

A Nomogram for Predicting the Risk of Tuberculosis Infection

Huanqing Liu

The First Affiliated Hospital of Jinan University

Tingting Li

The First Affiliated Hospital of Jinan University

Jun Lyn (✉ lyujun2020@jnu.edu.cn)

The First Affiliated Hospital of Jinan University

Research Article

Keywords: tuberculosis infection, nomogram, predictors

Posted Date: May 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-527122/v1>

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

[Read Full License](#)

Abstract

Background: Tuberculosis (TB) has become one of the main causes of deaths worldwide. Because of certain conditions prevent the early TB diagnosis and treatment to some extent. This study aimed to develop a tuberculosis (TB) infection risk model and validate the ability of nomogram to predict risk for TB infection in a Chinese population.

Methods: A prediction model based on the training dataset of 272 patients was established. Minimum absolute shrinkage and selection operator regression model were adopted to optimize the feature selection of the TB infection risk model. Using multivariate logistic regression analysis, a predictive model combining the features selected in the minimum absolute shrinkage and the selected operator regression model was constructed. The ability of this predictive model to discriminate and calibrate TB infection risk and its utility in clinical settings were assessed via concordance index (C-index), calibration plot, area under time-dependent receiver operating characteristic curve (AUC), and decision curve analysis (DCA). The clinical practicality of nomogram was evaluated via net reclassification index (NRI) and integrated discrimination improvement (IDI). Bootstrapping validation allowed internal validation.

Results: According to this predictive nomogram, the main predictors of TB infection risk were gender, age, smoking history, fever, hemoptysis, fatigue, emaciation, CD8, CD4/CD8, ESR, CRP, and abnormal liver function. The model exhibited superior risk calibration and discrimination with a C-index of 0.737 (95% CI: 0.685–0.789). The internal validation reached a C-index value of 0.688. The predictive model was able to produce an AUC of 0.729 (95% CI: 0.677–0.781). Analysis of the decision curve revealed the TB infection probability nomogram manifested its clinical usefulness on the condition that intervention was decided at the TB probability threshold of 13%. Moreover, results demonstrated that nomogram could be utilized as an effective prognostic tool according to NRI and IDI.

Conclusion: The new TB probability nomogram for predicting TB infection risk developed herein that combines various factors, such gender, age, smoking history, fever, hemoptysis, fatigue, emaciation, CD8, CD4/CD8, ESR, CRP, and abnormal liver function is convenient and useful in predicting individual TB risks among patients.

Background

Tuberculosis (TB), which is caused by a single pathogenic bacterium, is one of the main causes of deaths worldwide. Almost a fourth of the world population is infected with *Mycobacterium tuberculosis* (MTB) [1]. According to the World Health Organization (WHO), the number of new TB cases worldwide in 2018 was estimated to be about 10 million. In recent years, the incidence and mortality of TB in China have gradually decreased, but its prevalence rate remains second only to India; the number of TB cases in China accounts for 9% of the total globally [2]. Undoubtedly, TB is an urgent public health concern. Some of the traditional methods for detecting MTB are etiological diagnosis, such as acid-fast staining smear and culture of MTB. However, these methods require high-quality clinical specimens. Moreover, they have

a low detection rate, poor sensitivity, and time consuming; furthermore, their use may result in misdiagnosis or delayed diagnosis [3, 4]. These disadvantages prevent the early TB diagnosis and treatment to some extent. Molecular biology examination has recently an important supplement to traditional diagnostic methods. Molecular methods are rapid, accurate, efficient, and have a high flux. These methods represent a new avenue for TB diagnosis, treatment, and control. However, molecular methods are expensive. Thus, their application in grassroots medical institutions and far-flung areas is severely limited.

TB is primarily transmitted through the respiratory tract. In hospital settings or public places, people should be protected from pathogens. Thus, measures to prevent the airborne spread of TB must be taken to protect people at risk of TB infection. However, recognizing people at high risk of TB infection is difficult. Therefore, a practical risk prediction model for TB must be developed. A nomogram is an easy-to-use tool that can create a simple and intuitive prediction model chart. It can be used to quantify the risk of a clinical event by graphically representing the influence of each prediction factor on the results. Hence, a reader/clinician can more easily understand the impact of each predictor on the results. A clinician can predict 1-, 3-, and 5-year survival by looking at the sum of all predictors for a given patient [5].

