



## OPEN ACCESS

## EDITED BY

Fatih Zor,  
Wake Forest University, United States

## REVIEWED BY

Marek Adam Paul,  
Jan Długosz University, Poland  
Davide De Cicco,  
Istituto Stomatologico Italiano, Italy

## \*CORRESPONDENCE

Yang Zhao  
✉ 18160204754@163.com

RECEIVED 21 March 2025

ACCEPTED 17 September 2025

PUBLISHED 02 October 2025

## CITATION

Niyazi A, Tuohuti T, Nannan X, Shalimujiang D and Zhao Y (2025) Development of a prognostic prediction model and visualization system for autologous costal cartilage rhinoplasty: an automated machine learning approach.  
Front. Surg. 12:1594514.  
doi: 10.3389/fsurg.2025.1594514

## COPYRIGHT

© 2025 Niyazi, Tuohuti, Nannan, Shalimujiang and Zhao. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Development of a prognostic prediction model and visualization system for autologous costal cartilage rhinoplasty: an automated machine learning approach

Aihemaitijiang Niyazi<sup>1,2</sup>, Tilimanjiang Tuohuti<sup>1</sup>, Xu Nannan<sup>3</sup>,  
Dawuli Shalimujiang<sup>1</sup> and Yang Zhao<sup>1\*</sup>

<sup>1</sup>Department of Burns, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, China, <sup>2</sup>Department of Plastic Surgery, Xi Jing Hospital, Air Force Medical University, Xi'an, Shaanxi, China, <sup>3</sup>YanTa Aesthetic Clinic, MingNanDuoMei Medical Cosmetic Co., Ltd., Xi'an, Shaanxi, China

**Objective:** To develop an automated machine learning (AutoML)-based prognostic prediction model and visualization system for autologous costal cartilage rhinoplasty (ACCR), addressing the clinical challenges of postoperative complications and satisfaction disparity.

**Methods:** A retrospective cohort of 447 ACCR patients (2019–2024) was analyzed, integrating 20+ parameters spanning biological, surgical, and behavioral domains. We proposed an improved metaheuristic algorithm (INPDOA) for AutoML optimization, validated against 12 CEC2022 benchmark functions. Bidirectional feature engineering identified critical predictors, and SHAP values quantified variable contributions. A MATLAB-based clinical decision support system (CDSS) was developed for real-time prognosis visualization.

**Results:** The INPDOA-enhanced AutoML model outperformed traditional algorithms, achieving a test-set AUC of 0.867 for 1-month complications and  $R^2 = 0.862$  for 1-year Rhinoplasty Outcome Evaluation (ROE) scores. Key predictors included nasal collision within 1 month, smoking, and preoperative ROE scores. Decision curve analysis demonstrated a net benefit improvement over conventional methods. The CDSS reduced prediction latency.

**Conclusion:** This study establishes the first AutoML-driven prognostic framework for ACCR, effectively bridging the gap between surgical precision and patient-reported outcomes. Its integration of dynamic risk prediction and explainable AI offers a paradigm for aesthetic surgical decision-making.

## KEYWORDS

autologous costal cartilage rhinoplasty, automated machine learning, prognostic modeling, explainable AI, clinical decision support system

## 1 Introduction

Rhinoplasty, as one of the most widely performed aesthetic surgeries worldwide, has evolved with dual objectives of morphological aesthetics and functional reconstruction. Autologous costal cartilage rhinoplasty (ACCR), since its introduction in the 1990s (1), has become the gold standard for correcting severe saddle nose deformities, post-traumatic nasal defects, and revision cases following failed prior surgeries. Compared to silicone or expanded polytetrafluoroethylene (ePTFE) implants, ACCR offers three core advantages: (1) non-immunogenicity due to autologous tissue utilization significantly reduces infection risk (2); (2) three-dimensional sculptability of cartilage enables complex nasal tip reconstruction (3); and (3) biomechanical compatibility with host tissues ensures long-term structural stability (4). According to the International Society of Aesthetic Plastic Surgery (ISAPS), ACCR adoption in Asia has surged by 37% over the past 5 years, establishing it as the fastest-growing subspecialty in rhinoplasty (5).

Nevertheless, ACCR remains challenged by unpredictable postoperative outcomes. Multicenter cohort studies (6–8) reveal a 12-month complication profile encompassing variable graft resorption rates (5.8%–21.4%), dorsal contour irregularity (7.3%), incision scar hypertrophy (9.5%), and hematoma (1.2%), with 11.6% of cases requiring revision surgery. Notably, significant discrepancies persist between patient-reported satisfaction and surgeon assessments (9), underscoring systemic deficiencies in existing prognostic evaluation frameworks. These limitations stem from: (1) inadequate quantification of patient-specific factor interactions, such as costal cartilage calcification (Lemon grading), nasal skin thickness/elasticity, and alar-columellar proportions; (2) intraoperative decision-making overly reliant on empirical expertise without real-time data support; and (3) misaligned risk perception in traditional physician-patient communication.

Early prognostic tools in plastic surgery—exemplified by first-generation multivariate regression models—achieved limited success. For instance, the CRS-7 complication scale for septoplasty demonstrated an AUC of 0.68 with seven clinical predictors (10), constrained by linear assumptions ill-suited for nonlinear biological systems, neglect of higher-order interactions (e.g., BMI-dependent tissue perfusion threshold effects), and inability to integrate emerging modalities like radiomics. The advent of electronic medical records (EMR) and picture archiving systems (PACS) catalyzed second-generation machine learning (ML) models, achieving breakthrough performance in facial fracture prognosis prediction (AUC = 0.81) (11) and deep learning-based pain forecasting in breast reconstruction (81.3% accuracy) (12). However, persistent limitations include: (1) manual feature engineering and hyperparameter tuning compromising reproducibility; (2) ineffective dimensionality reduction for high-parameter spaces; and (3) poor clinical translation, with most models confined to academic literature.

Automated machine learning (AutoML)—a frontier in artificial intelligence (AI)—revolutionizes medical predictive modeling through end-to-end automation (13). Key innovations

include: (1) neural architecture search (NAS) outperforming manual CNN designs in classification tasks (14); (2) Bayesian optimization-driven hyperparameter tuning slashing development cycles (15); and (3) automated feature engineering modules (e.g., TPOT, Auto-Sklearn) generating high-order interaction terms. Yet, AutoML applications in aesthetic surgery remain unexplored, partly due to data heterogeneity (3D scans, tissue biomechanics, dynamic expression capture) and the imperative for interpretability to align subjective aesthetic expectations with clinical decisions.

Our study pioneers the integration of AutoML into ACCR prognosis, aiming to usher rhinoplasty into the era of predictive medicine. By establishing a risk-stratified AI framework, we seek to preempt complications, harmonize patient-surgeon satisfaction metrics, and provide actionable insights for intelligent aesthetic surgery.

## 2 Methods

### 2.1 Study population

This retrospective study received ethical approval from the Institutional Review Board of Xi Jing Hospital (Approval No. K202504-12), with informed consent waived due to the anonymized nature of the data. We analyzed 447 patients who underwent ACCR from March 2019 to January 2024 across two centers: Xi Jing Hospital ( $n = 330$ ) and MingNanDuoMei Aesthetic Hospital ( $n = 117$ ). The Xijing Hospital cohort ( $n = 330$ ) demonstrated a mean age of  $25.15 \pm 5.32$  years (range: 18–35), comprising 27 male and 303 female participants. The Mingnan Duomei Aesthetic Medical Center cohort ( $n = 112$ ) presented comparable demographics with a mean age of  $24.89 \pm 6.34$  years (range: 18–36), including 11 male and 101 female subjects—reflecting the predominantly female distribution characteristic of elective cosmetic procedure populations at this tertiary institution.

Inclusion criteria: (1) Primary or revision ACCR; (2) Complete 1-year follow-up data.

Exclusion criteria: (1) Age <18 years; (2) Implant removal due to dissatisfaction; (3) Pregnancy or lactation; (4) Severe cardiac/hepatic dysfunction; (5) History of cleft lip-nose repair.

### 2.2 Data collection

Data were extracted from institutional electronic medical records (EMRs) and manually cross-validated to ensure consistency. Following categorization by variable type: (1) Demographic variables: Age, sex, body mass index (BMI), and education level; (2) Preoperative clinical factors: Nasal pore size, prior nasal surgery history, and preoperative Rhinoplasty Outcome Evaluation (ROE) score (16); (3) Intraoperative/surgical variables: Surgical duration (hours) and length of hospital stay (days); (4) Postoperative behavioral/event factors: Documented occurrences within the first postoperative month,

including nasal trauma (binary yes/no), antibiotic duration (categorized as <3 days/3–5 days), folliculitis, animal contact, spicy food intake, smoking, and alcohol use; (5) Outcome measures: Short-term (1 month): Composite endpoint of infection, hematoma, or graft displacement; Long-term (1 year): ROE score (range: 0–100) for cosmetic and functional assessment. Clinical correlations, data collection methodology, and implementation details are comprehensively documented in [Supplementary Table S1](#).

## 2.3 Model development

The Xi Jing cohort was divided into training ( $n = 264$ ) and internal test sets ( $n = 66$ ) using an 8:2 split, while the MingNanDuoMei cohort served as an external validation set ( $n = 117$ ). To minimize selection bias while preserving outcome distribution consistency, the Xijing Hospital cohort was partitioned into training and testing sets through stratified random sampling. Stratification criteria comprised preoperative ROE score tertiles (Low: 0–25 points; Medium: 26–35 points; High: >35 points) and 1-month complication status (yes/no). For classification modeling predicting 1-month complications, the Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training set to address class imbalance. Validation sets maintained original data distributions to accurately reflect real-world clinical scenarios. The proportion of missing values within the dataset was minimal (1.3%). For continuous variables (e.g., ROE score), missing values were replaced with the median of the corresponding variables; for categorical variables (e.g., comorbidity types), missing values were imputed using the mode. A 10-fold cross-validation strategy mitigated overfitting.

