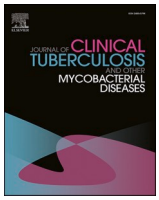
Journal of Clinical Tuberculosis and Other Mycobacterial Diseases 38 (2025) 100511

Contents lists available at ScienceDirect

Journal of Clinical Tuberculosis and Other Mycobacterial Diseases

journal homepage: www.elsevier.com/locate/jctube



Two-centers machine learning analysis for predicting acid-fast bacilli results in tuberculosis sputum tests

Jichong Zhu [a,b,1], Yong Zhao [c,1], Chengqian Huang [b], Chenxing Zhou [b], Shaofeng Wu [b], Tianyou Chen [b], Xinli Zhan [b,*]

[a] People's Hospital of Guilin, Guilin 541002, PR China
[b] First Affiliated Hospital of Guangxi Medical University, Nanning 530021, PR China
[c] Guangxi Hospital, the First Affiliated Hospital of Sun Yat-sen University, Nanning 530021, PR China

ARTICLE INFO

Keywords:
Machine learning
Tuberculosis
Acid-fast bacilli
Big data
SVM-RFE

ABSTRACT

Background: Tuberculosis (TB) is a chronic respiratory infectious disease caused by Mycobacterium tuberculosis, typically diagnosed through sputum smear microscopy for acid-fast bacilli (AFB) to assess the infectivity of TB.
Methods: This study enrolled 769 patients, including 641 patients from the First Affiliated Hospital of Guangxi Medical University as the training group, and 128 patients from Guangxi Hospital of the First Affiliated Hospital of Sun Yat-sen University as the validation group. Among the training cohort, 107 patients were AFB-positive, and 534 were AFB-negative. In the validation cohort, 24 were AFB-positive, and 104 were AFB-negative. Blood samples were collected and analyzed using machine learning (ML) methods to identify key factors for TB diagnosis.
Results: Several ML methods were compared, and support vector machine recursive feature elimination (SVM-RFE) was selected to construct a nomogram diagnostic model. The area under the curve (AUC) of the diagnostic model was 0.721 in the training cohort and 0.758 in the validation cohort. The model demonstrated clinical utility when the threshold was between 38% and 94%, with the NONE line above the ALL line in the decision curve analysis.
Conclusion: We developed a diagnostic model using multiple ML methods to predict AFB results, achieving satisfactory diagnostic performance.

1. Introduction

Tuberculosis (TB) refers to chronic respiratory infectious diseases caused by *Mycobacterium tuberculosis* infection in the lungs[1]. The source of TB infection, *Mycobacterium tuberculosis*, is mainly in the sputum, and the main route of transmission is through the respiratory tract[2]. The main clinical symptoms include cough, expectoration, fever, and fatigue; severe complications can lead to irreversible lung damage and even death[3]. TB is a common infectious disease in China. The proportion of *Mycobacterium tuberculosis* detected in the sputum smear of individuals aged 15 years and above is 66 per 100,000[4]. The susceptible groups mainly include the elderly, immunodeficiency patients, and infants.

Despite many decades of research, TB remains the leading cause of death from an infectious agent worldwide. However, the diagnosis of TB often needs to be performed by considering various aspects, combining clinical symptoms, imaging and bacteriology, and other tests, and no convenient detection method is available for making a clear diagnosis of TB[5,6]. According to the diagnostic criteria of pulmonary TB issued by China in 2017, the detection of AFB in sputum smear plays a crucial role in the diagnosis of pulmonary TB. Pulmonary TB diagnosis is made if two sputum standard smears are positive for AFB, one sputum smear is positive for AFB, and there is imaging support for pulmonary TB, or when one sputum smear is positive for AFB, and there is a positive culture of *Mycobacterium tuberculosis* in the sputum specimen[7]. The World Health Organization defines the diagnosis of infectious sputum smear-positive active pulmonary tuberculosis as 1, one positive sputum smear for AFB plus one positive sputum culture for Mycobacterium tuberculosis complex. 2, two or more sputum smears were positive for AFB[8].

* Corresponding author.
E-mail address: zhanxinli@stu.gxmu.edu.cn (X. Zhan).
[1] Dr. Jichong Zhu and Dr. Yong Zhao contributed equally to this work.

https://doi.org/10.1016/j.jctube.2025.100511

Available online 27 January 2025

Sputum smear testing for AFB is often lacking in many underdeveloped countries. AFB positivity is not limited to *Mycobacterium tuberculosis*; other mycobacteria can yield a positive result. This affects the doctor's diagnosis[9]. No in-depth study has been conducted on whether there is a difference between the blood of AFB-positive and AFB-negative TB patients.

Machine learning (ML) is a multifield interdisciplinary research subject. The development of ML in clinical data processing has brought rich insights and new possibilities to the field of medicine [10]. The use of highly efficient ML algorithms allows for tackling many previously intractable statistical problems[11,12]. Therefore, in this study, we compared the blood test results between AFB-positive and AFB-negative TB patients using ML methods to identify differences that may help clinicians better assess AFB status in confirmed TB cases.

2. Methods

2.1. Patients

All subjects in this study were from the First Affiliated Hospital of Guangxi Medical University and the Guangxi Hospital of the First Affiliated Hospital of Sun Yat-sen University. The Ethics Committee of the First Affiliated Hospital of Guangxi Medical University and the Guangxi Hospital of the First Affiliated Hospital of Sun Yat-sen University approved this study after discussion.

From July 2017 to June 2022, we collected data on patients with a confirmed diagnosis of TB from the First Affiliated Hospital of Guangxi Medical University and Guangxi Hospital of the First Affiliated Hospital of Sun Yat-sen University. Utilizing the information systems of these two hospitals, we retrieved patient blood samples based on their ID numbers. It is important to note that we collected blood samples only at the time of the patients' initial hospitalization. We did not group the patients based on specific TB disease stages or collect information regarding whether they had received treatment. The patient inclusion criteria were as follows: [1] according to Chinese 2017 edition WS288–2017 pulmonary TB, pulmonary TB was diagnosed by two or more experts; [2] patients who had complete data regarding erythrocyte sedimentation rate (ESR), blood routine examination, liver function test, renal function test, blood lipid test, blood electrolyte test, and coagulation function test; [3] patients for whom sputum smear AFB examination was performed on admission; [4] the diagnosis of pulmonary TB did not solely depend on the culture results of Mycobacterium tuberculosis but was based on a comprehensive assessment by the experts, considering clinical, radiological, and laboratory finding ; and [5] patients voluntarily participated in this study, signing a consent form prior to the study. The exclusion criteria were as follows: [1] patients who received multiple AFB examinations after admission with inconsistent results; [2] patients with missing clinical data; [3] patients who refused to participate in this study; and [4] patients who had severe cardiovascular, cerebrovascular or other inflammatory diseases.

A total of 768 patients were enrolled in this study. The subjects were 641 patients from the First Affiliated Hospital of Guangxi Medical University as the training group and 128 patients from the Guangxi Hospital of the First Affiliated Hospital of Sun Yat-sen University as the verification group. In the training cohort, 107 patients were positive for AFB, and 534 patients were negative for AFB. In the validation cohort, 24 patients were positive for AFB, and 104 patients were negative for AFB.

All patient information was collected from the electronic information system of the First Affiliated Hospital of Guangxi Medical University and the Guangxi Hospital of the First Affiliated Hospital of Sun Yat-sen University by using the patient's ID number. After collecting the information, we concealed the personal information of the patients and replaced their identities with serial numbers. Sex, ESR, blood routine examination, Liver function examination, kidney function examination, Lipid examination, Plasma electrolyte examination and Coagulation function tests and data were collected. All the results were complete.

2.2. Machine learning method

Support vector machine recursive feature elimination (SVM-RFE) was selected due to its effectiveness in handling high-dimensional data and performing feature selection by recursively eliminating the least important features. This method helps reduce the dimensionality of the data while preserving the most relevant features for prediction, ultimately improving the model's performance.

Lasso regression was chosen for its regularization property, which applies L1 regularization to shrink the coefficients of less relevant features to zero. This technique effectively selects the most predictive features while minimizing overfitting, making it particularly useful when dealing with large datasets where many features may not be relevant.

Random forest (RF) was selected due to its robustness in handling noisy data and its ability to model non-linear relationships between variables. It also provides an efficient way to evaluate feature importance through metrics like %IncMSE and IncNodePurity, making it ideal for identifying key blood biomarkers that contribute significantly to AFB status prediction.

In the data preprocessing stage, missing numerical data were handled using mean imputation, while features with a high proportion of missing values were removed; subsequently, the data were standardized to have a mean of 0 and a standard deviation of 1 to ensure fairness across features, and finally, the dataset was split into a training set and a validation set, maintaining the proportion of AFB-positive and AFB-negative samples to address class imbalance.