In this study, we identified a combination of variables that can be used to accurately predict the risk of TB infection. A nomogram for this purpose was established. This nomogram can aid physicians in diagnosing and applying appropriate TB treatment.

Patients And Methods

Patients

A retrospective review of 272 consecutive patients who came to our hospital from January 2014 to December 2014 was conducted. TB diagnosis was based on acid-fast staining smear of MTB. Secondary analysis was ethically permitted by the Xi'an Chest Hospital (Xi'an, China), and all subjects have received informed consent. All patient information was not made public.

The variables analyzed were gender, age, alcoholism, smoking history, cough, phlegm, fever, hemoptysis, night sweat, fatigue, emaciation, chest pain, short breath, cluster of differentiation 4(CD4), cluster of differentiation 8(CD8), CD4/CD8, erythrocyte sedimentation rate(ESR), C-reactive protein(CRP), abnormal liver function, albumin(ALB), and total protein(TP).

Statistics analysis

Count (%) was adopted to express all data comprising demographic, disease, and therapy features. Statistical analysis was performed using the R software (Version 3.1.1; <https://www.R-project.org>).

Least absolute shrinkage and selection operator (LASSO) regression technique was utilized in selecting the best prediction characteristics in risk factors of patients with TB. Characteristics with nonzero

coefficients in the LASSO regression model were selected. Using multivariable logistic regression analysis, a predictive model and a nomogram for TB was developed. The characteristics had been deemed to be odds ratio (OR), which possessed 95% confidence interval (CI) and as P-value. Statistically significant levels were reflected in two aspects. The model consisted of sociodemographic variables with a P-value of 0.05, whereas all variables related to disease and therapy features were also covered. Using cohort studies, all possible predictors were applied in establishing a predictive model of TB risk. The calibration of the forecast nomogram was evaluated by plotting calibrated curves. Harrell's C-index was also measured to quantify the prediction nomogram and assess its discriminatory performance. Area under the curve (AUC) was calculated to determine the discrimination ability of the model. Bootstrapping (resampling = 113) was performed to internally validate the model. The accuracy of the nomogram was compared with that of a doctor's judgment on the basis of his or her clinical experience by determining net reclassification improvement (NRI) and integrated discriminative improvement (IDI). As alternatives to AUC, NRI and IDI were employed to assess improvements in risk prediction and measure the practicality of this novel model. Decision curve analyses (DCA) were conducted to determine the practicality or feasibility and the advantages of this predictive model.

Results

Patients' characteristics

A total of 385 patients, including 243 patients with TB and 142 patients with other diseases, were treated in our hospital from January 2014 and December 2014. About 70% of the patients (272) were randomly selected and analyzed to build the model, and 30% of the patients (113) were evaluated to validate the model. Information of the patients' in the two groups, such as demographics, disease, and therapy characteristics, is summarized in Table 1.

Table 1. Distinctions amid \demographic and clinical features for TB and non-TB groups

Demographic characteristics	n(%)		
	TB(n=177)	non-TB(n=95)	Total(n=272)
gender			
male	133(75.14)	53(55.79)	186(68.38)
female	44(24.86)	42(44.21)	86(31.62)
alcoholism			
YES	16(9.04)	3(3.16)	19(6.99)
NO	161(90.96)	92(96.84)	253(93.01)
age			
<50	117(66.10)	54(56.84)	171(62.87)
50-59	31(17.51)	19(20.00)	50(18.38)
60-69	14(7.91)	16(16.84)	30(11.03)
70-79	11(6.21)	3(3.16)	14(5.15)
80+	4(2.26)	3(3.16)	7(2.57)
smoking			
YES	48(27.12)	24(25.26)	72(26.47)
NO	129(72.88)	71(74.74)	200(73.53)
cough			
YES	161(90.96)	86(90.53)	247(90.81)
NO	16(9.04)	9(9.47)	25(9.19)
phlegm			
YES	151(85.31)	77(81.05)	228(83.82)
NO	26(14.69)	18(18.95)	44(16.18)
fever			
YES	93(52.54)	51(53.68)	144(52.94)
NO	84(47.46)	44(46.32)	128(47.06)
hemoptysis			
YES	34(19.21)	21(22.11)	55(20.22)
NO	143(80.79)	74(77.89)	217(79.78)

