

Early Prediction of Hemotoxic Snakebite Based on Stratified Machine Learning: A Multicenter Retrospective Study

Qinfen Gao

First People's Hospital of Yunnan Province

Lei Cai

Kunming Medical University

Yajun Teng

First People's Hospital of Yunnan Province

Rui Zeng

People's Hospital of Linxiang District, Lincang

Xiaoyan Li

Baoshan People's Hospital, Baoshan

Ying Geng

Lincang People's Hospital

Shegang Li

Dali Bai Autonomous Prefecture People's Hospital

Chengyi Yang

Zhaotong First People's Hospital

Jianhai Wang

The First Affiliated Hospital of Dali University

Qinrong Gao

Eryuan County Hospital of Traditional Chinese Medicine

Ling Yang

Kunming Medical University

Wei Lu

Kunming Medical University

Jun Huang

Kunming Medical University

Xiao Luo

Kunming Medical University

Hua Xu

The First People's Hospital of Lancang Lahu Autonomous County

Zhonghui Ma

People's Hospital of Linxiang District, Lincang

Wei Li

The First People's Hospital of Lancang Lahu Autonomous County

Yuhong Feng

People's Hospital of Linxiang District, Lincang

Xia Zhang

Dali Bai Autonomous Prefecture People's Hospital

Jie Wang

People's Hospital of Linxiang District, Lincang

Ling Zhu

First People's Hospital of Yunnan Province

Zengzheng Li

First People's Hospital of Yunnan Province

Wei Zhang

zhangwei719719@126.com

First Affiliated Hospital of Kunming Medical University

Research Article

Keywords: snakebite, hemotoxicity, machine learning, prediction model, stratified modeling, multicenter study

Posted Date: October 12th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10702738/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Early Prediction of Hemotoxic Snakebite Based on Stratified Machine Learning: A Multicenter Retrospective Study

Authors

Qinfen Gao^{1,2‡}, Lei Cai^{2,3‡}, Yajun Teng^{1‡}, Rui Zeng⁴, Xiaoyan Li⁵, Ying Geng⁶, Shegang Li⁷, Chengyi Yang⁸, Jianhai Wang⁹, Qinrong Gao¹⁰, Ling Yang^{2,3}, Wei Lu^{2,3}, Jun Huang^{2,3}, Xiao Luo^{2,3}, Hua Xu¹¹, Zhonghui Ma⁴, Wei Li¹¹, Yuhong Feng⁴, Xia Zhang⁷, Jie Wang⁴, Ling Zhu^{1*}, Zengzheng Li^{1*}, Wei Zhang^{3*}

Affiliations

- 1.The First People's Hospital of Yunnan Province, Affiliated Hospital of Kunming University of Science and Technology, Kunming, Yunnan, China
- 2.Kunming Medical University, Kunming, Yunnan, China
- 3.The First Affiliated Hospital of Kunming Medical University, Kunming, Yunnan, China
- 4.People's Hospital of Linxiang District, Lincang, Yunnan, China
- 5.Baoshan People's Hospital, Baoshan, Yunnan, China
- 6.Lincang People's Hospital, Lincang, Yunnan, China
- 7.Dali Bai Autonomous Prefecture People's Hospital, Dali, Yunnan, China
- 8.Zhaotong First People's Hospital, Zhaotong, Yunnan, China
- 9.The First Affiliated Hospital of Dali University, Dali, Yunnan, China
- 10.Eryuan County Hospital of Traditional Chinese Medicine, Dali, Yunnan, China
- 11.The First People's Hospital of Lancang Lahu Autonomous County, Pu'er, Yunnan, China

Author Notes

‡ These authors contributed equally to this work.

* Corresponding authors. Wei Zhang is the lead corresponding author, email: zhangwei719719@126.com

Abstract

Background Snakebite is an important public health problem in tropical and subtropical regions, and the burden is particularly heavy in areas with rich venomous-snake diversity, such as Yunnan. Atypical clinical manifestations and difficulty in identifying the snake species often lead to delayed or inappropriate antivenom therapy for hemotoxicity-related conditions, such as venom-induced consumption coagulopathy. Developing early prediction models from routine clinical data with machine learning, and making them applicable across different diagnostic and treatment settings, is essential for precise classification of snakebite.

Methods This multicenter retrospective study included 2,094 patients with snakebite from 34 hospitals in Yunnan Province. Because the completeness of clinical data at admission varied substantially, we adopted a stratified modeling strategy: (1) a Class A basic model (n=2,094), incorporating a limited set of variables such as physical signs and coagulation indices to simulate primary-care settings; and (2) a Class B complete model (n=555), incorporating more comprehensive laboratory tests to simulate resource-rich settings. After multiple imputation and SMOTE oversampling, five algorithms, namely random forest (RF), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), LightGBM, and logistic regression (LR), were used to construct hemotoxicity prediction models. Model performance was comprehensively evaluated using AUC, calibration curves, and decision curve analysis, and feature importance was interpreted using SHAP.

Results In the Class A models, RF showed the best performance (AUC=0.838) and was superior to LR (AUC=0.746). In the Class B models, RF remained the best-performing model (AUC=0.894). Decision curve analysis showed that the RF model provided positive net benefit across a broad range of clinical decision thresholds. SHAP analysis indicated that high D-dimer, low fibrinogen, and specific bite locations (for example, grassland/shrubland increasing the risk of hemotoxicity, whereas bites near roads or residences increased the probability of non-hemotoxicity) were key predictors of hemotoxicity. In external validation with 51 patients, the AUCs of the two RF models were 0.709 and 0.731, respectively, indicating moderate external discrimination.

Conclusions This study developed and validated a set of early prediction models for hemotoxic snakebite based on stratified machine learning. This tool can provide precise risk stratification according to different levels of medical-resource availability and support clinical decision-making throughout the pathway from primary screening to in-hospital confirmation. It may optimize the early and precise use of antivenom and offer an innovative "Yunnan approach" for reducing the burden of snakebite.

Keywords snakebite; hemotoxicity; machine learning; prediction model; stratified modeling; multicenter study

Introduction

Snakebite envenoming is a serious but often neglected public health problem. Globally, approximately 6.85 billion people live within the distribution range of at least one common venomous snake species [1]. The global incidence of snakebite reaches 69.4 per 100,000 population, with a mortality rate of approximately 0.33 per 100,000; in Asia, these figures are even higher, reaching an incidence of 130.7 per 100,000 and a mortality rate of 0.7 per 100,000 [2]. Each year, approximately 5.4 million snakebite events occur worldwide, causing about 81,000 to 138,000 deaths and leaving around 400,000 people with permanent physical and psychological disability [3], thereby imposing a substantial economic burden on society.

Snake venom has complex components and can be broadly classified as neurotoxic, hemotoxic, cytotoxic, or mixed venom. Different venom types produce different clinical manifestations, but all may cause severe complications and even death [4][5]. Clinical progression after snakebite, especially venomous snakebite, can be rapid, and the therapeutic window is narrow. Treatment delay is the most important determinant of poor outcomes after snakebite [6], and timely and accurate administration of antivenom is crucial for improving prognosis [4][5]. Therefore, rapid and accurate identification of the offending snake species and venom type is a key step in targeted antivenom therapy and prognostic improvement. At present, diagnosis of snakebite mainly relies on the patient's description of the snake, local and systemic clinical manifestations, and laboratory testing [4]. In clinical practice, however, snake species are often difficult to identify, and venom type may be judged late because symptoms are delayed or atypical. This can result in inappropriate use, overuse, or delayed use of antivenom, increasing both medical burden and patient risk.

In recent years, with the development of artificial intelligence, machine learning has been increasingly applied in medicine. Machine learning can capture nonlinear relationships in multidimensional clinical data that are difficult to identify using traditional research methods, thereby supporting rapid clinical decision-making or rapid classification of multiclass samples [7][8]. Notably, although AI/ML has been widely used in medical imaging, chronic disease management, and other fields [9], its use in the early diagnosis of acute poisoning such as snakebite remains at an exploratory stage.

This study aimed to use machine-learning algorithms to construct stratified binary prediction models for venom type after snakebite (hemotoxic vs. non-hemotoxic) based on basic information such as symptoms and signs and simple laboratory tests. Through a data-driven approach, we sought to provide clinicians with an auxiliary diagnostic tool, optimize early precision intervention for snakebite, and provide new technical and theoretical support for improving patient prognosis and allocating medical resources rationally.

2 Methods

2.1 Data Collection

This retrospective multicenter study collected basic information, clinical examinations, diagnoses, and treatment courses for all snakebite patients from 34 hospitals across 16 prefectures and municipalities in Yunnan Province (see Supplementary List 1). A total of 2,816 clinical records were obtained, covering all prefecture- and municipal-level administrative regions in Yunnan Province.

After the collection was completed, the following patients were excluded: (1) those who did not clearly record the bite snake species or venom type in their medical records ($n=579$); (2) those who did not undergo any laboratory examination after admission ($n=133$). Finally, 2104 patients were included in subsequent analysis.

This retrospective study was conducted in accordance with the Declaration of Helsinki and has obtained approval from the ethics committees of the First Affiliated Hospital of Kunming Medical University, the First People's Hospital of Yunnan Province and other relevant hospitals. Considering the characteristics of the retrospective study, all ethics committees waived the requirement for written confirmation of patient informed consent.

2.2 Text Processing of Symptoms and Signs and Feature Encoding

Patients' symptoms and signs were obtained from free-text descriptions in electronic medical records. Before machine-learning feature encoding, these free-text symptoms and signs were standardized. First, R was used to extract high-frequency symptom/sign terms (top 50). Clinical experts then merged different expressions referring to the same symptom or sign and standardized their coding. Finally, six representative symptom/sign variables were selected for subsequent modeling analyses.

Continuous variables and non-mutually exclusive categorical variables were standardized after measurement units were harmonized. Mutually exclusive categorical variables, such as bite location, were one-hot encoded, and variables with more than 30% missing data were excluded. Laboratory values above an instrument's upper detection limit were replaced with the corresponding upper-limit value, whereas values below the lower detection limit were replaced with the corresponding lower-limit value, ensuring consistent and clinically reasonable processing across centers.

2.3 Dataset Characterization and Quality Control

After discussion among several clinical experts and assessment of dataset characteristics,

we noted first that approximately 86% of cases in the dataset were hemotoxic, whereas cytotoxic, neurotoxic, mixed, and other venom types together accounted for only about 14%. Some venom types were extremely rare; for example, neurotoxic cases accounted for only about 1%. Forcing supervised multiclass machine learning on these sparse categories would easily lead to severe overfitting. Therefore, we merged all venom types other than hemotoxic venom into a single non-hemotoxic category and performed supervised binary machine-learning modeling. Second, all patients in the dataset had clinical symptoms and signs, but most patients (approximately 73%) had only coagulation testing at admission, whereas a smaller proportion (approximately 27%) had relatively complete admission examination data. We therefore performed stratified analysis and machine-learning modeling, as described below.