2.3. Statistical analysis

IBM SPSS 23 and R software were used for statistical analysis and visualization of the data. Continuous variables were analyzed using the Student *t*-test, and the data were normally distributed and had homogeneity of variance[13]. Categorical variables were statistically analyzed using the chi-square test[14]. All statistical analysis data were repeatedly verified by two or more members of our team to avoid errors. For all analyzed data, a two-sided probability of less than 0.05 was considered statistically significant [15]. In the statistical analysis, we used the Student's *t*-test for comparing continuous variables and the chi-square test for categorical variables. All tests were two-sided, with a significance level set at $P < 0.05$. To account for multiple comparisons and reduce the risk of false positives, we applied the Bonferroni correction to adjust the P-values for all statistical tests conducted.

Due to the significant computational demands posed by incorporating 55 independent variables, each corresponding to a large number of patient records, we initially employed univariate logistic regression to streamline our analysis. Variables with univariate logistic regression p-value of < 0.05 were included in multivariate logistic regression, Lasso regression, RF, and SVM-RFE for further screening[16].

For logistic regression, we used the "rms," "glmnet," and "plyr" R packages in R software for the calculations. For Lasso regression, we utilized the "glmnet" R package for the calculations[17]. For RF, we used the "randomForest" R package; an increase in mean squared (% IncMSE) and an increase in node purity (IncNodePurity) are two screening methods in the randomForest package. In both methods, higher values represent more important variables[18]. SVM-RFE was calculated using the "e071" package in R software; a smaller AvgRank value indicates that the dependent variable has a greater influence on the independent variable[19].

3. Results

Pairwise correlations were computed among all features. As shown in Fig. 1, high correlations were observed between several pairs of variables, including apolipoprotein A (APOA) and apolipoprotein C (APOC), high-density lipoprotein (HDL) and APOA, Na + and Cl-,

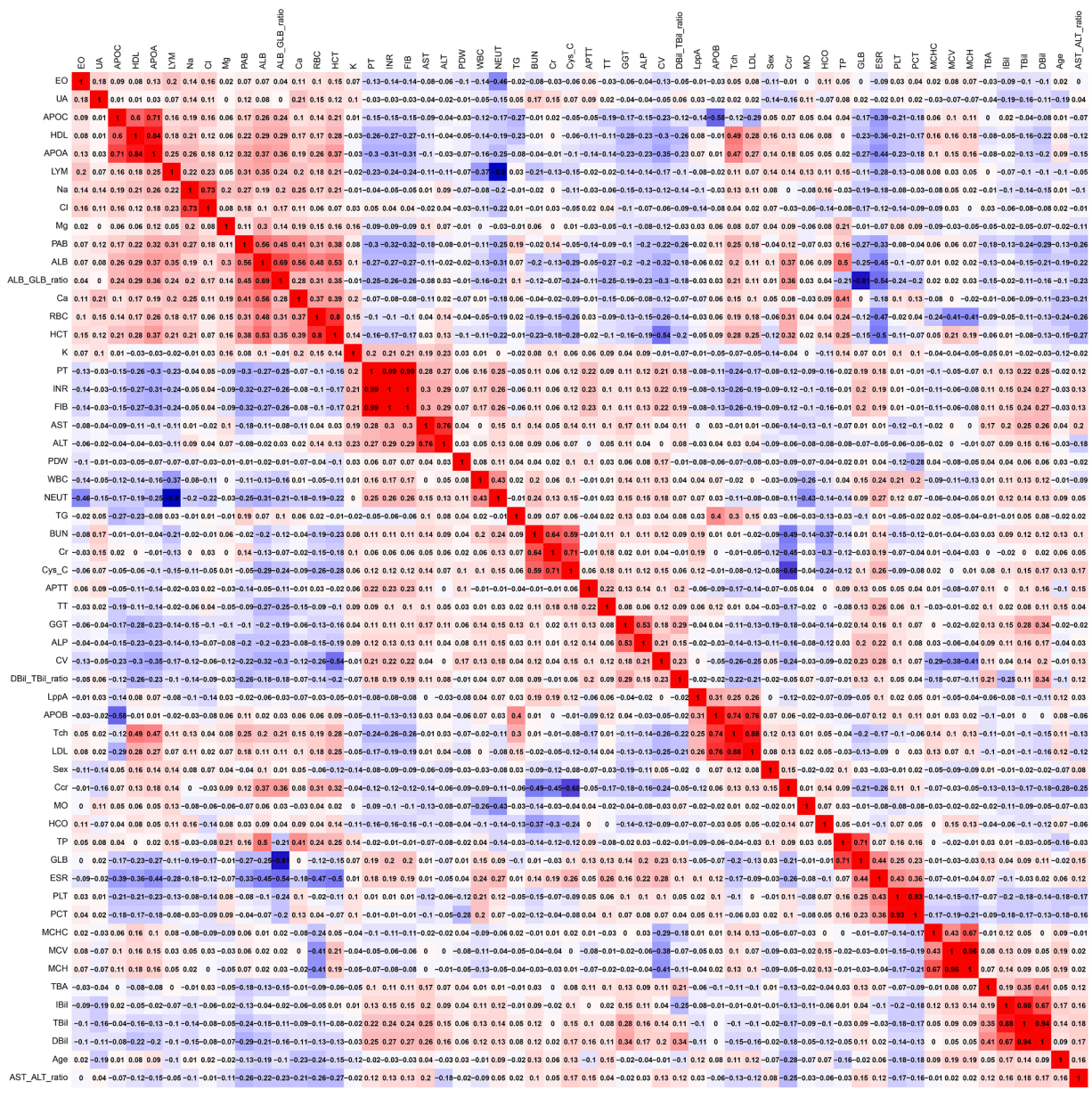


Fig. 1. Correlation between all data.

hematocrit value (HCT) and red blood cell (RBC) count, international normalized ratio (INR) and prothrombin time (PT), fibrinogen (FIB) and PT, alanine aminotransferase (ALT) and aspartate aminotransferase (AST), creatinine (Cr) and cysteine C (Cys-C), total cholesterol (Tch) and apolipoprotein B (APOB), low-density lipoprotein (LDL) and Tch, globulin (GLB) and total protein (TP), thrombocytocrit (PCT) and platelet count (PLT), as well as between total bilirubin (TBil) and indirect bilirubin (IBil). Additionally, a positive correlation was noted between TBil and direct bilirubin (DBil), while negative correlations were observed between neutrophil percentage (NEUT) and lymphocyte percentage (LYM), GLB and albumin/globulin ratio (ALB/GLB), and between creatinine clearance rate (Ccr) and Cys-C.

As can be seen in Table 1, in the training cohort, the RBC and PLT of AFB-positive patients were higher than those of AFB-negative patients, and the mean corpuscular hemoglobin concentration (MCHC) of AFB-positive patients was lower than that of AFB-negative patients, with statistically significant differences. As can be seen in Table 2, IBil, gamma-glutamyl transpeptidase (GGT), and alkaline phosphatase (ALP) decreased in AFB-positive patients, whereas uric acid (UA) and bicarbonate radical (HCO) decreased and exhibited statistically significant differences in AFB-negative patients. There was no significant difference in blood lipid examination and coagulation function between AFB-positive and AFB-negative patients (Table3) .

We included all the variables in the logistic regression. According to the results of univariate logistic regression (Table 4), ALP, ALT, GGT, HCO, HCT, IBil, MCHC, PCT, PLT, RBC, and UA were included in the analysis due to their p-values being less than 0.05, indicating statistical significance. To reduce the computational load of ML methods, we incorporated these 11 dependent variables into the subsequent multivariate logistic regression, Lasso regression, RF, and SVM-RFE.

The results of multivariate logistic regression are presented in Table 4; the p-values of ALT, MCHC, and UA were less than 0.05 and exhibited statistically significant differences. Fig. 2A illustrates the visualization results obtained using Lasso regression. As can be seen in Fig. 2B, the best performance was achieved when nine dependent variables (Table 5) were included in the regression model.

The results obtained using RF are depicted in Fig. 3. Fig. 3A shows the algorithm scores of %IncMSE and IncNodePurity for the 11

Table 1
The differences ESR,Sex and blood routine examination.