fatigue				
YES	57(32.20)	27(28.42)	84(30.88)	
NO	120(67.80)	68(71.58)	188(69.12)	
night sweat				
YES	56(31.64)	22(23.16)	78(28.68)	
NO	121(68.36)	73(76.84)	194(71.32)	
chest pain				
YES	33(18.64)	19(20.00)	52(19.12)	
NO	144(81.36)	76(80.00)	220(80.88)	
emaciation				
YES	35(19.77)	11(11.58)	46(16.91)	
NO	142(80.23)	84(88.42)	226(83.09)	
CD4				
<35%	19(10.73)	10(10.53)	29(10.66)	
35-55%	74(41.81)	44(46.32)	118(43.38)	
>55%	84(47.46)	41(43.16)	125(45.96)	
CD8				
<20%	15(8.47)	7(7.37)	22(8.09)	
20-30%	37(20.90)	16(16.84)	53(19.49)	
>30%	125(70.62)	72(75.79)	197(72.43)	
CD4/CD8				
<1.4	90(50.85)	49(51.58)	139(51.10)	
1.4-2.0	42(23.73)	28(29.47)	70(25.74)	
>2.0	45(25.42)	18(18.95)	63(23.16)	
ESR [mm/h]				
male:0-15; female:0-20	31(17.51)	38(40.00)	69(25.37)	
male:>15; female:>20	146(82.49)	57(60.00)	203(74.63)	

CRP[mg/L]			
0-6	49(27.68)	47(49.47)	96(35.29)
>6	128(72.32)	48(50.53)	176(64.71)
abnormal liver function			
YES	21(11.86)	7(7.37)	28(10.29)
NO	156(88.14)	88(92.63)	244(89.71)
ALB[g/L]			
≤40	62(35.03)	29(30.53)	91(33.46)
40-55	52(29.38)	40(42.11)	92(33.82)
≥55	63(35.59)	26(27.37)	89(32.72)
TP[g/L]			
≤65	92(51.98)	51(53.68)	143(52.57)
65-85	83(46.89)	44(46.32)	127(46.69)
≥85	2(1.13)	0(0.00)	2(0.74)

Feature selection

Among the 24 characteristics covering demographics, diseases, and therapy characteristics from 272 patients in the cohort (Figures 1A and 1B), only 12 possible predictive factors, namely, gender, age, smoking history, fever, hemoptysis, fatigue, emaciation, CD8, CD4/CD8, ESR, CRP, and abnormal liver function, remained (Table 2). These factors possessed nonzero coefficients in the LASSO regression model.

Table 2. Multivariable logistic regression analysis for constructing a predictive model through incorporation of nonzero coefficients chosen in the LASSO regression mode

Intercept and variable	Prediction model		
	β	Odds ratio (95% CI)	P-value
Intercept	-0.5390	0.5833(0.226-1.518)	0.264
gender			
male	1.1411	3.1301(1.802-5.526)	<0.001
age			
50-59	-0.7238	0.4849(0.258-0.911)	0.024
60-69	-1.2125	0.2975(0.142-0.612)	0.001
70-79	-0.5780	0.5610(0.222-1.461)	0.225
80+	-1.0385	0.3540(0.076-1.911)	0.192
smoking	-0.6870	0.5031(0.280-0.897)	0.020
fever	-0.6187	0.5387(0.326-0.878)	0.014
hemoptysis	-0.4445	0.6410(0.363-1.133)	0.124
fatigue	-0.3683	0.6919(0.408-1.171)	0.170
emaciation	0.5835	1.7924(0.916-3.657)	0.097
CD8			
20-30%	0.0620	1.0639(0.258-4.146)	0.930
>30%	-0.0354	0.9652(0.395-2.272)	0.936
CD4/CD8			
1.4-2.0	0.0626	1.0646(0.613-1.857)	0.824
>2.0	0.4970	1.6437(0.574-5.213)	0.371
ESR	0.9739	2.6483(1.506-4.701)	<0.001
CRP	0.6867	1.9871(1.150-3.452)	0.014
abnormal liver function	0.4042	1.4981(0.703-3.369)	0.309

Establishment of a personalized predictive model

The results of logistic regression analyses of gender, age, smoking history, fever, hemoptysis, fatigue, emaciation, CD8, CD4/CD8, ESR, CRP, and abnormal liver function are given in Table 2. A model incorporating the aforementioned independent predictive factors was constructed and displayed as a nomogram (Figure 2).

