 400 patients, who received impacted mandibular third molars extraction in Hospital of Stomatology, Wuhan University, were randomly selected in this study. There were 300 patients assigned as the training samples, which is the design set, and another 100 patients were made as the test samples. All samples were selected from a pool of patients admitted for mandibular third molars extraction between January 2014 and December 2016.

These selected patients had no known immune impairment, no contraindications for oral surgery and no medication taking. The surgical procedure was standardized and uniformly performed by one same oral surgeon, and patients strictly followed post-extraction instructions. Antibiotics were used for three days after extraction to prevent infection in all patients. A single examiner performed all clinical measurements prior to surgery as well as 72 hours after the operation.

In this study, the parameters, which related to the occurrence of facial swelling following impacted mandibular third molars extraction, were considered as the networks input, including 15 factors. These 15 factors were consisted of patient's age, gender, physique and oral hygiene as personal factors; relation of the third molars to the mandible ramus and second molar, relative depth of the third molars in bone, relationship of the long axis of the third molars in relation to the long axis of the second molar, relation of the third molars in mandibular dental arch, number of root as anatomy factors of third molars; and type of incision, location and quantity of bone removal, section into pieces or not, root fracture condition, fracture of lingual bone plate or not, surgical time as factors of surgical procedure. All the parameters and the data normalization are shown in Table 1.

According to Table 1, the parameters data of 400 samples were normalized and counted as the inputs. The patients' prognosis data were the output, in which severe swelling was recorded 1, moderate swelling was recorded 0.5, and mild swelling was recorded 0.

**Measurement of facial swelling.** The facial swelling measurements were taken using a 2–0 nylon thread and a millimeter ruler before the surgery and 72 hours after the operation. To evaluate the swelling, markings with permanent marker were made prior to the surgery on the following facial regions: the angle of the mandible, the tragus, the labial commissure, the nasal border, laterally to the external corner of the eye, and on the soft pogonion. The following measurements were taken: Distance I (from the angle of the mandible to tragus); Distance II (from the angle of the mandible to the external corner of the eye); Distance III (from the angle of the mandible to the nasal border); Distance IV (from the angle of the mandible to the labial commissure); Distance V (from the angle of the mandible to the soft pogonion)<sup>23,24</sup>. Differences between the measurements prior to surgery and those made 72 hours after the operation at distances were recorded, and the average of the five differences value was calculated. When the average of the five differences value was less than 10 mm, the patient was classified as mild swelling; 10–20 mm was moderate swelling; more than 20 mm was severe swelling. All postoperative data were recorded by the same independent blinded investigator in order to avoid observer bias.

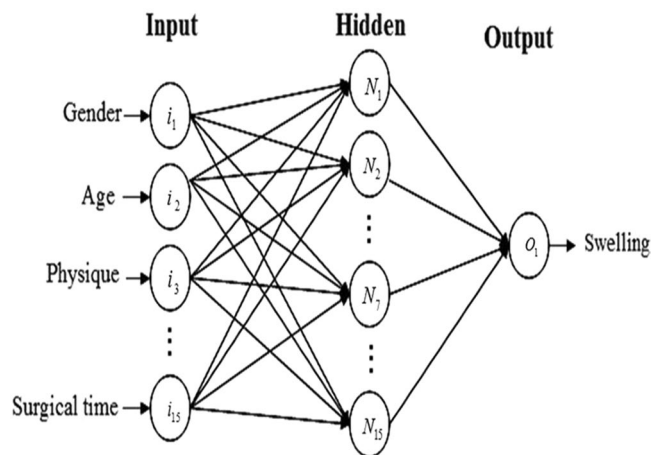
**Basic Principle of the neural networks model.** The neural networks model based on the improved conjugate gradient BP algorithm combining adaptive BP algorithm and conjugate gradient BP algorithm is shown in Fig. 1.

