

RESEARCH

Open Access



Development of a machine learning-based predictive model for maxillary sinus cysts and exploration of clustering patterns

Haoran Yang^{1,2,3}, Yuxiang Chen^{1,2}, Anna Zhao^{1,2}, Xianqi Rao^{1,2}, Lin Li^{1,2} and Ziliang Li^{1,2*}

Abstract

Background and objective There are still many controversies about the factors influencing maxillary sinus cysts and their clinical management. This study aims to construct a prediction model of maxillary sinus cyst and explore its clustering pattern by cone beam computerized tomography (CBCT) technique and machine learning (ML) method to provide a theoretical basis for the prevention and clinical management of maxillary sinus cyst.

Methods In this study, 6000 CBCT images of maxillary sinus from 3093 patients were evaluated to document the possible influencing factors of maxillary sinus cysts, including gender, age, odontogenic factors, and anatomical factors. First, the characteristic variables were screened by multiple statistical methods, and ML methods were applied to construct a prediction model for maxillary sinus cysts. Second, the model was interpreted based on the SHapley Additive exPlanations (SHAP) values, and the risk of maxillary sinus cysts was predicted by generating a web page calculator. Finally, the K-mean clustering algorithm further identified risk factors for maxillary sinus cysts.

Results By comparing the various metrics in the training and test sets of multiple ML models, eXtreme Gradient Boosting (XGBoost) is the best model. The average area under curve (AUC) values of the XGBoost model in the training, validation, and test sets, respectively, are 0.939, 0.923, and 0.921, which indicates its excellent classification and discrimination ability. The cluster analysis model further categorized maxillary sinus cysts into high-risk and low-risk groups, with apical lesions, severe periodontitis, and age ≥ 53 as high-risk factors for maxillary sinus cysts.

Conclusion These findings provide valuable insights into the etiology and risk stratification of maxillary sinus cysts, offering a theoretical basis for their prevention and clinical management. The integration of CBCT imaging and ML techniques holds the potential for prevention and personalized treatment strategies of maxillary sinus cysts.

Keywords Maxillary sinus cyst, Cone beam computerized tomography, Machine learning, Predictive modeling, Cluster analysis

*Correspondence:

Ziliang Li

1752114604@qq.com

¹Affiliated Stomatology Hospital of Kunming Medical University, Kunming, Yunnan, China

²Yunnan Provincial Key Laboratory of Stomatology, Kunming, Yunnan, China

³Kunming Medical University Haiyuan College, Kunming, Yunnan, China



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Maxillary sinus cysts are the most common sinus cysts, second only to maxillary sinusitis in incidence. Maxillary sinus cysts mainly include pseudocysts, retention cysts, and mucus cysts [1, 2]. Giotakis et al. found that the incidence of maxillary sinus cysts ranged from 3.6 to 35.6%, and 66% were located on the maxillary sinus floor [3]. Ren et al. analyzed 2,571 CBCT scans of 5,000 sinuses in a Chinese population and found that the incidence of maxillary sinus cysts at the level of the sinus was 15.46%, and the incidence at the patient level was 23.44% [4]. Yeung et al. evaluated 310 maxillary sinuses by CBCT and found that the incidence of maxillary sinus cysts at the level of the sinuses was 12.9% [5]. Maxillary sinus cysts are primarily asymptomatic in the early stages, and symptoms such as stuffiness and pain on pressure in the cheeks and toothache on the same side can occur when the cyst increases in size and fills the sinus cavity. If the cyst grows and compresses the nasal septum, it can lead to deviation of the nasal septum; if the cyst blocks the maxillary sinus orifice, symptoms related to sinusitis, such as nasal congestion and runny nose, can occur [6].

When oral implant surgery is performed in the maxillary posterior region, the problem of insufficient bone is often faced due to missing teeth, alveolar bone resorption, and maxillary sinus pneumatization [7]. Maxillary sinus floor elevation effectively increases the bone height in the maxillary posterior region, including external elevation with lateral wall openings and internal elevation via the top of the alveolar ridge [8]. However, performing a maxillary sinus floor lift in the presence of a maxillary sinus cyst may lead to a reduction in maxillary sinus volume, increase the risk of maxillary sinus orifice obstruction and subsequent complications, cause maxillary sinusitis, increase the risk of bone graft failure [9]. Therefore, it is essential to accurately diagnose maxillary sinus cysts and fully understand their causative factors to manage them when performing maxillary sinus floor elevation effectively.

CBCT is a widely used computed tomography imaging technique, and studies have shown that its diagnostic accuracy for maxillary sinus lesions is comparable to that of sinus endoscopy [10]. Meanwhile, some scholars have pointed out that CBCT technology has higher accuracy and sensitivity in assessing the anatomy of the maxillary sinus [11]. In diagnosing maxillary sinus lesions, CBCT can provide more accurate and detailed imaging information, which helps physicians comprehensively assess the nature and extent of the lesion and provides a reliable basis for developing diagnostic and treatment plans [12]. While CBCT imaging cannot directly diagnose maxillary sinus cysts, analyzing a large sample of CBCT data provides valuable insights into the characteristics and prevalence of maxillary sinus cysts.

There is no clear evidence regarding the aetiology of maxillary sinus cysts. Previous studies have reported anatomical and odontogenic factors associated with maxillary sinus cysts, among others, but have not reached a unified conclusion. Nascimento et al. evaluated CBCT scan images of 400 patients showing sinus disease (mucosal thickening, maxillary sinusitis, and retention cysts) in 1 or 2 maxillary sinuses and found that mucosal thickening was uniquely associated with odontogenic disease [13]. Curi et al. evaluated CBCT scan images of 4402 patients and found that odontogenic infections were associated with maxillary sinus pathology (mucosal thickening, sinus turbidity, and mucus retention cysts) and that the proximity between the palatal root and the floor of the maxillary sinus was a predisposing factor for maxillary sinus pathology [14]. In addition, the intricate nonlinear relationship between various influencing factors and maxillary sinus cysts poses a significant challenge to applying traditional linear statistical methods.

Considering the high incidence of maxillary sinus cysts, unclear influencing factors, impact on patient's quality of life, as well as trouble for physicians, there is an urgent need to identify risk factors for maxillary sinus cysts and construct a prediction model for maxillary sinus cysts using ML methods in order to improve the prediction accuracy of maxillary sinus cysts and the ability of early intervention, as well as to improve the quality of life of patients. Compared with traditional statistical methods, ML algorithms can adapt more flexibly to the nonlinear relationships of the data and do not require too many prior assumptions about the distribution and relationships of the data. ML has demonstrated the ability to process efficiently and deeply analyze large-scale, complex, and diverse clinical data in clinical research, providing new opportunities and challenges for clinical research and medical decision-making [15, 16].

This study evaluated 6000 maxillary sinus CBCT data from 3093 patients to document the possible influencing factors of maxillary sinus cysts. Characteristic variables were screened using various statistical methods, and ML methods were applied to construct a predictive model for maxillary sinus cysts. The study also predicted the risk of maxillary sinus cysts by generating a web calculator. Finally, the K-mean clustering algorithm categorized maxillary sinus cysts into high-risk and low-risk groups (Fig. 1).

Materials and methods

Selection of the study sample

A total of 6000 CBCT images of maxillary sinuses of 3093 patients who underwent CBCT for implant treatment at the stomatological hospital of Kunming Medical University from June 2016 to June 2024 were selected, of which CBCT images of bilateral maxillary sinuses were used in

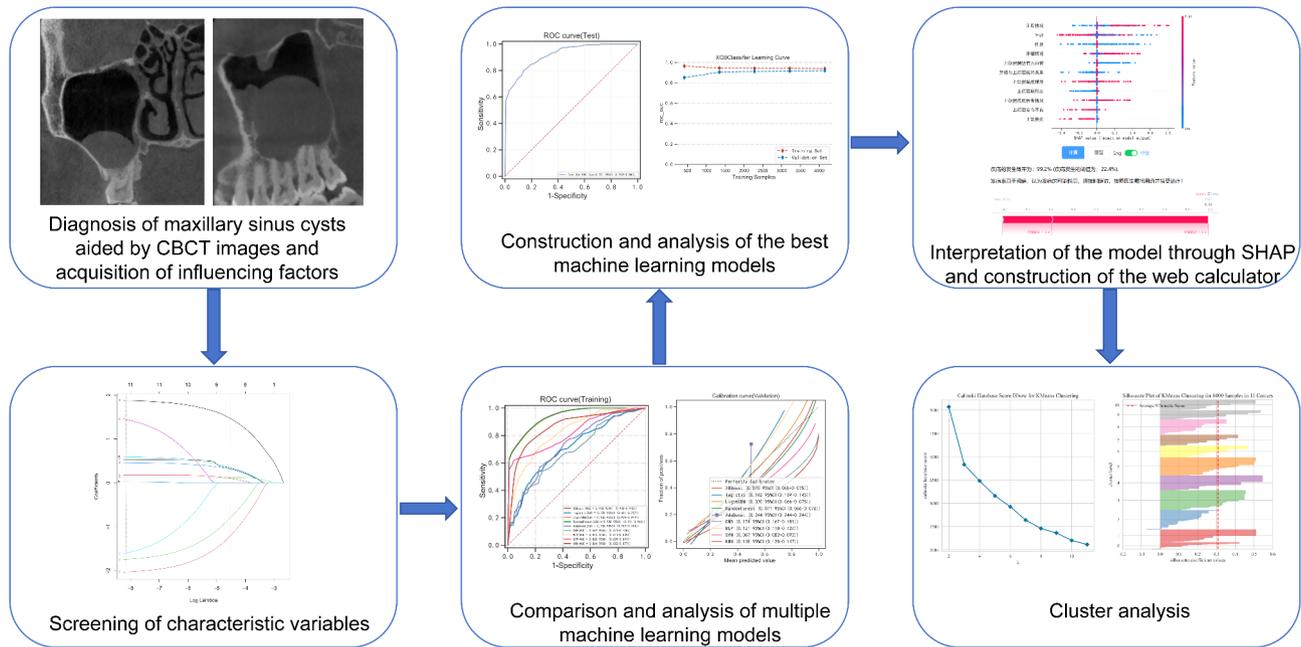


Fig. 1 Flowchart of this study

2907 patients simultaneously. CBCT images of unilateral maxillary sinuses were used in 186 patients only.