Performance of the TB risk nomogram in the cohort

According to the calibration curves, the performance of the infection risk nomogram for TB prediction was excellent and consistent with this cohort (Figure 3). The C-index for the predictive nomogram was 0.737 (95% CI: 0.685–0.789). Bootstrapping proved the C-index was 0.688. These results demonstrated the outstanding discriminative capability of this model. According to the TB risk nomogram, the predictive ability of the model was excellent.

Validation of the nomogram for TB risk prediction

The training data were utilized to plot ROC curves. The AUC of the training set was statistically significant (0.729, 95% CI: 0.677–0.781). Analysis of accuracy showed that the NRI and IDI values in the training cohort were 0.224 (95% CI: 0.087–0.553) and 0.007 (95% CI: –0.0096–0.0236), respectively.

Clinical use

The decision curve of the TB nomogram is plotted in Figure 5. According to the decision curve, when the threshold probabilities of patients and doctors were 13% and 83%, respectively. In this scope, the net benefit could be contrasted with multiple overlapping parts according to the TB risk nomogram.

Discussion

Nomogram is widely utilized in predicting tumors and various diseases [6–8]. It relies on a humanized digital interface for high precision and easy understanding of prognosis to facilitate a rational clinical decision-making [5]. Nomogram has higher clinical value than traditional predictive models [9]. In the present study, a nomogram for TB risk prediction was established. It consisted of 12 variables, namely, gender, age, smoking history, fever, hemoptysis, fatigue, emaciation, CD8, CD4/CD8, ESR, CRP, and abnormal liver function. This nomogram is a comparatively precise predictive tool for diagnosing patients with TB. Using an internally validated cohort, results showed that this model has excellent discriminative and calibration ability. In particular, the C-index indicated that the nomogram has a large sample size. Thus, it can be widely utilized in clinical settings.

TB is a prominent public health concern worldwide. According to previous studies, recognition of TB infections remains unsatisfactory. Physicians recognize that over half of patients have active TB, but they still fail to enact preventive measures [10]. As the cost of healthcare increases and as resources become scarce, suspected cases must be promptly and correctly identified. Moreover, therapy is commonly postponed for patients with TB who are not subjected to sputum or drug-resistance experiments or for patients with sputum-negative TB. In some cases, patients undergo improper therapy. Several studies examined the factors related to TB infection, some of which have varying degrees of success in terms of risk.

In the present study, a nomogram for TB risk prediction was established on the basis of 12 variables for patients with TB. LASSO regression analysis was performed to filter the variables involved in the

nomogram. Clinically important predictive factors were evaluated. Results that smoking history was one of the predictors in the nomogram. Systematic reviews revealed that smoking has an adverse relationship with global TB epidemic [11]. The role of smoking in TB pathogenesis is associated with ciliary dysfunction, decreased immune responses, defects in immunological reactions of macrophages, with or without reduced CD4 count, and increased susceptibility to MTB infection [12]. The TB epidemic disproportionately affects males and females. Borgdorff et al. [13] provided compelling evidence that males are predominantly infected with TB than females. However, other researchers argued that this conclusion is due to reporting bias as TB infection among females may be underreported in developing regions [14]. In the present study, gender was found to be one of the predictors. TB incidence was higher in males than in females probably because most men are the breadwinner of the family. Hence, they are more exposed to TB infection than females. Moreover, the elderly are more susceptible to infectious diseases than younger individuals, especially to respiratory tract infections [15]. According to the WHO Global Tuberculosis Report (2014), the greatest TB infection rates worldwide were recorded in people aged 45–55 years. In the Western Pacific region, Eastern Mediterranean, and Southeast Asia, TB infections are high among the elderly, especially among those over 65 years old [16]. This report was consistent with our nomogram prediction. Kara [17] reported that four clinical factors, including chronic symptoms, such as persistent fever, fatigue, and emaciation, are associated with TB. This finding was consistent with the results of our prediction model. Several laboratory inspection indicators were also inputted into the model. Results indicated that CD8, CD4/CD8, ESR, CRP, and abnormal liver function can be used as predictors to predict TB risk. CD4⁺ Th1 cells and CD8⁺ T cells generate IFN- γ , which is crucial to the protection of MTB-infected humans [18]. Among patients with reactive TB, CD4⁺ and CD8⁺ cell counts are low and the CD4/CD8 ratio is reduced [19]. Elevated levels of inflammatory markers (CRP and ESR) are commonly noted in patients with TB [20]. CRP may be used as an indicator of TB infection with good sensitivity [21]. However, it may not be able to detect less severe cases [22].