The neural network used in this experiment was a one-hidden-layer perception with the improved conjugate gradient BP algorithm that had 15 neurons in the input layer, 15 in the hidden layer, and 1 in the output layer. Figure 1 illustrates this neural network. Where in and  $N_n$  ( $n$  varies from 1 to 15) are the  $n^{\text{th}}$  input neuron and the  $n^{\text{th}}$  hidden-layer neuron, respectively, and  $Ok$  is the output neuron for  $k^{\text{th}}$  set of input data. These artificial neurons passed information from one layer to the other, in a process that imitated a human synapse, and the processed information produced an output signal. The number of interactions or learning cycles of the network was established upon trial and error in all the interactions, and used the reduction of the mean square error (MSE) as learning criterion.

Like conventional BP algorithm, conjugate gradient BP algorithm has the same process of propagating the input forward through the networks. The main difference between them is the process of propagating the error backward through the networks. The conjugate gradient BP algorithm reduces the amount of computation in the

Parameters	Normalization
X1: Gender	Male: 0.5; Female: 1
X2: Age	$\leq 20$ : 0.5; 20–30: 0.25; 30–40: 0.5; $\geq 40$ : 1
X3: Physique	Strong: 0; Middle: 0.5; Slim: 1
X4: Oral hygiene	Good: 0; Middle: 0.5; Bad: 1
X5: Relation of the wisdom teeth to the mandible ramus and second molar	Type I: 0; Type II: 0.5; Type III: 1
X6: Relative depth of the wisdom teeth in bone	High: 0; Middle: 0.5; Low: 1
X7: Relationship of the long axis of the wisdom teeth in relation to the long axis of the second molar	Vertical: 0; Buccalclination/ Lingualclination: 0.25; Mesialclination/ Distalclination: 0.50; Horizontal: 0.75; Inversion: 1
X8: Relation of the wisdom teeth in mandibular dental arch	Lingual displacement: 0.25; Normotopia: 0.5; Buccal displacement: 1
X9: Number of root	1: 0; 2: 0.5; $\geq 3$ : 1
X10: Type of incision	No incision: 0; Buccal incision: 0.5; Distal incision: 0.5; Buccal incision + Distal incision: 1
X11: Location and quantity of bone removal	No: 0; Buccal/Distal/Occlusal: 0.5; Buccal + Distal/Buccal + Occlusal/ Distal + Occlusal: 0.75; Buccal + Distal + Occlusal: 1
X12: Section into pieces or not	no section: 0; section into 2 pieces: 0.5; section into 3 or more pieces: 1
X13: Root fracture condition	No: 0; 1 root fracture: 0.25; 2 roots fracture: 0.5; $\geq 3$ roots fracture: 1
X14: Fracture of lingual bone plate or not	No: 0; Yes: 1
X15: Surgical time	$\leq 10$ minutes: 0; 10–20 minutes: 0.5; $\geq 20$ minutes: 1

**Table 1.** Normalization of the parameters data.



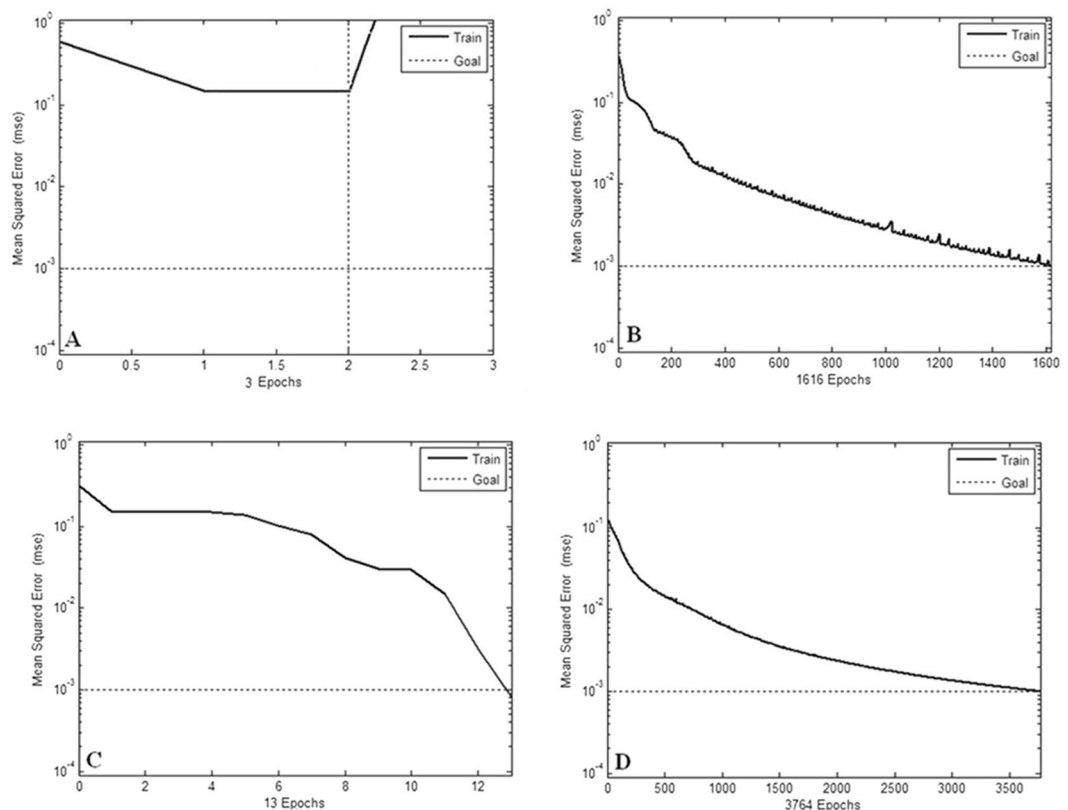
**Figure 1.** The architecture of the neural networks model.