Type	Training cohort			Validation cohort		
	AFB positive	AFB negative	P-value	AFB positive	AFB negative	P-value
	(N = 107)	(N = 534)		(N = 24)	(N = 104)	
Sex						
male	67	357	0.398	12	73	0.059
female	40	177		12	31	
Age						
Mean (SD)	57.53 (14.7)	59.27 (16.11)	0.301	53.33 (16.59)	58.05 (17.31)	0.228
ESR						
Mean (SD)	51.70 (29.05)	46.46 (30.1)	0.099	51.96 (24.15)	40.5 (30.27)	0.086
WBC						
Mean (SD)	7.98 (3.31)	8.33 (5.38)	0.509	7.95(3.5)	8.11 (3.77)	0.848
RBC						
Mean (SD)	4.19 (0.78)	3.98 (0.82)	**0.016**	4.23 (0.82)	4.02 (0.86)	0.280
HCT						
Mean (SD)	0.35 (0.06)	0.34 (0.06)	**0.048**	0.35 (0.04)	0.35 (0.06)	0.895
MCV						
Mean (SD)	84.11 (9.90)	85.25 (9.68)	0.267	83.22 (8.52)	86.9 (9.37)	0.080
MCH						
Mean (SD)	27.23 (3.82)	27.96 (3.73)	0.066	27.25 (3.48)	28.55 (4.14)	0.155
MCHC						
Mean (SD)	322.85 (11.92)	327.34 (13.94)	**0.002**	326.51 (13.48)	327.68 (19.91)	0.791
PLT						
Mean (SD)	335.46 (134.25)	300.16 (131.82)	**0.012**	320.43 (114.54)	289.43 (119.29)	0.250
PDW						
Mean (SD)	0.16 (0.02)	0.16 (0.02)	0.156	0.16 (0.01)	0.15 (0.03)	**0.037**
NEUT						
Mean (SD)	0.66 (0.11)	0.66 (0.13)	0.661	0.64 (0.12)	0.67 (0.13)	0.437
LYM						
Mean (SD)	0.2(0.09)	0.2(0.10)	0.932	0.21 (0.10)	0.21 (0.10)	0.813
MO						
Mean (SD)	0.1(0.03)	0.09 (0.04)	0.362	0.11 (0.04)	0.09 (0.04)	0.091
EO						
Mean (SD)	0.04 (0.03)	0.04 (0.04)	0.783	0.03 (0.03)	0.038 (0.04)	0.893
CV						
Mean (SD)	0.16 (0.03)	0.16 (0.03)	0.631	0.15 (0.02)	0.15 (0.03)	0.832
PCT						
Mean (SD)	0.26 (0.01)	0.24 (0.01)	**0.018**	0.25 (0.07)	0.24 (0.10)	0.660

The red text means that the p-value was < 0.05. SD, standard deviation; WBC, white blood cell; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MO, monocyte percentage; EO, eosinophil percentage; CV, RBC volume distributing width.

variables. As can be seen in Fig. 3B, a good prediction effect was achieved when seven factors were selected by RF.

The AvgRank values of SVM-RFE are shown in Table 6; the smaller the value, the more important it is. As can be seen in Fig. 4, in the SVM-RFE algorithm, the best diagnostic efficiency was achieved when all 11 factors were included in the model.

We calculated the area under the curve (AUC) for the models constructed by the four ML methods. As can be seen in Fig. 5, the AUC values of SVM-RFE were better than that of other ML methods in both the training cohort and the validation cohort. Additionally, the data selected through SVM, as depicted in Fig. 1, did not show any clear correlations, which is why we included all these variables in our model. The AUC and

Table 2
The differences liver function examination, renal function examination, and plasma electrolyte examination.

Type	Training cohort			Validation cohort		
	AFB positive	AFB negative	P-value	AFB positive	AFB negative	P-value
	(N = 107)	(N = 534)		(N = 24)	(N = 104)	
TBil						
Mean (SD)	8.36 (6.94)	9.96 (9.89)	0.109	7.00 (3.85)	9.73 (9.31)	0.163
DBil						
Mean (SD)	3.99 (5.33)	4.62 (6.26)	0.335	2.89 (1.98)	4.43 (6.60)	0.259
IBil						
Mean (SD)	4.36 (3.15)	5.36 (4.37)	**0.024**	4.11 (2.29)	5.29 (4.13)	0.18
DBil/TBil						
Mean (SD)	0.46 (0.15)	0.44 (0.16)	0.175	0.41 (0.10)	0.44 (0.17)	0.405
TP						
Mean (SD)	66.26 (7.43)	65.25 (7.85)	0.221	67.57 (6.45)	65.01 (9.06)	0.193
ALB						
Mean (SD)	35.17 (5.89)	34.58 (5.63)	0.33	36.33 (4.72)	34.46 (5.96)	0.155
GLB						
Mean (SD)	31.06 (6.04)	30.66 (7.08)	0.585	30.2 (7.92)	30.23 (6.10)	0.964
ALB/GLB						
Mean (SD)	1.19 (0.33)	1.20 (0.35)	0.808	1.20 (0.24)	1.21 (0.39)	0.876
GGT						
Mean (SD)	53.97 (42.27)	73.13 (78.01)	**0.014**	45.29 (64.84)	82.62 (103.14)	0.93
TBA						
Mean (SD)	10.92 (29.95)	10.79 (20.23)	0.956	4.21 (2.50)	11.37 (27.97)	0.214
AST						
Mean (SD)	24.93 (23.19)	32.15 (47.41)	0.124	22.13 (11.81)	32.83 (31.26)	0.102
ALT						
Mean (SD)	17.84 (12.65)	28.43 (59.20)	0.066	17.88 (17.93)	29.66 (31.59)	0.081
AST/ALT						
Mean (SD)	1.56 (0.83)	1.46 (0.89)	0.251	1.47 (0.43)	1.51 (1.30)	0.876
ALP						
Mean (SD)	90.24 (34.71)	108.25 (80.53)	**0.024**	89.29 (41.41)	113.13 (113.50)	0.314
PAB						
Mean (SD)	169.46 (71.56)	167.24 (66.87)	0.758	116.26 (60.71)	178.96 (74.87)	0.441
BUN						
Mean (SD)	4.63 (4.57)	5.11 (3.41)	0.212	3.89 (1.34)	5.34 (4.26)	0.102
Cr						
Mean (SD)	80.00 (72.95)	87.76 (92.05)	0.411	66.33 (17.6)	76.45 (45.82)	0.291
UA						
Mean (SD)	409.24 (214.86)	338.54 (180.01)	**<0.001**	338.21 (149.87)	303.51 (174.93)	0.371
HCO						
Mean (SD)	25.92 (4.51)	25.05 (3.81)	**0.037**	25.8 (3.96)	25.09 (3.69)	0.405
Ccr						
Mean (SD)	92.63 (29.99)	86.25 (32.55)	0.062	102.47 (26.08)	86.23 (30.20)	**0.018**
Cys_C						
Mean (SD	0.97 (0.64)	1.10 (0.87)	0.132	0.83 (0.35)	1.05 (0.64)	0.107
K^+						
Mean (SD)	3.94 (0.44)	4.01 (0.52)	0.235	3.93 (0.42)	4.01 (0.43)	0.437

(continued on next page)

Table 2 (*continued*)

Type	Training cohort AFB positive	Training cohort AFB negative	P-value	Validation cohort AFB positive	Validation cohort AFB negative	P-value
Na^+						
Mean (SD)	137.88 (3.91)	138.59 (3.95)	0.091	138.53 (4.61)	138.11 (4.14)	0.667
Cl^-						
Mean (SD)	102.02 (4.58)	102.72 (4.40)	0.136	102.03 (5.41)	101.42 (5.02)	0.602
Ca^{2+}						
Mean (SD)	2.20 (0.14)	2.18 (0.15)	0.175	2.19 (0.12)	2.17 (0.18)	0.693
Mg^{2+}						
Mean (SD)	0.87 (0.13)	0.86 (0.15)	0.622	0.82 (0.13)	0.85 (0.12)	0.198

The red text means that the p-value was < 0.05. SD, standard deviation; TBA, total bile acid; PAB, prealbumin; ChE, cholinesterase.

Table 3
Differences in lipid examination and coagulation examination.

Type	Training cohort AFB positive	Training cohort AFB negative	P-value	Validation cohort AFB positive	Validation cohort AFB negative	P-value
	(N = 107)	(N = 534)		(N = 24)	(N = 104)	
Tch						
Mean (SD)	4.19 (0.91)	4.16 (1.03)	0.74	4.10 (0.91)	4.12 (1.12)	0.941
TG						
Mean (SD)	1.26 (0.64)	1.17 (0.67)	0.195	1.19 (0.65)	1.19 (0.74)	0.991
HDL						
Mean (SD)	0.97 (0.41)	1.01 (0.41)	0.369	1.07 (0.35)	0.99 (0.37)	0.34
LDL						
Mean (SD)	2.54 (0.83)	2.50 (0.85)	0.668	2.40 (0.68)	2.43 (0.87)	0.867
APOA						
Mean (SD)	0.94 (0.32)	0.96 (0.31)	0.61	0.97 (0.25)	0.97 (0.32)	0.983
APOB						
Mean (SD)	0.92 (0.26)	0.88 (0.23)	0.061	0.86 (0.23)	0.86 (0.23)	0.97
APOC						
Mean (SD)	1.11 (0.47)	1.17 (0.52)	0.277	1.18 (0.40)	0.19 (0.49)	0.872
LppA						
Mean (SD)	0.40 (0.44)	0.36 (0.47)	0.43	0.30 (0.29)	0.30 (0.32)	0.964
PT						
Mean (SD)	11.71 (1.41)	12.03 (2.54)	0.202	11.61 (1.36)	12.40 (4.27)	0.376
INR						
Mean (SD)	0.99 (0.12)	1.02 (0.12)	0.159	0.99 (0.12)	1.05 (0.35)	0.387
FIB						
Mean (SD)	0.99 (0.12)	1.02 (0.19)	0.158	0.99 (0.12)	1.05 (0.35)	0.387
APTT						
Mean (SD)	33.57 (3.91)	33.23 (4.47)	0.467	32.62 (3.69)	32.91 (6.16)	0.824
TT						
Mean (SD)	12.35 (1.66)	12.19 (2.06)	0.731	12.78 (2.03)	12.63 (2.80)	0.598

SD, standard deviation; TG, triglyceride; HDL, high-density lipoprotein; LppA, lipoprotein A; APTT, activated partial thromboplastin time; TT plasma thrombin time.