In conclusion, this study validated the clinical variables that can be used to risk-stratify patients with suspected TB upon admission. The predictive nomogram developed herein can aid doctors in identifying patients both at low risk and at high risk for TB. However, this study has several limitations. First, it was a retrospective study conducted in the same medical institute. Therefore, sampling and selection bias due to the heterogeneous and small number of subjects might have affected our analyses. Therefore, the nomogram developed herein must be explored if it can be applied to other regions or nations and multiple centers. Finally, the unmeasured variables might have produced extra effects.

Declarations

Acknowledgments

We thank all colleagues involved in the study for their contributions.

Author Contributions

Huan-qing Liu performed statistical analysis and data interpretation, and was responsible for the quality control of data and algorithms. Ting-ting Li performed literature research. Jun Lyn contributed to the study concept and study design. All authors contributed to writing of the manuscript and approved the final version.

Funding

No funding.

Availability of data and material

The datasets analyzed during current study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Ethics, consent and permissions: All procedures performed were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This current study was approved by the ethics committee of the Xi'an Chest hospital (No. S2021-0001). This article does not contain any studies with animals performed by any of the authors. All participants participated in the study with informed consent (in writing, with parents or legal guardians for participants under 16 years of age).

Competing interests

The authors declare that there is no conflict of interest.

Consent for publication

Not applicable

References

1. Houben RM, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS Med.* 2016;13(10):e1002152. doi:10.1371/journal.pmed.1002152
2. World Health Organization. Global tuberculosis report 2019[M]. Geneva: World Health Organization, 2019.
3. Fukunaga H, Murakami T, Gondo T, Sugi K, Ishihara T. Sensitivity of acid-fast staining for *Mycobacterium tuberculosis* in formalin-fixed tissue. *Am J Respir Crit Care Med.* 2002;166(7):994–997. doi:10.1164/rccm.2111028.
4. Shrestha NK, Tuohy MJ, Hall GS, Reischl U, Gordon SM, Procop GW. Detection and differentiation of *Mycobacterium tuberculosis* and nontuberculous mycobacterial isolates by real-time PCR. *J Clin Microbiol.* 2003;41(11):5121–5126. doi:10.1128/jcm.41.11.5121-5126.2003.