To control data quality and identify potential outlier samples, we performed unsupervised K-means clustering on all samples to help detect individuals with potentially abnormal distributions. For samples whose cluster labels were inconsistent with the original clinical labels, clinical experts reviewed the original medical records one by one and made comprehensive judgments based on the diagnostic and treatment course to assess the reliability and authenticity of the clinical labels. Finally, 10 samples (approximately 0.47%) with highly suspected label errors were excluded, and 2,094 patients entered the machine-learning stage.

2.4 Stratified Modeling Design, Dataset Splitting, and Preprocessing

Considering the differences in completeness of laboratory tests between different patients in the original dataset, and in line with the principle of maximizing the mining of dataset information, this study decided to use a stratified modeling strategy to establish Class A models and Class B models respectively:

Class A Model: Basic Model.

The basic model was constructed using all included patients ($n=2,094$). Because 73% of patients had only coagulation tests and many other laboratory indices were missing, variables with a missing rate greater than 30% were removed according to the variable-processing principles described above (Section 2.2). Therefore, the basic model included only basic demographic information, clinical symptoms and signs, and coagulation-related indices. This model was intended to simulate clinical application scenarios in primary-care or resource-limited settings.

Class B Model: Complete Model

The complete model was constructed using patients in the complete dataset with

relatively comprehensive examination data (n=555). Variables added to the complete model included basic demographic information, clinical symptoms and signs, coagulation function, routine blood tests, liver and kidney function, blood gas analysis, and related tests. This model was intended to simulate application scenarios in which comprehensive laboratory testing and sufficient medical resources are available.

After the stratified modeling design was defined, all data were split into training and validation sets (8:2) using stratified sampling according to the proportions of hemotoxic and non-hemotoxic cases, ensuring that the class distribution was consistent with the real-world distribution. In the training set, missing values were handled by multiple imputation, and SMOTE was used to oversample the minority class (non-hemotoxic) to alleviate class imbalance.

2.5 Feature Selection and Model Training

For the Class A model, because the number of variables was small, all candidate variables were included in the initial training process, followed by structural optimization and screening based on model-derived feature importance. For the Class B model, because there were more variables, feature screening was performed using multiple strategies in parallel: (1) LASSO regression and univariable logistic regression were used to screen features ($P < 0.1$ was used as a liberal screening threshold to avoid missing potentially important variables); (2) clinical relevance and physiological plausibility were assessed with input from clinical experts; and (3) for laboratory indices of the same type, such as liver-function enzyme profiles, one to two representative variables were preferentially selected for model inclusion to avoid uninterpretable results caused by excessive collinearity among similar tests.

Both classes of models were trained and compared using the following algorithms: random forest (RF), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), LightGBM, and multivariable logistic regression (as an interpretable baseline model). The area under the receiver operating characteristic curve (AUC) was used as the optimization objective. Bayesian optimization (for RF) and Latin hypercube sampling random search (for gradient-boosting models) were used to tune the hyperparameters of tree-based models to achieve optimal performance.

2.6 Model Evaluation and Result Analysis

Internal and external validation were performed for all Class A and Class B models. Model performance was evaluated from multiple dimensions: ROC curves and AUC were used to assess discrimination between hemotoxic and non-hemotoxic cases; calibration curves

were plotted to evaluate agreement between predicted probabilities and observed incidence and to help assess prediction bias; prediction-probability density plots were used to analyze differences in prediction distributions between classes; SHAP (SHapley Additive exPlanations) was used to quantify the contribution of each feature to model predictions; and decision curve analysis was used to evaluate net benefit at different threshold probabilities, thereby assessing the potential clinical value of the models in real-world decision-making.

The research is conducted according to the flowchart (Fig 1), All statistical and visualization work was done using R (version: 4.5.3).

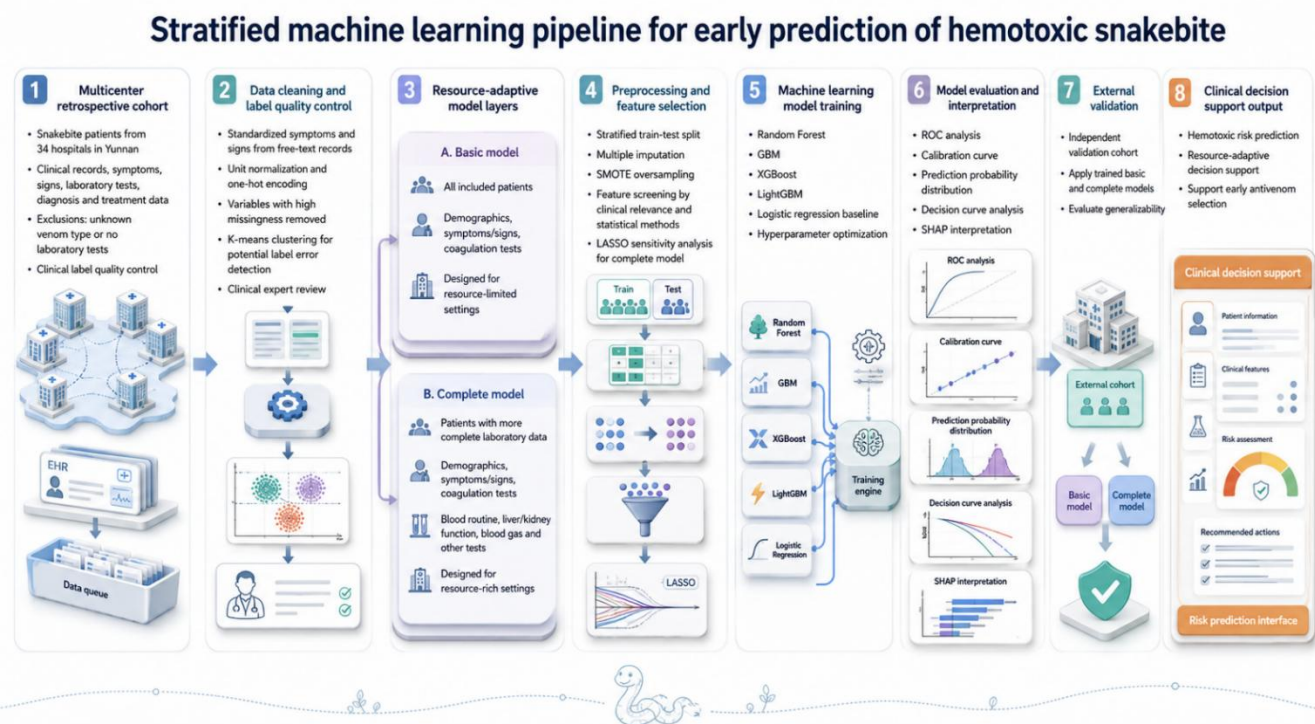


Fig. 1 Study Flowchart.

3 Results

3.1 Data Quality Control and Overall Patient Baseline Characteristics

After data collection and preliminary cleaning, we considered that snakebite is relatively uncommon compared with other critical illnesses and that patients may have confused recall during emergencies, while physicians may fail to accurately identify venom type. We therefore used K-means clustering for unsupervised clustering of all samples to help detect individuals with potentially abnormal distributions (Fig. 2). For patients whose cluster labels were inconsistent with clinical labels, clinical experts reviewed the original electronic medical records and clinical data one by one. Finally, 10 patients (approximately 0.47%) with highly suspected clinical-label errors were excluded, and 2,094 patients were included in the subsequent machine-learning stage.

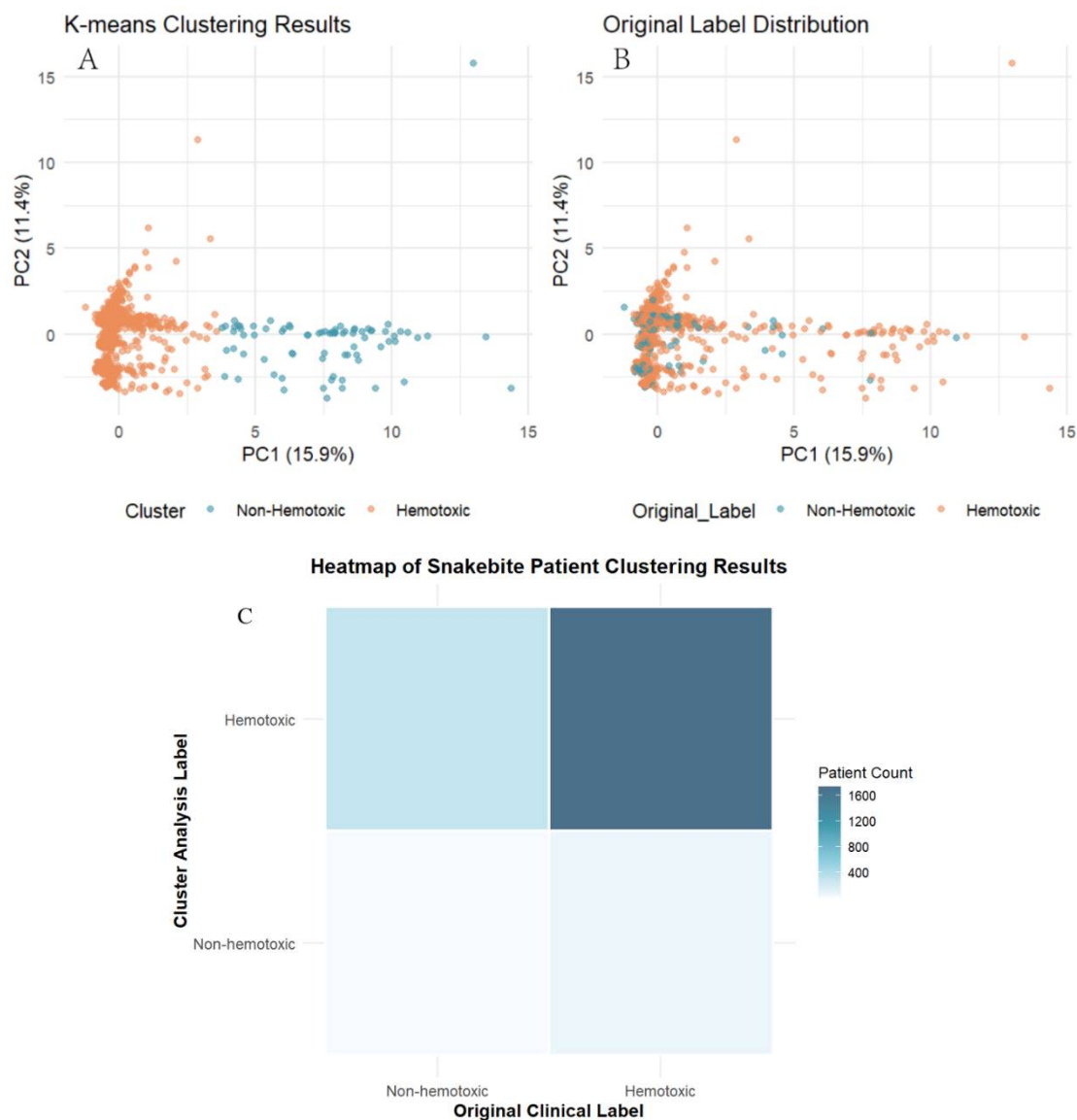


Fig. 2 K-means Clustering Plot. A, K-means clustering results; B, original clinical labels; C, clustering heatmap.

