

Predicting patient experience of Invisalign treatment: An analysis using artificial neural network

Lin Xu^a 
Li Mei^b
Ruiqi Lu^c
Yuan Li^a
Hanshi Li^a
Yu Li^a 

^aDepartment of Orthodontics, State Key Laboratory of Oral Diseases, National Clinical Research Center for Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China

^bDiscipline of Orthodontics, Department of Oral Sciences, Sir John Walsh Research Institute, Faculty of Dentistry, University of Otago, Dunedin, New Zealand

^cDepartment of Electronic Engineering, Tsinghua University, Beijing, China

Objective: Poor experience with Invisalign treatment affects patient compliance and, thus, treatment outcome. Knowing the potential discomfort level in advance can help orthodontists better prepare the patient to overcome the difficult stage. This study aimed to construct artificial neural networks (ANNs) to predict patient experience in the early stages of Invisalign treatment. **Methods:** In total, 196 patients were enrolled. Data collection included questionnaires on pain, anxiety, and quality of life (QoL). A four-layer fully connected multilayer perception with three backpropagations was constructed to predict patient experience of the treatment. The input data comprised 17 clinical features. The partial derivative method was used to calculate the relative contributions of each input in the ANNs. **Results:** The predictive success rates for pain, anxiety, and QoL were 87.7%, 93.4%, and 92.4%, respectively. ANNs for predicting pain, anxiety, and QoL yielded areas under the curve of 0.963, 0.992, and 0.982, respectively. The number of teeth with lingual attachments was the most important factor affecting the outcome of negative experience, followed by the number of lingual buttons and upper incisors with attachments. **Conclusions:** The constructed ANNs in this preliminary study show good accuracy in predicting patient experience (i.e., pain, anxiety, and QoL) of Invisalign treatment. Artificial intelligence system developed for predicting patient comfort has potential for clinical application to enhance patient compliance.

Key words: Computer algorithm, Pain, Compliance, Aligners

Received September 28, 2021; Revised January 27, 2022; Accepted March 2, 2022.

Corresponding author: Yu Li.

Professor, Department of Orthodontics, State Key Laboratory of Oral Diseases, National Clinical Research Center for Oral Diseases, West China Hospital of Stomatology, Sichuan University, 14#, 3rd Section, South Renmin Road, Chengdu 610041, China.

Tel +86-028-85503645 **e-mail** yuli@scu.edu.cn

How to cite this article: Xu L, Mei L, Lu R, Li Y, Li H, Li Y. Predicting patient experience of Invisalign treatment: An analysis using artificial neural network. Korean J Orthod 2022;52(4):268-277. https://doi.org/10.4041/kjod21.255

© 2022 The Korean Association of Orthodontists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Artificial intelligence (AI) has been developing rapidly and has made remarkable achievements in various domains of medicine and dentistry.¹ Compared with the traditional logistic regression models, artificial neural networks (ANNs) with multilayer perceptions (MLP) have shown an advantage in modelling complicated nonlinear relationships, with higher sensitivity, specificity, and accuracy for medical diagnosis.² In a study that determined the cervical vertebrae stages in orthodontics, ANN demonstrated high accuracy value and was the most stable among the algorithms evaluated, which included k-nearest neighbors, naive Bayes, decision tree, support vector machine, and random forest.³ Given its capabilities of modelling nonlinear relationships in a high-dimensional data set, ANN has provided new approaches for orthodontists in automated cephalometric analysis and cone beam computed tomography image segmentation, accurate diagnosis, and treatment planning.⁴

Invisalign, one of the fastest-developing orthodontic appliances in dentistry, translates orthodontic treatment plans into a series of clear aligners to align teeth.⁵ Although Invisalign is more esthetic and comfortable than traditional fixed appliances,^{6,7} in our clinical practice, some patients still complain about a varying degree of discomfort and anxiety.^{8,9} Both traditional appliances and Invisalign have caused, to some extent, oral dysfunction, mucosal irritation, difficulty in chewing, and swollen throat or tongue.^{10,11} This could reduce the wear time of aligners and compliance, thus influencing the treatment outcome,¹² and a small number of patients even give up treatment because of terrible experience.^{13,14} Therefore, attention to mental status should be considered in the treatment plan for the best possible patient-centered care.

Theoretically, the complexity of an appliance may directly affect a patient's comfort level. However, the impact of different aligner designs and the relationship among them are unclear. Clinical evidence for predicting patient experience using ANN is lacking. Therefore, an AI system was constructed for patient comfort prediction, which can be later applied in software to help orthodontists predict the comfort level of designed aligners. If significant discomfort is detected, some modifications and health education can be considered in advance to improve the patient's comfort level and compliance. This study aimed to construct ANNs to predict patient experience (i.e., pain, anxiety, and quality of life [QoL]) of Invisalign treatment based on different designs of Invisalign treatment to, ultimately, help clinicians identify individuals at risk of poor patient experience and reduced treatment compliance.

MATERIALS AND METHODS

Study design and subject selection

This prospective cohort study was approved by the Ethics Committee of the West China Hospital of Stomatology, Sichuan University (WCHSIRB-D-2019-073) and was conducted according to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all the patients.