5. Park SY. Nomogram: An analogue tool to deliver digital knowledge. *J Thorac Cardiovasc Surg.* 2018;155(4):1793. doi:10.1016/j.jtcvs.2017.12.107
6. Su TW, Zhong X, Ye L, et al. A nomogram for predicting the presence of germline mutations in pheochromocytomas and paragangliomas. *Endocrine.* 2019;66(3):666–672. doi:10.1007/s12020-019-02075-9
7. Jin C, Cao J, Cai Y, et al. A nomogram for predicting the risk of invasive pulmonary adenocarcinoma for patients with solitary peripheral subsolid nodules. *J Thorac Cardiovasc Surg.* 2017;153(2):462–469.e1. doi:10.1016/j.jtcvs.2016.10.019
8. Chung SM, Park JC, Moon JS, Lee JY. Novel nomogram for screening the risk of developing diabetes in a Korean population. *Diabetes Res Clin Pract.* 2018;142:286–293. doi:10.1016/j.diabres.2018.05.036
9. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* 2015;16(4):e173-e180. doi:10.1016/S1470-2045(14)71116-7
10. Sokolove PE, Rossman L, Cohen SH. The emergency department presentation of patients with active pulmonary tuberculosis. *Acad Emerg Med.* 2000;7(9):1056–1060. doi:10.1111/j.1553-2712.2000.tb02100.x
11. Bates MN, Khalakdina A, Pai M, Chang L, Lessa F, Smith KR. Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis. *Arch Intern Med.* 2007;167(4):335–342. doi:10.1001/archinte.167.4.335
12. van Zyl Smit RN, Pai M, Yew WW, et al. Global lung health: the colliding epidemics of tuberculosis, tobacco smoking, HIV and COPD. *Eur Respir J.* 2010;35(1):27–33. doi:10.1183/09031936.00072909
13. Borgdorff MW, Nagelkerke NJ, Dye C, Nunn P. Gender and tuberculosis: a comparison of prevalence surveys with notification data to explore sex differences in case detection. *Int J Tuberc Lung Dis.* 2000;4(2):123–132.
14. Holmes CB, Hausler H, Nunn P. A review of sex differences in the epidemiology of tuberculosis. *Int J Tuberc Lung Dis.* 1998;2(2):96–104.
15. Gardner ID. The effect of aging on susceptibility to infection. *Rev Infect Dis.* 1980;2(5):801–810. doi:10.1093/clinids/2.5.801
16. Organization WH. Global Tuberculosis Report 2014. In: World Health Organization, Geneva, Switzerland. 2014.
17. Rakoczy KS, Cohen SH, Nguyen HH. Derivation and validation of a clinical prediction score for isolation of inpatients with suspected pulmonary tuberculosis. *Infect Control Hosp Epidemiol.* 2008;29(10):927–932. doi:10.1086/590667
18. Winslow GM, Cooper A, Reiley W, Chatterjee M, Woodland DL. Early T-cell responses in tuberculosis immunity. *Immunol Rev.* 2008;225:284–299. doi:10.1111/j.1600-065X.2008.00693.x
19. Ashtekar MD, Samuel AM, Kadival GV, Sakhalkar V, Rajadhyaksha S, Viridi SS. T lymphocytes in pulmonary tuberculosis. *Indian J Med Res.* 1993;97:14–17.

20. Ding RD, Zhang HJ. Effect of linezolid on serum PCT, ESR, and CRP in patients with pulmonary tuberculosis and pneumonia. *Medicine (Baltimore)*. 2018;97(37):e12177. doi:10.1097/MD.00000000000012177
21. Sigal GB, Segal MR, Mathew A, et al. Biomarkers of Tuberculosis Severity and Treatment Effect: A Directed Screen of 70 Host Markers in a Randomized Clinical Trial. *EBioMedicine*. 2017;25:112–121. doi:10.1016/j.ebiom.2017.10.018
22. Ito K, Yoshiyama T, Wada M, Ogata H. C-reactive protein in patients with bacteriological positive lung tuberculosis[J]. *Kekkaku*, 2004, 79(4): 309–311.