Baseline characteristics and clinical features of the 2,094 snakebite patients included in this study were analyzed, and the results are shown in Table 1. Baseline data showed that PT, TT, and DD2 were higher in hemotoxic patients than in non-hemotoxic patients, whereas FIB was lower; all differences were statistically significant. Notably, although these differences were statistically significant, most values remained within the normal range. Regarding symptoms and signs, pain was more likely to occur after non-hemotoxic snakebite, whereas swelling was more likely after hemotoxic snakebite. Other symptoms and signs did not differ significantly between the two groups.

Table 1 Overall Patient Baseline Characteristics and Admission Laboratory Tests

Variable	Total	Non-hemotoxic	Hemotoxic	P value
N	2094	295	1799	
Age (years)	55.0 [44.0, 65.0]	54.0 [38.5, 62.5]	56.0 [45.0, 65.0]	0.015
Male	1183 (56.5%)	177 (60.0%)	1006 (55.9%)	0.212
Bite site: dry land	1446 (69.1%)	206 (69.8%)	1240 (68.9%)	0.808
Bite site: paddy field	467 (22.3%)	60 (20.3%)	407 (22.6%)	0.425
Swelling	1520 (72.6%)	197 (66.8%)	1323 (73.5%)	0.019
Pain	1190 (56.8%)	218 (73.9%)	972 (54.0%)	<0.001
Bleeding	149 (7.1%)	17 (5.8%)	132 (7.3%)	0.394
Numbness	60 (2.9%)	10 (3.4%)	50 (2.8%)	0.693
Digestive symptoms	70 (3.3%)	14 (4.7%)	56 (3.1%)	0.204
Neurological symptoms	46 (2.2%)	10 (3.4%)	36 (2.0%)	0.196
PT (s)	12.0 [11.1, 13.3]	11.6 [10.8, 12.9]	12.0 [11.2, 13.5]	<0.001
INR	1.0 [1.0, 1.1]	1.0 [0.9, 1.1]	1.0 [1.0, 1.1]	<0.001
TT (s)	18.3 [16.7, 21.2]	18.0 [16.8, 19.7]	18.4 [16.7, 21.7]	0.021
APTT (s)	27.8 [24.8, 32.5]	27.1 [24.9, 32.8]	27.9 [24.8, 32.4]	0.658
FIB (g/L)	2.1 [1.2, 2.7]	2.4 [1.8, 2.9]	2.0 [1.1, 2.6]	<0.001
FDP (mg/L)	2.2 [0.8, 5.3]	2.1 [0.6, 5.0]	2.2 [0.8, 5.3]	0.196
DD2 (mg/L)	1.8 [0.7, 4.3]	1.1 [0.4, 2.5]	1.9 [0.8, 4.6]	<0.001

3.2 Class A Model Development and Analysis

To predict the probability of hemotoxicity early after snakebite, we selected five machine-learning algorithms to build prediction models: RF, GBM, XGBoost, LightGBM, and LR. LR was used as an interpretable baseline model for comparison with the other models. Because the Class A model contained relatively few variables, all variables were included as features in machine-learning training, followed by structural optimization and screening based on model-derived feature importance.

Because the data distribution was imbalanced (hemotoxic cases accounted for approximately 85%), we used stratified sampling to split the dataset into training and test sets. Multiple imputation and SMOTE oversampling were applied to the training data to mitigate imbalance, and the models trained on the training set were then tested in the test set. AUC was used to compare model performance.

The test-set results are shown in Table 2 and Fig. 3. RF was the best-performing model, with the highest AUC (0.838) and the highest specificity (0.854). LR was the worst-performing model, with the lowest AUC (0.746), suggesting that under resource-limited conditions LR cannot capture the complex nonlinear relationships in the data. The AUCs of GBM and XGBoost were similar; GBM had higher sensitivity, whereas XGBoost had higher specificity.

Table 2 Details of Class A Models

Model	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
RF	0.838	0.705	0.854	0.722	0.974	0.273
GBM	0.811	0.795	0.688	0.782	0.951	0.303
XGBoost	0.776	0.624	0.833	0.648	0.966	0.223
LightGBM	0.830	0.800	0.792	0.799	0.967	0.339
LR	0.746	0.678	0.771	0.689	0.958	0.237

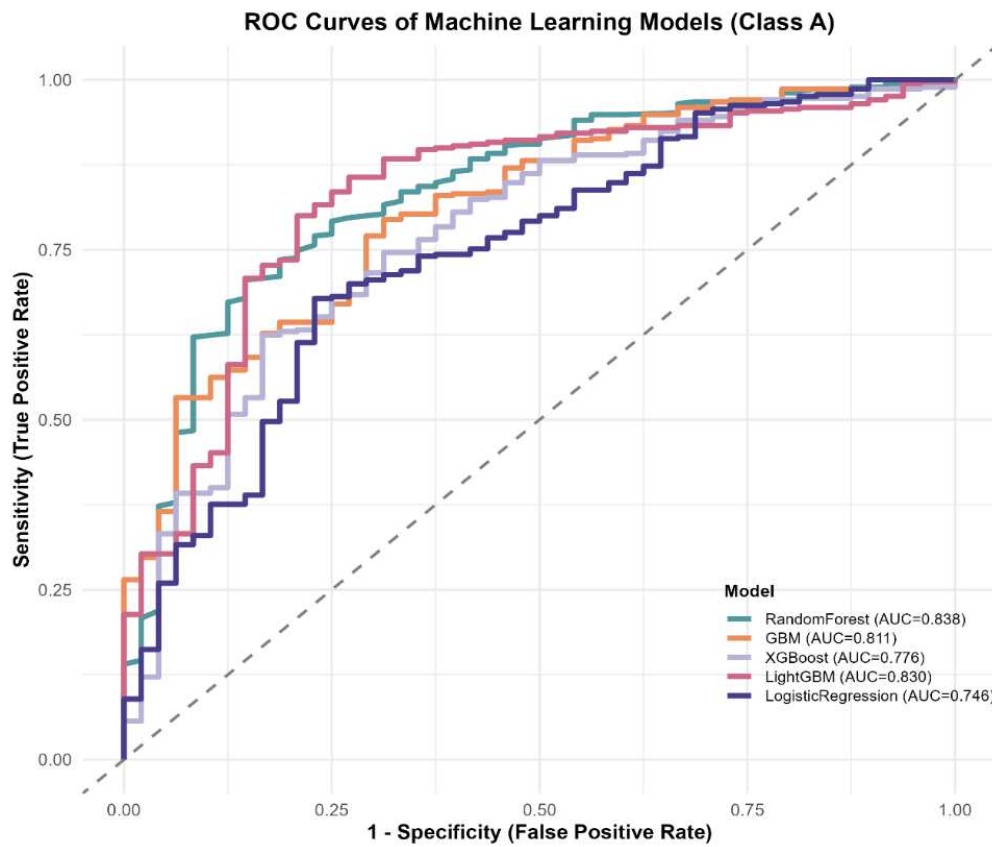


Fig. 3 ROC Curves of Class A Models

We plotted calibration curves, density plots, and decision curves for each model. The calibration curve shows the accuracy of the probability output for each model. As shown in Fig. 4, the ideal calibration curve is a diagonal line. Among all models, the calibration performance of RF is relatively the best because it is closest to the ideal curve most of the time. Similarly, the calibration performance of LR is the worst. The calibration curve results show that most models (RF, LightGBM, GBM and LR) show an overestimation of the actual risk throughout the entire process, while XGBoost shows an underestimation of the actual risk in the high probability prediction interval.

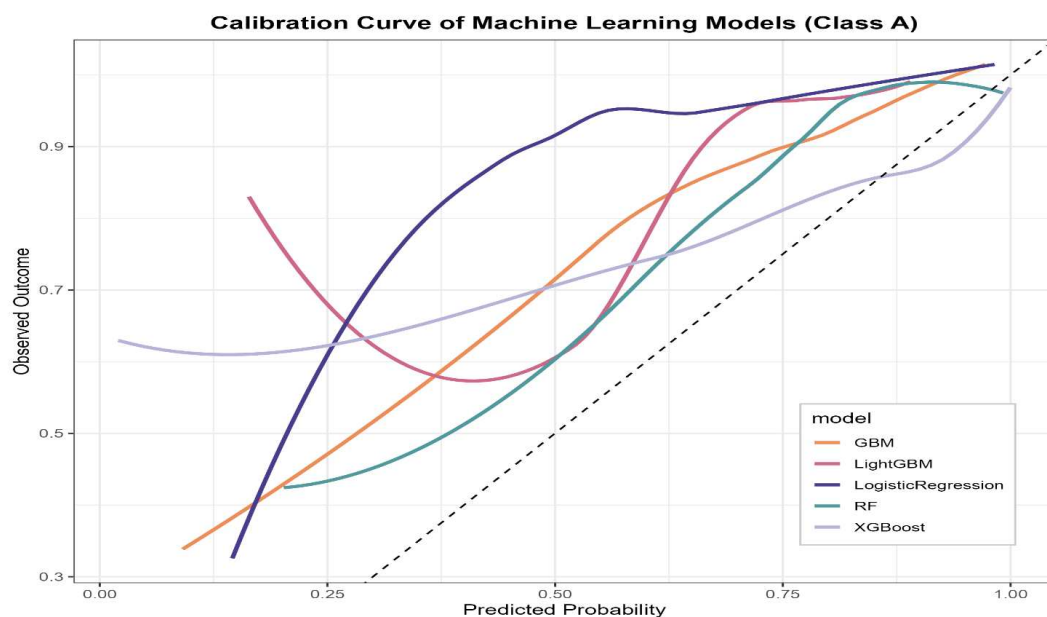


Fig. 4 Calibration Curves of Class A Models

We plotted density curves for each model with predicted probability on the x-axis and distribution density on the y-axis. Different models showed markedly different prediction patterns (Fig. 5). XGBoost showed an extreme unimodal distribution, suggesting that the model tended to make deterministic predictions. The predicted probabilities of RF, GBM, and LightGBM were more dispersed, with substantial overlap in the intermediate-probability range, reflecting uncertainty in classification decisions. Logistic regression showed obvious overlapping bimodality, suggesting that under basic conditions it could not accurately distinguish hemotoxic patients.