A total of 196 patients wearing Invisalign clear aligners (Align Technology, Phoenix, AZ, USA) were recruited at the Department of Orthodontics, West China Hospital of Stomatology, Sichuan University, Chengdu, China between 2018 and 2021. The sample size was decided based on practical grounds (existing study cohort) and in reference to similar studies concerning AI systems.¹⁵⁻¹⁷ The inclusion criteria were: (1) age 18–50 years; (2) planned wearing of Invisalign clear aligners; and (3) no history of major dentoalveolar diseases. The exclusion criteria were as follows: (1) dental and oral diseases, such as caries, periodontal diseases, and temporomandibular joint disorders; (2) severe systemic diseases; (3) psychological and mental disorders; and (4) current medications that treat or cause pain and mental diseases. Tooth extraction surgeries and the placement of temporary anchorage devices would interfere with, and even mask, the patients' self-reported discomfort from the aligner designs. Therefore, these surgeries were performed at least 1 week before or after the questionnaire investigation to avoid potential interference. All the patients wore clear aligners attached following the same protocol (22 h/day for 10 days).

Data collection and assessments

Clinical patient records were obtained from hospital databases. The animation scheme and treatment design of the Invisalign treatment (with anonymized personal information) were collected from ClinCheck (Align Technology). A total of 17 clinical features were collected from the medical records and ClinCheck (Table 1). Given the complexity of clinical practice, the features were not grouped together based on any experience, such as grouping the elastics and precision cut together, and were maintained as originally as possible.

The patients were asked to fill in the questionnaires daily for 8 days: the day before and the first 7 days after wearing the first set of aligners. The questionnaires included the visual analog scale (VAS) of pain, Self-Rating Anxiety Scale (SAS), and Oral Health Impact Profile-14 (OHIP-14) based on the literature.¹⁸⁻²⁰ In VAS, the degree of pain was assessed by marking the pain level on a 10-mm straight line, ranging from 0 mm (no pain) to 10 mm (worst pain). In SAS, the level of anxiety was assessed using 20 questions, with each question answered

Table 1. Input normalization in the three artificial neural networks

Categories	Data type		Criterion
Age (yr)	Integral		MMN
Treatment stage	Binary	First-time	0
		Refinement	1
Crowding (mm)	Discrete	No	0
		I°	1/3
		II°	2/3
		III°	1
With/without extraction	Binary	Yes	1
		No	0
Number of extractions	Integral		MMN
Wearing aligners and bonding attachments simultaneously or separately	Binary	Yes	1
		No	0
With/without molar distalization	Binary	Yes	1
		No	0
With/without elastics	Binary	Yes	1
		No	0
Number of elastics	Integral		MMN
With/without interproximal reduction	Binary	Yes	1
		No	0
Amount of interproximal reduction (mm)	Continuous		MMN
Number of teeth with attachments	Integral		MMN
Number of teeth with optimized attachments	Integral		MMN
Number of teeth with lingual attachments	Integral		MMN
Number of upper incisors with attachments	Integral		MMN
Number of lingual buttons	Integral		MMN
With/without precision cut	Binary	Yes	1
		No	0

The crowding data classification: I°, 0 mm ≤ crowding < 4 mm; II°, 4 mm ≤ crowding < 8 mm; and III°, crowding ≥ 8 mm. MMN, maximum minimum normalization.

as occasionally, sometimes, often, or always. In OHIP-14, the QoL was assessed using 14 items, with 5 options for responses: never, hardly, occasionally, fairly often, and very often. These questionnaires have been validated to have good reliability.

Dataset pre-processing

All elements of the input clinical features were normalized in the range of 0 to 1. The normalization of the input is presented in Table 1. Maximum minimum normalization, a linear and hyperparameter-free method, was utilized for normalization of integral and continuous features. This is because non-linear methods, such as dividing subgroups (analysis of data grouping) and non-linear curves, would introduce more hyperparam-

eters and increase the complexity of the ANN model and the risk of overfitting. For the binary feature, the raw value was already normalized and, therefore, remained unchanged. The label for patient experience was 0 or 1. After collecting the questionnaire scores of patient experience, the differences in pain, anxiety, and QoL between the highest and lowest scores were calculated to address inter-patient subjectivity. A higher difference indicated a more negative patient experience with the Invisalign treatment.

The differences were then binarized using thresholds, which were determined as the averages of the above differences (3.0 for pain, 6.5 for anxiety, and 7.0 for QoL). Patients with differences higher than the threshold were considered positive samples (label = 1) and those with

Table 2. Number of patients with positive and negative tags in the training, validation, and test sets in the three artificial neural networks

Dataset	Pain			Anxiety			Quality of life		
	Positive	Negative	Total	Positive	Negative	Total	Positive	Negative	Total
Training set	45	71	116	48	64	112	62	61	123
Validation set	16	20	36	16	15	39	17	18	35
Test set	16	28	44	25	20	45	18	28	38
Total set	77	119	196	83	113	196	97	99	196

The changes are calculated as the difference between the highest and lowest scores. Higher values indicate more negative patient experience with the Invisalign treatment. The values are then binarized using predefined thresholds to distinguish between positive and negative samples at the following cutoffs: 3.0 for pain, 6.5 for anxiety, and 7.0 for quality of life. The binarized values are the final prediction targets.