Figures

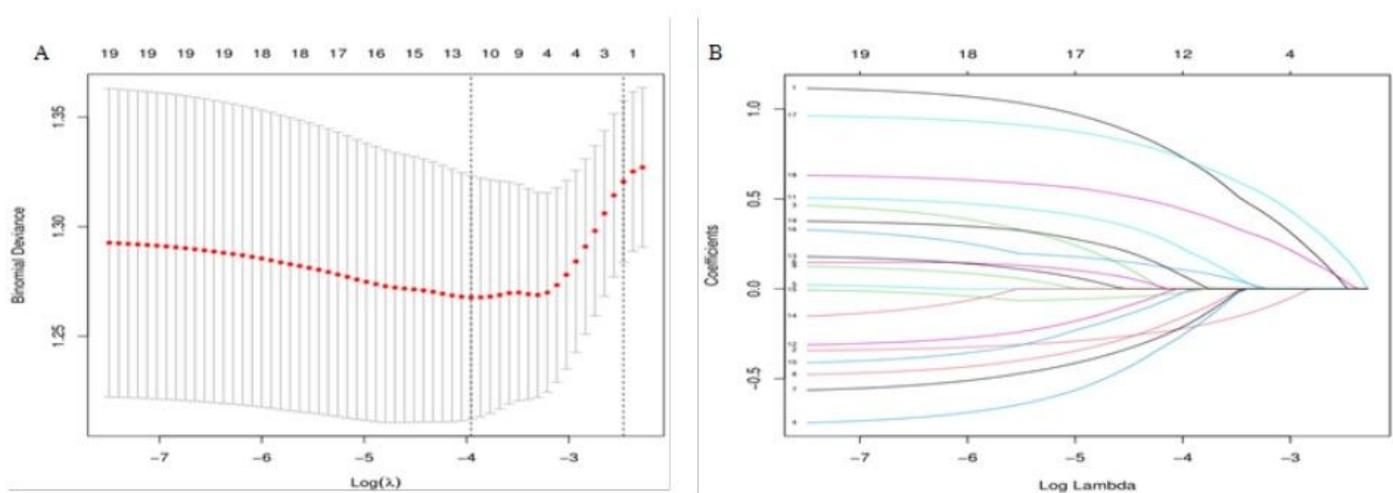


Figure 1

Demographic and clinical feature selection using the LASSO binary logistic regression mode. (A) Optimal parameter (λ) selection in the LASSO model used fivefold cross-validation via minimum criteria. The partial likelihood deviance (binomial deviance) curve was plotted versus $\log(\lambda)$. Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 SE of the minimum criteria (the 1-SE criteria). (B) LASSO coefficient profiles of the 24 features. A coefficient profile plot was produced against the $\log(\lambda)$ sequence. Vertical line was drawn at the value selected using fivefold cross-validation, where optimal λ resulted in 12 features with nonzero coefficients.

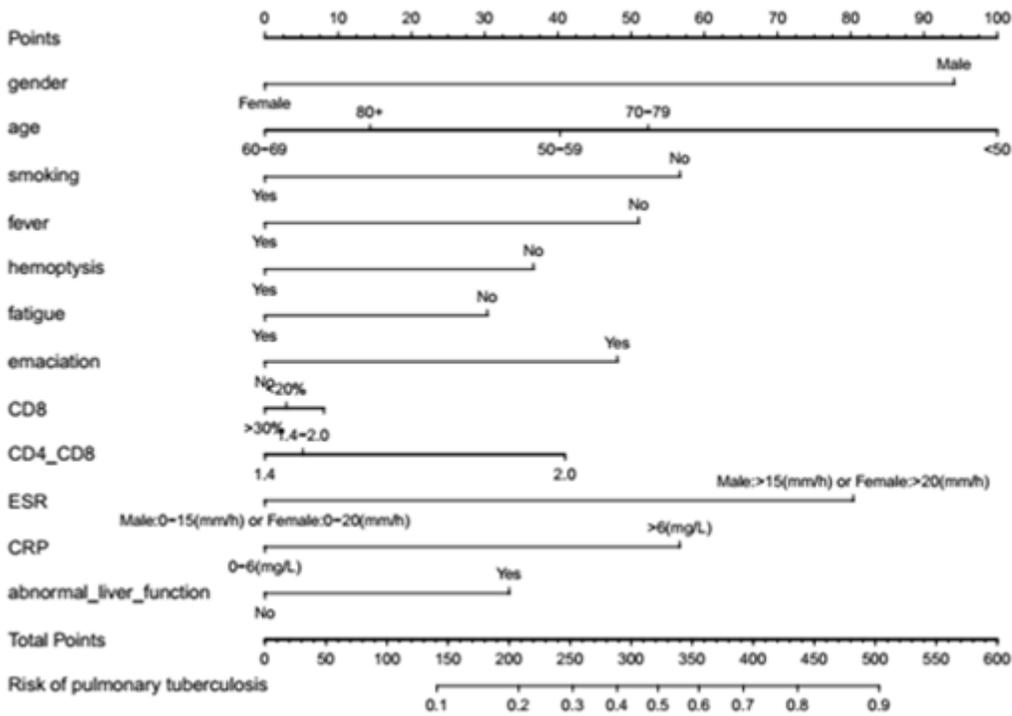


Figure 2

Established TB risk prediction nomogram. Note: The TB risk prediction nomogram was established in the cohort, including gender, age, smoking, fever, hemoptysis, fatigue, emaciation, CD8, CD4/CD8, ESR, CRP, abnormal liver function