Inclusion criteria included clear CBCT images, clear visualization of one or both maxillary sinuses, and complete visualization of the maxillary first premolar, second premolar, first molar, and second molar. Exclusion criteria included patients who had undergone maxillary sinus surgery (e.g., otorhinolaryngology, oral and maxillofacial surgery) or patients with a known history of trauma to the maxillary sinus region; artifacts in the maxillary sinus region or invisible maxillary sinuses [5, 17, 18].

This study was reviewed by the medical ethics committee of the stomatological hospital of Kunming medical university under the approval number KYKQ2024MEC0042.

CBCT data acquisition and analysis

CBCT (NewTom VG, Italy), scanning parameters: voltage 110 KV, current 4 mA, exposure time 3.6 S. All the ingested films were operated by a radiologist with more than 10 years of experience in the same conditions with the same CBCT machine. The image data were opened using its own NNT software to reconstruct cross-sectional, sagittal, and coronal images, and the contrast and brightness of the images were adjusted using the image processing tools in the software to ensure optimal visualization. All images were analyzed on a Xiaomi RedmiBook Pro 15 computer (Xiaomi, Beijing, China) with a 2560×1600 pixels resolution. In the CBCT images (Fig. 2), the maxillary sinus cyst was characterized on imaging as a dome-shaped, low-density blocked image

with clear boundaries and continuous smooth edges [4, 5].

Assessment of influencing factors

Factors associated with maxillary sinus cysts, including gender, age, odontogenic factors, anatomical factors, and other factors, were documented by CBCT image analysis. The dentition between the maxillary sinuses from the first premolar to the second molar was categorized as: (a) the presence of teeth, (b) the partial absence of teeth, and (c) the complete absence of teeth. Their endodontic and periodontal conditions are assessed if teeth are present (from the first premolar to the second molar). The endodontic situation was categorized as (if more than one situation was present, the most severe situation was selected): (a) healthy teeth; (b) deep caries with or without treatment; (c) endodontic lesions or endodontic treatment without apical lesions; (d) apical lesions [4, 5, 17]. The periodontal condition was expressed in terms of the degree of bone loss. It was categorized as (if more than one condition existed, the most severe condition was selected): (a) standard to mild periodontitis (bone loss < 15%); (b) moderate periodontitis (bone loss ≥ 15%; ≤ 33%); and (c) severe periodontitis (bone loss > 33%) [19]. The relationship between the roots and the floor of the maxillary sinus was assessed: (a) the presence of a gap between the apical portion of all roots and the floor of the sinus; (b) the presence of at least one root in contact with the lower wall of the maxillary sinus [4, 5, 18]. Morphology of the maxillary sinus floor: (a) flat maxillary sinus floor; (b) uneven or separated maxillary sinus floor

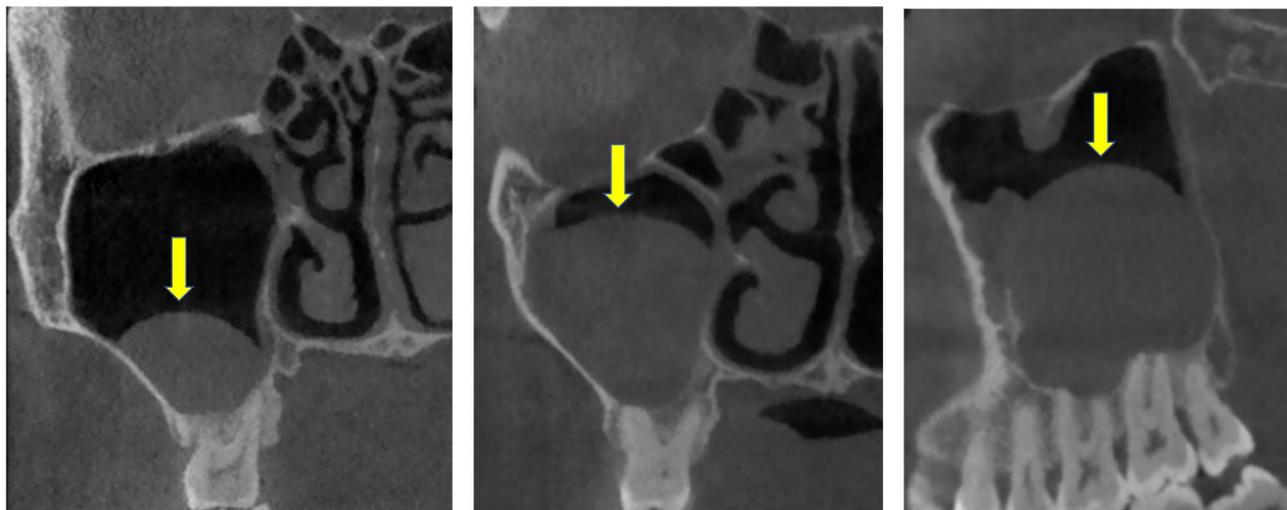


Fig. 2 CBCT image of maxillary sinus cyst (indicated by the yellow arrow)

Table 1 Assignment of variables

Characteristic variable	Assessment of influencing factors
Cyst or not	0 = no cysts, 1 = with cysts
Gender	0 = female, 1 = male
Age	0 = ≥ 18 years, < 35 years; 1 = ≥ 35 years, < 53 years; 2 = ≥ 53 years
Dental condition	0 = dentate, 1 = partially edentulous, 2 = completely edentulous
Pulp condition	0 = healthy tooth, 1 = with deep caries or treated deep caries 2 = having endodontic lesions or having undergone endodontic treatment, 3 = having apical lesions
Periodontal condition	0 = normal to mild periodontitis (< 15% bone loss) 1 = moderate periodontitis (bone loss $\geq 15\%$; $\leq 33\%$), 2 = severe periodontitis (bone loss > 33%)
Maxillary sinus different sides	0 = left side, 1 = right side
Relationship of the tooth root to the floor of the maxillary sinus	0 = no contact, 1 = at least one root in contact with sinus floor
Morphology of the floor of the maxillary sinus	0 = flat maxillary sinus floor, 1 = uneven or separated maxillary sinus floor
Vessels of the lateral wall of the maxillary sinus	0 = no blood vessels in the lateral wall of the maxillary sinus, 1 = blood vessels in the lateral wall of the maxillary sinus
Mucosal condition of the maxillary sinus	0 = no mucosal thickening (mucosal thickness less than 2 mm) 1 = with mucosal thickening (mucosal thickness greater than 2 mm)
Presence of maxillary sinusitis	0 = without maxillary sinusitis, 1 = with maxillary sinusitis
Development of the maxillary sinus	0 = normal maxillary sinus development, 1 = maxillary sinus hypoplasia
Cortical bone condition of the floor of the maxillary sinus	0 = cortical bone at the floor of the maxillary sinus is intact, 1 = cortical bone at the floor of the maxillary sinus is incomplete

[34]. Other influencing factors included gender (male and female), age (≥ 18 years, < 35 years; ≥ 35 years, < 53 years; ≥ 53 years), different sides of the maxillary sinus (left and right), whether or not blood vessels were detected in the lateral wall of the maxillary sinus in the CBCT images, condition of the mucosa of the maxillary sinus, presence or absence of maxillary sinusitis, development of the maxillary sinus, and condition of the cortical bone at the floor of the maxillary sinus. To address the concern regarding the criteria used for diagnosing maxillary sinusitis in our study, we utilized specific CBCT imaging features that are commonly associated with this condition. These criteria include the following: (a) Fluid Levels: The presence of air-fluid levels within the maxillary sinus visibly delineated on CBCT scans suggests acute or sub-acute sinusitis, typically caused by impaired sinus drainage and ventilation. (b) Opacity or Turbidity: Increased radiopacity or complete turbidity of the sinus on CBCT images often indicates sinusitis. This finding reflects the presence of fluid, inflammatory tissue, or polypoid degeneration, which are hallmarks of sinus inflammation. Detailed variable characterization and assessment of influencing factors are shown in Table 1.