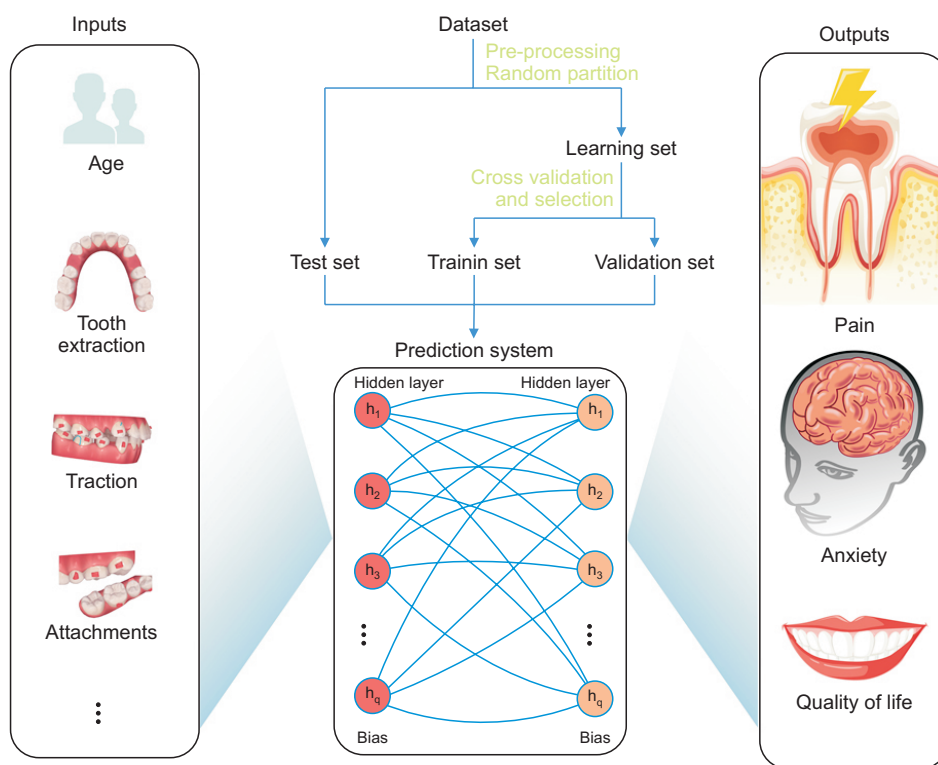


Figure 1. Flow diagram of the construction of artificial neural networks. The three artificial neural networks are fully connected and includes two hidden layers with a hidden size of nine.

differences lower than the threshold were considered negative (label = 0). Table 2 shows the number of patients with positive and negative labels distributed in the training, validation, and test sets. Given that the number of positive and negative samples varied for different prediction targets of pain, anxiety, and QoL, samples with different labels were randomly divided among the datasets. Thereafter, the positive and negative samples were almost equally distributed in the training, validation, and test sets for each ANN. The binarized values were the final prediction targets.

Construction of artificial neural networks

Figure 1 shows the ANN analysis process. The 17 clinical

features (Table 1) for each patient were collected as inputs. Three ANNs were constructed to predict whether negative experiences would occur in patients receiving their first set of aligners. The ANNs were four-layer fully connected MLPs, with 17 input nodes, two hidden layers with nine hidden nodes per layer, and one output node. The rectified linear units (ReLU) function was chosen as the activation function for nonlinearity after each hidden layer. It is calculated as follows:

$$\text{ReLU}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where x is the value calculated by linear operations before the activation function.²¹

A positive probability was obtained by applying a nonlinear sigmoid function to the value of the output node. It is calculated as follows:

$$\text{sigmoid}(x) = \frac{1}{1 + \exp(-x)} \quad (1)$$

where x is the result of the output node.²² The other two ANNs for anxiety and QoL prediction shared the same model structure but were trained separately; therefore, the parameter values were different. During the training stage, random dropout with a probability of 0.5 was adopted in the hidden layer, which randomly set the activation values of a certain number of hidden nodes to 0 to increase the training stability. The binary cross-entropy (BCE) loss was used to calculate the difference between the ground truth and the predicted result. It is calculated as follows:

$$\text{BCE}(x) = - [(1-y) \log(1-x) + y \log(x)] \quad (2)$$

where y is the ground truth label, and x is the predicted result of the ANN. The backpropagation algorithm was used to update the parameters of the neural network based on equation (2). The learning rate was set to 0.1, according to recent literature.²³ An adaptive moment (Adam) estimation optimizer was used to update the ANN parameters.²⁴

Training and evaluation of ANNs

The dataset of the 196 patients was divided into the training, validation, and test sets in a ratio of 3:1:1.^{23,25} Although the back propagation method could be used to train the ANN parameters, there were still some untrainable hyper-parameters, such as the learning rate, total number of training steps, and number of nodes in the hidden layer. To determine these hyperparameters, a four-fold cross-validation method was used in the training and validation sets.^{17,26} For each validation fold, the samples within the training and validation sets were first combined and then randomly divided into two in a 3:1 ratio.