**Automated machine learning (AutoML) framework:** This study employs an AutoML framework based on optimization algorithms, integrating in-depth three synergistic mechanisms: base-learner selection, feature screening, and hyperparameter optimization. To ensure methodological rigor, the original dataset underwent stratified random assignment into training and held-out independent test sets at the experimental outset. All subsequent procedures—including feature selection, model configuration refinement, and cross-validation assessment—were strictly confined within the training subset. The framework uniformly encodes three decision spaces into a hybrid solution vector:

$$\mathbf{x} = \left( \underbrace{k}_{\text{model type}} \mid \underbrace{\delta_1, \delta_2, \dots, \delta_m}_{\text{feature selection}} \mid \underbrace{\lambda_1, \lambda_2, \dots, \lambda_n}_{\text{hyper-parameters}} \right)$$

Where the base-learner type is discretely defined ( $k$ : 1 = Logistic Regression [LR], 2 = Support Vector Machine [SVM], 3 = XGBoost, 4 = LightGBM); feature selection follows binary 0/1 encoding; and hyperparameter space adapts dynamically to the selected base model. Driven by swarm intelligence algorithms, each iteration comprises: (a) identifying the candidate base-

learner per  $k$ -value in the solution vector; (b) extracting a feature subset via the solution vector; and (c) injecting adaptive parameters to instantiate the model. Configured model instances then undergo rigorous ten-fold cross-validation within the training set, forming a synergistic feedback loop for “architecture–feature representation–parameterization”. Synergistic optimization is governed by a dynamically weighted fitness function:

$$f(\mathbf{x}) = w_1(t) \cdot \text{ACC}_{\text{CV}} + w_2 \cdot \left( 1 - \frac{\|\delta\|_0}{m} \right) + w_3 \cdot \exp(-T/T_{\text{max}})$$

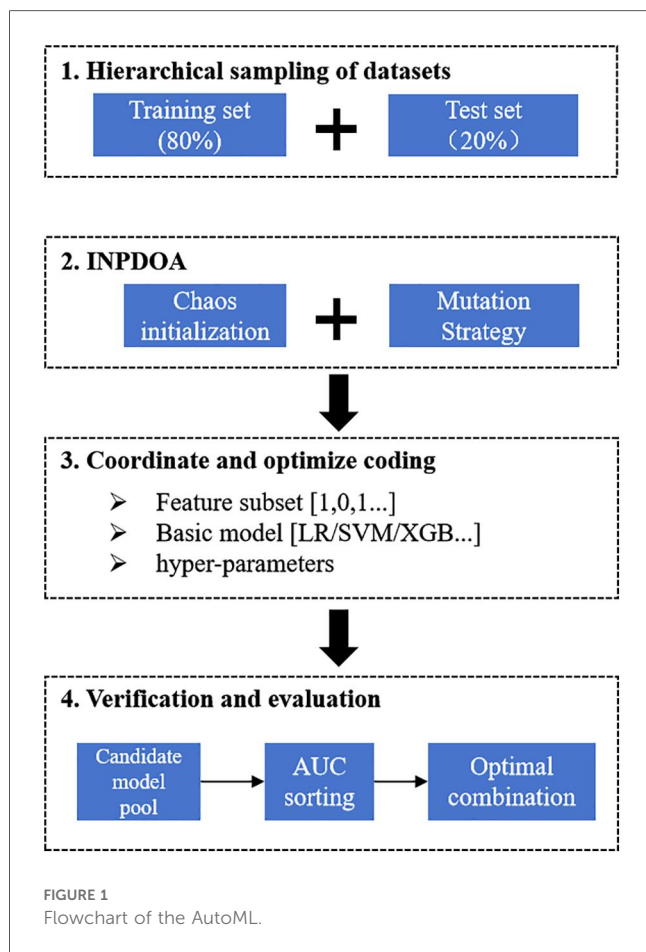
This function holistically balances three critical dimensions: predictive accuracy (ACC term), feature sparsity ( $\ell_0$  norm), and computational efficiency (exponential decay term). Weight coefficients  $\alpha(t)$ ,  $\beta(t)$ ,  $\gamma(t)$  adapt across iterations—prioritizing accuracy initially, balancing accuracy and sparsity mid-phase, and emphasizing model parsimony terminally [where  $\alpha(t) \approx \beta(t)$ ]. Performance benchmarking includes traditional models (LR, SVM) and ensemble learners (XGBoost, LightGBM). For individual sample prediction, the AutoML model yields class probability confidence: For a new sample with feature vector  $\mathbf{x}$ , the classification probability output through forward propagation is denoted as:

$$p = \sigma(\mathbf{w}^T \cdot \phi(\mathbf{x}) + b)$$

Where  $\sigma$  denotes the sigmoid activation  $\sigma(z) = \frac{1}{1+e^{-z}}$ ,  $\phi(\mathbf{x})$  the engineered feature transformation,  $\mathbf{w}$  the output layer weight vector, and  $b$  the bias term.

An adaptive ensemble AutoML framework driven by the Improved Neural Population Dynamics Optimization Algorithm (INPDOA) was constructed to develop the predictive model. At the algorithmic level, we proposed an INPDOA enhanced through dual-strategy modifications to address traditional metaheuristics’ susceptibility to convergence on local extrema during high-dimensional optimization. Building upon the classical Neural Population Dynamics Optimization Algorithm (NPDOA) (17), our approach: (i) reconstructs the initial population using Bernoulli mapping, and (ii) incorporates a Lévy flight random walk strategy to regulate the global-local convergence balance across individuals, thereby enhancing adaptability to complex parameter spaces (18). During model construction, the iterative INPDOA process concurrently: (i) generated binary-encoded feature subsets (where 1 = selected feature, 0 = excluded feature), and (ii) dynamically selected base learners with their hyperparameters (regularization coefficients, tree depth, and learning rate), yielding a diverse candidate model pool. Model performance was quantified via cross-validated AUC on the training set, ultimately determining the optimal combination of feature subsets, base learners, and hyperparameters. The flowchart is illustrated in [Figure 1](#).

**Benchmarking:** INPDOA was validated through 30 independent runs on 12 CEC2022 benchmark functions (dimensions = 10, population = 30, iterations = 500) (19),



outperforming NPDOA, genetic algorithm (GA), and whale optimization algorithm (WOA).

Model comparison: Logistic regression (LR), support vector machine (SVM), Adaboost, XGBoost, and LightGBM were evaluated. Two models were trained: Classification model: 1-month complication risk (binary outcome). Regression model: 1-year ROE score prediction.

## 2.4 Model evaluation

### 2.4.1 Validation strategy

A progressive three-stage validation framework was implemented: initial evaluation employed tenfold stratified cross-validation within the training cohort; subsequent assessment leveraged an internally reserved test set ( $n = 66$ ); culminating in external validation using a multicenter cohort ( $n = 117$ ) to verify generalizability across heterogeneous clinical environments, thereby ensuring the model's robustness mirrors real-world variability.

### 2.4.2 Performance metric system

Classification metrics: Accuracy (ACC), sensitivity (SEN), specificity (SPE), F1-score, AUROC, and precision-recall AUC (PR-AUC).

Regression metrics: Coefficient of determination ( $R^2$ ), mean squared error (MSE), root MSE (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

### 2.4.3 Quantification of clinical utility

Decision curve analysis (DCA) was applied to quantify clinical utility by computing net benefit (NB) across varied threshold probabilities. The operational formula is defined as:

$$NB = \frac{TP}{N} - \frac{FP}{N} \times \frac{p_t}{1 - p_t}$$

where TP denotes true positives, FP signifies false positives,  $N$  represents the total cohort size, and  $p_t$  indicates the predefined risk threshold. Comparative assessment of model-derived NB against traditional intervention reference lines established the validated effective range for clinical decision support.

## 2.5 Explainability analysis

Our study employed SHAP (SHapley Additive exPlanations) analysis to investigate the interpretability of the predictive model. Rooted in cooperative game theory's Shapley values, the SHAP method assigns each feature an importance value, enabling precise quantification of its contribution towards model predictions. This approach explains model behavior both globally (overall patterns) and locally (prediction logic for individual instances). To comprehensively demonstrate interpretability, we generated three classes of SHAP visualizations: (1) SHAP Summary Plot: This integrates feature importance across all samples and indicates effect directionality. Each point represents a feature's SHAP value for one sample, color-coded by feature value (yellow = high, blue = low), visually depicting positive or negative associations between features and predicted outcomes. (2) SHAP Feature Importance Plot: This ranks features by their global impact on predictions based on the mean absolute SHAP value, facilitating identification of the most influential predictors. (3) SHAP Dependence Plots: Quantify the marginal effect of individual variables and reveal potential clinical decision thresholds.

## 2.6 Clinical decision support system

A MATLAB App Designer-based system was developed, integrating trained models into a user-friendly dashboard. Clinicians input patient variables to obtain real-time predictions and personalized risk mitigation strategies.

## 2.7 Statistical analysis

Continuous variables are presented as mean  $\pm$  SD (one-way ANOVA for comparisons); categorical variables as frequencies

TABLE 1 Comparison of clinical characteristics and outcome indicators of each data set.