Characteristic variable screening

All the influencing factors were expressed as categorical variables. The outcome indicator was a dichotomous variable (i.e., the presence of maxillary sinus cyst). Categorical variables were expressed as numbers and percentages and were screened using the chi-square test, with a two-sided P-value of less than 0.05 considered statistically significant. The chi-square test was analyzed using SPSS (version 27.0). The characteristic variables were further screened using Least Absolute Shrinkage and Selection Operator (LASSO) regression, which compresses the coefficients of the variables in the regression model

by generating a penalty function to prevent overfitting; at the same time, LASSO regression can reduce the effect of multicollinearity on the regression results and solve the problem of severe covariance by making the coefficients of correlated independent variables zero through the correlation between independent variables [20]. The LASSO regression was set up with 10-fold cross-validation and executed by the R package “glmnet4.1.2”.

Comparison and analysis of multiple ML models

In this study, 6000 maxillary sinuses were divided into a training set and a validation set in a ratio of 8:2. Nine ML classifiers were fitted to the training set: Logistic Regression, eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), RandomForest, Adaptive Boosting (AdaBoost), Gaussian Naive Bayes (GNB), Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), and K-Nearest Neighbors (KNN). XGBoost was built with Python “xgboost1.2.1”, LightGBM was built using Python “lightgbm3.2.1”, and the other machine learning algorithms were built using Python “sklearn0.22.1”. Multiple interpolation was used to fill in missing values. The Synthetic Minority Oversampling Technique (SMOTE) handled the unbalanced data such that the ratio of cysts to no cysts was 1:1. The optimal model is selected by analyzing the importance of each metric in the training and testing sets of multiple models. The receiver operating characteristic (ROC) curve was constructed using Python “sklearn0.22.1” to measure the performance of the models by comparing the AUC values [21]. The AUC value is the most crucial index for model selection and performance comparison, which ranges from 0 to 1, with higher values meaning better model performance. The R package “rmda1.6” was used to generate decision curve analysis (DCA) to help determine the clinical utility and application value of the model [22]. Python “sklearn0.22.1” was used to construct calibration curves to assess the accuracy of the model’s predictive probabilities [23]. Python “scikit0.22.1” was used to plot precision-recall (PR) curves, which were used to measure the accuracy and omission rate of the model in predicting positive case samples [24].

Construction and analysis of the best ML models

After screening the best-performing model through comprehensive comparison, 6000 maxillary sinuses were reclassified into training, validation, and test sets according to the ratio of 6:2:2. The optimal model was used for 10-fold cross-validation on the training set, and the validation and test sets were evaluated. To assess the fit and stability of the model on the training and validation sets, learning curves were generated using Python “scikit0.22.1” [25].

Interpretability of the best ML models

Explaining black-box models has become a research hotspot in ML. Some methods can extract feature importance from a black-box model or visualize some of its operational mechanisms, but fully explaining its internal decision-making process remains challenging. Lundberg et al. proposed a unified framework for explaining predictions in 2017: SHAP. The SHAP is derived from the game-theoretical Shapley value, which quantifies the contribution of each feature in a model to the final prediction of an observation. Since SHAP calculates the Shapley value, which can be sampled to estimate the contribution of each feature to the prediction, SHAP explains force diagrams that visualize attributions to features like the Shapley value as “forces,” each feature value being a force that increases or decreases the prediction. Predictions start at the baseline, where the baseline Shapley value is the average of all predictions. SHAP feature importance plots show the importance of each feature in order of magnitude. SHAP summary plots combine feature importance with feature effects, where each point is a feature and an instance of the Shapley value. The above results were visualized using Python “Shap0.39.0” [26].

Building a web calculator

The web calculator is an interactive graphical application interface constructed using the predictor variables in the predictive model. By entering the values of the corresponding predictor variables in Table 1, medical professionals can quickly assess the probability of the predictive outcomes of the predictive model better to understand the likelihood of maxillary sinus cyst development. This study develops a web calculator for the maxillary sinus cyst prediction model based on the best-performing machine learning model through the extreme intelligence analytics platform (<https://www.xsmartanalysis.com/>).

Cluster analysis

K-means clustering algorithm is a commonly used unsupervised learning algorithm that can classify patients with similar disease risks into the same group. In this study, the distance between features was calculated using the K-mean clustering algorithm and automatically aggregated to classify patients with maxillary sinus cysts [27]. The optimal number of clusters (K) was determined by comparing the bent elbow method [28, 29]. All the above statistical analyses were done using the R package (version 4.2.3) and Python (version 3.11.4). The performance of the clustering model is commonly evaluated by Silhouette Coefficient and the Calinski-Harabasz index [30, 31]. The Silhouette Coefficient is a metric used to measure the quality of clustering. It evaluates how similar an object is to its own cluster compared to other clusters. The coefficient ranges from -1 to 1, where a higher

value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters. The Calinski-Harabasz Index, also known as the Variance Ratio Criterion, assesses the overall goodness of fit for a clustering algorithm. It is defined as the ratio of the sum of between-cluster dispersion to the sum of within-cluster dispersion. A higher Calinski-Harabasz score indicates better-defined clusters.

Results

Results of characteristic variables screening

In the present study, CBCT images of 6000 maxillary sinuses were analyzed. The influencing factors included age, gender, dental condition, pulpal condition, periodontal condition, different sides of the maxillary sinus, the relationship between the roots of the teeth and the floor of the maxillary sinus, the morphology of the floor of the maxillary sinus, the vascularity of the lateral walls of the maxillary sinus, the condition of the mucous membranes of the maxillary sinus, the presence of maxillary sinusitis, development of the maxillary sinus, and the condition of

the cortical bone of the floor of the maxillary sinus. The incidence of maxillary sinus cysts at the level of the maxillary sinus was 19.767%. The results of the chi-square test analysis showed that there was a significant difference ($P < 0.05$) in gender, age, pulp condition, periodontal condition, relationship between the root of the tooth and the floor of the maxillary sinus, morphology of the floor of the maxillary sinus, vascularity of the lateral wall of the maxillary sinus, condition of the mucous membranes of the maxillary sinus, the presence of maxillary sinusitis, maxillary sinus development, and condition of the cortical bone of the floor of the maxillary sinus (Table 2). Characteristic variables were further screened by LASSO regression, in which the λ of the minor mean square error in LASSO regression was 0.000. The final characteristic variables selected were gender, age, pulpal condition, periodontal condition, relationship between the root of the tooth and the floor of the maxillary sinus, the morphology of the floor of the maxillary sinus, the vasculature of the lateral wall of the maxillary sinus, the mucosal condition of the maxillary sinus, the presence

Table 2 Results of chi-square test analysis

Variable n (%)		No cysts (n = 4814)	With cysts (n = 1186)	χ^2	P
Gender	0	2215 (46.012)	408 (34.401)	52.131	<0.001
	1	2599 (53.988)	778 (65.599)		
Age	0	1792 (37.225)	344 (29.005)	28.35	<0.001
	1	1424 (29.580)	388 (32.715)		
	2	1598 (33.195)	454 (38.280)		
Dental condition	0	3588 (74.533)	852 (71.838)	3.67	0.16
	1	1035 (21.500)	284 (23.946)		
	2	191 (3.968)	50 (4.216)		
Pulp condition	0	2948 (63.768)	617 (54.313)	116.442	<0.001
	1	591 (12.784)	100 (8.803)		
	2	284 (6.143)	62 (5.458)		
	3	800 (17.305)	357 (31.426)		
Periodontal condition	0	1065 (23.037)	201 (17.694)	57.78	<0.001
	1	2093 (45.274)	441 (38.820)		
	2	1465 (31.689)	494 (43.486)		
Maxillary sinus different sides	0	2429 (50.457)	571 (48.145)	2.035	0.154
	1	2385 (49.543)	615 (51.855)		
Relationship of the tooth root to the floor of the maxillary sinus	0	2447 (50.831)	466 (39.292)	50.724	<0.001
	1	2367 (49.169)	720 (60.708)		
Morphology of the floor of the maxillary sinus	0	2144 (44.537)	427 (36.003)	28.295	<0.001
	1	2670 (55.463)	759 (63.997)		
Vessels of the lateral wall of the maxillary sinus	0	1843 (38.284)	415 (34.992)	4.395	0.036
	1	2971 (61.716)	771 (65.008)		
Mucosal condition of the maxillary sinus	0	2746 (57.042)	571 (48.145)	30.469	<0.001
	1	2068 (42.958)	615 (51.855)		
Presence of maxillary sinusitis	0	4647 (96.531)	1176 (99.157)	22.918	<0.001
	1	167 (3.469)	10 (0.843)		
Development of the maxillary sinus	0	4529 (94.080)	1165 (98.229)	33.854	<0.001
	1	285 (5.920)	21 (1.771)		
Cortical bone condition of the floor of the maxillary sinus	0	4628 (96.136)	1021 (86.088)	174.451	<0.001
	1	186 (3.864)	165 (13.912)		