The larger part was used to optimize the parameters of the ANNs, and the smaller part was used to monitor the training process and check for overfitting. Normally, the loss in the smaller part will first decrease for some training steps and then start to increase at a certain point, producing a minimum. This procedure was repeated four times for each set of hyperparameters, and the average loss was calculated. The set of hyperparameters with the smallest loss was selected. After all hyperparameters

were determined through validation, the training and validation sets were combined again to train the final model. The test set was held-out and not available during all the processes stated above and was only used to evaluate the success rate of the final model.

Each sample was labeled either 0 or 1, but the predicted probability of the ANN using equation (1) ranged from 0 to 1. Therefore, a threshold was needed to determine whether the sample was influenced by bad experiences.

We defined a determination of pain, anxiety, and decreased QoL for each patient as the predicted probability being higher than the corresponding optimum diagnostic cutoff value derived from the receiver operating characteristic (ROC) curves.²⁷ Using the cutoff value, the predicted probability of each sample was binarized into 0 and 1. If the prediction of one sample was equal to its label, then it was counted as a success sample; otherwise, it was considered as a failure sample. The success rate of prediction, sensitivity, specificity, and area under the ROC curve (AUC) were used to evaluate the ANN performance. As the training and validation sets were all used during training, the success rate was high. However, the test set was held-out during the training process to simulate a real situation in which the model predicts experiences for new patients. Therefore, if the success rate was also high on the test set, we can assume with certainty that the model could perform well in real scenes.

Analysis of contributions of input features

The partial derivatives method, widely applied in providing the contribution profile of input factors, was used to calculate the relative contributions of inputs and rank them in order.²⁸ The contribution of each input to each ANN was calculated, indicating the influence of each feature on the output of pain, anxiety, and QoL. The total contribution of the three ANNs for each input was also calculated, with higher values of the total contribution denoting a higher influence of each input factor on the output of the overall negative experience with equal weights.

RESULTS

Accuracy of ANN prediction of patient experience

The learning curves of the three ANNs during cross-validation are shown in Figures 2A-2C. For the training and validation losses, the training loss was optimized during training, and the validation loss was used to determine when the learning was sufficient and whether to stop the training process. The training loss decreased slowly with fluctuation, while the validation loss decreased relatively quickly in the early stage and quickly