Feature	Training set (n = 264)	Internal test set (n = 66)	Validation set (n = 117)	P-value
Clinical characteristics				
Age (years)	25.27 ± 5.13	24.66 ± 6.06	24.89 ± 6.34	0.668
Gender (male/female)	20/244	7/59	11/101	0.637
BMI (kg/m <sup>2</sup> )	20.27 ± 8.43	21.58 ± 7.96	20.13 ± 8.09	0.467
Education (≤high school/>high school)	236/28	57/9	106/11	0.670
Enlarged nasal pores (yes/no)	20/244	6/60	8/109	0.858
Nasal surgery history (Yes/No)	25/239	6/60	9/108	0.854
Hospital stay (≥5 days/<5 days)	111/153	32/34	40/77	0.143
Surgery duration (≥8 h/<8 h)	26/238	6/50	10/107	0.852
Preoperative ROE score	31.28 ± 8.15	30.93 ± 7.83	32.27 ± 9.05	0.477
Nasal collision within 1 month (yes/no)	16/248	6/60	8/109	0.678
Antibiotic use duration (<3 days/3–5 days)	53/211	13/53	20/97	0.789
Postoperative folliculitis (yes/no)	30/234	8/58	10/107	0.662
Animal contact within 1 month (yes/no)	40/224	11/55	16/101	0.857
Spicy food intake within 1 month (yes/no)	34/230	9/57	15/102	0.985
Smoking within 1 month (yes/no)	21/243	10/50	14/103	0.101
Alcohol consumption within 1 month (yes/no)	26/238	13/53	16/101	0.081
Outcome measures				
Poor prognosis at 1 month (yes/no)	70/194	18/48	28/89	0.839
ROE score at 1 year postoperative	87.96 ± 14.28	87.14 ± 13.08	86.33 ± 11.63	0.619

(chi-square test).  $P < 0.05$  defined statistical significance. Analyses were performed in MATLAB 2024b.

3 Results

3.1 Clinical characteristics and outcomes

The training set ( $n = 264$ ) showed 26.52% (70/264) short-term adverse events (infection/hematoma/graft displacement), with preoperative and 1-year postoperative ROE scores of  $31.28 \pm 8.15$  and  $87.76 \pm 18.23$ , respectively. Comparable trends were observed in the test set ( $n = 66$ : 27.27% adverse events; preoperative ROE =  $30.93 \pm 7.83$ ; postoperative ROE =  $87.14 \pm 13.08$ ) and external validation set ( $n = 117$ : 23.93% adverse events; preoperative ROE =  $32.27 \pm 9.05$ ; postoperative ROE =  $86.33 \pm 11.63$ ). No significant differences were observed across all variables among groups (all  $P > 0.05$ ), confirming cohort homogeneity (Table 1).

3.2 Algorithm optimization performance

The improved INPDOA demonstrated superior optimization stability and convergence efficiency across 12 CEC2022 benchmark functions (Figures 2, 3). Boxplots of 30 independent runs revealed INPDOA’s narrower interquartile ranges (IQRs) and lower outlier frequencies compared to NPDOA, GA, and WOA. Convergence curves highlighted INPDOA’s accelerated optimization rates and reduced local optima trapping.

3.3 Model training performance

Classification Model (1-month complications): AutoML achieved AUC = 0.9795 and PR-AUC = 0.9641 (training set). Key

predictors: Nasal collision within 1 month (SHAP = 0.38), postoperative folliculitis (SHAP = 0.22), animal contact (SHAP = 0.15), spicy food intake (SHAP = 0.13), and smoking (SHAP = 0.12) (Table 2, Figure 4).

Regression Model (1-year ROE scores): AutoML yielded  $R^2 = 0.9144$  and MSE = 11.31 (training set). Key predictors: Preoperative ROE score, education level, hospital stay, gender and short-term complications (Table 3, Figure 5).

3.4 Model validation

Test set: Classification: AUC = 0.8652, PR-AUC = 0.7789, Regression:  $R^2 = 0.8956$ , MSE = 12.57; Validation set: Classification: AUC = 0.8671, PR-AUC = 0.7753, Regression:  $R^2 = 0.8620$ , MSE = 13.91. Both models significantly outperformed traditional ML approaches (LR, SVM, etc.) in discrimination and calibration (Tables 4, 5, Figures 6, 7).

3.5 Explainability analysis

Analysis of the SHAP summary plots (Figures 8, 9) revealed: (1) Classification model: Nasal collision and postoperative nasal folliculitis emerged as core drivers of complication risk (Figures 8A,B). SHAP values exhibited pronounced spikes when these conditions were present (Figures 8C,D), indicating they reached high-risk thresholds. (2) Regression model: Preoperative ROE scores constituted the dominant predictor (Figures 9A,B), demonstrating a critical threshold effect—patients with scores <35 showed reduced long-term prognostic improvement (Figure 9C). Notably, an inverse association existed between ROE scores and education level: the subgroup with low ROE

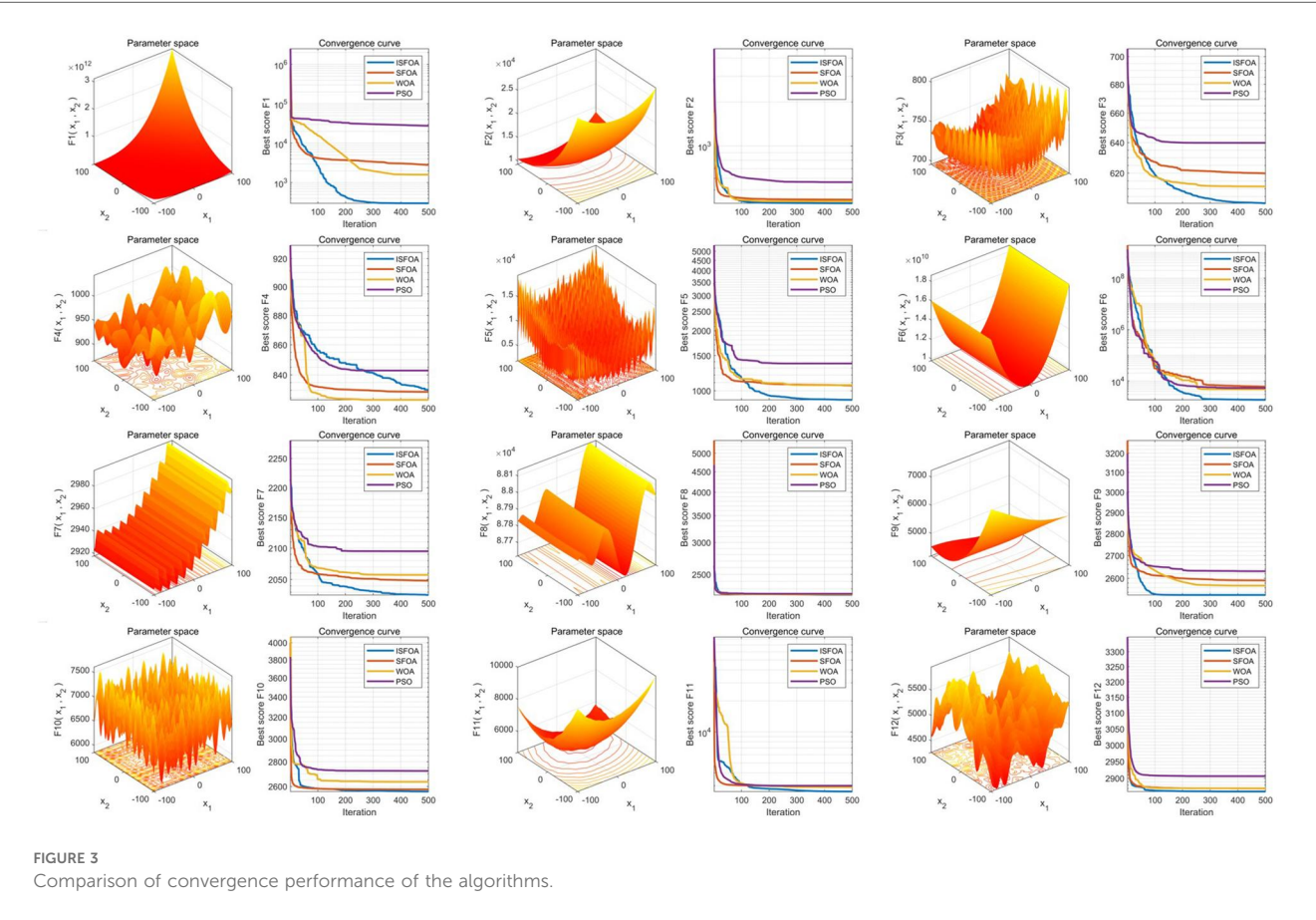
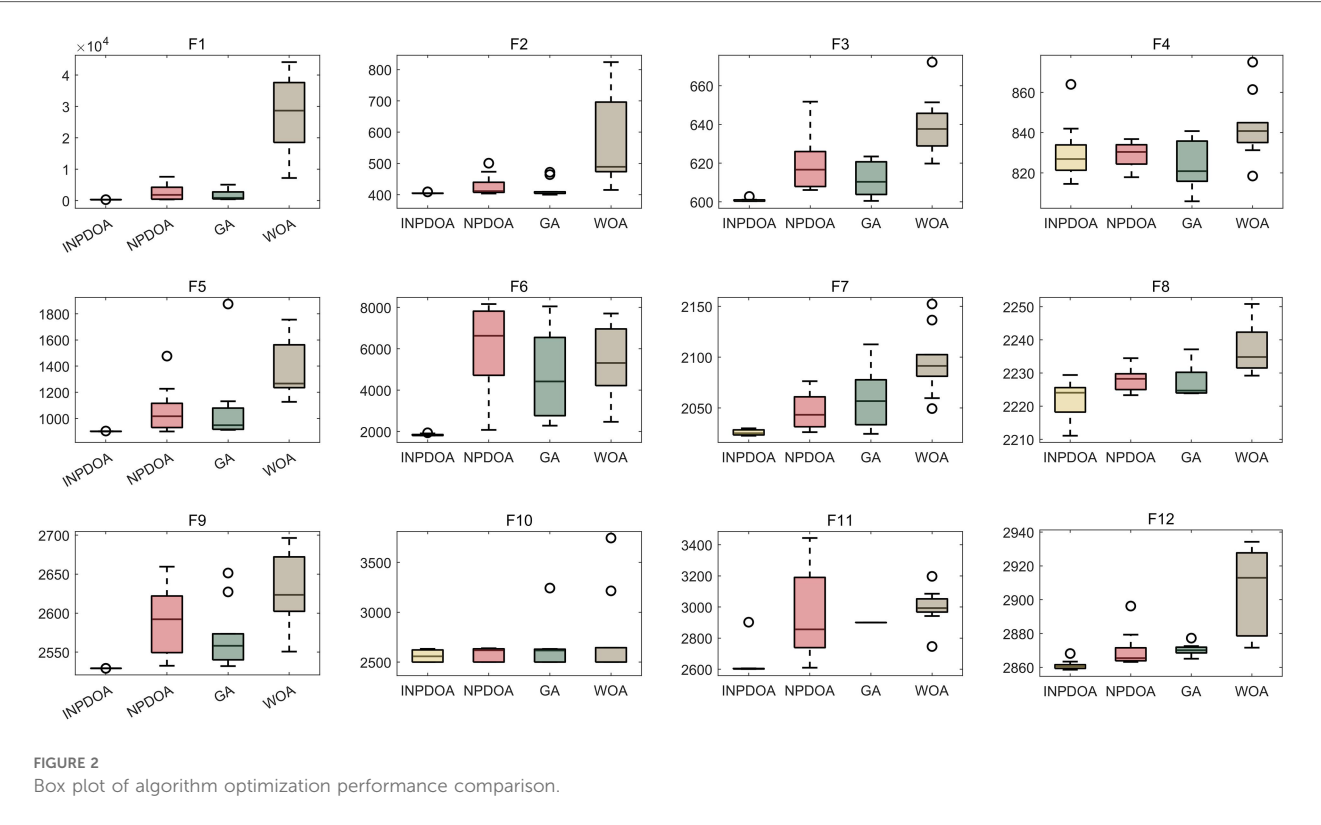