training process, thus there is no need to calculate the search direction in every step of the iterative process in the conventional BP algorithm.

**Algorithm process.** The process of the improved BP Algorithm is as follows. At the first step, the networks weights are initialized by random values, and the training sample values are inputted, in which the required error precision is set as  $10^{-3}$ . At the second step, propagating the input forward through the networks, and judging whether MSE of sample meets the required error precision ( $10^{-3}$ ). If it does, then the learning will end. At the third step, it is judgment that whether the networks training enters extreme point. If it does, then the conjugate gradient method will be used in the backpropagating of networks. Otherwise, the adaptive BP algorithm will be used in the backpropagating of networks. At the fourth step, they are propagating the optimized weights forward through the networks, and judging whether the MSE of sample meets the error precision. If it does not, then it will return to the third step. After the above four steps, one epoch of the improved BP Algorithm is completed, in which epoch is an iteration of subset of the samples for training algorithm.

Judging values of extreme point	> 0.1	0.01	0.001	0.0001
Training epochs	overflow	1616	13	3764

**Table 2.** Judging value of extreme point has influence on networks training speed.



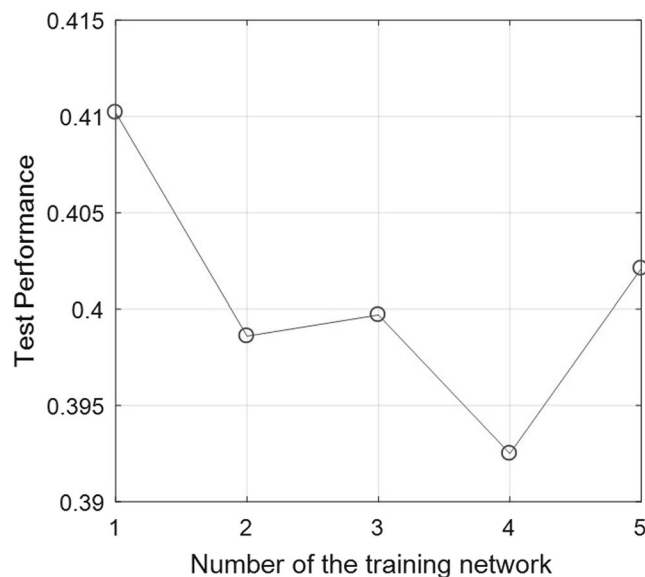
**Figure 2.** (A–D) The different networks training epochs under four different judging value of extreme point (A: 0.1; B: 0.01; C: 0.001; D: 0.0001.).