C value of the SVM −RFE diagnostic model were 0.721 and 0.721, respectively. Table 7 lists the AUC values of the models constructed using different ML algorithms.

To construct a nomogram diagnostic model, we used the results

Table 4
Univariate logistic regression and multivariate logistic regression were used to predict tuberculosis sputum smear positive acid-fast bacilli.

Type	Univariate OR (95 % CI)	P-value	Multivariate OR (95 % CI)	P-value
Age	1.0069 (0.9938–1.02)	0.3010	/	/
ALB	0.9817 (0.9456–1.0185)	0.3291	/	/
ALB/GLB	1.0769 (0.595–1.971)	0.8081	/	/
ALP	1.0064 (1.0017–1.0119)	0.0144	1.0044 (0.9978–1.0112)	0.1935
ALT	1.0297 (1.0118–1.0509)	0.0026	1.0274 (1.0065–1.0489)	0.0102
APOA	1.1926 (0.6127–2.3717)	0.6095	/	/
APOB	0.4368 (0.1833–1.0507)	0.0623	/	/
APOC	1.2751 (0.8401–2.0099)	0.2757	/	/
APTT	0.9829 (0.9829–0.9387)	0.4668	/	/
AST	1.0117 (1.0117–1.0010)	0.0791	/	/
AST/ALT	0.8825 (0.7172–1.1063)	0.2528	/	/
BUN	1.0475 (0.9816–1.1371)	0.2148	/	/
Ca^{2+}	0.3864 (0.0971–1.5180)	0.1746	/	/
Ccr	0.9939 (0.9939–0.9876)	0.0622	/	/
Cl^-	1.0351 (0.9886–1.0830)	0.1367	/	/
Cr	1.0013 (0.9988–1.0053)	0.4170	/	/
CV	5.4158 (0.0072–7346.00)	0.6308	/	/
Cys-C	1.3498 (0.9692–2.1275)	0.1351	/	/
DBil	1.0223 (0.9842–1.0785)	0.3400	/	/
Dbil/TBil	0.4101 (0.1138–1.5016)	0.1748	/	/
EO	0.4530 (0.0021–75.32.2)	0.7829	/	/
ESR	0.9943 (0.9877–1.0011)	0.0996	/	/
FIB	3.0660 (0.7769–15.64)	0.1506	/	/
GGT	1.0053 (1.0014–1.0099)	0.0145	1.0009 (0.9957–1.0061)	0.7288
GLB	0.9918 (0.9634–1.0222)	0.5845	/	/
HCO	0.9444 (0.8943–0.9962)	0.0373	0.9615 (0.9079–1.0182)	0.1789
HCT	0.0345 (0.0012–0.9527)	0.0489	0.0038 (0.0000–2.5473)	0.0933
HDL	1.2655 (0.7634–2.1333)	0.3684	/	/
IBil	1.0813 (1.0157–1.1619)	0.0232	1.0204 (0.9493–1.0967)	0.5836
INR	3.0619 (0.7763–15.611)	0.1509	/	/
K^+	1.2964 (0.8553–2.0069)	0.2337	/	/
LDL	0.9480 (0.7440–1.2133)	0.6678	/	/
LppA	0.8526 (0.5764–1.3252)	0.4344	/	/
LYM	0.9121 (0.1143–7.7399)	0.9317	/	/
MCH	1.0509 (0.9958–1.1079)	0.0671	/	/
MCHC	1.0258 (1.0096–1.0425)	0.0018	1.0266 (1.0062–1.0475)	0.0105

(*continued on next page*)

Table 4 (*continued*)

Type	Univariate OR (95 % CI)	P-value	Multivariate OR (95 % CI)	P-value
MCV	1.0118 (0.9907–1.0328)	0.2671	/	/
Mg^{2+}	0.7054 (0.1804–2.9001)	0.6216	/	/
MO	0.0954 (0.0007–16.9341)	0.3615	/	/
Na^{+}	1.0447 (0.9924–1.0988)	0.0910	/	/
NEUT	1.4543 (0.2722–7.7255)	0.6603	/	/
PAB	0.9995 (0.9965–1.0026)	0.7574	/	/
PCT	0.0873 (0.0114–0.6847)	0.0191	3.2119 (0.0089–163.45)	0.6979
PDW	551.6 (0.6066–2803839)	0.1573	/	/
PLT	0.9981 (0.9967–0.9996)	0.0130	0.9970 (0.9928–1.0013)	0.1698
PT	1.0841 (0.9754–1.2346)	0.1907	/	/
RBC	0.7300 (0.5635–0.9429)	0.0163	1.0645 (0.6461–1.7536)	0.8063
Sex	0.8305 (0.5418–1.2858)	0.3983	/	/
TBA	0.9997 (0.9914–1.0107)	0.9563	/	/
TBil	1.0274 (0.9983–1.0658)	0.1080	/	/
Tch	0.9668 (0.9729–1.1816)	0.7398	/	/
TG	0.8288 (0.6271–1.1179)	0.1976	/	/
TP	0.9835 (0.9577–1.0101)	0.2214	/	/
TT	0.9816 (0.9185–1.0640)	0.5996	/	/
UA	0.9982 (0.9971–0.9992)	0.0005	0.9984 (0.9973–0.9995)	0.0051
WBC	1.01612 (0.9749–1.0723)	0.5090	/	/

obtained using SVM-RFE (Fig. 6A). As can be seen from the fitting curve of the nomogram (Fig. 6B), the proposed diagnostic model achieved a good fit. A decision curve was plotted to analyze the clinical utility of the model; the model exhibited clinical utility when the threshold of the model was in the range of 38 %–94 %, and the NONE line of the decision curve was above the ALL line (Fig. 6C). Fig. 6D–F show the AUC values for each factor in the nomogram diagnostic model.

In the validation cohort, we verified the nomogram diagnostic model, and the diagnostic model achieved a good fit (Fig. 7A), with an AUC value of 0.758 (Fig. 7B) and a C value of 0.758.

4. Discussion

AFB-positive sputum smear is very important for the diagnosis and treatment of pulmonary TB[20,21]. AFB are found in the sputum of only 27 % of patients with TB because TB bacilli are difficult to eliminate from the bronchi deep in the lungs [22]. However, AFB-positive TB patients are considered to be highly infectious in clinical practice and usually need to be isolated and treated in time to avoid serious complications[23]. In this study, we compared the blood samples of 769 patients with pulmonary TB who had undergone AFB examination. We used ML methods to process the data and found multiple differences in the blood examination data between AFB-positive and AFB-negative patients. We selected multiple machine learning methods (SVM-RFE, LASSO, and Random Forest) for feature selection and model construction. The main objective was to fully explore the relationship between blood biomarkers and AFB status while improving the model's predictive performance.

In terms of gender, we found that male patients with TB were greater in number than female patients; however, the AFB positivity rate was

Table 5
Lasso regression results.