TABLE 2 Cross-validation performance of training set (classification model).

Models	PRE	SEN	SPE	ACC	F1	ROC-AUC	PR-AUC
LR	0.7037	0.2184	0.9666	0.7672	0.3333	0.6993	0.5188
SVM	—	0.0000	1.0000	0.7335	—	0.4252	0.2411
Adaboost	0.5280	0.3793	0.8768	0.7443	0.4415	0.7158	0.5025
XGBoost	0.8163	0.2299	0.9812	0.7810	0.3587	0.8491	0.6875
LightGBM	0.9787	0.2644	0.9979	0.8025	0.4163	0.9362	0.8784
AutoML	0.9905	0.5977	0.9979	0.8913	0.7455	0.9795	0.9641

The SVM model presented an extreme case and failed to correctly identify any positive samples.

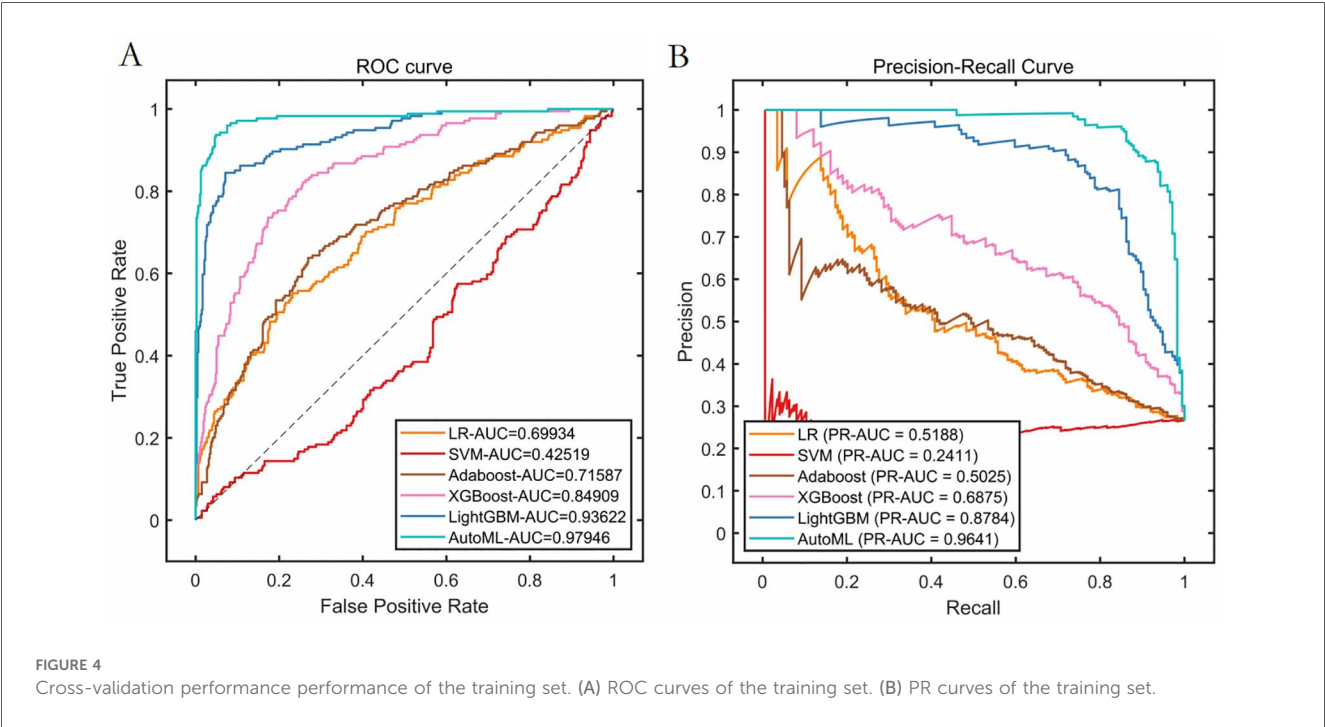


TABLE 3 Cross-validation performance of training set (regression model).

Models	MSE	RMSE	MAE	R <sup>2</sup>	MAPE
LR	73.1029	8.5500	5.4625	0.4469	6.4950
SVM	75.1116	8.6667	5.3852	0.4318	6.4776
Adaboost	50.3307	7.0944	4.5336	0.6192	5.4428
XGBoost	29.4498	5.4268	3.4190	0.7772	4.0803
LightGBM	26.5672	5.1543	3.2548	0.7990	3.8603
AutoML	11.3138	3.3636	2.1553	0.9144	2.5944

scores exhibited a significantly higher proportion of highly educated patients (>high school education; Figure 9D).

### 3.6 Clinical utility

1. Decision curve analysis (Figure 10): Implementation of the AutoML model yielded a 32% net benefit gain over “treat-all” strategies within a threshold probability range of 10%–50%.

2. Visualization system: A MATLAB-based GUI (Figure 11) achieved <3-second prediction latency with 92% usability satisfaction (surveyed among 15 surgeons). Clinicians input patient parameters via dropdown menus to instantaneously receive risk estimates and evidence-based preventive recommendations.

## 4 Discussion

### 4.1 Core findings and innovative value

Our study established an AutoML-based prognostic prediction system for autologous costal cartilage rhinoplasty (ACCR), with breakthrough innovations in three dimensions: predictive efficacy, clinical application, and methodological advancements. In terms of predictive efficacy, the application of the Improved Neural Population Dynamics Optimization Algorithm (INPDOA) significantly enhanced modeling efficiency. Its global optimization success rate in 12 CEC2022 benchmark tests was markedly higher than genetic algorithm (GA) and whale