**Judging whether the networks training enters extreme value.** At the third step of the algorithm process, it is necessary to judge whether the networks training enters extreme point. In the present study, 10 epochs training is the interval, and the difference value of training errors in the intervals is evaluated. If the difference value is less than the judging value of extreme point, the networks training is considered as entering the extreme point. The judging value of extreme point will influence the speed of networks training. If the judging value of extreme point is too high, error function of the sample will not meet the conditions to further use conjugate gradient method, which will have trouble with overflow. If the judging value of extreme point is too low, conjugate gradient method will be used until networks training is trapped in the local minimum for a long time, which will then result in remarkably slow networks training speed.

## Results

**Networks training.** When adaptive BP algorithm was applied, the learning curve would be apt to be trapped in infinitesimal, and there would be still no convergence until training epoch was over 100000. When conjugate gradient BP algorithm was used in networks training of the same sample, it would be apt to overflow. However, when improved BP algorithm was used, the problem of local minimum would be avoided effectively by networks training, and the speed of networks training would be improved greatly. The networks training epoch would be 13 when networks training of this sample converged to the condition that MSE was 0.001.

**Judging value of extreme point has influence on networks training speed.** When the required error precision was set at  $10^{-3}$ , networks training epochs were represented as different judging values of extreme point (Fig. 2A–D). From the Fig. 2, we can obtain the Table 2, which shows the convergence epoch of networks training when different judging values of extreme point were set. When judging value of extreme point was set at 0.001, networks training speed would be the fastest.



**Figure 3.** The test performance of the training network in 5-fold cross-validation.

**Improving the predictive accuracy of the neural network and improving generalization.** To get a better sense of the predictive accuracy and improve generalization, 5-fold cross-validation and early stopping were used. In this technique the available data (400 samples) is divided into two subsets. The first subset (300 samples) is the designing dataset, which is used for training and validating the neural network, the second subset (100 samples) is the test dataset.

5-fold cross-validation and early stopping technique were used in order to get a better sense of the predictive accuracy and avoid overfitting. 5-fold cross-validation splits the designing dataset into five parts at random. It trains five new neural networks, each one on four parts of the designing dataset, it then examines the predictive accuracy of each new neural network on the data not included in training dataset, which is named the validation set. The method for improving generalization is called early stopping. In this technique, the error on the validation set is monitored during the training process. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to overfit the data, the error on the validation set typically begins to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned. In this case, every training fold contains roughly 4/5 of the designing dataset and every validation fold contains roughly 1/5 of the designing dataset.

The test dataset which is not used during training, but it is used to compare different models.

**Results of prediction test.** When the neural network output value was lower than 0.4, it meant mild or no swelling; when it was higher than 0.6, it meant severe swelling; when it was between 0.4 and 0.6, it meant moderate swelling. When 5-fold cross-validation and early stopping technique are used, the test performance of each training network in 5-fold cross-validation is showed in Fig. 3. The best test performance is the fourth time of cross-validation, its value is 0.3925, which was computed using matlab's function `perform`<sup>22</sup>.

The test dataset (the diagnosis 100 samples) was performed with prediction test by the fourth trained network of 5-fold cross-validation. The confusion matrix of the test dataset, the validation dataset, the test dataset and all of the dataset was shown in Fig. 4, the test dataset and the validation dataset are both fitted, while two cases misfit in the test dataset. The network's performance was measured according to the mean of squared errors, error vs. epoch for the training, validation, and test performances of the training record was shown in Fig. 5, the training stops after 28 consecutive increases in validation error, and the best performance is taken from the epoch with the lowest validation error. Error histogram was shown in Fig. 6. As is shown in Table 3, in the prediction test dataset, 98 cases did match the reality, and the prediction accuracy was 98.00%.

## Discussion

Postoperative facial swelling is a common complaint after impacted third molar removal, and the swelling degree is different among the patients. The swelling differs depending on the patients' characteristics, preoperative difficulty index, and intraoperative factors<sup>1-8</sup>. All clinicians stress the need for swelling control in patients who undergo third molar surgery, and many investigations focused on reduction of the swelling<sup>23-27</sup>. However, there is no special intelligent method to predict the swelling. The inaccuracy in the prediction of the swelling response always affects the quality of medical services and the follow-up treatment plan.