HCT	MCHC	PLT
IBil	GGT	ALT
ALP	HCO	UA

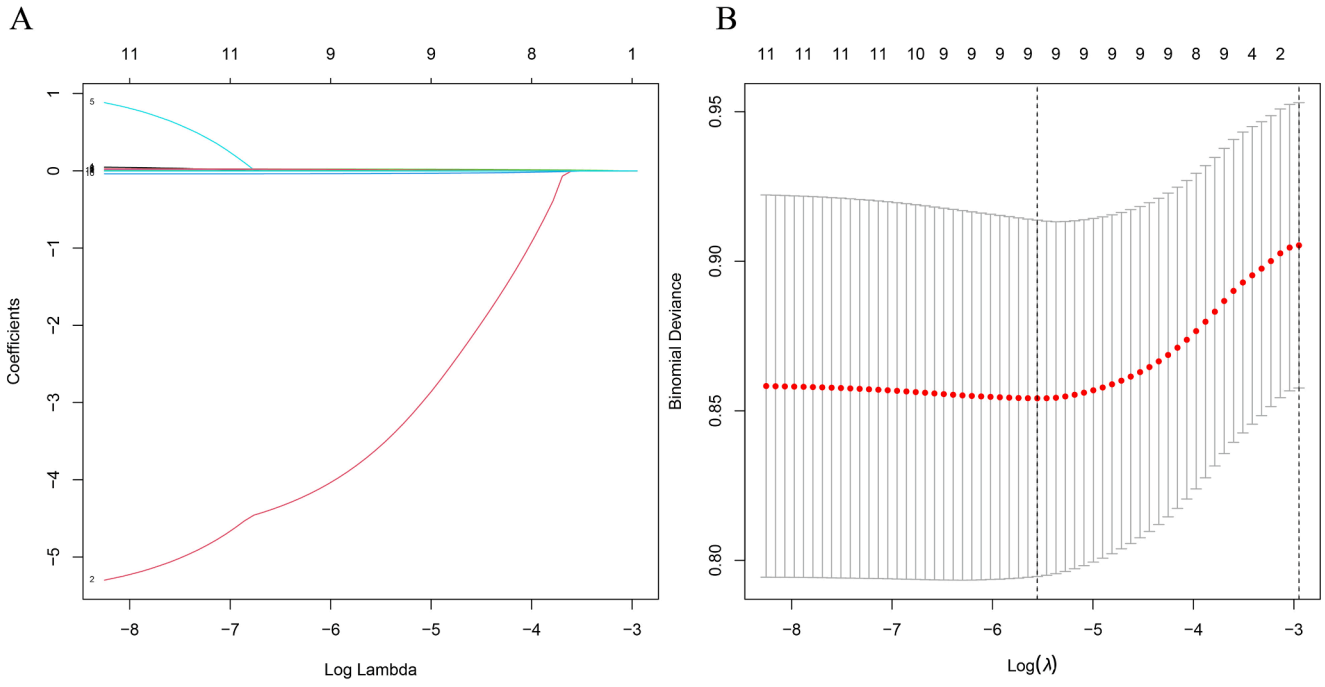


Fig. 2. Lasso regression results. (A) Results of LASSO regression for all variables. (B) There were significant differences in nine factors between bone and joint TB and TB.

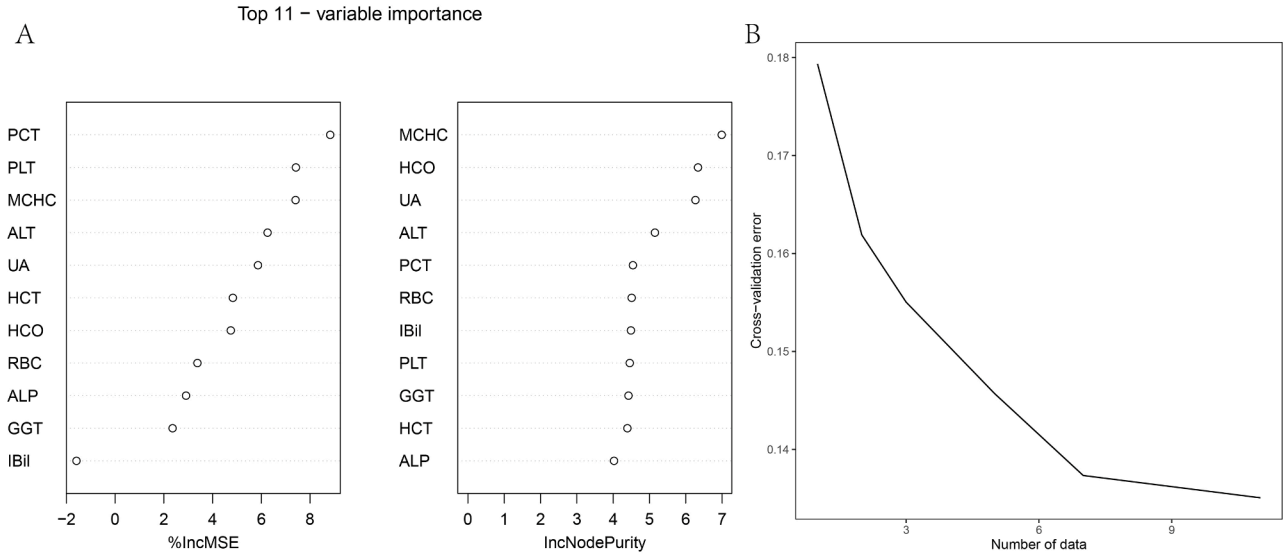


Fig. 3. Random forest. (A) Ranking of two different algorithms for random forest. (B) The regression curve of random forest.

Table 6
Importance ranking of SVM-RFE screening.

TYPE	AvgRank	TYPE	AvgRank
ALT	2.4	PCT	3.0
PLT	4.8	RBC	5.4
MCHC	5.8	HCO	6.0
ALP	6.2	HCT	7.0
IBil	7.4	GGT	8.8
UA	9.2		

not statistically different between the genders. In some studies, it has been suggested that the risk factors of TB include engaging in manual labor, complex outdoors working environments, and smoking and drinking. The elderly and people with weakened immunity are also at a high risk of TB[24,25]. The ESR of AFB-positive patients was higher than that of AFB-negative patients in both the training cohort and the validation cohort. ESR is an indicator of infection and tissue damage and is used to evaluate the effect of TB treatment, indicating that the degree of TB infection and damage in AFB-positive patients is higher than that in AFB-negative patients[26].

RBC, PLT, PCT, MCHC, and HCT are items of blood routine examination. The mean values of RBC and HCT in AFB −positive patients were higher than those in AFB- negative patients. RBCs are the largest blood cells in the blood and contain hemoglobin which transports oxygen and expels carbon dioxide[27]. HCT is the relative ratio of the volume occupied by RBCs in a given volume of blood. When the number of RBCs goes up, so does the percentage of whole blood. RBC count is often slightly elevated in patients with diarrhea, vomiting, and dehydration. platelet distribution width (PDW) did not differ between AFB-positive and AFB-negative patients. No in-depth study has been conducted on whether AFB-positive patients have increased RBC and HCT due to blood concentration or whether other reasons promote the relative increase of RBC [28].

MCHC was significantly increased in AFB-negative patients. MCHC reflects the average hemoglobin concentration per RBC. There are a variety of possibilities for high MCHC, such as RBC aggregation, hemolytic diseases, inflammatory diseases, and electrolyte abnormalities [29]. Anemia was considered when both MCHC and total hemoglobin were low[30]. Due to the lack of hemoglobin collection for technical

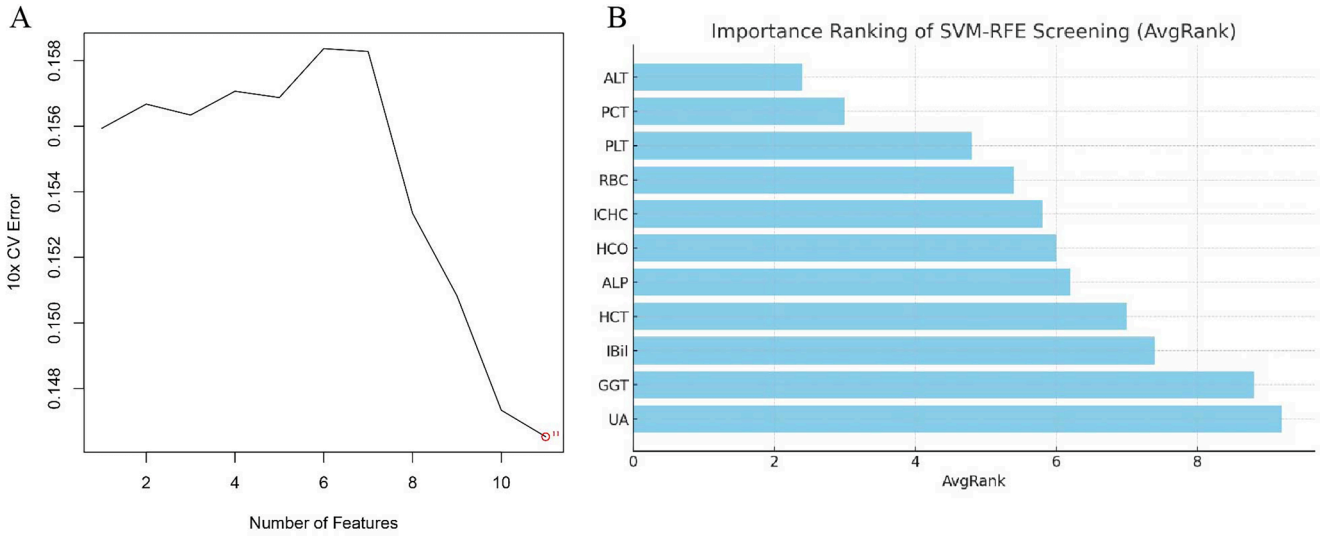


Fig. 4. (A)The optimal diagnostic efficiency in SVM-RFE was achieved with all 11 factors in the model. (B) Bar chart visualizing the importance ranking of the SVM-RFE screening based on the "AvgRank" values.