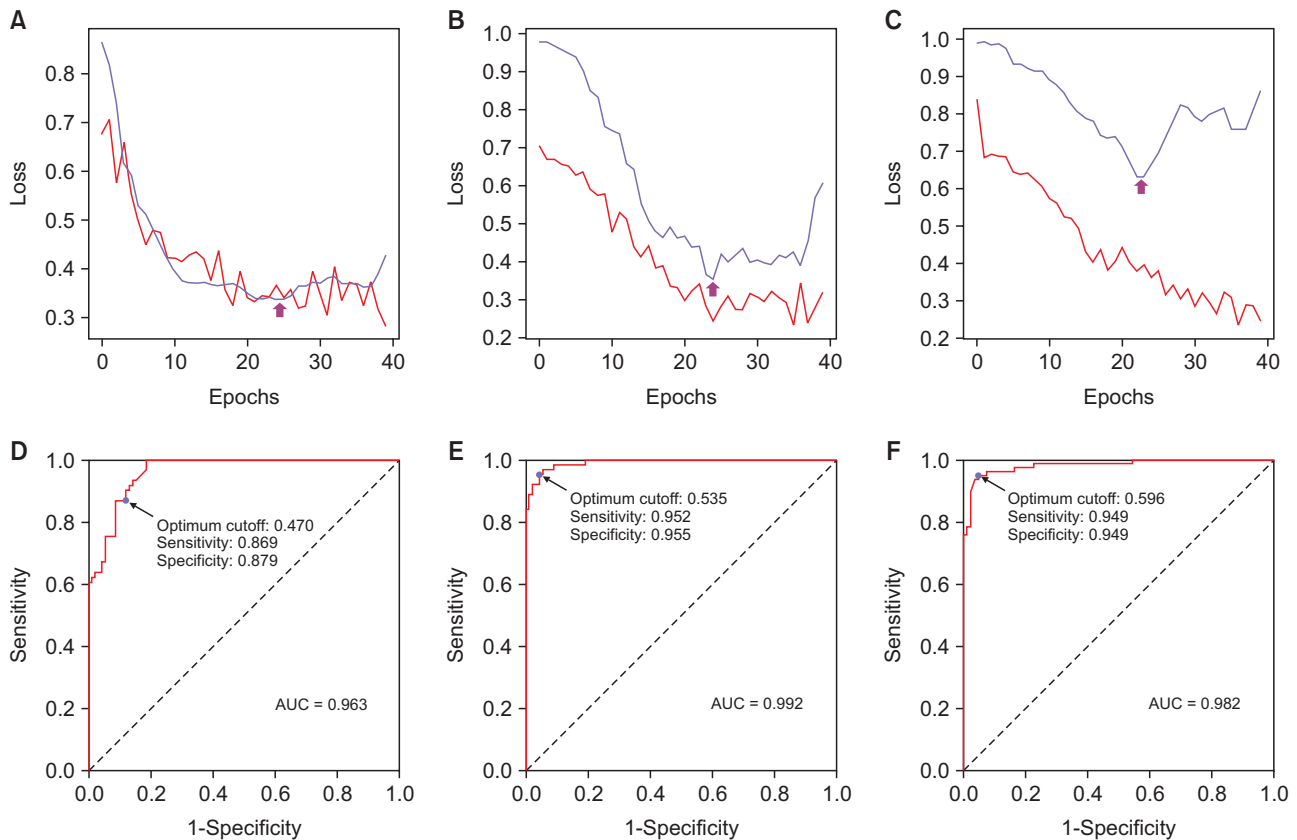


Figure 2. Prediction performance of the artificial neural networks (ANNs). The learning curves of ANNs for pain (A), anxiety (B), and quality of life (C). Red lines represent train loss curve; purple lines, validation loss curve. Arrows indicate the lowest point of validation loss curve, which means the training procedure for pain, anxiety, and quality of life are stopped at 25, 24, and 22 epochs, respectively. The ROC curves of ANNs for pain (D), anxiety (E), and quality of life (F). The optimum diagnostic cutoff value is marked as purple points, where the sensitivity and specificity are shown upon the arrows.

ROC, receiver operating characteristic; AUC, area under the curve.

Table 3. Performance of artificial neural networks for patient experience

Performance	Pain	Anxiety	Quality of life
AUC	0.963 (0.904, 0.972)	0.992 (0.983, 0.995)	0.982 (0.950, 0.990)
Sensitivity	0.885 (0.803–0.984)	0.952 (0.921–0.968)	0.937 (0.899–0.975)
Specificity	0.890 (0.813–0.934)	0.955 (0.920–0.977)	0.937 (0.873–0.962)

Data are presented as the median (95% confidence interval). AUC, area under the curve.

saturated. After training for certain epochs, the validation loss stopped decreasing and started to increase. This means that although the model behaved better on the training set, its accuracy on the validation set did not improve. This was a sign of overfitting, and the training procedure was therefore stopped at the lowest point of the validation loss curve. It could also be seen that the validation loss was consistently higher than the training loss for the prediction of anxiety and QoL because the

model was optimized only on the training set. Based on the learning curves, training for pain, anxiety, and QoL was stopped at 25, 24, and 22 epochs, respectively.

The ROC and AUC are effective and comprehensive measures for assessing the inherent validity of a diagnostic test and the overall performance of the ROC curve. The AUC, sensitivity, and specificity of predicting pain, anxiety, and QoL are shown in Figures 2D–2F and Table 3. The results demonstrated satisfactory

performance of the three ANNs in predicting patient discomfort. The success rates of the ANNs were calculated according to the ROC curves. The overall success rate of ANN for pain prediction was 87.7%, and the success rates of the training, validation, and test sets were 87.9% (95% confidence interval [CI]: 83.6–90.5%), 86.1% (95% CI: 83.3–91.7%), and 88.6% (95% CI: 84.1–93.2%), respectively. The total success rate of anxiety prediction was 93.4%, and the success rates of the training, validation, and test sets were 94.6% (95% CI: 87.5–97.3%), 94.9% (95% CI: 89.7–97.4%), and 88.9% (95% CI: 80.0–91.1%), respectively. The overall success rate of the ANN for predicting QoL was 92.4%, and the success rates of the training, validation, and test sets were 91.9% (95% CI: 83.7–96.2%), 94.3% (95% CI: 80.0–97.1%), and 92.1% (95% CI: 81.6–97.4%), respectively. The accuracy of the test set was consistent with that of the training and validation sets, indicating negligible overfitting. Notably, the test set was not available during the learning process until the final evaluation of success rate. This demonstrated that the constructed ANNs could prospectively predict the discomfort level of new patients with clinical features of the treatment plan.

Predictors and their influence on patient experience

The contributions of the inputs to the output target were analyzed using the partial derivatives method. The results of the contribution of the inputs to each ANN are illustrated in Table 4, and the total contributions are ranked in order, as shown in Figure 3. The number of teeth with lingual attachments was the most important factor affecting the outcome of negative experiences, followed by the number of lingual buttons and the number of upper incisors with attachments. Wearing the first pair of aligners and bonding attachments simultaneously or separately had a minimal impact on overall patient experience. The treatment stage was a negligible feature in predicting pain, had a mild impact on anxiety, and had a moderate impact on QoL. This means that the treatment stage has a variable impact on the patient experience.