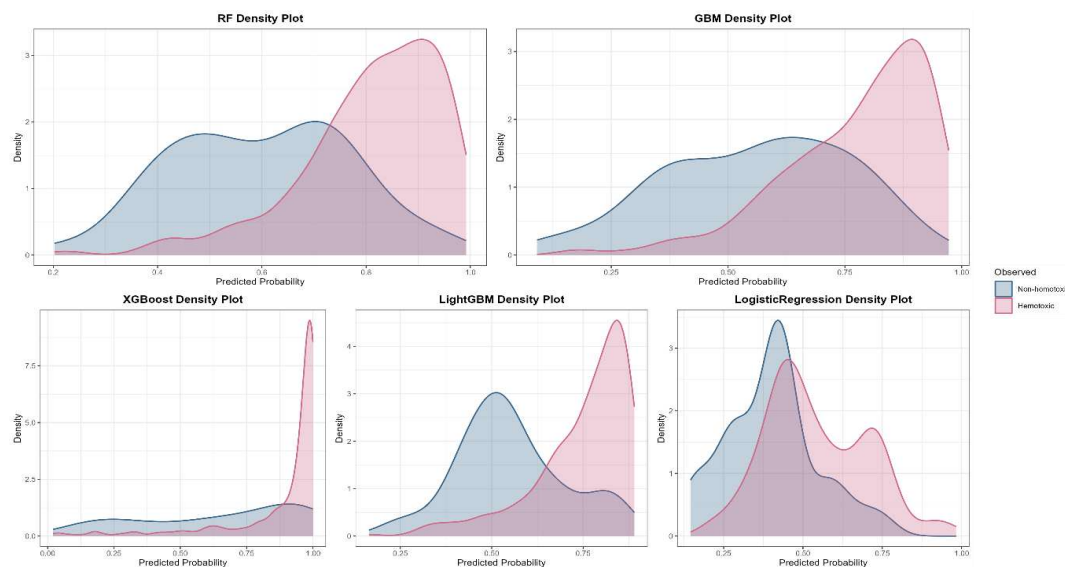


Fig. 5 Density Distribution Plots of Class A Models

We also plotted decision curves for each model (Fig. 6). Decision curves use decision threshold as the x-axis and net benefit as the y-axis, showing the net benefit of each model compared with treating all patients or treating none, thereby assisting clinical decision-making. In the low-threshold range (0-0.25), none of the models provided substantial benefit compared with treating all patients. This finding was attributable to the imbalanced sample distribution (hemotoxic cases accounted for 85%); at low decision thresholds, treating all patients already covered most hemotoxic cases, so prediction models could not show obvious clinical net benefit. As the decision threshold increased, most models showed net benefit compared with treating all patients. Notably, logistic regression showed clear clinical net benefit only when the decision threshold exceeded approximately 0.8, whereas RF showed greater clinical net benefit than treating all patients throughout the range.

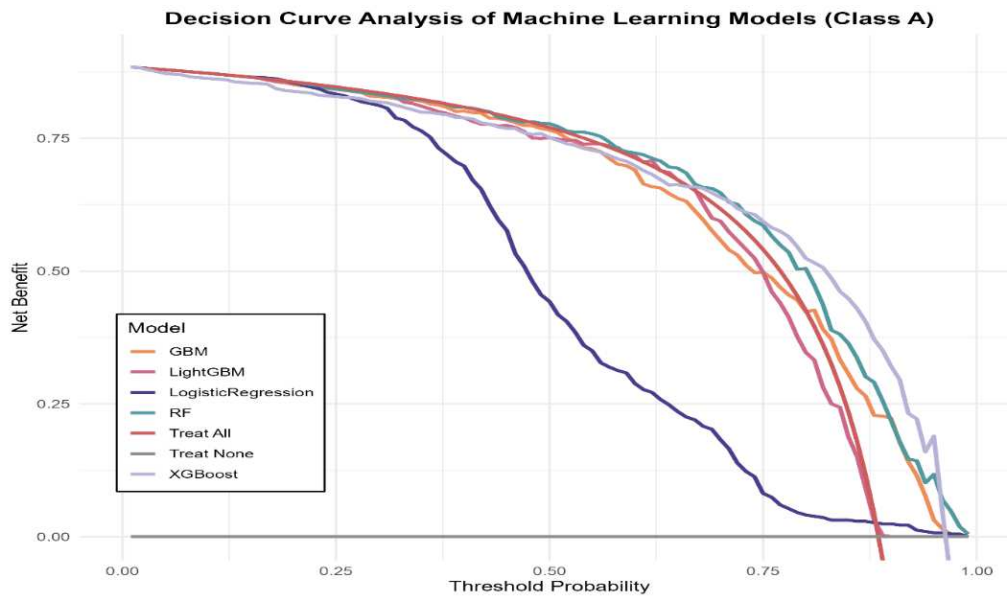
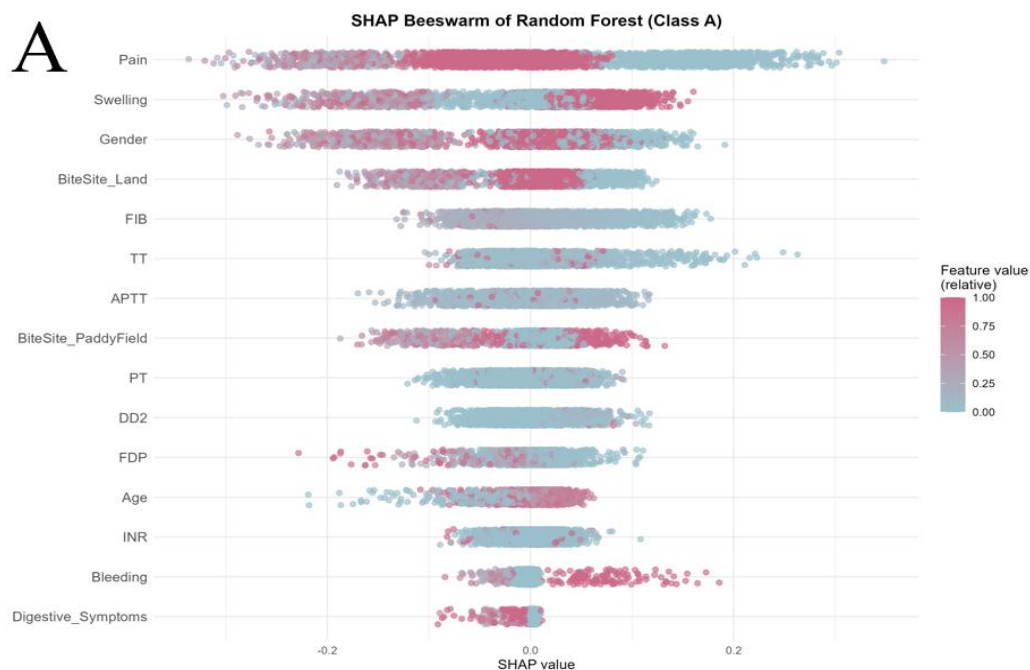


Fig. 6 Decision Curves of Class A Models

Based on the AUC, calibration curve, and decision curve of each model, we considered RF to be the best Class A model. It had the highest AUC and the best calibration curve, and the decision curve indicated that it provided better clinical net benefit than treating all patients over almost the entire threshold range. We performed SHAP analysis for RF and summarized the importance contribution of each feature in the model. The results are shown in the beeswarm plot (Fig. 7A) and dependence plot (Fig. 7B). Among basic clinical features, pain, swelling, bite location_land, and bite location_paddy field were the most important contributors. Among objective laboratory data, DD2, APTT, FIB, TT, and FDP were the most important laboratory tests.



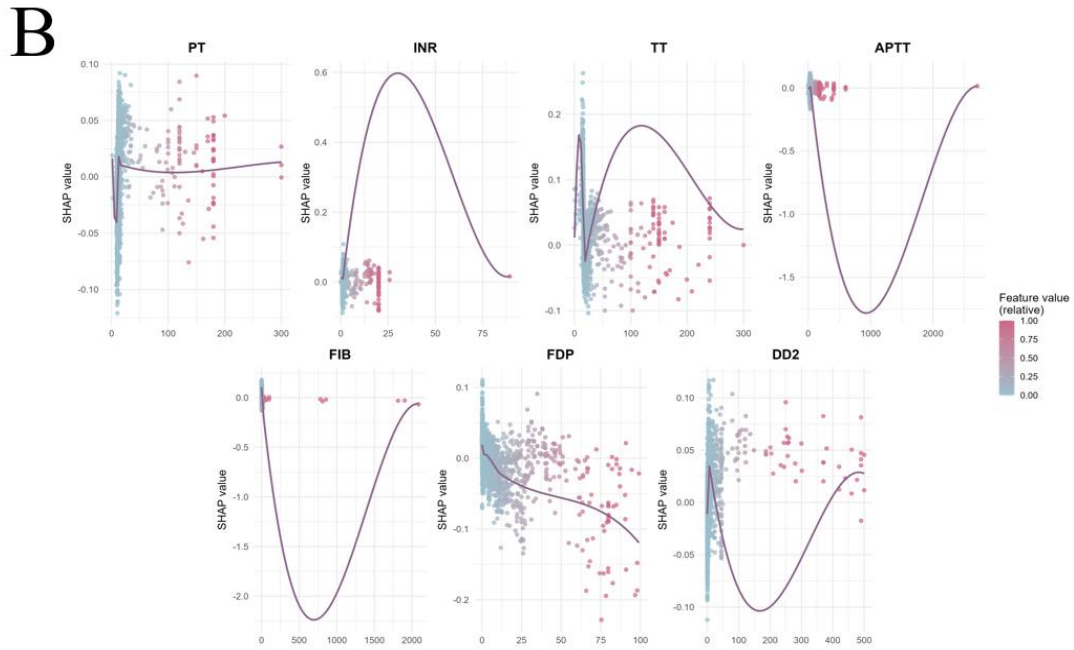


Fig. 7 SHAP Analysis of the Class A Model: Beeswarm Plot (A) and Dependence Plot (B)

3.3 Baseline Analysis of the Class B Model Dataset

We performed a separate subgroup analysis of patients with more detailed clinical data and laboratory tests on admission to fully exploit the potential of the dataset. The results of baseline data analysis are shown in Table 3. Due to too much clinical data, only part of the information is shown in the text. For detailed information, see Table 1 in the supplementary material. Baseline data showed that forests or orchards, grass or shrubs were always the more common injury locations for patients with hemotoxicity, while roads or near residences were the most common injury locations for patients with non-hemotoxicity.