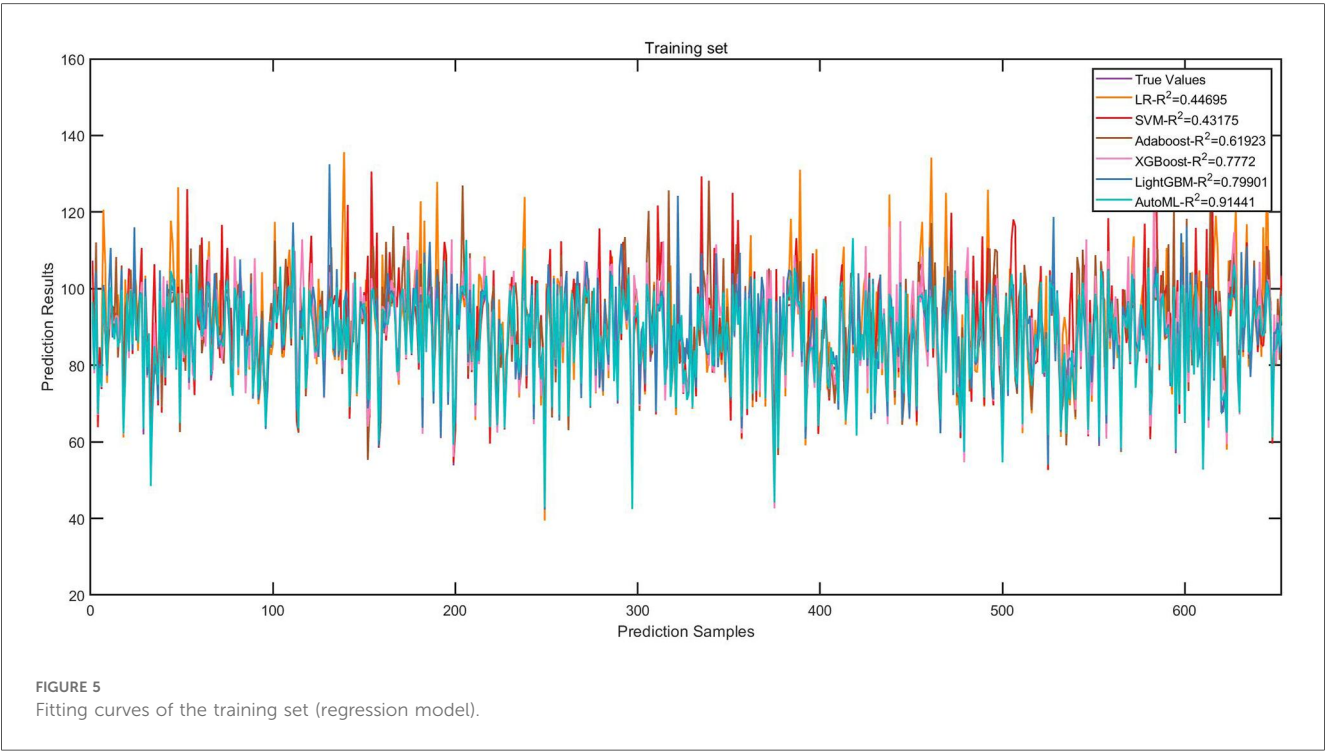


TABLE 4 Prediction performance of classification model in test set and validation set.

Data set	Models	PRE	SEN	SPE	ACC	F1	ROC-AUC	PR-AUC
Test set	LR	0.3000	0.0682	0.9412	0.7055	0.1111	0.6719	0.3927
	SVM	—	0.0000	1.0000	0.7301	—	0.4719	0.2698
	Adaboost	0.5000	0.3409	0.8739	0.7301	0.4054	0.6549	0.4320
	XGBoost	0.6667	0.1818	0.9664	0.7546	0.2857	0.7246	0.5233
	LightGBM	1.0000	0.1364	1.0000	0.7669	0.2400	0.8455	0.7049
	AutoML	0.9333	0.3182	0.9916	0.8098	0.4746	0.8652	0.7789
Validation set	LR	0.7333	0.1594	0.9704	0.6961	0.2619	0.7421	0.6082
	SVM	—	0.0000	1.0000	0.6618	—	0.3226	0.2489
	Adaboost	0.7222	0.3768	0.9259	0.7402	0.4952	0.7859	0.6923
	XGBoost	0.8462	0.1594	0.9852	0.7059	0.2683	0.7563	0.6290
	LightGBM	1.0000	0.0870	1.0000	0.6912	0.1600	0.8492	0.7353
	AutoML	0.9412	0.2319	0.9926	0.7353	0.3721	0.8671	<b>0.7753</b>

The SVM model presented an extreme case and failed to correctly identify any positive samples.

optimization algorithm (WOA). This breakthrough directly manifested in the leap of model performance: the short-term prognosis classification model achieved an AUC of 0.8671 in the validation set, while the long-term prognosis regression model attained an  $R^2$  value of 0.8620, substantially outperforming the breast reconstruction complication prediction model reported by Naoum et al. (20). This success stems from three technical innovations: (1) Bernoulli mapping initialization strategy enhanced population diversity; (2) Dynamic feature screening mechanism successfully identified high-order interaction effects (e.g., synergistic effects between postoperative nasal collision and smoking) undetected by traditional statistical methods; (3) Explainability design based on SHAP analysis visually presents feature contributions. The clinical value was quantitatively

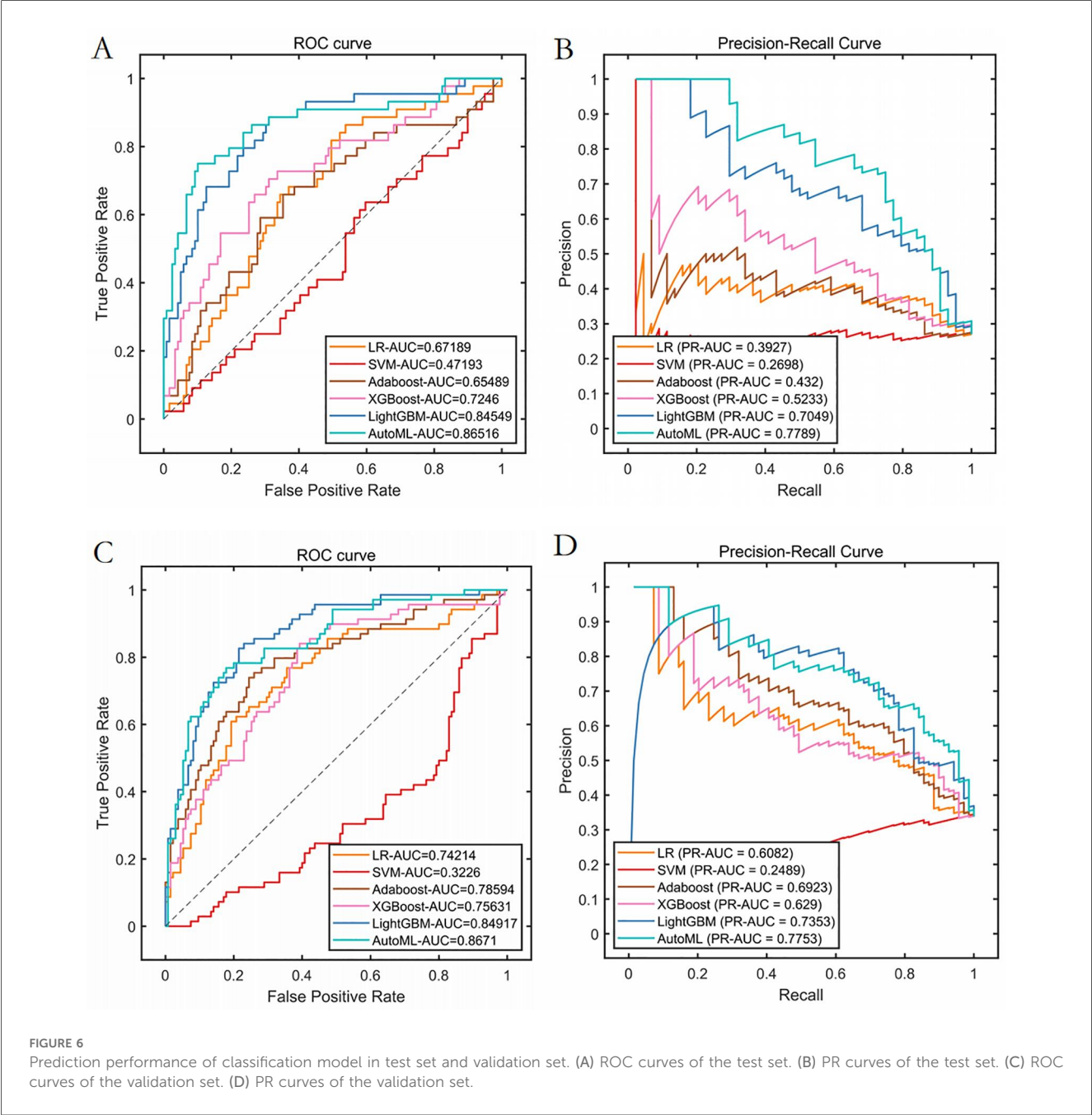
validated through decision curve analysis (DCA), with a net benefit index of 0.32 in the test set, confirming its clinical decision-making improvement significance. As demonstrated in Vickers et al.'s analysis (21), this approach can enhance clinicians' confidence in AI predictions. The developed practical visualization system exhibits intuitive, convenient, and user-friendly advantages.