DISCUSSION

To the best of our knowledge, this study is the first to construct ANNs to predict patient experience of Invisalign treatment. A patient’s treatment experience is clinically important. In general, although orthodontists might mainly consider the treatment outcome, they are

Table 4. Contributions of the 17 inputs for target prediction

Input categories	Contribution		
	Pain	Anxiety	Quality of life
Number of teeth with lingual attachments	30.495 (3.164, 64.621)	13.557 (3.018, 33.124)	414.976 (190.285, 684.003)
Number of lingual buttons	6.139 (1.226, 12.541)	263.655 (145.051, 419.175)	71.548 (28.348, 131.841)
Number of upper incisors with attachments	0.462 (0.068, 4.223)	8.710 (3.520, 15.665)	127.323 (53.239, 223.390)
Crowding (mm)	0.173 (0.004, 1.253)	40.256 (20.196, 66.327)	44.515 (11.818, 84.555)
Amount of interproximal reduction (mm)	0.670 (0.113, 2.692)	39.468 (16.395, 64.707)	0.942 (0.080, 4.211)
Treatment stage	1.451 (0.216, 5.189)	9.740 (5.509, 16.898)	28.250 (13.301, 50.561)
With/without precision cut	3.495 (0.615, 9.066)	32.480 (16.052, 51.128)	1.266 (0.084, 5.304)
Age (yr)	9.140 (1.622, 19.141)	0.904 (0.078, 3.351)	25.386 (10.694, 42.722)
With/without interproximal reduction	1.680 (0.300, 7.595)	2.0378 (0.161, 5.573)	27.993 (9.141, 69.337)
Number of teeth with optimized attachments	17.884 (5.501, 42.235)	9.216 (5.308, 17.528)	4.456 (0.872, 12.977)
With/without elastics	0.937 (0.160, 3.142)	14.724 (7.662, 23.509)	13.357 (5.600, 27.314)
Number of extractions	23.077 (7.160, 48.523)	2.976 (0.518, 8.308)	0.827 (0.084, 4.143)
With/without extraction	15.276 (5.932, 38.595)	0.204 (0.020, 1.557)	3.132 (0.441, 8.072)
Number of teeth with attachments	6.380 (1.382, 16.947)	0.544 (0.035, 2.550)	10.554 (3.485, 24.553)
With/without molar distalization	0.962 (0.355, 3.817)	0.925 (0.158, 3.657)	7.119 (1.743, 15.734)
Number of elastics	1.334 (0.198, 5.679)	0.224 (0.039, 1.006)	6.383 (1.064, 20.720)
Wearing aligners and bonding attachments simultaneously or separately	0.594 (0.063, 2.863)	2.012 (0.418, 5.371)	0.996 (0.054, 4.459)

Data are presented as the median (95% confidence interval).

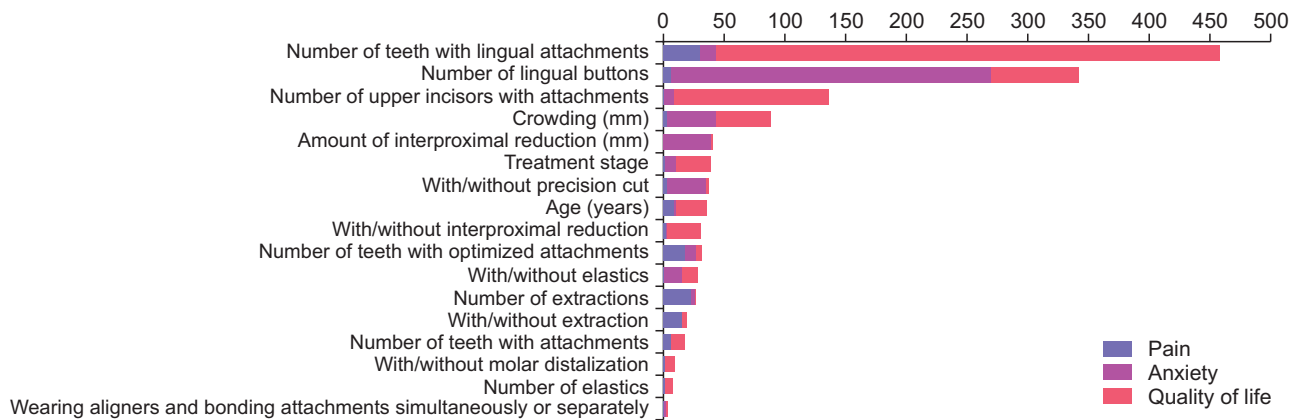


Figure 3. Total contribution of the 17 input features in descending order.

not particularly clear about the impact of appliance designs on patient comfort. Patient experience and responses to different designs of orthodontic appliances are complicated, and thus evaluating by experience alone could be inaccurate. Aligner designs have a non-linear relationship (e.g., extraction always requires elastics and precision cuts); as such, it is difficult to measure using traditional multiparameter linear models, such as correlation analysis and logistic regression. AI can be used to investigate the nonlinear relationships in a high-dimensional dataset and determine the potential role of each feature on patient comfort.