Table 3 Subgroup Baseline Characteristics and Admission Laboratory Tests (Partial)

Variable	Total	Non-hemotoxic	Hemotoxic	P value
N	555	118	437	
Age (years)	54.0 [43.0, 63.0]	52.0 [39.0, 61.0]	55.0 [45.0, 63.0]	0.134
Male	303 (54.6%)	73 (61.9%)	230 (52.6%)	0.092
Time to hospital (h)	4.0 [2.0, 7.8]	4.0 [2.0, 8.0]	4.0 [2.0, 7.0]	0.824
Bite site: residence or road	75 (13.5%)	31 (26.3%)	44 (10.1%)	<0.001
Bite site: mountain forest or orchard	67 (12.1%)	7 (5.9%)	60 (13.7%)	0.032
Bite site: grass or bushes	66 (11.9%)	6 (5.1%)	60 (13.7%)	0.016
Tooth mark: double point	459 (82.7%)	82 (69.5%)	377 (86.3%)	<0.001

Variable	Total	Non-hemotoxic	Hemotoxic	P value
Tooth mark: double rows	43 (7.7%)	16 (13.6%)	27 (6.2%)	0.014
Heart rate (bpm)	76.0 [65.0, 88.0]	80.0 [65.0, 92.2]	74.5 [65.0, 87.0]	0.006
WBC (x10 ⁹ /L)	8.8 [6.9, 11.2]	8.5 [6.7, 11.7]	8.9 [6.9, 11.0]	0.679
Hematocrit (%)	0.4 [0.4, 0.5]	0.4 [0.4, 0.5]	0.4 [0.4, 0.5]	0.452
Platelet count (x10 ⁹ /L)	233.0 [189.0, 282.0]	228.0 [189.0, 273.5]	234.0 [189.0, 282.0]	0.735
APTT (s)	30.9 [25.8, 35.5]	30.2 [25.4, 35.8]	31.1 [26.0, 35.1]	0.683
PT (s)	12.1 [11.0, 13.8]	12.0 [11.0, 13.7]	12.1 [11.0, 13.9]	0.652
INR	1.0 [1.0, 1.1]	1.0 [0.9, 1.1]	1.0 [1.0, 1.1]	0.093
TT (s)	18.4 [17.1, 19.9]	18.2 [16.9, 19.5]	18.4 [17.1, 20.4]	0.070
FIB (g/L)	2.2 [1.6, 2.7]	2.4 [1.9, 3.0]	2.1 [1.5, 2.6]	<0.001
FDP (mg/L)	0.8 [0.4, 3.1]	0.6 [0.4, 2.3]	0.9 [0.4, 3.2]	0.022
DD2 (mg/L)	4.5 [1.8, 12.4]	1.8 [1.4, 4.2]	5.1 [2.5, 16.8]	<0.001
AST (U/L)	24.0 [20.0, 29.9]	25.0 [20.9, 33.3]	24.0 [20.0, 29.4]	0.121
TBIL (umol/L)	11.4 [8.0, 15.0]	11.1 [8.5, 15.1]	11.4 [7.8, 15.0]	0.939
Globulin (g/L)	27.3 [24.8, 30.3]	27.1 [23.9, 29.9]	27.4 [24.9, 30.4]	0.123
Cr (umol/L)	66.0 [56.0, 79.5]	64.0 [52.1, 77.0]	66.0 [57.0, 80.0]	0.125
BUN (mmol/L)	5.9 [4.8, 7.2]	5.8 [4.6, 7.1]	5.9 [4.8, 7.3]	0.556
eGFR (mL/min)	87.0 [65.2, 107.3]	89.4 [69.0, 109.5]	84.4 [64.8, 106.9]	0.198
Ca (mmol/L)	2.3 ± 0.1	2.3 ± 0.2	2.3 ± 0.1	0.087
pH	7.4 [7.4, 7.4]	7.4 [7.4, 7.4]	7.4 [7.4, 7.5]	0.028
PaCO2 (mmHg)	34.3 [29.9, 37.0]	36.1 [31.7, 38.0]	33.0 [29.6, 36.0]	0.026

3.4 Class B Model Development and Analysis

We used patients with more complete laboratory data on admission to conduct separate subgroup analysis and modeling. The machine learning algorithms are the same as Class A modeling, namely: RF, GBM, XGBoost, LightGBM and LR. LR serves as an interpretable baseline model for comparison with other models.

Because the subgroup contained many variables, feature selection was performed using multiple strategies in parallel. First, all candidate features were screened by univariable logistic regression, with $P < 0.1$ as the preliminary inclusion threshold (see Supplementary Table 2), followed by manual screening under the following constraints: (1) priority was

given to indicators routinely available in clinical practice; (2) no more than two indicators were selected from the same laboratory category; (3) collinearity was avoided; and (4) biological plausibility was prioritized. To verify the robustness of feature selection, LASSO regression with 10-fold cross-validation was also performed as a sensitivity analysis, with the results shown in Supplementary Table 3. The consistency between LASSO-selected and manually selected features was 85.7% (12/14), as detailed in Supplementary Table 4.

After applying these criteria, 14 variables were selected as model features: bite near a road or residence, bite in grass or bushes, two fang puncture marks, two rows of tooth marks, heart rate, TT, calcium, DD2, FIB, bite-to-presentation time, age, globulin, sex, and APTT.

Because the subgroup still had sample imbalance (hemotoxic cases accounted for approximately 79%), we continued to use stratified sampling to split the dataset into training and test sets, ensuring that both sets had sample proportions close to those of the original dataset and better reflected real clinical conditions. Multiple imputation and SMOTE oversampling were then applied to the training data, the models were trained, and model performance was evaluated in the test set. The evaluation results are shown in Fig. 8 and Table 4. RF had the highest AUC (0.894), followed by GBM and LightGBM. LR had the highest specificity among all models, and its AUC improved substantially after detailed laboratory tests were added.

Table 4 Details of Class B Models

Model	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
RF	0.894	0.724	0.917	0.766	0.969	0.478
GBM	0.873	0.759	0.875	0.784	0.957	0.500
XGBoost	0.842	0.655	0.917	0.712	0.966	0.423
LightGBM	0.856	0.805	0.792	0.802	0.933	0.528
LR	0.840	0.506	1.000	0.613	1.000	0.358

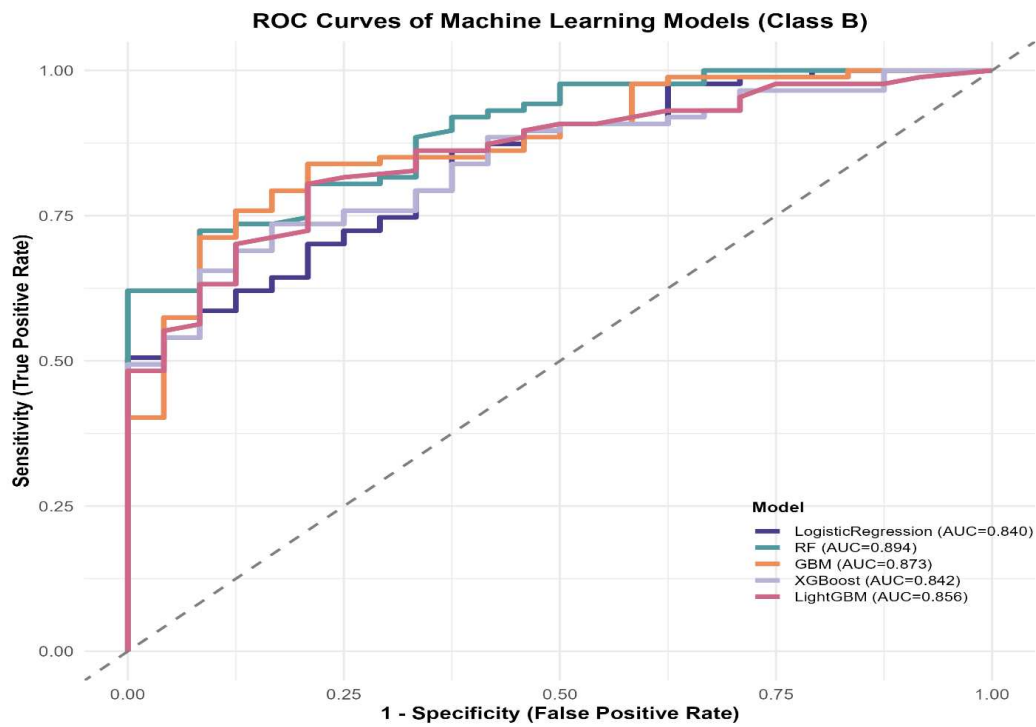


Fig. 8 ROC Curves of Class B Models

We also plotted the calibration curves, density plots and decision curves of each model in Class B samples. The calibration curves show (Fig. 9) that all models exhibit overestimation or underestimation of risk at different probability intervals.

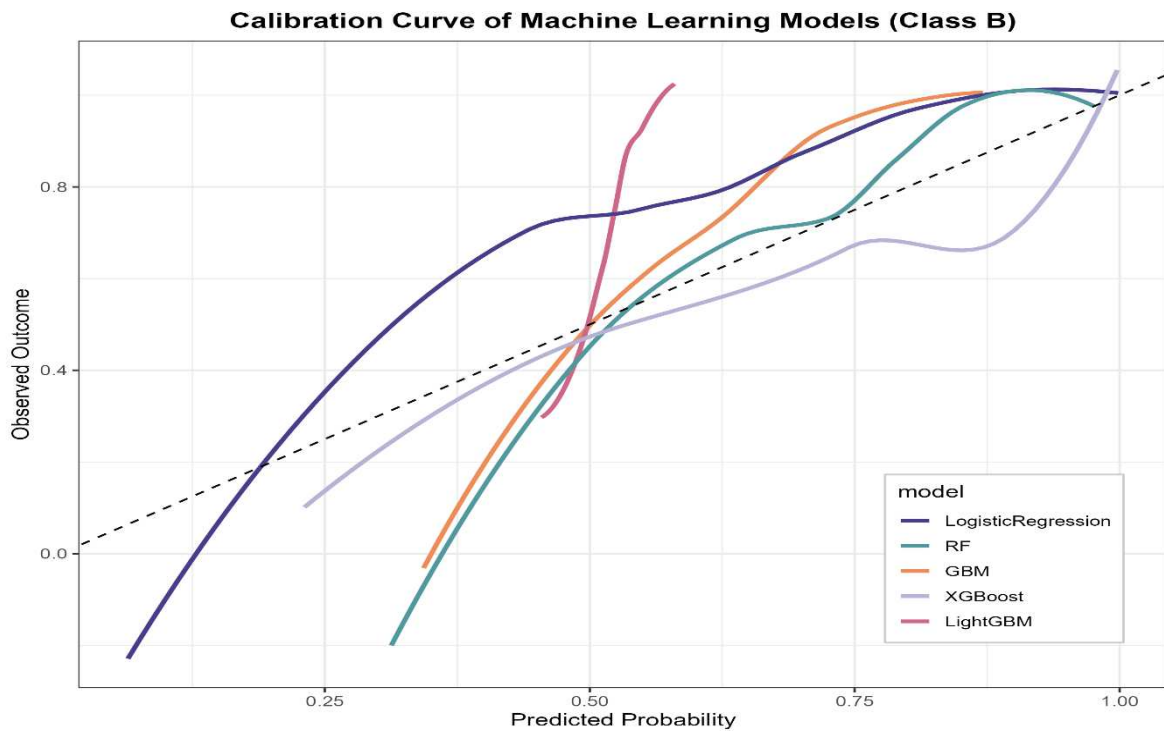


Fig. 9 Calibration Curves of Class B Models

The density plots of Class B models are shown in Fig. 10, with predicted probability on the x-axis and density on the y-axis, reflecting differences in prediction patterns among

models. Similar to the density plots in Class A samples, XGBoost remained the model most inclined to make deterministic predictions, as confirmed by the prominent unimodal distribution. RF, GBM, and LightGBM had more dispersed predicted probabilities, indicating uncertainty in the predicted probabilities. In the Class B models, LightGBM showed an obvious trimodal pattern in the high-probability range, with overlap between two peaks, suggesting poor discrimination of samples in the high-probability interval.