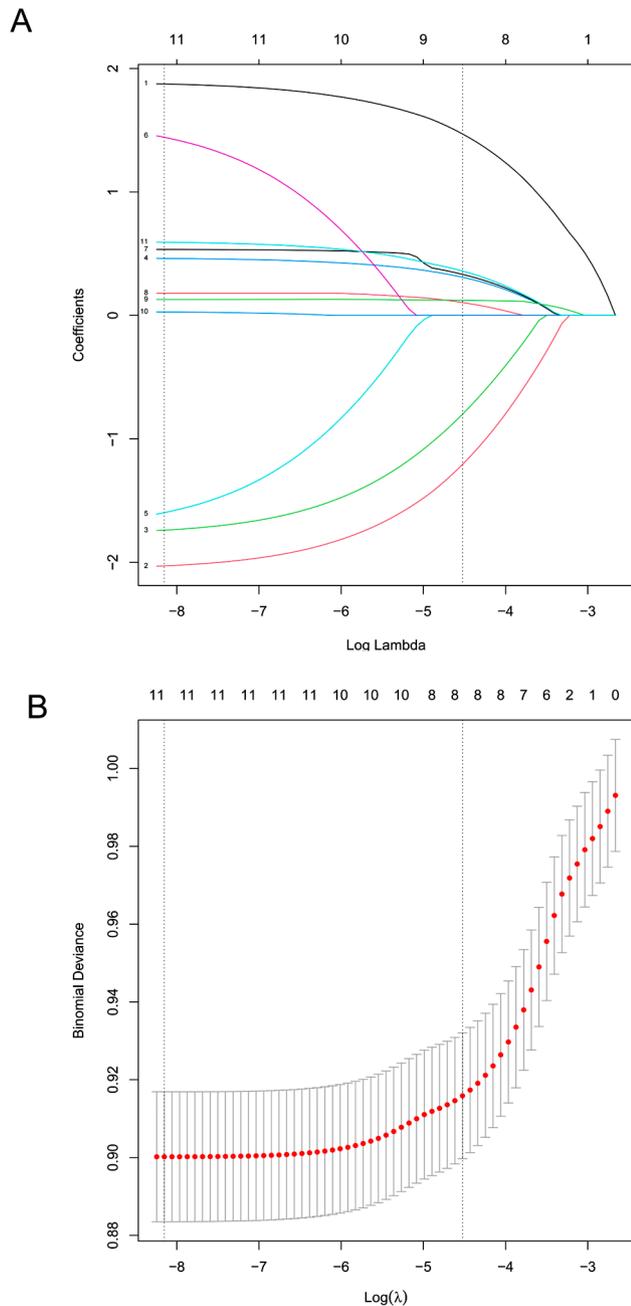


Fig. 3 LASSO regression screening for characteristic variables. **A** The upper horizontal coordinate indicates the number of non-zero coefficients in the model; the vertical coordinate indicates the values of the coefficients; the lower horizontal coordinate indicates the standardized coefficient vectors; the 11 different coloured lines represent the 11 variables, and each curve represents the trajectory of the coefficients of each independent variable. **B** Vertical coordinates indicate the error of cross-validation (the smaller the vertical axis indicates a better lasso fit), the upper horizontal coordinate indicates the number of variables corresponding to different λ , and the lower horizontal coordinate indicates the parameter corresponding to the left dashed line in the logarithmic plot of the lambda penalty coefficients (lambda. min)

of maxillary sinusitis, the development of the maxillary sinus, and the floor of the maxillary sinus cortical bone condition (Fig. 3).

Multi-model synthesis and analysis results

6000 maxillary sinus samples were classified using ML models, including Logistic Regression, XGBoost, LightGBM, RandomForest, AdaBoost, GNB, SVM, MLP, and KNN. The ROC curves of all models were cross-validated by a factor of 10. The predictive performance of the models was evaluated based on the AUC values, and XGBoost, LightGBM, and RandomForest had the best performance in the training set; XGBoost and LightGBM had the best performance in the validation set (Fig. 4A-B). The AUC values mainly assess the accuracy of the model’s predictions. However, it needs to directly provide information about the usability of the model in real clinical applications. Therefore, we analyzed the PR curves, calibration curves, and DCA. The PR curves showed that the XGBoost model produced the highest average precision (AP) values in both the training and validation sets (Fig. 4C-D). The calibration curves showed that LightGBM and XGBoost models predicted more accurately (Fig. 4E). The DCA showed that XGBoost had good clinical applicability (Fig. 4F). The comprehensive analysis showed that XGBoost was the optimal model.

Optimal model construction and analysis results

The XGBoost model was constructed and validated by 10-fold cross-validation. The results show that the average AUC value of the training set is 0.939 (0.930–0.947), the average AUC value of the validation set is 0.923 (0.893–0.952), and the average AUC value of the test set is 0.921 (0.902–0.940) (Fig. 5A-C). The average AUC value of the training, validation, and test sets is about 0.928, indicating excellent performance of the XGBoost model. Since the AUC value of the validation set does not exceed 10% of the test set, the fit can be considered successful, and the XGBoost model can be applied to the classification of maxillary sinus cysts. The calibration curves show that the XGBoost model is accurately predictive (Fig. 5D). The DCA results show that the XGBoost model has good clinical applicability (Fig. 5E). The learning curve demonstrated the fit and stability of the XGBoost model between the training and validation sets (Fig. 5F). In summary, XGBoost is suitable for constructing a prediction model for maxillary sinus cysts.

Interpretability of the model

The SHAP summary plot demonstrates the nonlinear relationship between each of the influencing factors and maxillary sinus cysts, where periodontal condition and pulpal condition are positively correlated with the occurrence of maxillary sinus cysts, i.e., the more severe the

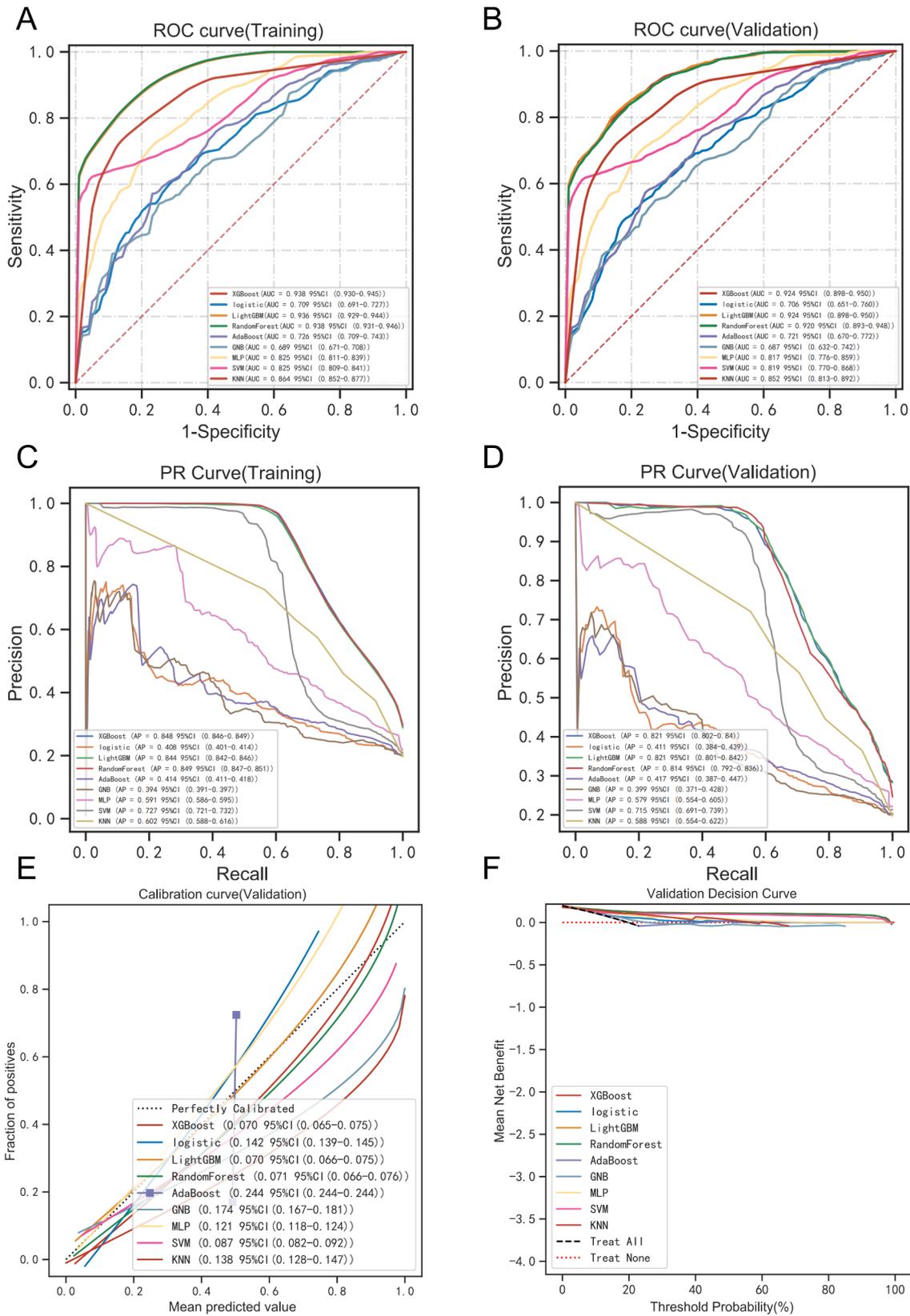


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Comparison and analysis of multiple machine learning models. **A-B** The ROC curve is obtained by graphing the true and false positive rates, which can be used to reflect the relationship between sensitivity and specificity. The AUC value indicates the area under the ROC curve enclosed with the coordinate axis. The larger the AUC value is, the better the effect of the machine learning model is, i.e., the higher the discriminative ability is. **C-D** The PR curve takes the recall rate as the horizontal coordinate and the precision rate as the vertical coordinate, and the area under the PR curve is used to evaluate the model differentiation ability. **E** In the calibration curve, the horizontal coordinate indicates the probability predicted by the model, and the vertical axis indicates the probability of the actual event occurring. If the model is perfectly calibrated, the curve will coincide with the diagonal. **F** The horizontal coordinate of the decision curve is the threshold probability, and the vertical coordinate is the average net return value. For a given threshold probability, a decision curve for a predictive model above the reference line indicates that the model has superior clinical utility