AI systems developed for predicting patient comfort has potential to enhance patient compliance. If a high risk of discomfort is predicted, orthodontists can avoid or delay the use of uncomfortable accessories that do not affect treatment outcomes. Importantly, a comprehensive explanation of possible discomfort and timely follow-up are recommended, with medical advice such as replacing aligners less frequently (2/3-week per pair), less wearing time (12-hour per day), and planning more appointments or telephone calls in the early stage of treatment.^{29,30} Advance patient preparation ensures better patient-specific outcomes in orthodontic treatment.

The ANNs achieved comparable prediction accuracies to those presented in the literature.³¹ For example, a convolutional neural network that was incorporated into a one-step, end-to-end diagnostic system to diagnose skeletal classification with lateral cephalograms achieved a prediction accuracy between 89% and 96%.³² A 23-13-1 back propagation ANN constructed to determine the need for dental extractions prior to orthodontic treatment achieved a success rate of 80% and identified two contributing indices that should be first considered.¹⁵ A neural network machine learning for the diagnosis of extraction patterns achieved an accuracy of 84%.¹⁷ With respect to the predictive performance of the ANNs for pain, anxiety, and QoL, the success rates

were 87.7%, 93.4%, and 92.4%, respectively, indicating satisfactory performance. To further improve accuracy, we could increase the training and evaluation sets and collect more diagnostic features that may be related to patient experience. A larger sample size will create a more sophisticated model structure (e.g., more layers and nodes) with better performance.

The current study found the number of teeth with lingual attachments and buttons as the most important factor affecting negative experiences. One explanation is that these lingual devices could aggravate the irritation of the mucosa and tongue during treatment. The same is true about the attachments on upper incisors, which compromise dental esthetics during treatment.³³ Age and crowding were found to influence QoL, and they have been previously reported during fixed appliance treatment.^{34,35} Patient experience at the different stages of Invisalign treatment (i.e., the initial treatment or refinement) is poorly understood. The present study found that the treatment stage was a negligible feature for predicting pain and had a minimal impact on anxiety, while it had a moderate impact on QoL. The influencing factors of patient experience found in the current study can guide orthodontists to closely monitor patients with high-ranking designs and provide earlier care.

Sex was not included as an input feature in the present study. Although the majority of patients were female, consistent with the real-world setting, we did not deliberately change the existing sex distribution in the recruitment process. In addition, there is generally no difference in QoL between male and female.³⁴ Many studies on ANNs have combined male and female patients in their analysis.^{15,17,23}

This study had some limitations. First, the dataset was relatively small for an ANN analysis, and scores fluctuated minimally around the baseline. Thus, the model may have had limited training intensity. However, the test set had consistent accuracy with that of the training and

validation sets, indicating negligible overfitting. Furthermore, substantial efforts have been made to enhance the predictive performance. Random dropout, cross-validation, and Adam optimizer reduced overfitting and increased the learning efficiency of the ANNs, thus compensating for the relatively small sample size.¹⁷ Another limitation is that the comfort level of wearing clear aligners was relatively subjective as it was self-reported and influenced by multiple factors. However, we defined the difference between the highest and lowest scores as the output, reducing the impact of subjective variability. In the future, we will consider involving objective measures, such as physical and laboratory examinations, to evaluate patient experience. The ANN models would be further stabilized with more clinical data.

CONCLUSIONS

The three constructed ANNs demonstrate good success rates in predicting pain (87.7%), anxiety (93.4%), and QoL (92.4%) during Invisalign treatment. The number of teeth with lingual attachments is the most important influencing factor of negative experiences, followed by the number of lingual buttons and the number of upper incisors with attachments. AI systems developed for predicting patient comfort has potential for clinical application to enhance patient compliance.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (NSFC) (grant number 31971247).

The repository of the source code underlying this article is available at <https://github.com/lurqTHU/clinic>.