The factors that contribute toward the postoperative facial swelling following third molar surgery are sophisticated. In this study, 15 variables were recorded as the parameters in the artificial neural networks. The major advantage of artificial neural networks is their ability to synthesize a considerable amount of factors (or variables) without requiring statistical modeling of the problem. In addition, artificial neural networks learn to recognize



Figure 4. The confusion matrix of the dataset.

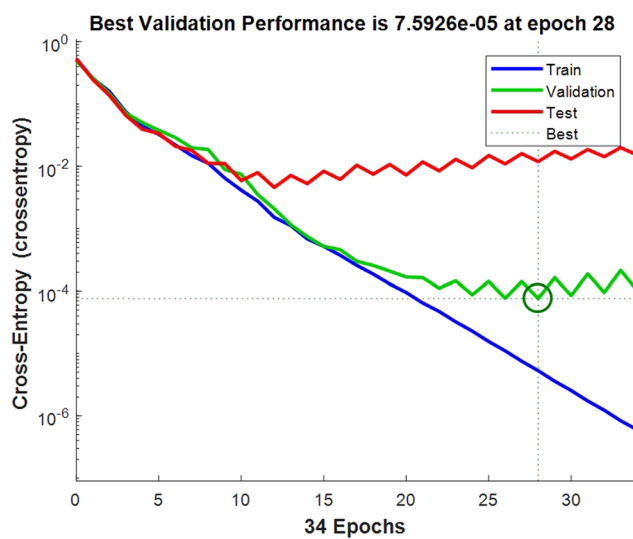
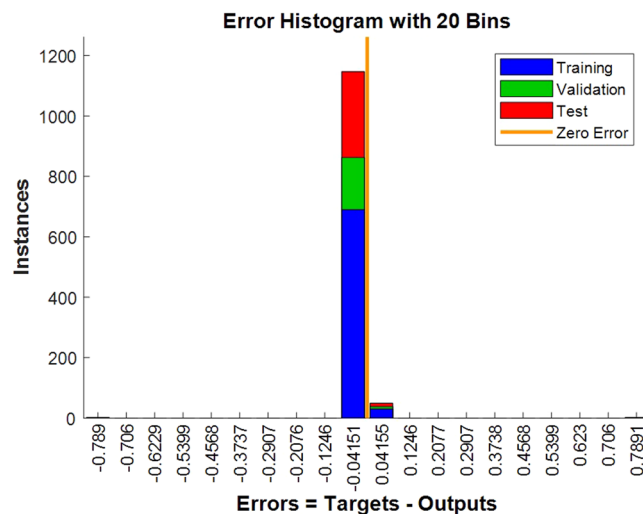


Figure 5. The error vs. epoch for the training, validation, and test performances of the training record.

The output values of the network	< 0.4 (mild or no swelling)	0.4–0.6 (moderate swelling)	> 0.6 (severe swelling)	Total
The number of the reality	35	47	16	98
The number of correct prediction	33	46	15	94

**Table 3.** Results of prediction test.



**Figure 6.** The error histogram.

patterns from entered data through training. Once it has been trained on entered data, the network is able to recognize similarities when presented with a new input pattern, which will result in a predicted output pattern<sup>28</sup>.

It is worth mentioning that we have employed computerized system based on artificial neural networks to predict the swelling following impacted mandibular third molars extraction, which shows satisfying results. The improved conjugate grads BP algorithm combining adaptive BP algorithm and conjugate gradient BP algorithm together can converge to a global optimal solution in certain conditions, which not only avoid the problem of local minimum effectively, but also improve networks training speed greatly<sup>29</sup>.