periodontal and pulpal conditions are, the higher the incidence of maxillary sinus cysts (Fig. 6A). The order of importance of the features showed that periodontal condition, age, gender, and endodontic condition characterized the XGBoost model, with the periodontal condition being the most crucial feature, changing the predicted absolute probability of maxillary sinus cysts by about 13% points on average (Fig. 6B). We further explained the model with two samples. The first of these samples had a predicted probability of 0.988, with age, maxillary sinus mucosal condition, and periodontal condition contributing to the results (Fig. 6C). The second sample had a predictive probability of 0.995, with vascularization of the lateral wall of the maxillary sinus contributing to the results and the presence of maxillary sinusitis inhibiting the results (Fig. 6D).

Online predictive modeling

This study successfully developed a web-based web calculator (<http://www.xsmartanalysis.com/model/list/predict/model/html?mid=13479&symbol=7171144uq38LqPnMx556>) for clinical physicians to use the XGBoost model to predict the probability of maxillary sinus cyst occurrence.

Results of cluster analysis

The total number of valid samples included in the cluster analysis was 6000 maxillary sinuses. The clustering model categorized all samples into 2 data cluster categories (Fig. 7). The clustering model's silhouette coefficient and calinski harabasz index were 0.422 and 5068.251, respectively, indicating excellent model performance (Table 3). Cluster 1 ($n=4123$, 68.7%) was dominated by healthy teeth (86.563%), mild (29.978%) and moderate (57.167%) periodontitis, and age < 35 years (51.686%), and it had a low incidence of maxillary sinus cysts (16.881%). Cluster 2 ($n=1877$, 31.3%) was dominated by apical lesions (74.108%), severe periodontitis (88.492%), and age \geq 53 years (81.779%), which had a higher incidence of maxillary sinus cysts (26.105%). Between the two subtypes of cluster 1 and cluster 2, there were significant differences in the incidence of maxillary sinus cysts between gender, age, pulpal condition, periodontal condition, relationship of the root of the tooth to the floor of the maxillary sinus, morphology of the floor of the maxillary sinus, vascularity of the lateral wall of the maxillary sinus, mucosal

condition of the maxillary sinus, and the presence of maxillary sinusitis ($P < 0.001$) (Table 4).

Discussion

This study evaluated 6,000 maxillary sinus CBCT data sets from 3093 patients and documented the possible influencing factors of maxillary sinus cysts, including gender, age, odontogenic factors, and anatomical factors. Characteristic variables were screened using various statistical methods, a prediction model for maxillary sinus cysts was constructed by applying various machine learning methods, and high-risk factors were identified by cluster analysis. In this study, with the help of large-sample CBCT image data and advanced ML techniques, we deeply analyzed the characteristic variables of maxillary sinus cysts. We constructed a prediction model, thus providing new ideas and tools for preventing and personalized treatment of maxillary sinus cysts.

In this study, gender, age, pulpal condition, periodontal condition, relationship between the root of the tooth and the floor of the maxillary sinus, the morphology of the floor of the maxillary sinus, the vasculature of the lateral wall of the maxillary sinus, the mucosal condition of the maxillary sinus, the presence of maxillary sinusitis, the development of the maxillary sinus, and the floor of the maxillary sinus Cortical bone condition were found to be characteristic variables for maxillary sinus cysts. The incidence of maxillary sinus cysts is significantly higher in males than in females, which may be related to the higher prevalence of smoking, apical lesions and periodontitis in males [32, 33]. In addition, there may be differences in hormone levels and lifestyle habits between genders, which may indirectly affect the incidence of cysts. The age factor should also not be ignored, as older patients are more likely to develop periodontitis and apical lesions [34]. The anatomical structures associated with the maxillary sinus may change with age, which alters the environment within the maxillary sinus and increases the incidence of maxillary sinus cysts [35]. In terms of odontogenic factors, apical lesions and severe periodontitis are associated with a higher prevalence of maxillary sinus cysts. When a tooth root is in contact with the sinus floor or extends into the maxillary sinus, it increases the likelihood of mucosal injury and inflammatory infection of the sinus floor [4, 14]. Apical lesions are often accompanied by chronic inflammation and bacterial infection,