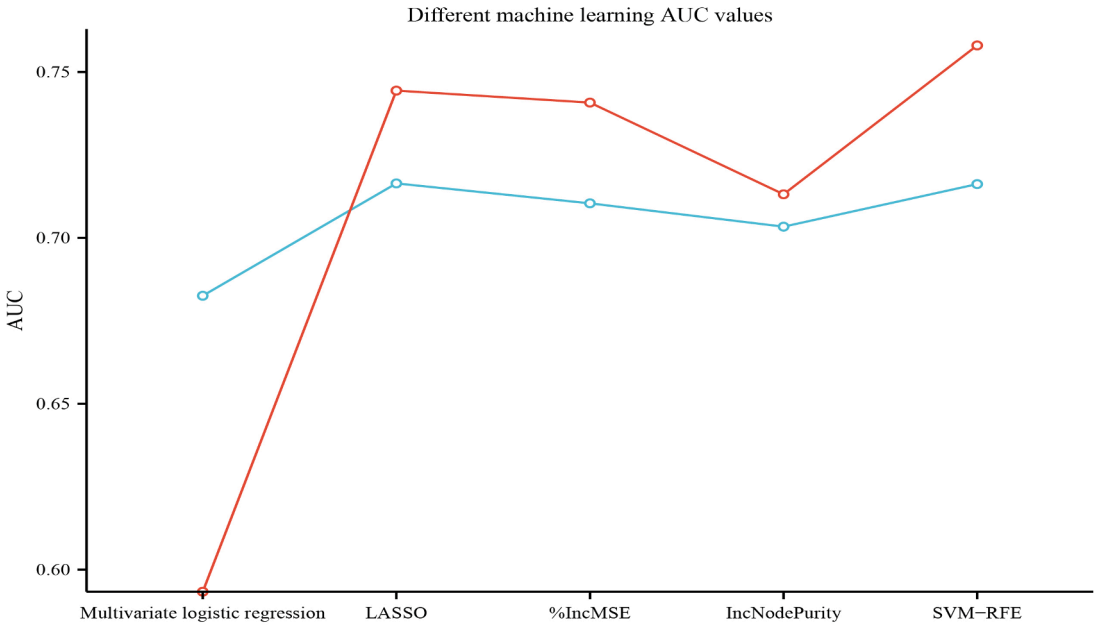


Fig. 5. AUC values of different ML in the training and validation cohort.

Table 7
AUC (95 %CI) values of different machine learning models.

Type	Training cohort	Validation cohort
Multivariate logistic regression analysis	0.683 (0.612–0.737)	0.714 (0.540–0.826)
Lasso regression	0.720 (0.640–0.760)	0.744 (0.571–0.847)
%IncMSE	0.714 (0.633–0.760)	0.741 (0.585–0.838)
IncNodePurity	0.703 (0.624–0.752)	0.713 (0.583–0.807)
SVM-RFE	0.721 (0.644–0.758)	0.758 (0.645–0.824)

reasons, we could not deeply explore whether AFB-positive patients were relatively more anemic compared with AFB-negative patients.

PLT and PCT were significantly increased in AFB-positive patients. Platelets are small cytoplasm shed from mature megakaryocyte cytoplasm in the bone marrow. Platelets play a crucial role in coagulation, repair of damaged blood vessels, thrombosis, and inflammatory response in the human body[31]. PCT is the percentage of platelet volume in peripheral blood to total blood volume. An increase in PLT leads to an increase in PCT, which can also be seen in the correlation diagram. In the study by Chen et al., PLT was higher in patients with pulmonary TB than in those with spinal TB[32]. PLT is significantly associated with liver fibrosis in HIV/HCV co-infected patients; pulmonary fibrosis caused by TB is one of the outcomes of TB[33]. We suspect that pulmonary fibrosis in AFB-positive patients may be more severe than in AFB-negative patients, which is the direction of our further research in the future.

IBil, GGT, ALT, and ALP are part of the liver function test, and they are decreased in AFB-positive patients. IBil is bilirubin that is not tuberculated with glucuronic acid in humans[34]. IBil and DBil were decreased in AFB-positive patients, indicating that less bilirubin was produced in the body. Low bilirubin can be caused by anemia, fatigue, and poor appetite; loss of appetite can also cause low bilirubin. GGT mainly exists in the liver cell membrane and microsomes. Serum GGT mainly comes from the liver and is often used to identify liver system diseases[35]. Elevated GGT is often observed in acute hepatitis, active chronic hepatitis, liver cancer, and obstructive jaundice. High GGT is also associated with coronary heart disease risk[36]. ALT is an enzyme involved in human metabolism and is mainly concentrated in the mitochondria of the liver. When the liver is damaged, ALT is released into the blood. ALP is a zinc-containing protein that is widely distributed in the human body, with the highest content in the liver. ALP is an important indicator of liver diseases. ALP in the sputum and blood of patients with pulmonary TB is higher than that in patients with lung cancer[37]. The ALP content in tuberculous pleural effusions is also higher than that in other pleural effusions[38]. Patients with TB have a 20 % increase in GGT values in the first week after using anti-TB drugs [39]. AFB-positive patients had less hepatotoxicity than those AFB-negative patients; however, we did not collect information regarding the use of antituberculosis drugs, preventing further analysis.

In renal function examinations, UA and HCO were elevated in AFB-positive patients. UA is the end product of purine metabolism and is the main cause of gout. Excessive purine intake increases endogenous purine production, and increased purine metabolism can lead to elevated UA. Uric acid excretion is reduced when the antituberculosis drug ofloxacin is used[40]. Elevated uric acid is a common adverse reaction of antituberculosis drugs in elderly patients with TB[41]. HCO is a commonly used indicator of acid-base balance in the human body. There are no studies on HCO in patients with TB. We found that blood urea nitrogen (BUN) and creatinine (Cr), which are commonly used to evaluate renal function in clinical practice, were higher in AFB-negative patients than in AFB-positive patients, and the renal burden of AFB-negative patients was higher than that of AFB-positive patients, as was the liver burden.

Furthermore, we compared the blood lipid and coagulation function between AFB-positive and AFB-negative patients and found no significant difference between the two tests. Finally, after screening and comparing various ML methods, we selected SVM-RFE as the best ML method. We used univariate logistic regression and SVM-RFE to screen out 11 factors to construct a nomogram diagnostic model and achieved good diagnostic efficiency. The nomogram diagnostic model developed using SVM-RFE has proven to be a valuable tool in predicting AFB positivity, demonstrating robust AUC values of 0.721 and 0.758 in the training and validation cohorts respectively. Its implementation offers a practical approach for streamlining TB diagnosis by quantifying risk based on predictive biomarkers, which is particularly beneficial in resource-limited settings. The model's decision curve analysis confirms