The 'feed-forward, back propagation' neural networks usually trained by BP of errors is perhaps the most popular networks architecture today<sup>30</sup>, in which the neural networks based on adaptive BP algorithm is a new widely used diagnosis and prediction method. Adaptive BP algorithm, however, may generally be trapped in a local minimum. Another variation on BP algorithm, the conjugate gradient BP algorithm allows the searching the direction of deterioration as a certain probability in each step, because each iteration is not a single negative gradient direction but a conjugate direction. Hence, conjugate gradient BP algorithm can converge to a global optimal solution in certain conditions, which can solve the local minimum problem of adaptive BP algorithm. However, conjugate gradient method requires the error function is convex, and the direction of each iteration requires exact linear search. The practical problems often can not meet the conditions. In general, error function of extreme value shows characteristics of convex function, and the conjugate gradient method being applied near the extreme value will improve convergence speed and learning accuracy. Therefore, the algorithm combining adaptive BP algorithm and conjugate gradient BP algorithm together may be an ideal neural networks model, which is able not only to avoid the problem of local minimum effectively, but also to improve networks training speed greatly. We believe this improved conjugate grads BP algorithm could serve as a valid neural networks model to predict the swelling following impacted mandibular third molars extraction. In this present study, the improved algorithm combining adaptive BP algorithm and conjugate gradient BP algorithm is used to establish a valid neural networks model applied in the prediction of facial swelling following impacted mandibular third molars extraction.

In this study, 5-fold cross-validation and early stopping were used to get a better sense of the predictive accuracy and improve generalization. The prediction accuracy of this method is 98.00%, which can meet the needs of actual diagnosis or prediction in clinic. There are two reasons may explain the fact that the prediction unable to achieve higher accuracy. First, there are too many related parameters ( $n = 15$ ) involved in this study. In the follow-up study, some minor parameters should be abandoned, in order to optimize the networks training and improve the prediction accuracy. Second, the normalization process of these parameters data is lack of objective evidence. In future, the mathematical formulae of normalization for these parameters data should be sought to improve the scientificity and accuracy of sample.

In conclusion, this artificial neural networks based on the improved conjugate grads BP algorithm has a high prediction accuracy, which is a new developed system that can help to predict the swelling following impacted mandibular third molars extraction.