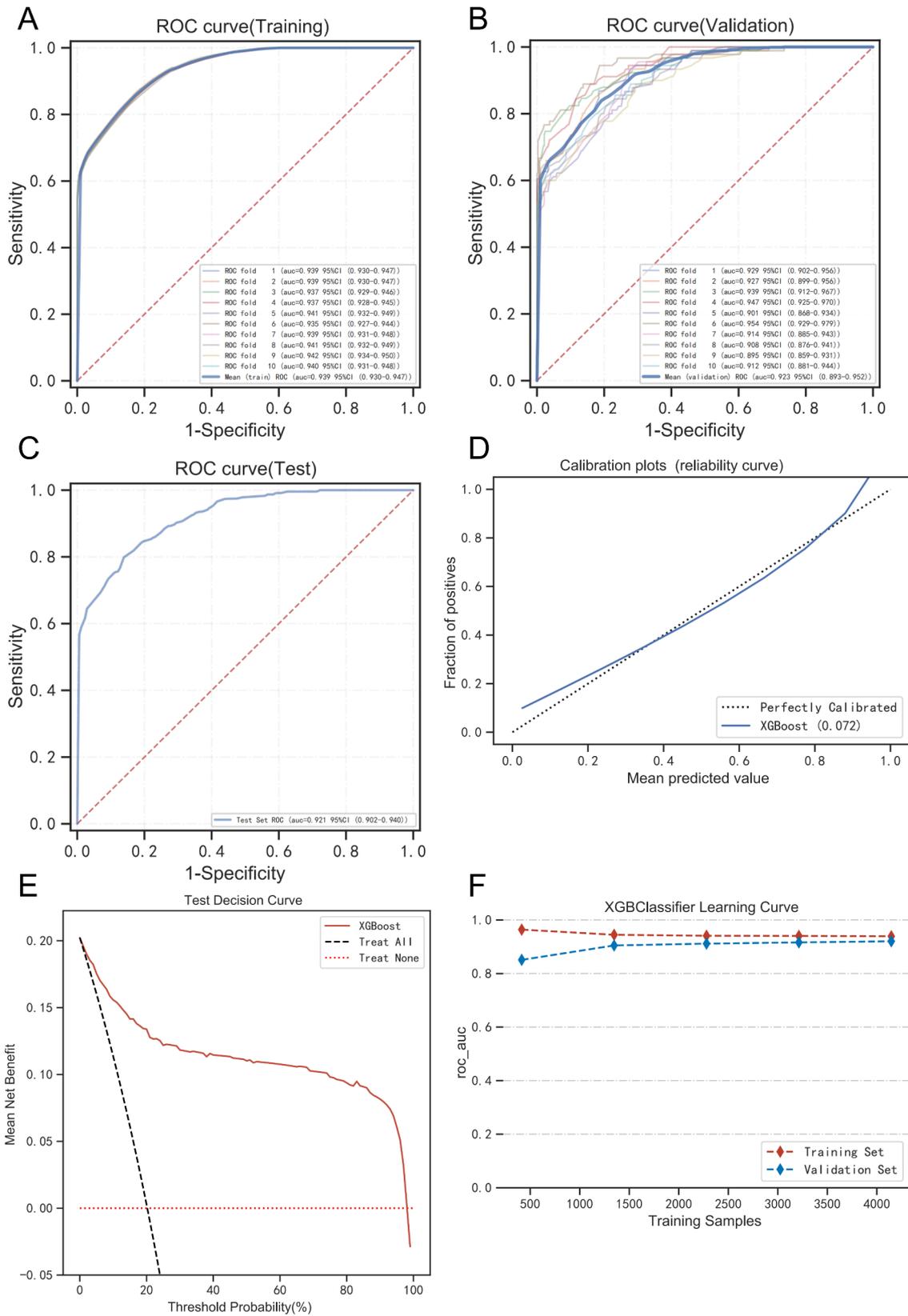


Fig. 5 Construction and analysis of the XGBoost model

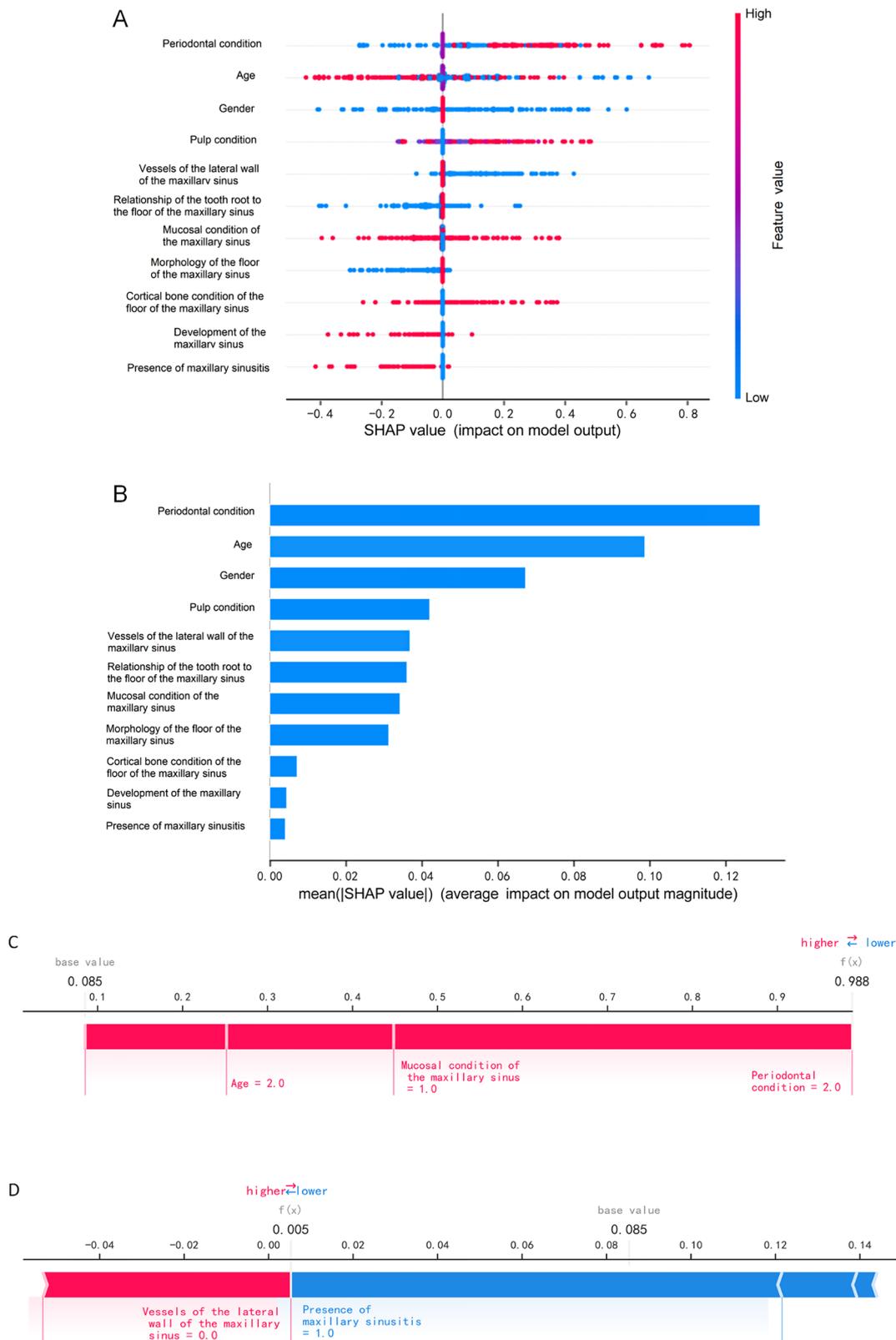
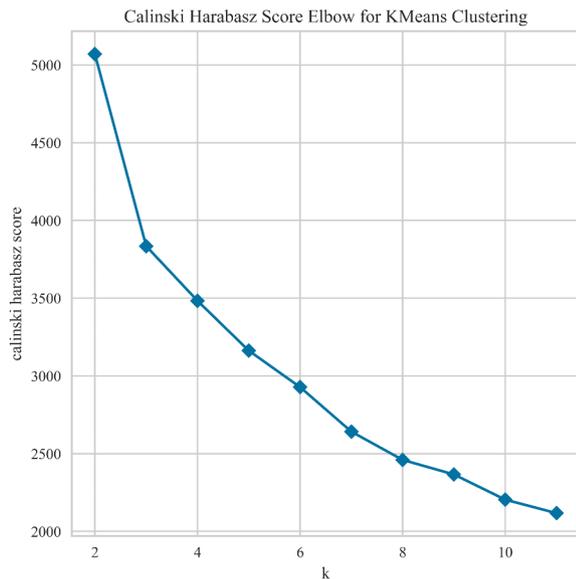


Fig. 6 Interpretability of the XGBoost model. **A** SHAP summary plot. **B** SHAP feature importance plot. **C** SHAP values explain the probability of occurrence of the first sample maxillary sinus cyst. **D** SHAP values explain the probability of occurrence of the second sample maxillary sinus cyst

A



B

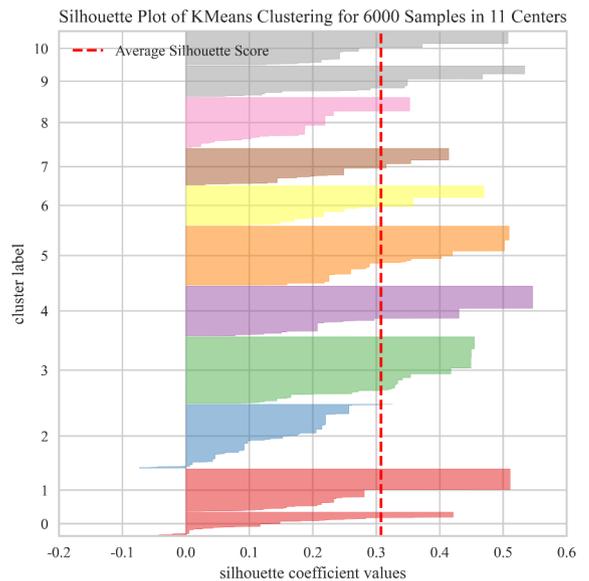


Fig. 7 Plot of clustering analysis results. (a) Elbow diagram of clusters. The horizontal axis represents the number of clusters K. The vertical axis represents the evaluation index of cluster quality. (b) Silhouette Plot. The vertical axis represents the number of clusters K, and the horizontal axis represents the silhouette coefficient value

Table 3 Summary of cluster analysis performance

Silhouette Coefficient	Calinski Harabasz Score
0.422	5068.251

and their pathogens and toxins can enter the floor of the maxillary sinus through the apical foramen, causing acute and chronic inflammation of the mucosa of the maxillary sinus. The deeper the root penetrates the floor of the maxillary sinus, the more pronounced the physical stimulation of the mucosa of the maxillary sinus, and the long-term stimulation can promote cystogenesis. Severe alveolar bone resorption can form periodontal pockets, in which pathogenic bacteria and their products may spread along the roots toward the maxillary sinus, inducing inflammation and even cysts in the maxillary sinus mucosa [36]. A septum at the maxillary sinus floor or an uneven floor of the maxillary sinus may lead to poor drainage of the maxillary sinus, allowing secretions to accumulate and bacteria to thrive, increasing the likelihood of cyst formation [37]. Abnormal distribution of blood vessels in the lateral wall of the maxillary sinus may underlie the microcirculation of inflammation and cyst formation, and a high density of distribution of blood vessels means that more inflammatory cells and factors rush to the mucosa of the maxillary sinus inducing cysts [38]. Chronic inflammatory stimulation often leads to thickening of the maxillary sinus mucosa, and the thickened mucosa is more likely to lead to poor ventilation and drainage, allowing pathogenic bacteria to grow and fueling cyst formation. Although mucosal thickening

can lead to cyst formation by obstructing sinus drainage, the maxillary sinus cyst identified on CBCT imaging has already progressed beyond the stage of simple mucosal thickening. The absence of mucosal thickening in the presence of a cyst does not imply a better clinical outcome; on the contrary, it may indicate a more severe pathology. Mucosal thickening could be masked by the cyst, making its identification challenging. In maxillary sinus hypoplasia, a relatively thick bony barrier between the maxillary sinus and tooth root is usually observed, making it less susceptible to odontogenic factors [39]. The results of this study revealed a negative correlation between maxillary sinusitis and maxillary sinus cysts, indicating that patients with sinusitis are less likely to develop cysts. The inflammation associated with maxillary sinusitis often leads to mucosal thickening and edema, which can obscure or imitate cystic structures on CBCT images, thus reducing the accuracy of cyst identification. This diagnostic challenge may lead to an underrepresentation of cysts in individuals with active sinusitis when relying solely on imaging. Additionally, the chronic inflammation and altered mucus flow seen in sinusitis may result in a distinct pathogenesis that decreases the likelihood of cyst formation due to the continuous clearance and drainage of cystic material. Together, these factors suggest that sinusitis might inversely correlate with visible cysts - not necessarily signifying an actual absence of cysts, but rather reflecting limitations in detection under inflammatory conditions. The cortical bone in the wall of the maxillary sinus floor may act as a barrier to