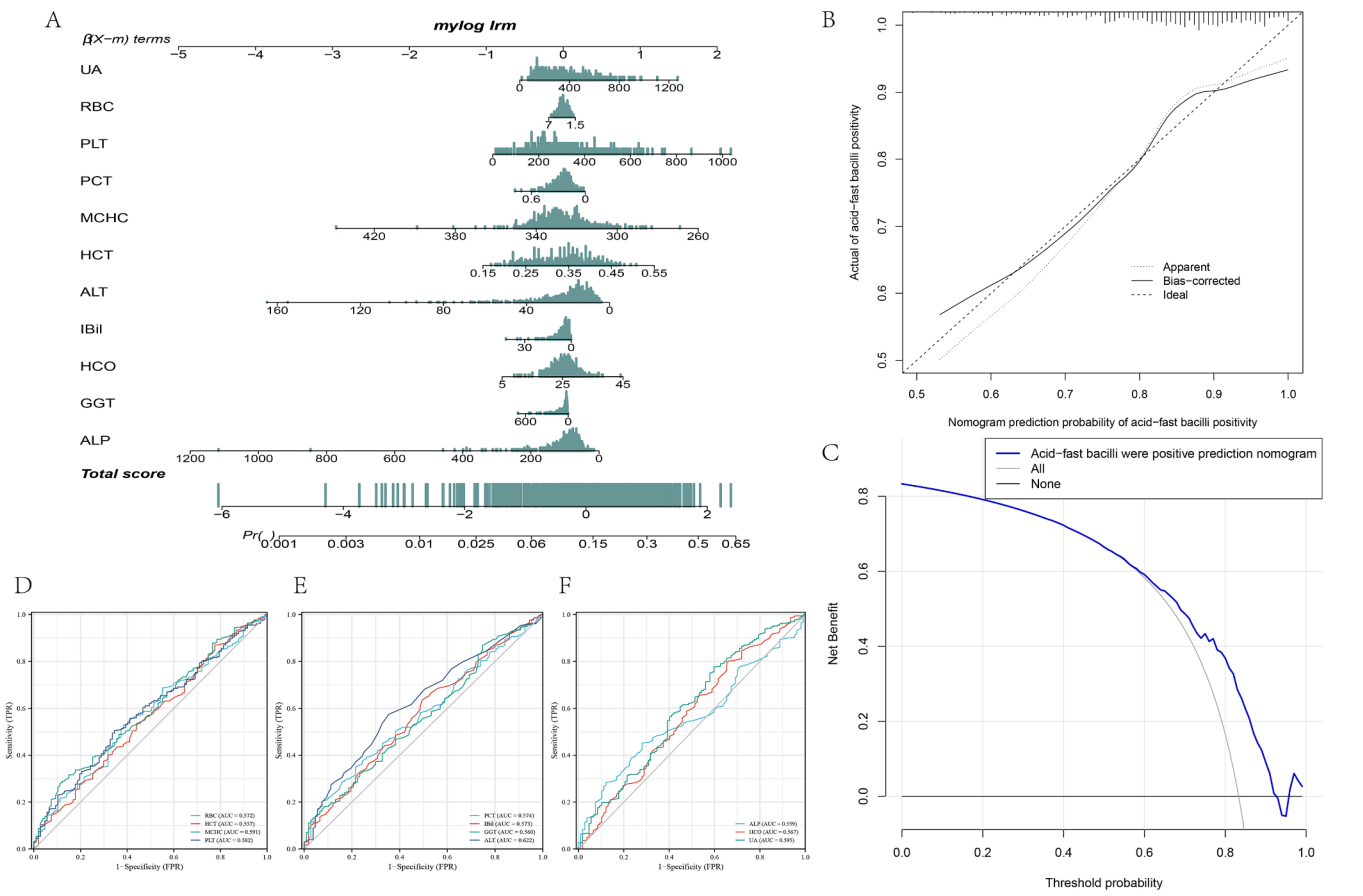


Fig. 6. SVM-RFE diagnosis model. (A) Nomogram to predict the probability of AFB result. (B) Calibration curves for predicting AFB result. (C) Decision curve analysis for the SVM-RFE prediction model. (D) The AUC values of RBC, HCT, MCHC and PLT. (E) The AUC values of PCT, IBil, GGT and ALT. (F) The AUC values of ALP, HCO and UA.

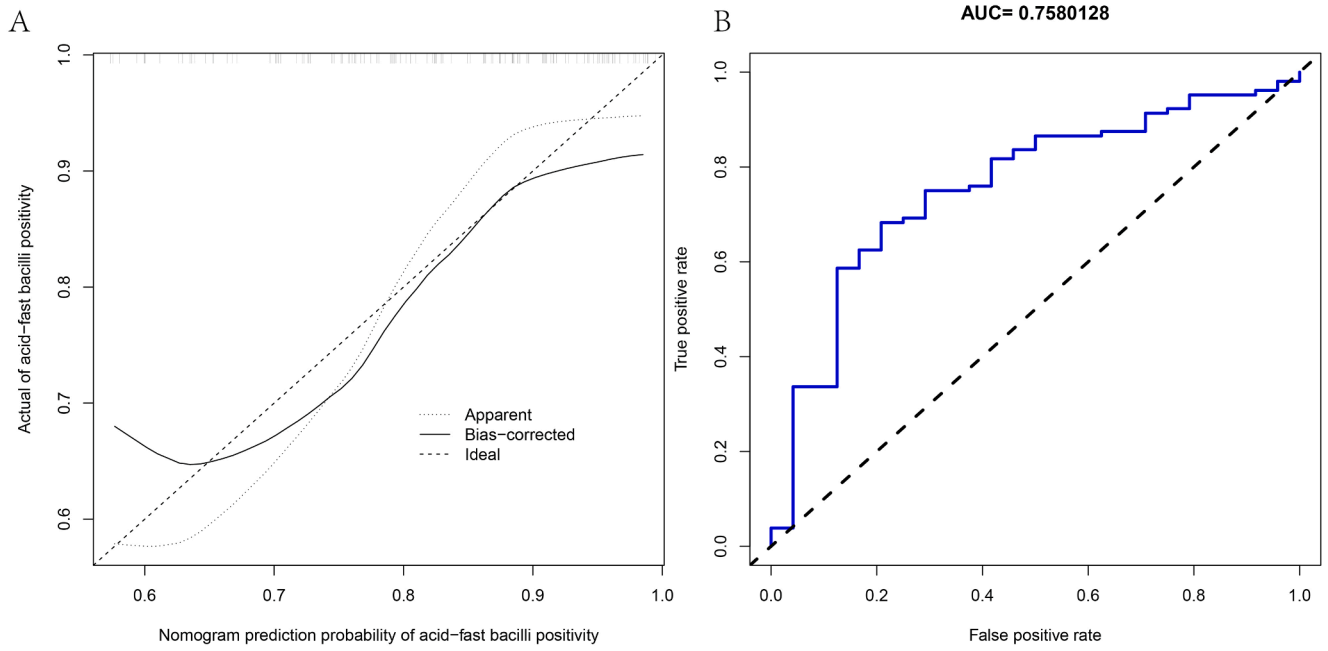


Fig. 7. Validation cohort. (A) Calibration curves for predicting AFB in the validation cohort. (B) The AUC value of SVM-RFE diagnostic model in the cohort was verified.

its clinical utility, supporting its integration into routine diagnostic workflows to enhance the accuracy and efficiency of TB management.

In this study, we collected a large amount of blood data from patients with TB. Liver function test, renal function test, blood lipid test, and coagulation function have rarely been included in previous research on TB. We collected data from different hospitals to validate this study. Nevertheless, this study has some limitations; First, not all potential confounding factors, such as comorbidities and lifestyle, were controlled, which may affect the model's performance. Additionally, the small sample size may limit the robustness and generalizability of the results. Future studies should consider larger sample sizes and control for confounding factors to further validate the model.

And, the number of AFB-positive patients was small. In addition, we did not collect information regarding the use of anti-TB drugs in patients with TB, and the diagnostic efficacy of the final model was not good enough. The patient data in this study were collected from two hospitals in the same city, which may introduce selection bias due to differences in clinical practices and patient demographics, affecting the representativeness of the sample. Additionally, the study did not control for disease stage or prior treatment history, which may introduce bias and impact the generalizability of the results. Future studies should validate these models in different hospitals and regions while controlling for these factors.

To provide more context, we compared our results with other studies that applied machine learning to TB diagnostics. Many studies using models like SVM and Random Forest achieved strong AUC values[42]. However, the generalizability of their findings may be limited due to differences in sample size or data quality. Similarly, while our study showed strong predictive ability with a smaller sample, the limited sample size and potential confounding factors may affect the generalizability. Future research should validate our model on larger datasets to improve its generalizability. We compared the difference between AFB-positive and AFB-negative patients by using different ML methods and screened blood data to construct a nomogram diagnostic model to help doctors more easily infer the AFB results in patients with TB in the absence of testing equipment and help patients with TB to obtain timely treatment and avoid complications. In our study, various machine learning algorithms were explored to construct the most effective diagnostic model. After evaluating the diagnostic performance of each model, the SVM emerged as the optimal choice due to its superior ability to handle high-dimensional data and its robustness in classification tasks. The SVM model was specifically selected for its accuracy in distinguishing between AFB-positive and AFB-negative patients, which is critical in guiding timely treatment decisions. The incorporation of SVM into our nomogram enhances the reliability of the diagnostic tool, making it a valuable resource in clinical settings where traditional testing may be limited.

The model can assist in the rapid screening of TB patients, particularly in resource-limited areas. For real-world application, it needs to be validated on multi-center and diverse datasets to ensure stability across different populations. Additionally, clinical factors such as patient history and disease stage must be considered, making further data expansion and analysis of false positives/negatives essential. Lastly, collaboration with clinicians for real-time feedback and improvements is needed to ensure its effectiveness and broader clinical applicability. The final diagnosis and medication for TB should, however, be based on established guidelines and the clinical judgment of healthcare professionals.

5. Conclusion

In this study, we used ML to construct a diagnostic model to infer AFB results and achieved good diagnostic performance.

CRediT authorship contribution statement

Jichong Zhu: Writing – original draft. **Yong Zhao:** Writing – original draft. **Chengqian Huang:** Data curation. **Chenxing Zhou:** Methodology. **Shaofeng Wu:** Resources. **Tianyou Chen:** Software. **Xinli Zhan:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Professor Xinli Zhan for their selfless assistance at all stages of the research.

Ethics approval and consent to participate

The names and ID numbers of all subjects are hidden to protect data confidentiality. The study obtained the informed consent of the participants. The code of ethics is consistent with the Declaration of Helsinki. The Ethics Committee of the First Affiliated Hospital of Guangxi Medical University and the First Affiliated Hospital of Sun Yat-sen University discussed and approved this study (2022-E344-01) (S upplement 2).