## References

- de Santana-Santos, T. *et al.* Prediction of postoperative facial swelling, pain and trismus following third molar surgery based on preoperative variables. *Med Oral Patol Oral Cir Bucal*. **18**, e65–70 (2013).
- Bouloux, G. F., Steed, M. B. & Perciaccante, V. J. Complications of third molar surgery. *Oral Maxillofac Surg Clin North Am*. **19**, 117–128 (2007).
- Larrazábal, C., García, B., Peñarrocha, M. & Peñarrocha, M. Influence of oral hygiene and smoking on pain and swelling after surgical extraction of impacted mandibular third molars. *J Oral Maxillofac Surg*. **68**, 43–46 (2010).
- Capuzzi, P., Montebugnoli, L. & Vaccaro, M. A. Extraction of impacted third molars. A longitudinal prospective study on factors that affect postoperative recovery. *Oral Surg Oral Med Oral Pathol*. **77**, 341–343 (1994).
- Haug, R. H., Perrott, D. H., Gonzalez, M. L. & Talwar, R. M. The American association of oral and maxillofacial surgeons age-related third molar study. *J Oral Maxillofac Surg*. **63**, 1106–1114 (2005).
- Suarez-Cunqueiro, M. M., Gutwald, R., Reichman, J., Otero-Cepeda, X. L. & Schmelzeisen, R. Marginal flap versus paramarginal flap in impacted third molar surgery: a prospective study. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. **95**, 403–408 (2003).
- MacGregor, A. J. The radiological assessment of ectopic lower third molars. *Ann R Coll Surg Engl*. **61**, 107–113 (1979).
- Berge, T. I. & Bøe, O. E. Predictor evaluation of postoperative morbidity after surgical removal of mandibular third molars. *Acta Odontol Scand*. **52**, 162–169 (1994).
- Stulp, F. & Sigaud, O. Many regression algorithms, one unified model: A review. *Neural Netw*. **69**, 60–79 (2015).
- Banerjee, A. K., Ravi, V., Murty, U. S., Sengupta, N. & Karuna, B. Application of intelligent techniques for classification of bacteria using protein sequence-derived features. *Appl Biochem Biotechnol*. **170**, 1263–1281 (2013).
- Muddle, J. *et al.* Predicting the fine particle fraction of dry powder inhalers using artificial neural networks. *J Pharm Sci*. **106**, 313–321 (2017).
- Sidhu, A. & Yang, Z. R. Prediction of signal peptides using bio-basis function neural networks and decision trees. *Appl Bioinformatics*. **5**, 13–19 (2006).
- Llorca, J., Dierssen-Sotos, T., Gómez-Acebo, I., Gonzalez-Castro, A. & Miñambres, E. Artificial neural networks predict mortality after lung transplantation better than logistic regression. *J Heart Lung Transplant*. **28**, 1237–1238 (2009).
- Mobley, B. A., Schechter, E., Moore, W. E., McKee, P. A. & Eichner, J. E. Neural network predictions of significant coronary artery stenosis in men. *Artif Intell Med*. **34**, 151–161 (2005).
- Saritas, I., Ozkan, I. A. & Sert, I. U. Prognosis of prostate cancer by artificial neural networks. *Expert Systems with Applications*. **37**, 6646–6650 (2010).
- Sheppard, D. *et al.* Predicting cytomegalovirus disease after renal transplantation: an artificial neural network approach. *Int J Med Inform*. **54**, 55–76 (1999).
- Paydar, K., Niakan Kalthori, S. R., Akbarian, M. & Sheikhtaheri, A. A clinical decision support system for prediction of pregnancy outcome in pregnant women with systemic lupus erythematosus. *Int J Med Inform*. **97**, 239–246 (2017).
- Raith, S. *et al.* Artificial Neural Networks as a powerful numerical tool to classify specific features of a tooth based on 3D scan data. *Comput Biol Med*. **80**, 65–76 (2016).
- Goodey, R. D., Brickley, M. R., Hill, C. M. & Shepherd, J. P. A controlled trial of three referral methods for patients with third molars. *Br Dent J*. **189**, 556–560 (2000).
- Devito, K. L., de Souza Barbosa, F. & Felipe Filho, W. N. An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. **106**, 879–884 (2008).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature*. **521**, 436–444 (2015).
- Beale, M. H., Hagan, M. T. & Demuth, H. B. Neural Network Toolbox™ User's Guide How to Contact MathWorks.: 406 (2015).
- Neupert, E. A. 3<sup>rd</sup>, Lee, J. W., Philput, C. B. & Gordon, J. R. Evaluation of dexamethasone for reduction of postsurgical sequelae of third molar removal. *J Oral Maxillofac Surg*. **50**, 1177–1182 (1992).
- De Menezes, S. A. & Cury, P. R. Efficacy of nimesulide versus meloxicam in the control of pain, swelling and trismus following extraction of impacted lower third molar. *Int J Oral Maxillofac Surg*. **39**, 580–584 (2010).
- Sağlam, A. A. Effects of tube drain with primary closure technique on postoperative trismus and swelling after removal of fully impacted mandibular third molars. *Quintessence Int*. **34**, 143–147 (2003).
- Satılmış, T., Garip, H., Arpacı, E., Sener, C. & Göker, K. Assessment of combined local anesthesia and ketamine for pain, swelling, and trismus after surgical extraction of third molars. *J Oral Maxillofac Surg*. **67**, 1206–1210 (2009).
- Van Eeden, S. P. & Bütow, K. Post-operative sequelae of lower third molar removal: a literature review and pilot study on the effect of Covomycin D. *SADJ*. **61**, 154–159 (2006).
- Bas, B. *et al.* Use of artificial neural network in differentiation of subgroups of temporomandibular internal derangements: a preliminary study. *J Oral Maxillofac Surg*. **70**, 51–59 (2012).
- Charalambous, C. Conjugate gradient algorithm for efficient training of artificial neural networks. *J IEE Proceedings G* **139**(3), 301–310 (1992).
- Loukas, Y. L. Artificial neural networks in liquid chromatography: Efficient and improved quantitative structure–retention relationship models. *J Chromatogr A*. **904**, 119–129 (2000).

## Acknowledgements

This work was supported by the Fundamental Research Fund for the Central Universities of China (2042017kf0206) and the Scientific Research Project of the Health and Family Planning Commission of Hubei Province (WJ2017M047).

## Author Contributions

Z.L. and W.Z. performed experiments, wrote the manuscript, analyzed the data, and validated the results. J.L. and Z.-B.L. served as consultants and assisted in the experiments.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018