Table 4 Results of cluster analysis

Variable n (%)	Cluster 1 (n=4123)	Cluster 2 (n=1877)	P
Gender	0 1916 (46.471)	707 (37.666)	<0.001
	1 2207 (53.529)	1170 (62.334)	
Age	0 2131 (51.686)	5 (0.266)	<0.001
	1 1475 (35.775)	337 (17.954)	
	2 517 (12.539)	1535 (81.779)	
Pulp condition	0 3569 (86.563)	0 (0.000)	<0.001
	1 531 (12.879)	163 (8.684)	
	2 23 (0.558)	323 (17.208)	
	3 0 (0.000)	1391 (74.108)	
Periodontal condition	0 1236 (29.978)	31 (1.652)	<0.001
	1 2357 (57.167)	185 (9.856)	
	2 530 (12.855)	1661 (88.492)	
Relationship of the tooth root to the floor of the maxillary sinus	0 1850 (44.870)	1063 (56.633)	<0.001
	1 2273 (55.130)	814 (43.367)	
Morphology of the floor of the maxillary sinus	0 1637 (39.704)	934 (49.760)	<0.001
	1 2486 (60.296)	943 (50.240)	
Vessels of the lateral wall of the maxillary sinus	0 1442 (34.975)	816 (43.474)	<0.001
	1 2681 (65.025)	1061 (56.526)	
Mucosal condition of the maxillary sinus	0 2398 (58.162)	919 (48.961)	<0.001
Presence of maxillary sinusitis	0 3989 (96.750)	1834 (97.709)	0.042
	1 134 (3.250)	43 (2.291)	
Development of the maxillary sinus	0 3900 (94.591)	1794 (95.578)	0.107
	1 223 (5.409)	83 (4.422)	
Cortical bone condition of the floor of the maxillary sinus	0 3895 (94.470)	1754 (93.447)	0.117
	1 228 (5.530)	123 (6.553)	
Cyst or not	0 3427 (83.119)	1387 (73.895)	<0.001
	1 696 (16.881)	490 (26.105)	

prevent the direct spread of odontogenic infection to the maxillary sinus. However, once the cortical bone of the floor wall of the maxillary sinus is defective, the infection is more likely to penetrate the maxillary sinus. The individual influencing factors analyzed above do not act independently; they may interact and work together in maxillary sinus cyst formation. In clinical practice, awareness and attention to these influencing factors should be increased, especially the treatment and intervention of odontogenic factors, to reduce the risk of maxillary sinus cysts.

XGBoost model is an integrated learning algorithm based on Gradient Boosting Decision Tree (GBDT), which reduces the risk of overfitting by controlling the complexity and regularization terms of the tree [40].

Meanwhile, the XGBoost model has a strong ability to handle large datasets and is robust in dealing with missing values and various types of data in the data [41]. As an advanced explanatory tool, SHAP analysis can help us to deeply understand the reasons behind the model's prediction results, identify the important features, and provide intuitive explanations, enhancing the interpretability and transparency of the model [42]. Based on SHAP's summary and force diagram methods, the XGBoost model can provide global and local explanations of influencing factors, helping to reveal the key factors affecting the occurrence of maxillary sinus cysts. In machine learning, model interpretability is crucial for improving trust, understanding decision-making, identifying potential problems, and driving model application and deployment. The maxillary sinus cyst prediction model of the web calculator constructed by the XGBoost model can help clinicians assess the risk of potential patients, take intervention and preventive measures in advance to reduce the incidence of maxillary sinus cysts and provide strong support for clinical decision-making [43].

K-mean clustering algorithm is a commonly used unsupervised learning algorithm, which divides the data points into K different clusters by iterative way, and realizes clustering according to the similarity between the data points. When using the K-mean clustering algorithm, it is necessary to weigh its advantages and disadvantages according to the specific application scenarios and the data characteristics and choose the appropriate parameters and initialization methods to obtain a better clustering effect [44]. In this study, 6000 maxillary sinuses were divided into two data cluster categories by the K-mean clustering algorithm, in which all feature variables were used as categorical variables, and data cleansing and preprocessing were performed on each categorical variable, including missing value processing, outlier processing, and data standardization to ensure the integrity and consistency of the data. The K-mean clustering model identified apical lesions, severe periodontitis, and age ≥ 53 as high-risk factors for maxillary sinus cysts. Apical lesions and severe periodontitis are crucial in cystogenesis [45]. Some studies have shown that odontogenic factors such as periodontal and apical lesions may affect the maxillary sinus through bacterial infection or inflammatory response and become primary triggers for cyst development [46]. Bacterial infections often accompany apical lesions, and these bacteria and their toxins may affect the healthy state of the mucosa of the maxillary sinus using blood circulation or direct invasion, which in turn conditionally triggers cyst formation. Severe periodontitis may not only lead to resorption of the alveolar bone and change the relationship between the root of the tooth and the floor of the maxillary sinus, but it may also induce an inflammatory response in the

mucosa of the maxillary sinus, exacerbating the pathological changes in the maxillary sinus [47]. According to cluster analysis results, the incidence of maxillary sinus cysts was significantly higher in patients with apical lesions and severe periodontitis than in patients with healthy teeth and only mild to moderate periodontitis. These findings are a good indication that the maxillary sinus health of patients with apical lesions and severe periodontitis should be given adequate attention and necessary examination to prevent the occurrence of maxillary sinus cysts in clinical practice.

However, there are some limitations to this study. First, because maxillary sinus retention cysts and pseudocysts cannot be differentiated on imaging, the maxillary sinus cysts detected in this study may contain either or both. Second, the definitive diagnosis of maxillary sinus cysts relies on pathologic examination, and CBCT may carry some degree of misdiagnosis risk when diagnosing maxillary sinus cysts. Finally, the pathogenesis and factors affecting maxillary sinus cysts are complex, diverse, and influenced by genetic, environmental, and lifestyle factors. Therefore, future studies should combine pathological examinations and molecular biology techniques to distinguish different types of maxillary sinus cysts and improve the accuracy and reliability of CBCT in diagnosing maxillary sinus cysts. In addition, the pathogenesis of maxillary sinus cysts should be studied in depth. Systematic studies should be conducted by combining the effects of genetics, biology, environmental factors and other aspects in order to comprehensively understand the developmental process of maxillary sinus cysts and the key factors in the diagnosis and treatment of maxillary sinus cysts, to provide more effective guidance and decision-making support for clinical practice.

Conclusion

The prediction model of maxillary sinus cysts constructed based on the XGBoost model can predict the incidence of maxillary sinus cysts and provide a theoretical basis for preventing maxillary sinus cysts and clinical decision-making. The cluster analysis model further identifies the high-risk factors of maxillary sinus cysts and guides the personalized treatment of maxillary sinus cysts.

Author contributions

YH and LZ conceived and designed this study. YH and CY were responsible for data acquisition, analysis, and interpretation. YH, ZA, and CY participated in writing the manuscript. YH, LL, and RX helped revise the manuscript. All the authors have read and approved the final manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (82360185).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 20 September 2024 / Accepted: 22 February 2025