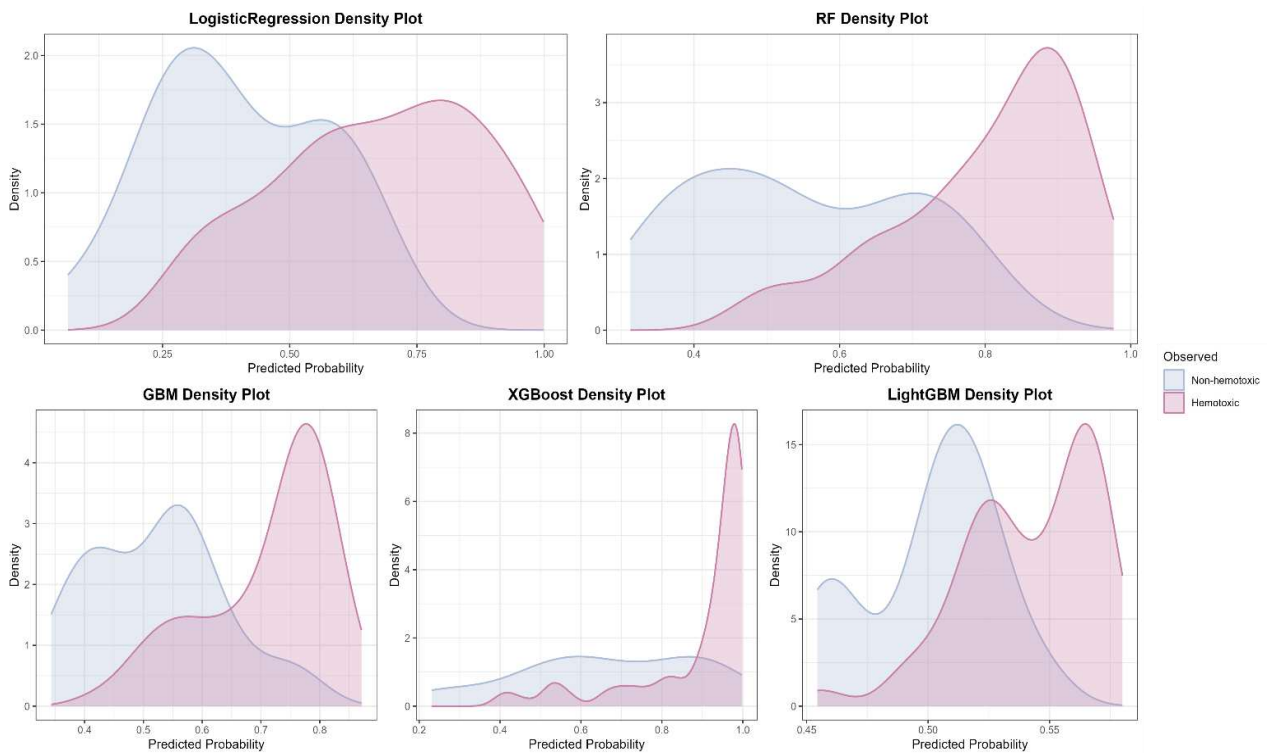


Fig. 10 Density Distribution Plots of Class B Models

To enhance clinical usability, we drew a decision curve (Fig. 11). Since both the overall dataset and the subgroup dataset have similar data imbalance problems, in the low decision threshold interval (0-0.25), none of the models showed significant clinical net benefit compared with treating all patients. In the medium-to-high decision threshold interval (0.25-0.5), the clinical net benefits of RF, GBM, and XGBOOST begin to become increasingly significant, but the clinical net benefits of LR are not as good as those of treating all patients. In the high decision threshold interval (0.5-1), the clinical net benefit of RF, GBM and XGBOOST is significantly higher than that of treating all patients. LR shows the same net benefit in a higher interval (about 0.62), but the clinical net benefit of LightGBM continues to decline and is always inferior to the clinical net benefit of treating all patients.

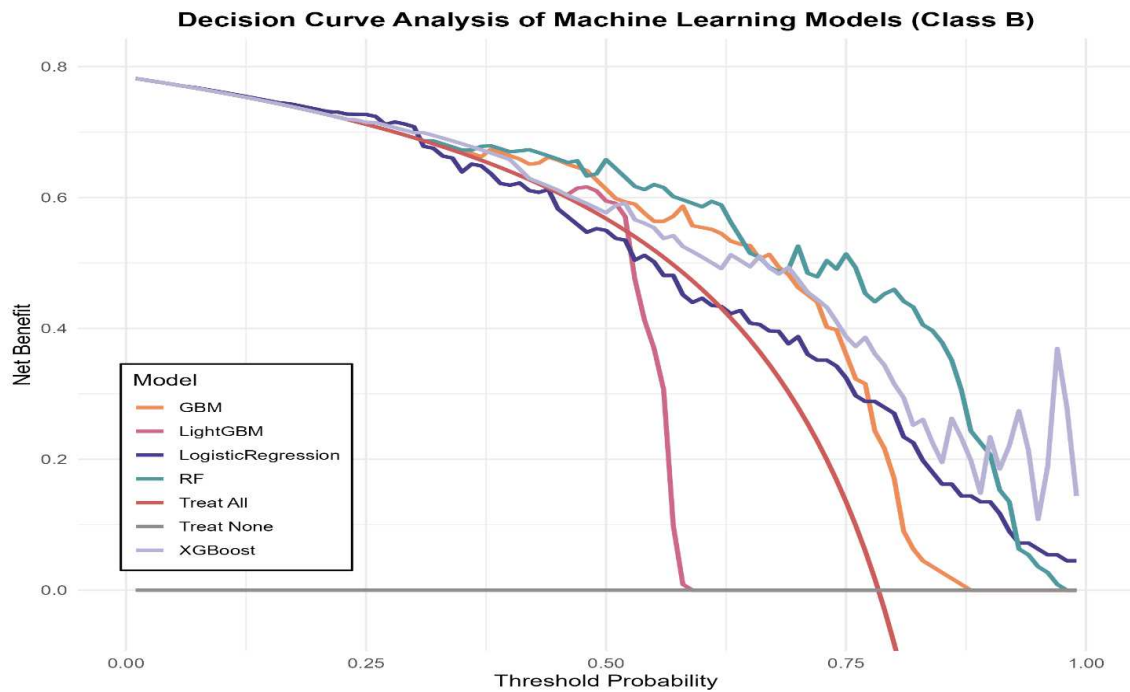
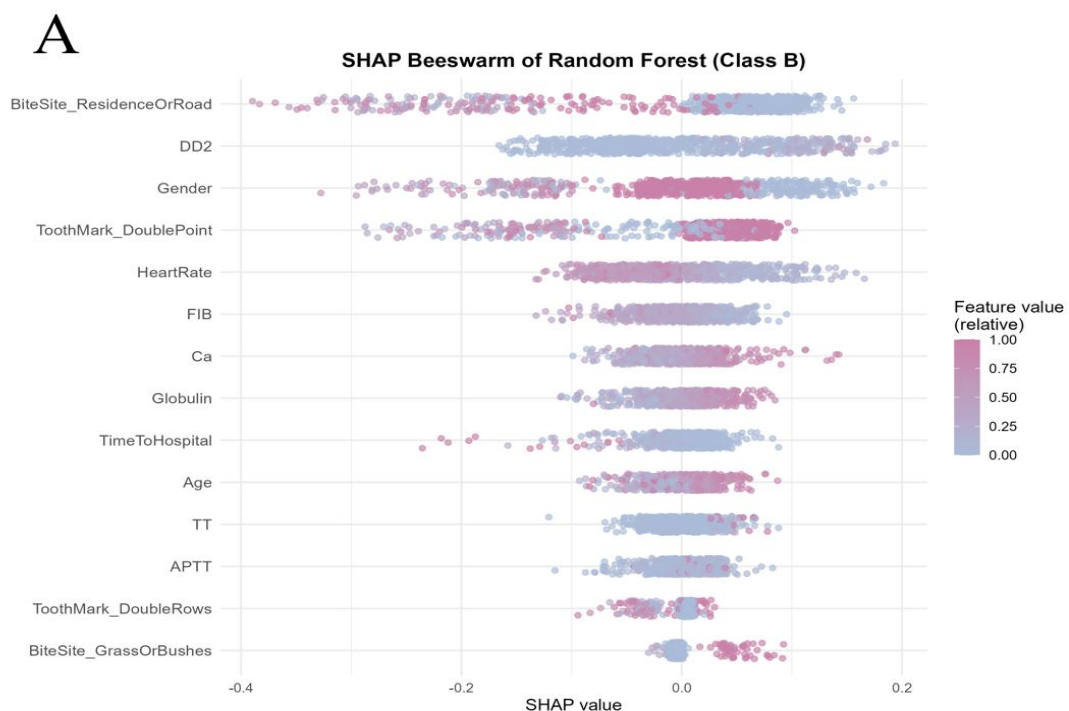


Fig. 11 Decision Curves of Class B Models

3.5 SHAP Analysis of RF in the Class B Models

Based on the Class B model analyses, we considered RF to be the optimal machine-learning model because it had the highest AUC, better calibration, and decision curves indicating clinical net benefit over almost the entire threshold range. Therefore, we performed SHAP analysis for RF and plotted a beeswarm plot (Fig. 12A) and a dependence plot (Fig. 12B) to interpret the contribution of each variable in the model. As shown below, bite location_near residence or road, sex, and fang marks_two points were the most important features, whereas DD2, FIB, and Ca were the most important laboratory tests.



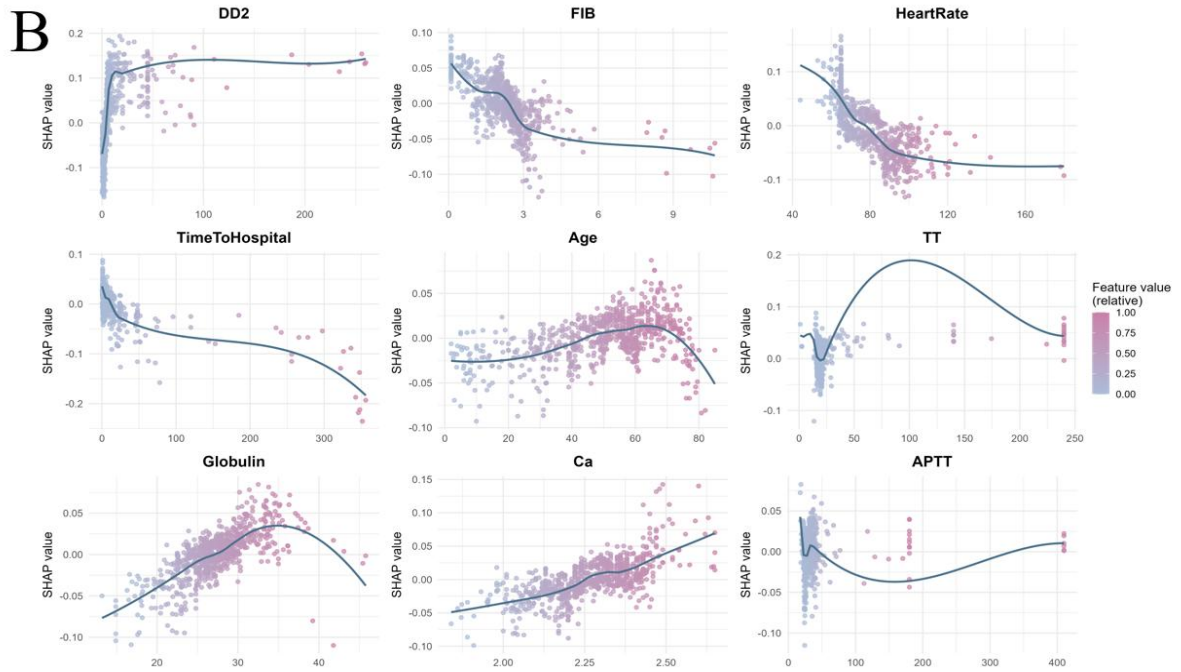


Fig. 12 SHAP Analysis of the RF Model in the Class B Models. Beeswarm Plot (A) and Dependence Plot (B)

3.6 External Validation

We collected data from snakebite patients treated between October 2025 and April 2026 as an external validation dataset. A total of 51 snakebite patients were included, comprising 43 hemotoxic patients (84.31%) and 8 non-hemotoxic patients (15.69%). The previously trained Class A and Class B models were then externally validated using these data, and the results are shown in Table 5 and 6, respectively, and the corresponding ROC curves are shown in Fig. 13.