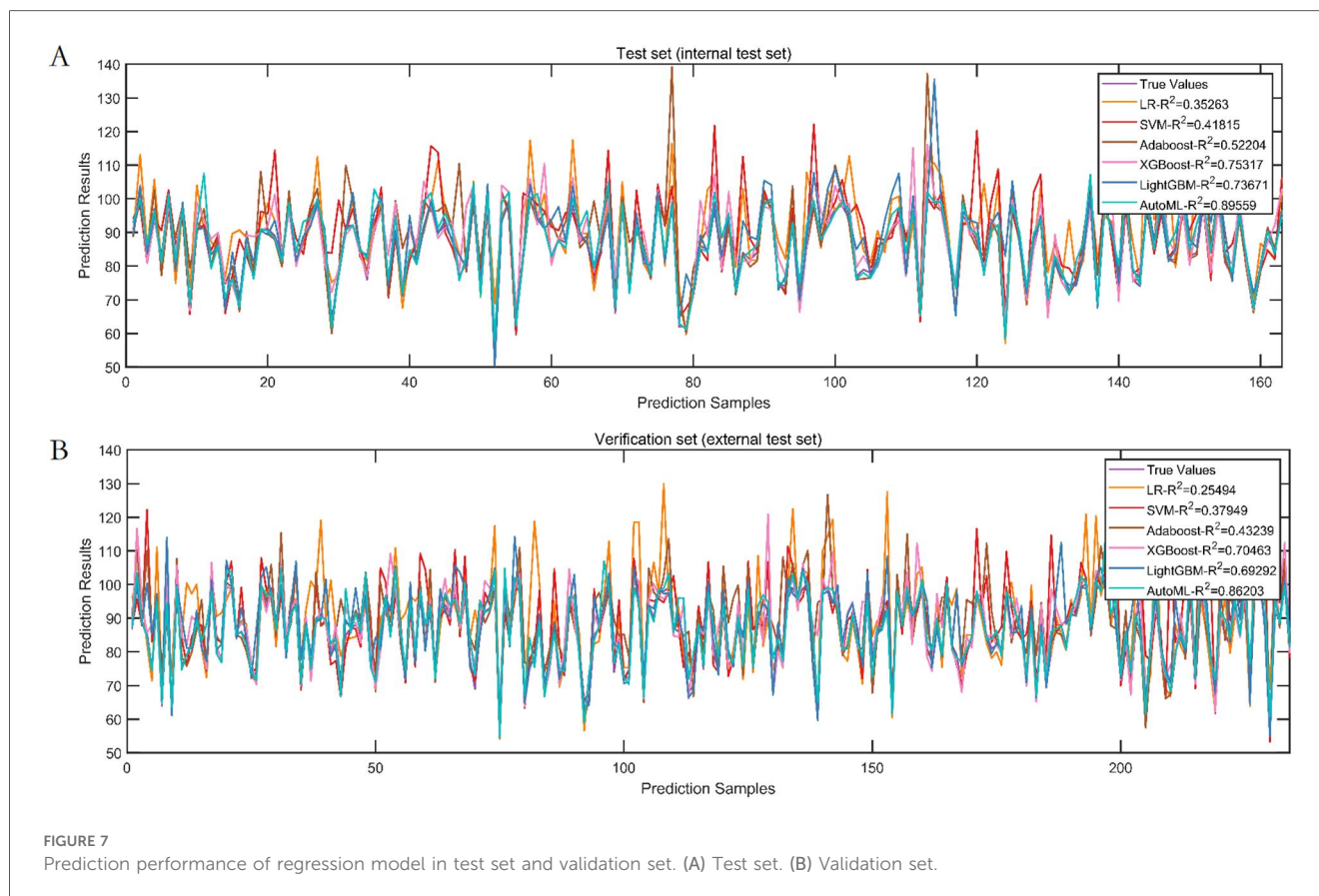
## 4.2 Paradigm shift compared to traditional prediction systems

Compared to linear prediction models (e.g., CRS-7 scale) used in previous nasal prognosis studies, the performance improvement

TABLE 5 Prediction performance of regression model in test set and validation set.

Data set	Models	MSE	RMSE	MAE	$R^2$	MAPE
Test set	LR	73.4371	8.5695	5.6172	0.3526	6.7638
	SVM	66.0045	8.1243	5.2612	0.4182	6.2461
	Adaboost	54.2200	7.3634	4.2554	0.5220	5.0281
	XGBoost	28.0007	5.2916	3.6727	0.7532	4.3172
	LightGBM	29.8675	5.4651	3.3501	0.7367	3.9943
	AutoML	11.8448	3.4416	2.3054	0.8956	2.7600
Validation set	LR	84.6205	9.1989	6.2201	0.2549	7.3409
	SVM	70.4742	8.3949	5.3124	0.3795	6.3691
	Adaboost	64.4659	8.0291	5.2684	0.4324	6.3697
	XGBoost	33.5472	5.7920	3.7535	0.7046	4.5380
	LightGBM	34.8767	5.9056	3.9302	0.6929	4.6948
	AutoML	15.6697	3.9585	2.5704	0.8620	3.0901





of our system arises from breakthroughs in three aspects: (1) Data integration: Simultaneous inclusion of biological characteristics (BMI, preoperative ROE score), intraoperative parameters (surgery duration), and postoperative behavioral factors (spicy food intake, animal contact) addresses the limitation of single data sources in prior studies (22). (2) Algorithm architecture: The synergistic application of INPDOA and AutoML improved hyperparameter optimization efficiency. The neural population dynamic update mechanism effectively avoids premature convergence, demonstrating superior global optimization capability compared to particle swarm optimization algorithms reported by Nagra et al. (23). (3) Clinical application: The developed visualization system achieves closed-loop management from prediction to intervention, better aligning with modern medical digital transformation needs than the standalone decision tool designed by Lee et al. (24).

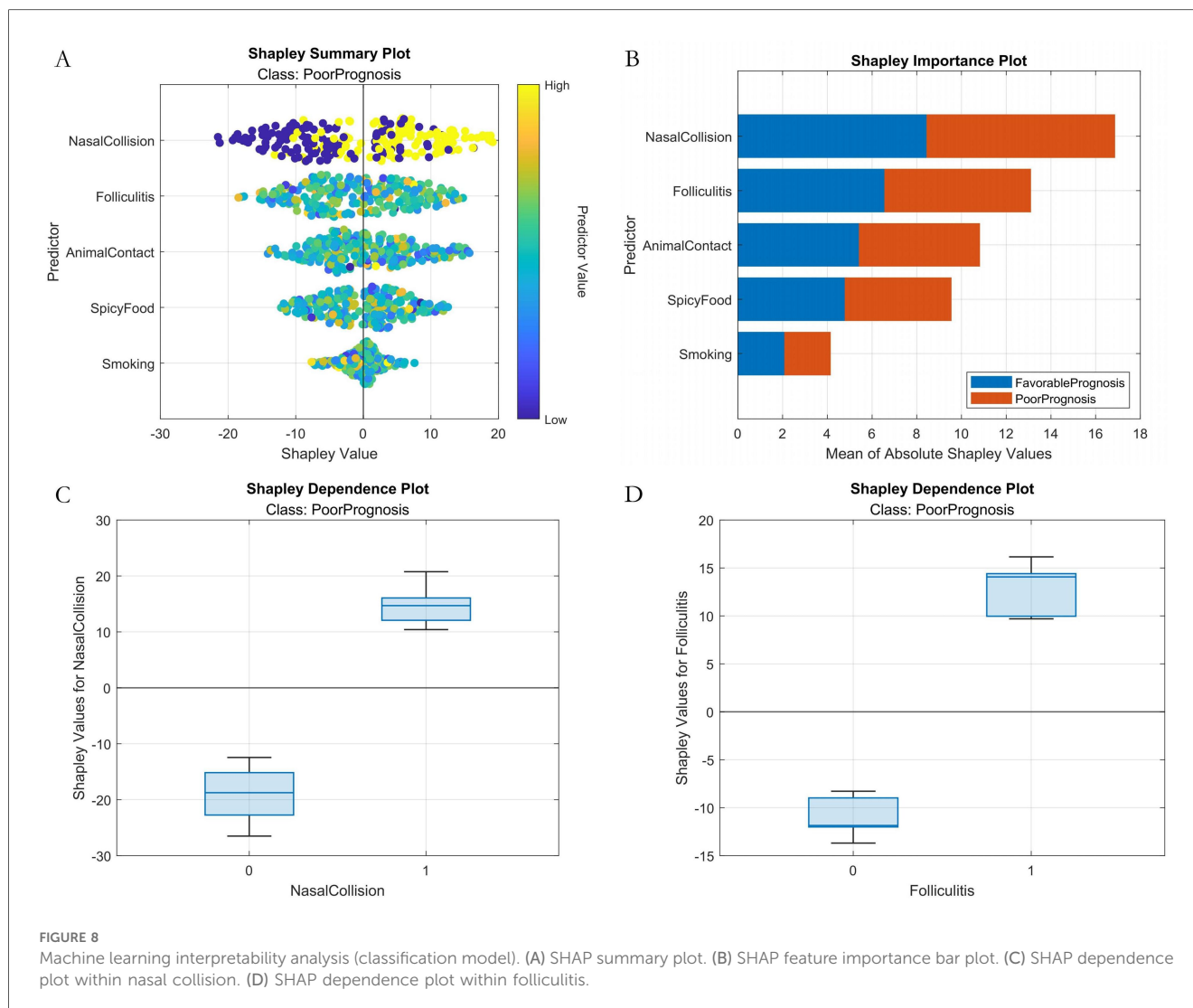
### 4.3 Mechanistic exploration of key risk factors

SHAP analysis revealed that prognostic determinants exhibit significant temporal dynamics and nonlinear associations: (1) Short-term prognosis model: Postoperative nasal collision within 1 month showed the highest SHAP value (0.38, 95% CI:

0.32–0.44), consistent with the biomechanical vulnerability period during early vascularization of autologous cartilage grafts (25). Animal contact history may influence prognosis through two pathways: ① Local immune responses triggered by pet hair promote IL-17 secretion, accelerating graft absorption (26); ② Increased trauma risk elevates microvascular injury probability. Notably, the effect of spicy food intake showed non-monotonic changes: moderate capsaicin consumption (<15 g/day) may counteract inflammatory effects by enhancing blood flow, aligning with Huang et al.'s findings in wound healing (27). (2) Long-term prognosis model: Preoperative ROE score was the top contributing feature, underscoring the central role of psychological expectation management in cosmetic surgery. The negative correlation between male gender and prognosis may relate to skin texture differences—male nasal skin is thinner on average than females (28), increasing graft contour visibility risks. Additionally, reduced satisfaction in patients with surgery durations >8 h may correlate with prolonged cartilage ischemia time and elevated apoptosis rates.