Consent for publication

NA.

Funding

The research was funded by the National Natural Science Foundation of China (81560359 and 81860393), chaired by Professor Xinli Zhan.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jctube.2025.100511.

References

[1] Gagneux S. Ecology and evolution of Mycobacterium tuberculosis. Nat Rev Microbiol 2018;16(4):202–13.
[2] Elkington PT, Ugarte-Gil CA, Friedland JS. Matrix metalloproteinases in tuberculosis. Eur Respir J 2011;38(2):456–64.
[3] Luies L, du Preez I. The Echo of Pulmonary Tuberculosis: Mechanisms of Clinical Symptoms and Other Disease-Induced Systemic Complications. Clin Microbiol Rev 2020;33(4).
[4] 王黎霞, 成诗明, 陈明亭, 赵雁林, 张慧, 姜世闻, et al. 2010年全国第五次结核病流行病学抽样调查报告. 中国防痨杂志. 2012;34(08):485-508.
[5] Furin J, Cox H, Pai M. Tuberculosis. Lancet 2019;393(10181):1642–56.
[6] Brodie D, Schluger NW. The diagnosis of tuberculosis. Clin Chest Med 2005;26(2).
[7] 陆普选. 中国最新肺结核诊断标准要点解读(附视频). 新发传染病电子杂志. 2018;3(01):57-8.
[8] Yeh J. Validation of a model for predicting smear-positive active pulmonary tuberculosis in patients with initial acid-fast bacilli smear-negative sputum. Eur Radiol 2018;28(1):243–56.
[9] Mbeha B, Mine M, Motswaledi MS, Dewar J. Nontuberculous Mycobacteria, Botswana, 2011-2014. Emerg Infect Dis 2019;25(7):1401–3.
[10] Deo RC. Machine Learning in Medicine. Circulation 2015;132(20):1920–30.
[11] Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. J Intern Med 2018;284(6):603–19.
[12] Zhu J, Lu Q, Zhan X, Huang S, Zhou C, Wu S, et al. To infer the probability of cervical ossification of the posterior longitudinal ligament and explore its impact on cervical surgery. Sci Rep 2023;13(1):9816.
[13] Kim TK, Park JH. More about the basic assumptions of t-test: normality and sample size. Korean J Anesthesiol 2019;72(4):331–5.
[14] Neely JG, Hartman JM, Forsen Jr JW, Wallace MS. Laryngoscope 2003;113(10):1719–25.
[15] Barber KE, Wagner JL, Larry RC, Stover KR. Frequency of and risk factors for carbapenem-resistant Enterobacteriaceae. J Med Microbiol 2021;70(2).

[16] Zhu J, Lu Q, Liang T, JieJiang LH, Zhou C, et al. Development and Validation of a Machine Learning-Based Nomogram for Prediction of Ankylosing Spondylitis. Rheumatology and therapy 2022;9(5):1377–97.
[17] Chen H, Zhang J, Sun X, Wang Y, Qian Y. Mitophagy-mediated molecular subtypes depict the hallmarks of the tumour metabolism and guide precision chemotherapy in pancreatic adenocarcinoma. Front Cell Dev Biol 2022;10:901207.
[18] Zhu J, Lu Q, Liang T, JieJiang LH, Zhou C, et al. Development and Validation of a Machine Learning-Based Nomogram for Prediction of Ankylosing Spondylitis. Rheumatology and therapy 2022.
[19] Yang Y, Yi X, Cai Y, Zhang Y, Xu Z. Immune-Associated Gene Signatures and Subtypes to Predict the Progression of Atherosclerotic Plaques Based on Machine Learning. Front Pharmacol 2022;13:865624.
[20] Zhang ZX, Sng LH, Yong Y, Lin LM, Cheng TW, Seong NH, et al. Delays in diagnosis and treatment of pulmonary tuberculosis in AFB smear-negative patients with pneumonia. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease 2017;21 (5):544–9.
[21] Koo HK, Min J, Kim HW, Lee J, Kim JS, Park JS, et al. Prediction of treatment failure and compliance in patients with tuberculosis. BMC Infect Dis 2020;20(1): 622.
[22] Ju Q, Li MX, Chen G, Wang HX, Shi QM, Ge X, et al. Evaluation of the New Automatic Mycob.T Stainer and Scanner for Detecting Acid-fast Bacilli in China. Biomedical and environmental sciences : BES 2018;31(8):572–8.
[23] Lohmann EM, Koster BF, le Cessie S, Kamst-van Agterveld MP, van Soolingen D, Arend SM. Grading of a positive sputum smear and the risk of Mycobacterium tuberculosis transmission. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease 2012;16(11):1477–84.
[24] Zhu QQ, Wu Q, Wang AM, Bao FJ, Zhang YZ, Liu J, et al. Epidemiological characteristics of pulmonary tuberculosis in Anhui Province, Eastern China from 2013 to 2018. PLoS One 2020;15(8):e0237311.
[25] Shaweno D, Horton KC, Hayes RJ, Dodd PJ. Assortative social mixing and sex disparities in tuberculosis burden. Sci Rep 2021;11(1):7530.
[26] Schulman H, Niward K, Abate E, Idh J, Axenram P, Bornefall A, et al. Sedimentation rate and suPAR in relation to disease activity and mortality in patients with tuberculosis. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease 2019;23(11):1155–61.
[27] Rogers S, Doctor A. Red Blood Cell Dysfunction in Critical Illness. Crit Care Clin 2020;36(2):267–92.
[28] Barath B, Somogyi V, Tanczos B, Varga A, Bereczky Z, Nemeth N, et al. Examination of the relation between red blood cell aggregation and hematocrit in human and various experimental animals. Clin Hemorheol Microcirc 2021;78(2): 187–98.
[29] Increased CS, Concentration MCH. Clin Chem 2022;68(6):861–2.
[30] Demeke G, Mengistu G, Abebaw A, Toru M, Yigzaw M, Shiferaw A, et al. Effects of intestinal parasite infection on hematological profiles of pregnant women attending antenatal care at Debre Markos Referral Hospital, Northwest Ethiopia: Institution based prospective cohort study. PLoS One 2021;16(5):e0250990.
[31] van der Meijden PEJ, Heemskerk JWM. Platelet biology and functions: new concepts and clinical perspectives. Nat Rev Cardiol 2019;16(3):166–79.
[32] Chen L, Liu C, Liang T, Ye Z, Huang S, Chen J, et al. Monocyte-to-Lymphocyte Ratio Was an Independent Factor of the Severity of Spinal Tuberculosis. Oxid Med Cell Longev 2022;2022:7340330.
[33] Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, Montaner J, et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology (Baltimore, MD) 2006;43(6): 1317–25.
[34] Wu TW, Zumbulyadis N, Gross S, Gohlke RS. Clin Chem 1980;26(9):1323–35.
[35] Newton KP, Lavine JE, Wilson L, Behling C, Vos MB, Molleston JP, et al. Alanine Aminotransferase and Gamma-Glutamyl Transpeptidase Predict Histologic Improvement in Pediatric Nonalcoholic Steatohepatitis. Hepatology (Baltimore, MD) 2021;73(3):937–51.
[36] Ndrepepa G, Colleran R, Kastrati A. Gamma-glutamyl transferase and the risk of atherosclerosis and coronary heart disease. Clinica chimica acta; international journal of clinical chemistry 2018;476:130–8.
[37] Jadhav AA, Jain A. Sputum adenosine deaminase and alkaline phosphatase activity in pulmonary tuberculosis. Arch Physiol Biochem 2012;118(1):6–9.
[38] Jadhav AA, Bardapurkar JS, Jain A. Alkaline phosphatase: Distinguishing between tuberculous and nontuberculous pleural effusion. Lung India : official organ of Indian Chest Society 2009;26(3):77–80.
[39] Enoh JE, Cho FN, Manfo FP, Ako SE, Akum EA. Abnormal Levels of Liver Enzymes and Hepatotoxicity in HIV-Positive, TB, and HIV/TB-Coinfected Patients on Treatment in Fako Division, Southwest Region of Cameroon. Biomed Res Int 2020; 2020:9631731.
[40] Kumar AK, Gurumurthy P. Disposition of uric acid upon administration of ofloxacin alone and in combination with other anti-tuberculosis drugs. Indian J Exp Biol 2004;42(3):323–5.
[41] Choi H, Park HA, Hyun IG, Kim JH, Hwang YI, Jang SH, et al. Incidence and outcomes of adverse drug reactions to first-line anti-tuberculosis drugs and their effects on the quality of life: a multicenter prospective cohort study. Pharmacoepidemiol Drug Saf 2022.
[42] Chen L, Yuan L, Sun T, Liu R, Huang Q, Deng S. The performance of VCS(volume, conductivity, light scatter) parameters in distinguishing latent tuberculosis and active tuberculosis by using machine learning algorithm. BMC Infect Dis 2023;23 (1):881.