Table 5 External Validation Results of Class A Models

Model	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
RF	0.709	0.814	0.625	0.784	0.921	0.385
GBM	0.701	0.791	0.750	0.784	0.944	0.400
XGBoost	0.701	0.814	0.625	0.784	0.921	0.385
LightGBM	0.576	1.000	0.250	0.882	0.878	1.000
LR	0.721	0.930	0.625	0.882	0.930	0.625

Table 6 External Validation Results of Class B Models

Model	AUC	Sensitivity	Specificity	Accuracy	PPV	NPV
RF	0.711	0.512	1.000	0.588	1.000	0.276
GBM	0.747	0.651	0.875	0.686	0.966	0.318
XGBoost	0.686	0.488	0.875	0.549	0.955	0.241
LightGBM	0.656	0.465	0.875	0.529	0.952	0.233
LR	0.733	0.419	1.000	0.510	1.000	0.242

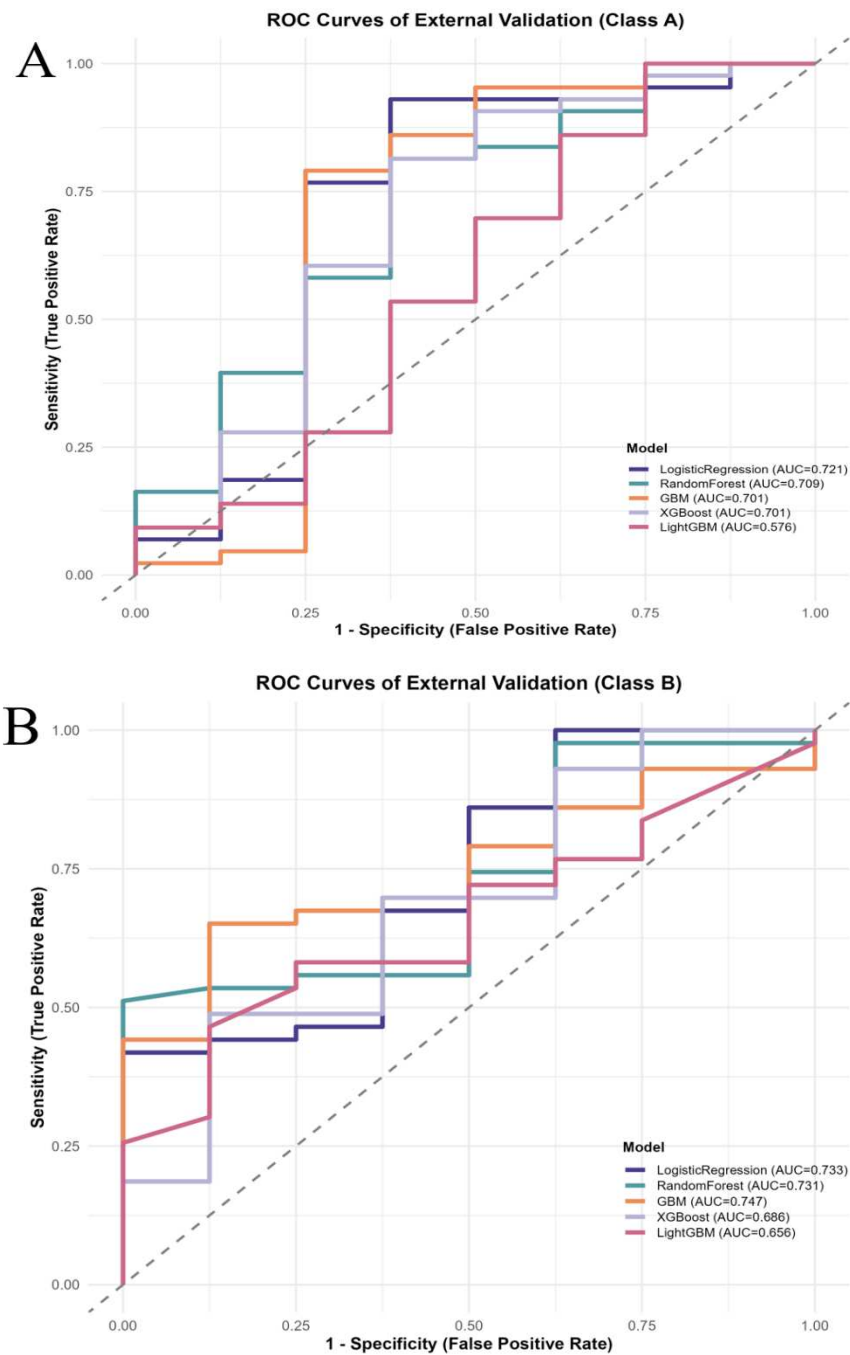


Fig. 13 External Validation of Class A Models (A) and Class B Models (B)

Discussion

In this study, we developed machine-learning models using symptoms, signs, and basic laboratory tests recorded at admission to estimate whether the envenoming phenotype was hemotoxic. The intended use is to support selection of appropriate antivenom and improve clinical outcomes. To our knowledge, relatively few studies have applied machine learning to snakebite classification.

Many clinicians may consider early diagnosis of hemotoxic snakebite relatively straightforward because such patients are expected to develop severe coagulation disorders. However, in this study, baseline analysis (Table 1) showed that among hemotoxic patients at admission, the median PT was 12 s, with the 25th percentile at 11.2 s and the 75th percentile at 13.5 s; the median APTT was 27.9 s, with the 25th percentile at 24.8 s and the 75th percentile at 32.4 s. These findings suggest that most patients did not have obvious coagulation disorders at admission. Possible reasons include the time from bite to medical visit, whether the venomous snake injected venom, the amount of venom injected, and individual constitutional differences[10]. In fact, atypical early clinical manifestations of snakebite and their incomplete correspondence with bite severity [10] are global challenges that can lead to clinical misjudgment. In our study, many hemotoxic patients may not have had obvious coagulation abnormalities at presentation and may even have shown clinical features similar to those of non-hemotoxic patients [11]. This was also reflected in our K-means clustering analysis (Fig. 2A and Fig. 2B). Fig. 2B shows clustering based on the original clinical labels, and the results indicate that many non-hemotoxic patients were mixed with hemotoxic patients, suggesting similar clinical features. In the real world, these similar features may lead to misdiagnosis or treatment delay, thereby increasing the medical burden on patients. Therefore, we aimed to use machine learning in a large sample to capture subtle differences in admission features among hemotoxic patients, enabling rapid diagnosis and classification of snakebite and improving patient prognosis.

In this study, we established two classes of models: a Class A basic model containing only selected symptoms/signs and coagulation function, and a Class B complete model containing comprehensive laboratory tests and a more detailed medical history. In public health and medical care, designing diagnostic and intervention pathways adapted to institutions with different resource levels is a key strategy for improving health-system efficiency and service accessibility[12], and AI technology is an important tool for addressing diagnostic-capacity gaps in areas with limited medical resources [13]. The design of the Class A and Class B models in this study represents an exploration of this concept in the specific emergency setting of snakebite. In resource-limited areas, simple symptoms/signs plus coagulation testing can be used for preliminary assessment of venom type, whereas hospitals with richer medical resources can conduct more comprehensive laboratory testing for more

accurate diagnosis, thereby making full use of local resources and supporting tiered diagnosis and treatment.

Comparison of the Class A and Class B ROC curves (Figs. 3 and 8) showed that adding a more complete medical history and laboratory testing increased AUC by 0.026 for LightGBM to 0.094 for LR. SHAP analysis of the Class B model (Figs. 12A and 12B) showed that bite locations contributed in different, sometimes opposite, directions. Bites near roads or residences contributed to prediction of non-hemotoxic envenoming, whereas bites in grass or bushes contributed to prediction of hemotoxic envenoming. High DD2, globulin, and calcium and low FIB were also important predictors. The pattern of elevated DD2 and reduced FIB is consistent with fibrinogen consumption and activation of fibrinolysis by thrombin-like and other venom components [14]. Comparison of SHAP analyses across the two model classes (Figs. 7A-B and 12A-B) suggests that clinicians should consider bite-to-presentation time, bite location, fang marks, vital signs, and relevant laboratory findings in addition to basic history and coagulation testing.

We also clarified the value of the two classes of prediction models by plotting decision curves. Overall, in both Class A and Class B models, RF and XGBoost achieved clinical net benefit no worse than the "treat-all" strategy across all decision thresholds. At moderate-to-high thresholds (>0.5), the net benefit of the models became more obvious, especially in the Class B models. However, it should be noted that LR in the Class A models and LightGBM in the Class B models did not show better clinical net benefit than the "treat-all" strategy in most situations. This finding has important clinical implications: when clinicians complete basic history taking and laboratory tests and estimate that the probability of hemotoxicity exceeds 50%, appropriate model predictions may help initiate targeted treatment more precisely, avoiding delays caused by waiting for typical coagulation dysfunction to appear and potentially reducing unnecessary antivenom use in low-probability patients [15][16].

During feature selection, we explored T/FIB as a composite variable intended to combine bite-to-presentation time with fibrinogen concentration. This ratio was an exploratory time-adjusted marker related to fibrinogen consumption, not a direct measure of the rate of fibrinogen consumption. Similar time-adjusted variables were constructed for other markers, but none outperformed the original variables. Taking T/FIB as an example, although a decrease in FIB among patients with hemotoxic envenoming would theoretically increase this ratio, the observed results did not consistently support this assumption. Further review showed that many patients had not yet developed a substantial reduction in FIB at admission, resulting in considerable distortion of the ratio. Therefore, we did not use these composite variables and instead retained the original variables for model development.. We therefore retained the original variables rather than these composite features.