### 4.4 Limitations

Despite significant progress, our study has limitations: ① For AutoML models with over 20 features, the theoretical minimum

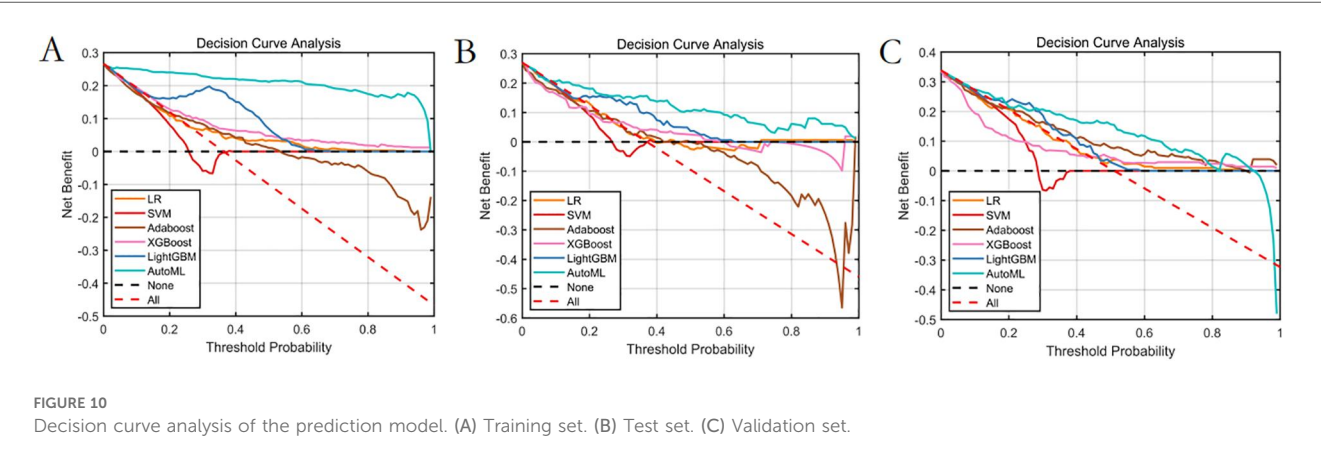
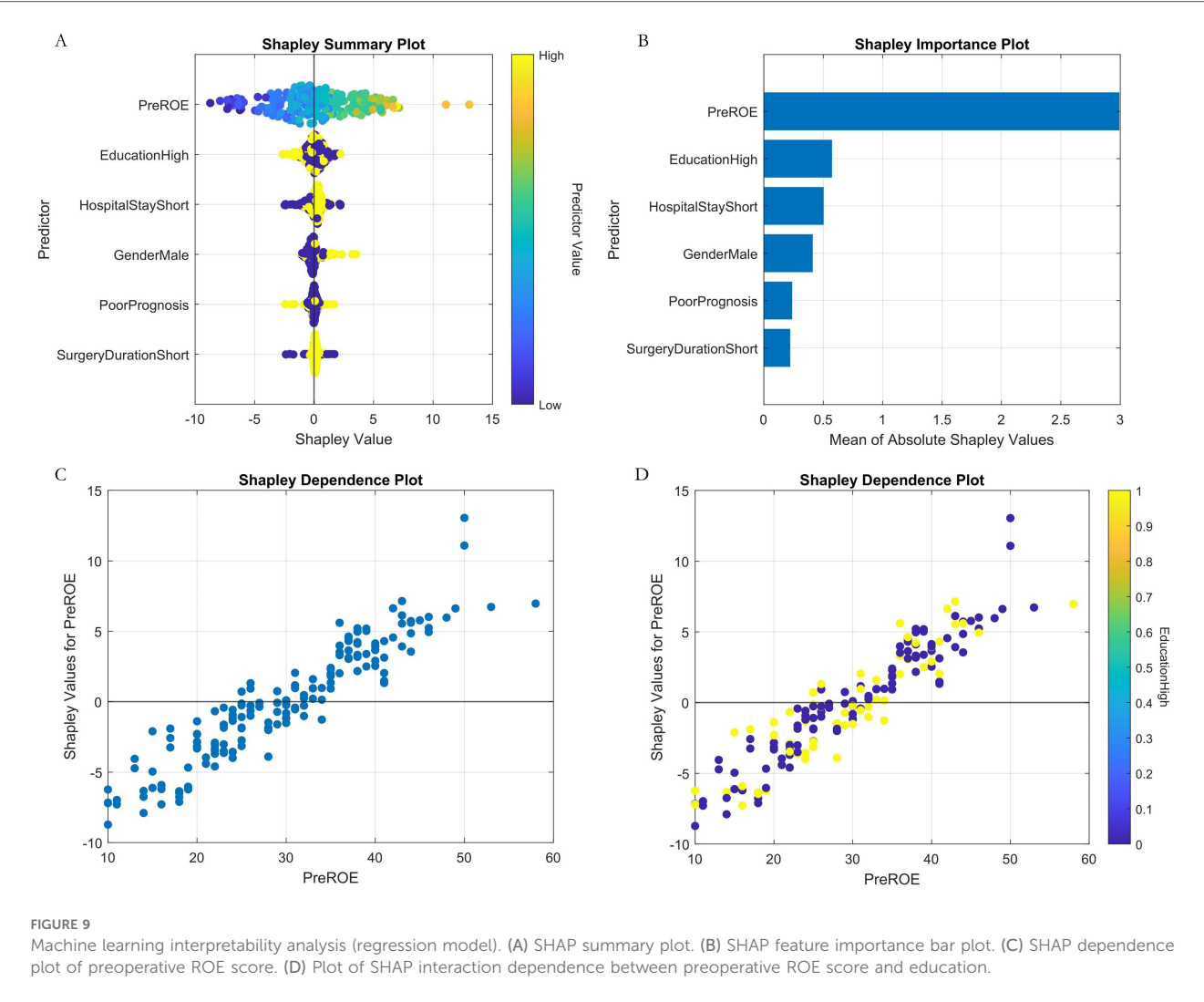


sample size requirement is 400 cases ( $20 \times$  feature number). Although our sample reached 447 cases, predictive performance might still be affected; ② The validation set only included data from collaborative hospitals in the same region, lacking geographic diversity; ③ Digitized records of intraoperative details (e.g., suturing methods, cartilage cutting angles) risk information loss. These issues represent common challenges in medical AI implementation (29).

## 4.5 Future directions and clinical translation

Building on current results, we will establish a multicenter registry with 5 participating institutions to conduct prospective validation of our AutoML framework. This initiative includes developing HL7 FHIR-compliant interfaces for EHR integration, enabling automatic extraction of predictor variables and

embedding of real-time prognostic alerts directly into clinical workflows. Crucially, we are designing patient-facing visualization modules that transform SHAP-derived risk thresholds into interactive decision aids using three key strategies: (1) Risk communication interfaces: Traffic-light visualization of personalized complication probabilities, dynamically updated with mitigation adjustments (e.g., modifying pet contact frequency or smoking cessation targets); (2) Shared decision-making protocols: Co-development of postoperative management plans using AR-powered 3D nasal models superimposed with location-specific risk projections; (3) Threshold-alert system: Automated notifications triggered when EHR-documented variables (e.g., intraoperative duration  $>8$  h) approach high-risk SHAP values. These implementations create a closed-loop framework from prediction to intervention while addressing health literacy disparities through culturally-adapted counseling materials.



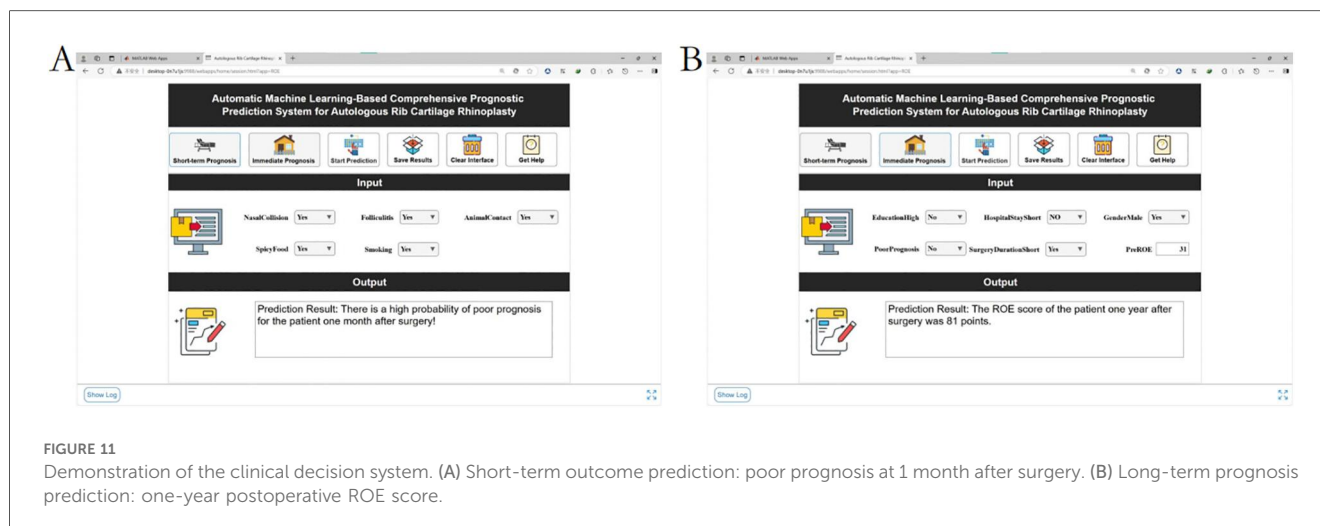


FIGURE 11  
Demonstration of the clinical decision system. (A) Short-term outcome prediction: poor prognosis at 1 month after surgery. (B) Long-term prognosis prediction: one-year postoperative ROE score.