REFERENCES

- Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. *J Dent Res* 2020;99:769-74.
- Shahmoradi L, Safdari R, Mirhosseini MM, Arji G, Jannat B, Abdar M. Predicting risk of acute appendicitis: a comparison of artificial neural network and logistic regression models predicting risk of acute appendicitis: a comparison of artificial neural network and logistic regression models. *Acta Med Iran* 2018;56:784-95.
- Kök H, Acilar AM, İzgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod* 2019;20:41.
- Wang H, Minnema J, Batenburg KJ, Forouzanfar T, Hu FJ, Wu G. Multiclass CBCT image segmentation for orthodontics with deep learning. *J Dent Res* 2021;100:943-9.
- Djeu G, Shelton C, Maganzini A. Outcome assessment of Invisalign and traditional orthodontic treatment compared with the American Board of Orthodontics objective grading system. *Am J Orthod Dentofacial Orthop* 2005;128:292-8; discussion 8.
- Cardoso PC, Espinosa DG, Mecenas P, Flores-Mir C, Normando D. Pain level between clear aligners and fixed appliances: a systematic review. *Prog Orthod* 2020;21:3.
- Rossini G, Parrini S, Castroflorio T, Deregibus A, Debernardi CL. Periodontal health during clear aligners treatment: a systematic review. *Eur J Orthod* 2015;37:539-43.
- Gao M, Yan X, Zhao R, Shan Y, Chen Y, Jian F, et al. Comparison of pain perception, anxiety, and impacts on oral health-related quality of life between patients receiving clear aligners and fixed appliances during the initial stage of orthodontic treatment. *Eur J Orthod* 2021;43:353-9.
- Flores-Mir C, Brandelli J, Pacheco-Pereira C. Patient satisfaction and quality of life status after 2 treatment modalities: Invisalign and conventional fixed appliances. *Am J Orthod Dentofacial Orthop* 2018;154:639-44.
- Nedwed V, Miethke RR. Motivation, acceptance and problems of invisalign patients. *J Orofac Orthop* 2005;66:162-73.
- Allareddy V, Nalliah R, Lee MK, Rampa S, Allareddy V. Adverse clinical events reported during Invisalign treatment: analysis of the MAUDE database. *Am J Orthod Dentofacial Orthop* 2017;152:706-10.
- Aljudaibi S, Duane B. Non-pharmacological pain relief during orthodontic treatment. *Evid Based Dent* 2018;19:48-9.
- Al-Moghrabi D, Salazar FC, Pandis N, Fleming PS. Compliance with removable orthodontic appliances and adjuncts: a systematic review and meta-analysis. *Am J Orthod Dentofacial Orthop* 2017;152:17-32.
- Charavet C, Le Gall M, Albert A, Bruwier A, Leroy S. Patient compliance and orthodontic treatment efficacy of Planas functional appliances with TheraMon microsensors. *Angle Orthod* 2019;89:117-22.
- Xie X, Wang L, Wang A. Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment. *Angle Orthod* 2010;80:262-6.
- Tanikawa C, Yamashiro T. Development of novel

- artificial intelligence systems to predict facial morphology after orthognathic surgery and orthodontic treatment in Japanese patients. *Sci Rep* 2021;11:15853.
17. Jung SK, Kim TW. New approach for the diagnosis of extractions with neural network machine learning. *Am J Orthod Dentofacial Orthop* 2016;149:127-33.
 18. He SL, Wang JH. Reliability and validity of a Chinese version of the Oral Health Impact Profile for edentulous subjects. *Qual Life Res* 2015;24:1011-6.
 19. Zhang Y, Liu R, Li G, Mao S, Yuan Y. The reliability and validity of a Chinese-version Short Health Anxiety Inventory: an investigation of university students. *Neuropsychiatr Dis Treat* 2015;11:1739-47.
 20. Katz J, Melzack R. Measurement of pain. *Surg Clin North Am* 1999;79:231-52.
 21. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv*. 1409.1556 [Preprint]. 2014 [cited 2020 Apr 1]. Available from: <https://doi.org/10.48550/arXiv.1409.1556>.
 22. Wanto A, Windarto AP, Hartama D, Parlina I. Use of binary sigmoid function and linear identity in artificial neural networks for forecasting population density. *IJISTECH* 2017;1:43-54.
 23. Li P, Kong D, Tang T, Su D, Yang P, Wang H, et al. Orthodontic treatment planning based on artificial neural networks. *Sci Rep* 2019;9:2037.
 24. Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv*. 1412.6980 [Preprint]. 2014 [cited 2020 Apr 1]. Available from: <https://doi.org/10.48550/arXiv.1412.6980>.
 25. Guyon I. A scaling law for the validation-set training-set size ratio. Murray Hill: AT & T Bell Laboratories; 1997.
 26. Sug H. The effect of training set size for the performance of neural networks of classification. *WSEAS Trans Comput* 2010;9:1297-306.
 27. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3:32-5.
 28. Gevrey M, Dimopoulos I, Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecol Model* 2003;160:249-64.
 29. Keith DJ, Rinchuse DJ, Kennedy M, Zullo T. Effect of text message follow-up on patient's self-reported level of pain and anxiety. *Angle Orthod* 2013;83:605-10.
 30. Bartlett BW, Firestone AR, Vig KW, Beck FM, Marucha PT. The influence of a structured telephone call on orthodontic pain and anxiety. *Am J Orthod Dentofacial Orthop* 2005;128:435-41.
 31. Liu JL, Li SH, Cai YM, Lan DP, Lu YF, Liao W, et al. Automated radiographic evaluation of adenoid hypertrophy based on VGG-lite. *J Dent Res* 2021;100:1337-43.
 32. Yu HJ, Cho SR, Kim MJ, Kim WH, Kim JW, Choi J. Automated skeletal classification with lateral cephalometry based on artificial intelligence. *J Dent Res* 2020;99:249-56.
 33. Thai JK, Araujo E, McCray J, Schneider PP, Kim KB. Esthetic perception of clear aligner therapy attachments using eye-tracking technology. *Am J Orthod Dentofacial Orthop* 2020;158:400-9.
 34. Masood Y, Masood M, Zainul NN, Araby NB, Husain SF, Newton T. Impact of malocclusion on oral health related quality of life in young people. *Health Qual Life Outcomes* 2013;11:25.
 35. Wang J, Tang X, Shen Y, Shang G, Fang L, Wang R, et al. The correlations between health-related quality of life changes and pain and anxiety in orthodontic patients in the initial stage of treatment. *Biomed Res Int* 2015;2015:725913.