Published online: 12 March 2025

References

- Gardner DG. Pseudocysts and retention cysts of the maxillary sinus. *Oral Surg Oral Med Oral Pathol.* 1984;58(5):561-7.
- Anitua E, Alkhraisat MH, Torre A, Eguia A. Are mucous retention cysts and pseudocysts in the maxillary sinus a risk factor for dental implants? A systematic review. *Med Oral Patologia Oral Y Cir Bucal.* 2021;26(3):e276-83.
- Giotakis EI, Weber RK. Cysts of the maxillary sinus: a literature review. *Int Forum Allergy Rhinology.* 2013;3(9):766-71.
- Ren L, Chen C, Li N, Hu J, Jiang Z, Yang G. Prevalence of and factors associated with maxillary sinus cyst in a Chinese population. *J Oral Sci.* 2022;64(1):22-7.
- Yeung AWK, Tanaka R, Khong PL, et al. Frequency, location, and association with dental pathology of mucous retention cysts in the maxillary sinus. A radiographic study using cone beam computed tomography (CBCT). *Clin Oral Invest.* 2018;22(3):1175-83.
- Anitua E, Alkhraisat M-H, Torre A, Eguia A. Are mucous retention cysts and pseudocysts in the maxillary sinus a risk factor for dental implants? A systematic review. *Med Oral Patol Oral Cir Bucal.* 2021;26:e276-83.
- Tassoker M. What are the risk factors for maxillary sinus pathologies? A CBCT study. *Oral Radiol.* 2020;36:80-4.
- Kim K, Lim CY, Shin J, Chung MJ, Jung YG. Enhanced artificial intelligence-based diagnosis using CBCT with internal denoising: clinical validation for discrimination of fungal ball, sinusitis, and normal cases in the maxillary sinus. *Comput Methods Programs Biomed.* 2023;240:107708.
- Whyte A, Boeddinghaus R. The maxillary sinus: physiology, development and imaging anatomy. *Dentomaxillofac Radiol.* 2019;48:20190205.
- Zojaji R, Naghibzadeh M, Mazloum Farsi Baf M, Nekooei S, Bataghva B, Noorbakhsh S. Diagnostic accuracy of cone-beam computed tomography in the evaluation of chronic rhinosinusitis. *ORL J Otorhinolaryngol Relat Spec.* 2015;77(1):55-60.
- Shahidi S, Zamiri B, Momeni Danaei S, et al. Evaluation of Anatomic Variations in Maxillary Sinus with the Aid of Cone Beam Computed Tomography (CBCT) in a Population in South of Iran. *J Dent (Shiraz).* 2016;17(1):7-15.
- Shahbazian M, Vandewoude C, Wyatt J, Jacobs R. Comparative assessment of panoramic radiography and CBCT imaging for radiodiagnostics in the posterior maxilla. *Clin Oral Invest.* 2014;18(1):293-300.
- Nascimento EH, Pontual ML, Pontual AA, Freitas DQ, Perez DE, Ramos-Perez FM. Association between odontogenic conditions and maxillary sinus disease: A study using Cone-beam computed tomography. *J Endod.* 2016;42(10):1509-15.
- Curi FR, Pelegrine RA, Nascimento MDCC, Monteiro JCC, Junqueira JLC, Panzarella FK. Odontogenic infection as a predisposing factor for pathologic disorder development in maxillary sinus. *Oral Dis.* 2020;26(8):1727-35.
- Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med.* 2023;388(13):1201-8.
- Sammur SJ, Crispin-Ortuzar M, Chin SF, Provenzano E, Bardwell HA, Ma W, Cope W, Dariush A, Dawson SJ, Abraham JE, Dunn J, Hiller L, Thomas J, Cameron DA, Bartlett JMS, Hayward L, Pharoah PD, Markowitz F, Rueda OM, Earl HM, Caldas C. Multi-omic machine learning predictor of breast cancer therapy response. *Nature.* 2022;601(7894):623-9.
- Alghofaily M, Alsufyani N, Althumairy RI, AlSuhailani A, Alfawzan F, AlSadhan L. Odontogenic factors associated with maxillary sinus Schneiderian membrane thickness and their relationship to chronic sinonasal symptoms: an ambispective cohort study. *Diagnostics (Basel Switzerland).* 2023;13(16):2710.
- Aksoy U, Orhan K. Association between odontogenic conditions and maxillary sinus mucosal thickening: a retrospective CBCT study. *Clin Oral Invest.* 2019;23(1):123-31.
- 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(Suppl 20):S149-61.

20. Guler H, Guler EO. Mixed LASSO estimator for stochastic restricted regression models. *J Applied Statistics*. 2021;48(13–15):2795–808.
21. Obuchowski NA, Bullen JA. Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine. *Phys Med Biol*. 2018;63(7):07TR01.
22. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak*. 2006;26(6):565–74.
23. Fenlon C, O'Grady L, Doherty ML, Dunnion J. A discussion of calibration techniques for evaluating binary and categorical predictive models. *Prev Vet Med*. 2018;149:107–14.
24. Li W, Guo Q. Plotting receiver operating characteristic and precision-recall curves from presence and background data. *Ecol Evol*. 2021;11(15):10192–206.
25. Belkin M, Hsu D, Ma S, Mandal S. Reconciling modern machine-learning prediction and the classical bias-variance trade-off. *Proc Natl Acad Sci USA*. 2019;116(32):15849–54.
26. Yi F, Yang H, Chen D, Qin Y, Han H, Cui J, Bai W, Ma Y, Zhang R, Yu H. XGBoost-SHAP-based interpretable diagnostic framework for Alzheimer's disease. *BMC Med Inf Decis Mak*. 2023;23(1):137.
27. Seymour CW, Kennedy JN, Wang S, Chang CH, Elliott CF, Xu Z, Berry S, Clermont G, Cooper G, Gomez H, Huang DT, Kellum JA, Mi Q, Opal SM, Talisa V, Van Der Poll T, Visweswaran S, Vodovotz Y, Weiss JC, Yealy DM, Yende S, Angus DC. Derivation, validation, and potential treatment implications of novel clinical phenotypes for Sepsis. *JAMA*. 2019;321(20):2003–17.
28. Sendi MSE, Salat DH, Miller RL, Calhoun VD. Two-step clustering-based pipeline for big dynamic functional network connectivity data. *Front NeuroSci*. 2022;16:895637.
29. Petegrosso R, Li Z, Kuang R. Machine learning and statistical methods for clustering single-cell RNA-sequencing data. *Brief Bioinform*. 2020;21(4):1209–23.
30. Borzooei S, Miranda GHB, Abolfathi S, Scibilia G, Meucci L, Zanetti MC. Application of unsupervised learning and process simulation for energy optimization of a WWTP under various weather conditions. *Water Sci Technol*. 2020;81(8):1541–51.
31. Ikotun AM, Ezugwu AE. Boosting k-mean.s clustering with symbiotic organisms search for automatic clustering problems. *PLoS ONE*. 2022;17(8):e0272861.
32. Darby I. Risk factors for periodontitis & peri-implantitis. *Periodontol* 2000. 2022;90:9–12.
33. Gaeta C, Malvicini G, Di Lascio D, Martignoni M, Ragucci G, Grandini S et al. Lifestyle, caries, and apical periodontitis: results from a university-based cross-sectional study. *Int Endod J*. 2024.
34. Dos Santos VC, Kublitski PMO, Da Silva BM, Gabardo MCL, Tomazinho FSF. Periapical lesions associated with demographic variables, dental conditions, systemic diseases, and habits. *J Contemp Dent Pract*. 2023;24(11):864–70.
35. Whyte A, Boeddinghaus R. The maxillary sinus: physiology, development and imaging anatomy. *Dento Maxillo Fac Radiol*. 2019;48(8):20190205.
36. Kim SM. Definition and management of odontogenic maxillary sinusitis. *Maxillofacial Plast Reconstr Surg*. 2019;41(1):13.
37. Kim S, Ward LA, Butaric LN, Maddux SD. Ancestry-based variation in maxillary sinus anatomy: Implications for health disparities in sinonasal disease. *Anat Rec (Hoboken)*. 2022;305(1):18–36.
38. Sheikhi M, Pozve NJ, Khorrami L. Using cone beam computed tomography to detect the relationship between the periodontal bone loss and mucosal thickening of the maxillary sinus. *Dent Res J*. 2014;11(4):495–501.
39. Nunes CA, Guedes OA, Alencar AH, Peters OA, Estrela CR, Estrela C. Evaluation of periapical lesions and their association with maxillary sinus abnormalities on Cone-beam computed tomographic images. *J Endod*. 2016;42(1):42–6.
40. Lin X, Chen L, Zhang D, Luo S, Sheng Y, Liu X, Liu Q, Li J, Shi B, Peng G, Zhong X, Huang Y, Li D, Qin G, Yin Z, Xu J, Meng C, Liu Y. Prediction of surgical approach in mitral valve disease by XGBoost algorithm based on echocardiographic features. *J Clin Med*. 2023;12(3):1193.
41. Li S, Dou R, Song X, Lui KY, Xu J, Guo Z, Hu X, Guan X, Cai C. Developing an interpretable machine learning model to predict in-Hospital mortality in Sepsis patients: A retrospective Temporal validation study. *J Clin Med*. 2023;12(3):915.
42. Liu X, Hu P, Yeung W, Zhang Z, Ho V, Liu C, Dumontier C, Thorat PJ, Mao Z, Cao D, Mark RG, Zhang Z, Feng M, Li D, Celi LA. Illness severity assessment of older adults in critical illness using machine learning (ELDER-ICU): an international multicentre study with subgroup bias evaluation. *Lancet Digit Health*. 2023;5(10):e657–67.
43. Al-Zaiti SS, Martin-Gill C, Zègre-Hemsey JK, et al. Machine learning for ECG diagnosis and risk stratification of occlusion myocardial infarction. *Nat Med*. 2023;29(7):1804–13.
44. Chen Y, Han H, Meng X, Jin H, Gao D, Ma L, Li R, Li Z, Yan D, Zhang H, Yuan K, Wang K, Zhang Y, Zhao Y, Jin W, Li R, Lin F, Chao X, Lin Z, Hao Q, Wang H, Ye X, Kang S, Li Y, Sun S, Liu A, Wang S, Zhao Y, Chen X. Development and validation of a scoring system for hemorrhage risk in brain arteriovenous malformations. *JAMA Netw Open*. 2023;6(3):e231070.
45. Galiè M, Guelli S, Ciorba A, Bianchini C, Iannella G, Stomeo F, Valente L, Pelucchi S. Unilateral sinus disease: not just Odontogenic! - A retrospective study. *Annals Maxillofacial Surg*. 2020;10(2):397–401.
46. Peñarrocha-Oltra S, Soto-Peñaloza D, Bagán-Debón L, Bagan JV, Peñarrocha-Oltra D. Association between maxillary sinus pathology and odontogenic lesions in patients evaluated by cone beam computed tomography. A systematic review and meta-analysis. *Med Oral Patologia Oral Y Cir Bucal*. 2020;25(1):e34–48.
47. Nair AK, Jose M, Sreela LS, Prasad TS, Mathew P. Prevalence and pattern of proximity of maxillary posterior teeth to maxillary sinus with mucosal thickening: A cone beam computed tomography based retrospective study. *Ann Afr Med*. 2023;22(3):327–32.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.