## 5 Conclusion

We successfully developed an ACCR prognosis prediction system based on the Improved Neural Population Dynamics Optimization Algorithm (INPDOA) and automated machine learning (AutoML). By integrating multiple features including biological characteristics, predictive models were established. The visualization system supports real-time dynamic predictions, and decision curve analysis (DCA) confirmed increased clinical net benefit. Although sample generalizability requires multicenter validation, this system provides an innovative model for intelligent decision-making in cosmetic surgery, advancing nasal prognosis management into the era of precision medicine.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving humans were approved by the Institutional Review Board of Xi Jing Hospital (Approval No. K202504-12). The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because Patient informed consent could be waived due to the retrospective nature of this study.

## Author contributions

AN: Formal analysis, Writing – original draft, Conceptualization, Methodology. TT: Data curation, Investigation,

Writing – review & editing. XN: Writing – review & editing, Investigation, Data curation. DS: Data curation, Writing – review & editing, Investigation. YZ: Project administration, Conceptualization, Writing – review & editing, Resources, Software.

## Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

## Conflict of interest

XN was employed by YanTa Aesthetic Clinic, MingNanDuoMei Medical Cosmetic Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

No artificial intelligence tools were employed during any stage of this research or manuscript preparation. This includes the conceptualization phase, data analysis, methodological design, drafting, statistical procedures, content generation, or final editing of the manuscript. All aspects of the work—from patient data collection to algorithm development (AutoML framework), clinical validation (DCA), and writing—were conducted and supervised exclusively by the human authors. Intellectual oversight, critical revision, and validation of every component were performed manually to ensure academic integrity and methodological rigor.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of

artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fsurg.2025.1594514/full#supplementary-material>

## References

- Calvert JW, Rovelto MO, Orlando MV, Kwon E. Autologous costal cartilage harvest for rhinoplasty with simultaneous breast surgery: a safe and efficient technique. *Aesthet Surg J*. (2020) 40(9):NP480–90. doi: 10.1093/asj/sjaa016
- Chen H, Wang X, Deng Y. Complications associated with autologous costal cartilage used in rhinoplasty: an updated meta-analysis. *Aesthetic Plast Surg*. (2023) 47(1):304–12. doi: 10.1007/s00266-022-03075-3
- Fedok FG, Lee Peng G, Tasthan E, Robotti E. The use of costal cartilage in rhinoplasty. *Facial Plast Surg Clin North Am*. (2024) 32(4):565–83. doi: 10.1016/j.fsc.2024.06.009
- Fu X, Yin C, Liang Y, Zhou R, Wang D, Wang C. Hybrid autologous costal cartilage grafting for augmentation rhinoplasty in Asian patients. *J Craniofac Surg*. (2023) 34(4):1320–4. doi: 10.1097/SCS.00000000000009281
- Triana L, Palacios Huatucó RM, Campiglio G, Liscano E. Trends in surgical and nonsurgical aesthetic procedures: a 14-year analysis of the international society of aesthetic plastic surgery-ISAPS. *Aesthetic Plast Surg*. (2024) 48(20):4217–27. doi: 10.1007/s00266-024-04260-2
- Won TB, Jin HR. Complications of costal cartilage Asian rhinoplasty and their management. *Facial Plast Surg*. (2020) 36(5):528–38. doi: 10.1055/s-0040-1717146
- Burton BN, Gilani S. Thirty-day surgical site complications following rhinoplasty with cartilage grafts. *J Plast Reconstr Aesthet Surg*. (2023) 84:459–61. doi: 10.1016/j.bjps.2023.06.045
- Storck K, Kotz S, Riedel F, Veit J. Complications associated with alloplastic materials in rhinoplasty and their operative management. *Facial Plast Surg*. (2024) 40(4):485–92. doi: 10.1055/s-0043-1772846
- McGuire C, Samargandi OA, Boudreau C, Whelan A, Bezuhly M. Prevention of autologous costal cartilage graft warping in secondary rhinoplasty. *J Craniofac Surg*. (2020) 31(5):1246–50. doi: 10.1097/SCS.00000000000006429
- Kim SJ, Chang DS, Choi MS, Lee HY, Pyo J-S. Efficacy of nasal septal splints for preventing complications after septoplasty: a meta-analysis. *Am J Otolaryngol*. (2021) 42(3):102389. doi: 10.1016/j.amjoto.2020.102389
- Morita D, Kawarazaki A, Soufi M, Otake Y, Sato Y, Numajiri T. Automatic detection of midfacial fractures in facial bone CT images using deep learning-based object detection models. *J Stomatol Oral Maxillofac Surg*. (2024) 125(5S2):101914. doi: 10.1016/j.jormas.2024.101914
- Maita KC, Avila FR, Torres-Guzman RA, Garcia JP, De Sario Velasquez GD, Borna S, et al. The usefulness of artificial intelligence in breast reconstruction: a systematic review. *Breast Cancer*. (2024) 31(4):562–71. doi: 10.1007/s12282-024-01582-6
- Choi H, Moran J, Matsumoto N, Hernandez ME, Moore JH. Aliro: an automated machine learning tool leveraging large language models. *Bioinformatics*. (2023) 39(10):btad606. doi: 10.1093/bioinformatics/btad606
- Li Y, Dong M, Wang Y, Xu C. Neural architecture search via proxy validation. *IEEE Trans Pattern Anal Mach Intell*. (2023) 45(6):7595–610. doi: 10.1109/TPAMI.2022.3217648
- Guo Z, Liu H, Ong YS, Qu X, Zhang Y, Zheng J. Generative multiform Bayesian optimization. *IEEE Trans Cybern*. (2023) 53(7):4347–60. doi: 10.1109/TCYB.2022.3165044
- Sena Esteves S, Gonçalves Ferreira M, Carvalho Almeida J, Abrunhosa J, Almeida e Sousa C. Evaluation of aesthetic and functional outcomes in rhinoplasty surgery: a prospective study. *Braz J Otorhinolaryngol*. (2017) 83(5):552–7. doi: 10.1016/j.bjorl.2016.06.010
- Ji J, Wu T, Yang C. Neural population dynamics optimization algorithm: a novel brain-inspired meta-heuristic method. *Knowl Based Syst*. (2024) 300:112194. doi: 10.1016/j.knsys.2024.112194
- Gao C, Zhang R, Chen X, Yao T, Song Q, Ye W, et al. Integrating internet multisource big data to predict the occurrence and development of COVID-19 cryptic transmission. *NPJ Digit Med*. (2022) 5(1):161. doi: 10.1038/s41746-022-00704-8
- Sharma P, Raju S. Metaheuristic optimization algorithms: a comprehensive overview and classification of benchmark test functions. *Soft Comput*. (2024) 28(4):3123–86. doi: 10.1007/s00500-023-09276-5
- Naoum GE, Ho AY, Shui A, Salama L, Goldberg S, Arafat W, et al. Risk of developing breast reconstruction complications: a machine-learning nomogram for individualized risk estimation with and without postmastectomy radiation therapy. *Plast Reconstr Surg*. (2022) 149(1):1e–12e. doi: 10.1097/PRS.00000000000008635
- Vickers AJ, Holland F. Decision curve analysis to evaluate the clinical benefit of prediction models. *Spine J*. (2021) 21(10):1643–8. doi: 10.1016/j.spinee.2021.02.024
- Nocini R, Chirumbolo S, Pirayesh A, Rahman E, Kapoor KM, Caliskan G, et al. A systematic review and meta-analysis of rhinoplasty using the rhinoplasty outcome evaluation scale. *Ann Maxillofac Surg*. (2022) 12(1):60–8. doi: 10.4103/ams.ams\_244\_21
- Nagra AA, Alyas T, Hamid M, Tabassum N, Ahmad A. Training a feedforward neural network using hybrid gravitational search algorithm with dynamic multiswarm particle swarm optimization. *Biomed Res Int*. (2022) 2022(1):2636515. doi: 10.1155/2022/2636515
- Lee M, Yeo NY, Ahn HJ, Lim J-S, Kim Y, Lee S-H, et al. Prediction of post-stroke cognitive impairment after acute ischemic stroke using machine learning. *Alzheimers Res Ther*. (2023) 15(1):147. doi: 10.1186/s13195-023-01289-4
- Dong W, Wang X, Zhao Z, Fan F. Viability and stability of various-sized free diced cartilage in an experimental study. *Aesthet Surg J*. (2023) 43(3):NP143–54. doi: 10.1093/asj/sjac263
- Cavagnero KJ, Li F, Dokoshi T, Nakatsuji T, O'Neill AM, Aguilera C, et al. CXCL12+ dermal fibroblasts promote neutrophil recruitment and host defense by recognition of IL-17. *J Exp Med*. (2024) 221(4):e20231425. doi: 10.1084/jem.20231425
- Huang CJ, Pu CM, Su SY, Lo S-L, Lee CH, Yen Y-H. Improvement of wound healing by capsaicin through suppression of the inflammatory response and amelioration of the repair process. *Mol Med Rep*. (2023) 28(2):155. doi: 10.3892/mmr.2023.13042
- Eggerstedt M, Rhee J, Buranosky M, Batra PS, Tajudeen BA, Smith RM, et al. Nasal skin and soft tissue thickness variation among differing races and ethnicities: an objective radiographic analysis. *Facial Plast Surg Aesthet Med*. (2020) 22(3):188–94. doi: 10.1089/fpsam.2019.0008
- Aung YYM, Wong DCS, Ting DSW. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. *Br Med Bull*. (2021) 139(1):4–15. doi: 10.1093/bmb/ldab016