When developing the Class B model, we deliberately avoided some relatively subjective

information, such as symptoms, and some information likely to be recorded inaccurately, such as respiratory rate. Although univariable logistic regression showed that respiratory rate was strongly associated with prediction, we considered that respiratory rate records were likely inaccurate in real clinical practice, because many physicians do not deliberately record respiratory rate for patients without obvious respiratory-rate abnormalities. Instead, we used more objective indicators, such as laboratory tests. The purpose was to make the model more accurate and more clinically generalizable, allowing potential application and dissemination beyond Yunnan Province because these model features also exist in other regions. Nevertheless, this concept still requires validation using clinical data from specific regions.

Yunnan Province is one of the provinces with the highest incidence of snakebite in China. Our previous epidemiological investigation showed that the annual incidence of snakebite in Yunnan is still increasing [17]. This public health challenge is closely related to the region's complex climate, including coexisting tropical and subtropical climates, diverse terrain such as plateaus, mountains, and river valleys, and extremely high biodiversity. Together, these factors create rich snake populations and increase exposure risk among people engaged in outdoor activities such as farming and forestry. More seriously, this natural environmental risk is amplified by inherent vulnerabilities in the regional public health system. In some areas of Yunnan, especially remote rural areas, real-world bottlenecks include insufficient capacity of primary medical institutions for rapid differential diagnosis of snakebite [18], limited accessibility of specific antivenoms, and incomplete standardized treatment pathways [19]. These factors often lead to diagnostic delays and inappropriate treatment, thereby increasing the risk of disability and death.

The World Health Organization has classified snakebite as a highest-priority neglected tropical disease and has emphasized that developing accessible diagnostic tools and standardized treatment pathways is essential for reducing disease burden [20][21]. Our previous study showed that fibrinogen supplementation together with antivenom can more effectively correct coagulation dysfunction in hemotoxic patients and shorten their hospital stay [22]. The value of the present study lies in providing tools spanning primary screening (Class A model) to in-hospital precision classification (Class B model), thereby forming a complete diagnostic-to-treatment loop for snakebite patients and offering a "Yunnan approach" to this major public health challenge.

Limitations

In this study, we used machine-learning algorithms to establish supervised binary prediction models to assist clinicians in estimating the probability that a snakebite patient has

hemotoxicity. This design was determined by the characteristics of snakebite patients in Yunnan Province, where the vast majority of snakebite patients are hemotoxic and the sample sizes of other venom types are too small to support multiclass prediction models. In addition, this study inferred venom type based on clinical phenotype rather than identifying specific snake species. Therefore, if validated in other regions in the future, this phenotype-based prediction approach may have some transportability, but its specific performance still requires confirmation in prospective studies.

Fund : This study is supported by the National Natural Science Foundation of China. Award number: 40123045.

Conflict of Interest Statement : On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- [1] Longbottom J, Shearer FM, Devine M, Alcoba G, Chappuis F, Weiss DJ, Ray SE, Ray N, Warrell DA, Ruiz de Castañeda R, Williams DJ, Hay SI, Pigott DM. Vulnerability to snakebite envenoming: a global mapping of hotspots. *Lancet*. 2018 Aug 25;392(10148):673-684. doi: 10.1016/S0140-6736(18)31224-8. Epub 2018 Jul 17. PMID: 30017551; PMCID: PMC6115328.
- [2] Afroz A, Siddiquea BN, Chowdhury HA, Jackson TN, Watt AD. Snakebite envenoming: A systematic review and meta-analysis of global morbidity and mortality. *PLoS Negl Trop Dis*. 2024 Apr 4;18(4):e0012080. doi: 10.1371/journal.pntd.0012080. PMID: 38574167; PMCID: PMC11020954.
- [3] Bawaskar HS, Bawaskar PH, Bawaskar PH. The global burden of snake bite envenoming. *J R Coll Physicians Edinb*. 2021 Mar;51(1):7-8. doi: 10.4997/JRCPE.2021.102. PMID: 33877125.
- [4] Lai R, Yan S, Wang S, Yang S, Yan Z, Lan P, Wang Y, Li Q, Wang J, Wang W, Ma Y, Liang Z, Zhang J, Zhou N, Han X, Zhang X, Zhang M, Zhao X, Zhang G, Zhu H, Yu X, Lyu C. The Chinese guideline for management of snakebites. *World J Emerg Med*. 2024;15(5):333-355. doi: 10.5847/wjem.j.1920-8642.2024.076. PMID: 39290598; PMCID: PMC11402871.
- [5] Ahmed SM, Ahmed M, Nadeem A, Mahajan J, Choudhary A, Pal J. Emergency treatment of a snake bite: Pearls from literature. *J Emerg Trauma Shock*. 2008 Jul;1(2):97-105. doi: 10.4103/0974-2700.43190. PMID: 19561988; PMCID: PMC2700615.
- [6] Kumar N, Ishaqui AA, Rajpoot PL, Maqsood MB, Sultana R, Khaskheli MS, Iqbal A, Tabassum S, Muhammad S. Determinants of Poor Treatment Outcomes Among Snakebite Envenoming Patients. *Ann Glob Health*. 2026 Apr 10;92(1):30. doi: 10.5334/aogh.5147. PMID: 41970964; PMCID: PMC13068097.
- [7] Deo RC. Machine Learning in Medicine. *Circulation*. 2015 Nov 17;132(20):1920-30. doi: 10.1161/CIRCULATIONAHA.115.001593. PMID: 26572668; PMCID: PMC5831252.
- [8] Guo W, Liu J, Dong F, Song M, Li Z, Khan MKH, Patterson TA, Hong H. Review of machine learning and deep learning models for toxicity prediction. *Exp Biol Med (Maywood)*. 2023 Nov;248(21):1952-1973. doi: 10.1177/15353702231209421. Epub 2023 Dec 6. PMID: 38057999; PMCID: PMC10798180.
- [9] Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med*. 2019 Apr 4;380(14):1347-1358. doi: 10.1056/NEJMr1814259. PMID: 30943338.
- [10] Warrell DA, Williams DJ. Clinical aspects of snakebite envenoming and its treatment in low-resource settings. *Lancet*. 2023 Apr 22;401(10385):1382-1398. doi: 10.1016/S0140-6736(23)00002-8. Epub 2023 Mar 14. PMID: 36931290.
- [11] Gutiérrez JM, Calvete JJ, Habib AG, Harrison RA, Williams DJ, Warrell DA. Snakebite envenoming. *Nat Rev Dis Primers*. 2017 Sep 14;3:17063. doi: 10.1038/nrdp.2017.63. Erratum in: *Nat Rev Dis Primers*. 2017 Oct 05;3:17079. doi: 10.1038/nrdp.2017.79. PMID: 28905944.
- [12] Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, Adeyi O, Barker P, Daelmans B, Doubova SV, English M, García-Elorrio E, Guanais F, Gureje O, Hirschhorn LR, Jiang L, Kelley E, Lemango ET, Liljestrand J, Malata A, Marchant T, Matsoso MP, Meara JG, Mohanan M, Ndiaye Y, Norheim OF, Reddy KS, Rowe AK, Salomon JA, Thapa G, Twum-Danso NAY, Pate M. High-quality health systems

- in the Sustainable Development Goals era: time for a revolution. *Lancet Glob Health*. 2018 Nov;6(11):e1196-e1252. doi: 10.1016/S2214-109X(18)30386-3. Epub 2018 Sep 5. Erratum in: *Lancet Glob Health*. 2018 Nov;6(11):e1162. doi: 10.1016/S2214-109X(18)30438-8. Erratum in: *Lancet Glob Health*. 2018 Nov;6(11):e1162. doi: 10.1016/S2214-109X(18)30456-X. Erratum in: *Lancet Glob Health*. 2021 Aug;9(8):e1067. doi: 10.1016/S2214-109X(21)00250-3. PMID: 30196093; PMCID: PMC7734391.
- [13] Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Health*. 2018 Aug 29;3(4):e000798. doi: 10.1136/bmjgh-2018-000798. PMID: 30233828; PMCID: PMC6135465.
- [14] Isbister GK. Snakebite doesn't cause disseminated intravascular coagulation: coagulopathy and thrombotic microangiopathy in snake envenoming. *Semin Thromb Hemost*. 2010 Jun;36(4):444-51. doi: 10.1055/s-0030-1254053. Epub 2010 Jul 7. PMID: 20614396.
- [15] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006 Nov-Dec;26(6):565-74. doi: 10.1177/0272989X06295361. PMID: 17099194; PMCID: PMC2577036.
- [16] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, Pencina MJ, Kattan MW. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010 Jan;21(1):128-38. doi: 10.1097/EDE.0b013e3181c30fb2. PMID: 20010215; PMCID: PMC3575184.
- [17] Gao Q, Teng Y, Xiao C, Zeng R, Han B, Gao H, Wang J, Li X, Yang C, Dai J, Li C, Huang Q, Li Z, Zhang W. Epidemiological investigation of venomous snakebites in Yunnan Province. *Front Toxicol*. 2025 Aug 6;7:1609487. doi: 10.3389/ftox.2025.1609487. PMID: 40842969; PMCID: PMC12364869.
- [18] Hu Y, Lv C, Song X, Gan Y, Wang J, Hao W, He L, Chen Y, Han X, Yan S. Current Status and Influencing Factors of Snakebite Diagnosis and Treatment Knowledge Among Medical Staff in China: A Cross-Sectional Study. *Int J Public Health*. 2023 Dec 11;68:1606601. doi: 10.3389/ijph.2023.1606601. PMID: 38146482; PMCID: PMC10749458.
- [19] Teng Y, Cao Y, Feng S, Han B, Gao Q, Li Z. Vulnerability characteristics and adaptation strategies of public health systems in different regions of China. *Front Public Health*. 2025 Sep 9;13:1609377. doi: 10.3389/fpubh.2025.1609377. PMID: 40994751; PMCID: PMC12454412.
- [20] Minghui R, Malecela MN, Cooke E, Abela-Ridder B. WHO's Snakebite Envenoming Strategy for prevention and control. *Lancet Glob Health*. 2019 Jul;7(7):e837-e838. doi: 10.1016/S2214-109X(19)30225-6. Epub 2019 May 22. PMID: 31129124.
- [21] Williams DJ, Faiz MA, Abela-Ridder B, Ainsworth S, Bulfone TC, Nickerson AD, Habib AG, Junghanss T, Fan HW, Turner M, Harrison RA, Warrell DA. Strategy for a globally coordinated response to a priority neglected tropical disease: Snakebite envenoming. *PLoS Negl Trop Dis*. 2019 Feb 21;13(2):e0007059. doi: 10.1371/journal.pntd.0007059. PMID: 30789906; PMCID: PMC6383867.
- [22] Gao Q, Wang Y, Teng Y, Huang Q, Wang Y, Feng S, Han B, Li Z. Fibrinogen supplementation enhances antivenom efficacy in treating coagulopathy due to hemotoxic snakebite envenoming by *Trimeresurus* and *Agkistrodon* species in Yunnan, China. *Front Med (Lausanne)*. 2025 Dec 18;12:1667800. doi: 10.3389/fmed.2025.1667800. PMID: 41488074; PMCID: PMC12756120.