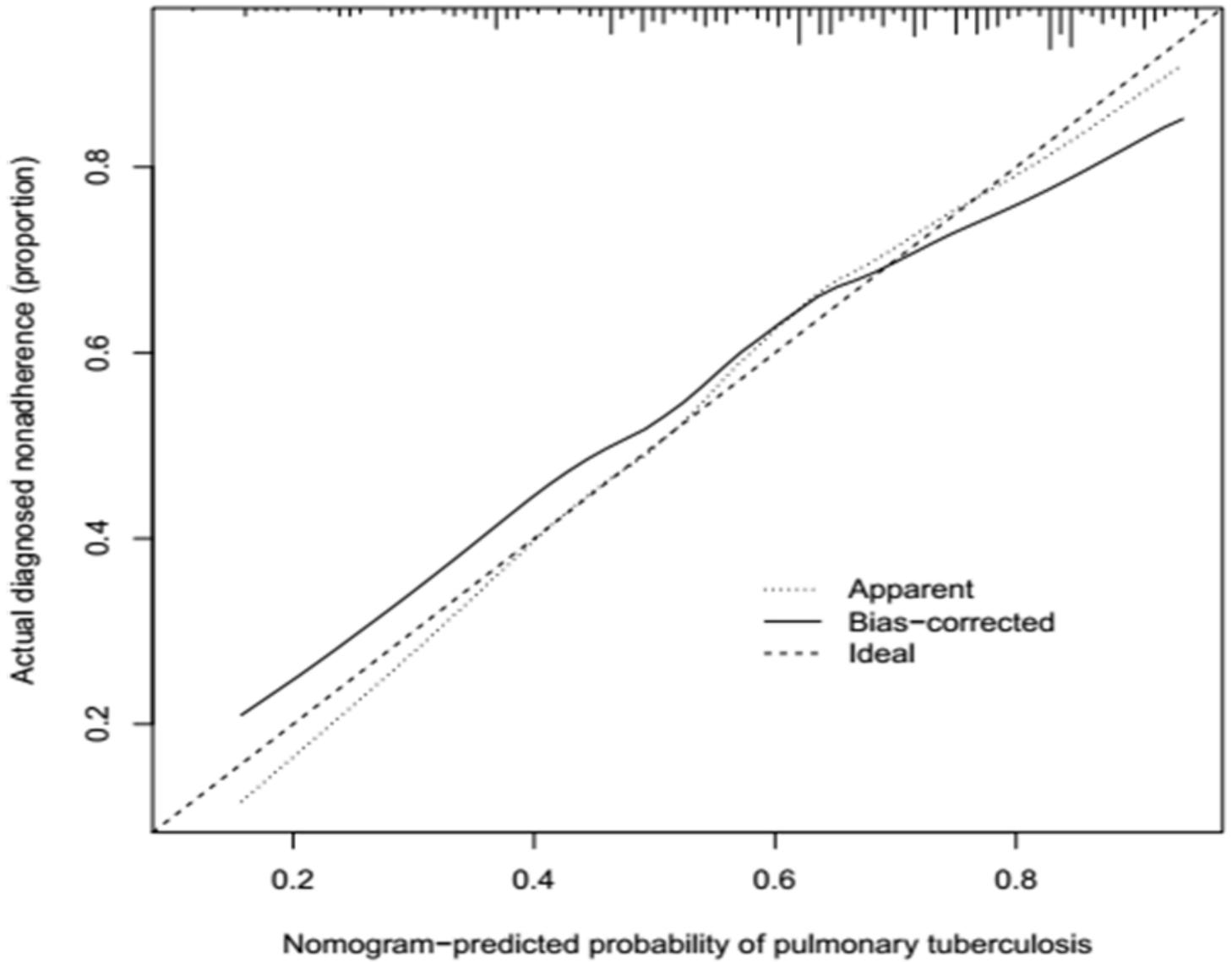


Figure 3

calibration curves of the TB probability nomogram prediction in the cohort. Notes: The x-axis refers to the predicted TB risk. The y-axis refers to the actual diagnosed TB. The diagonal dotted line represents a perfect prediction by an ideal model. The solid line represents the performance of the nomogram, of which a closer fit to the diagonal dotted line represents a better prediction.

AUC= 0.729

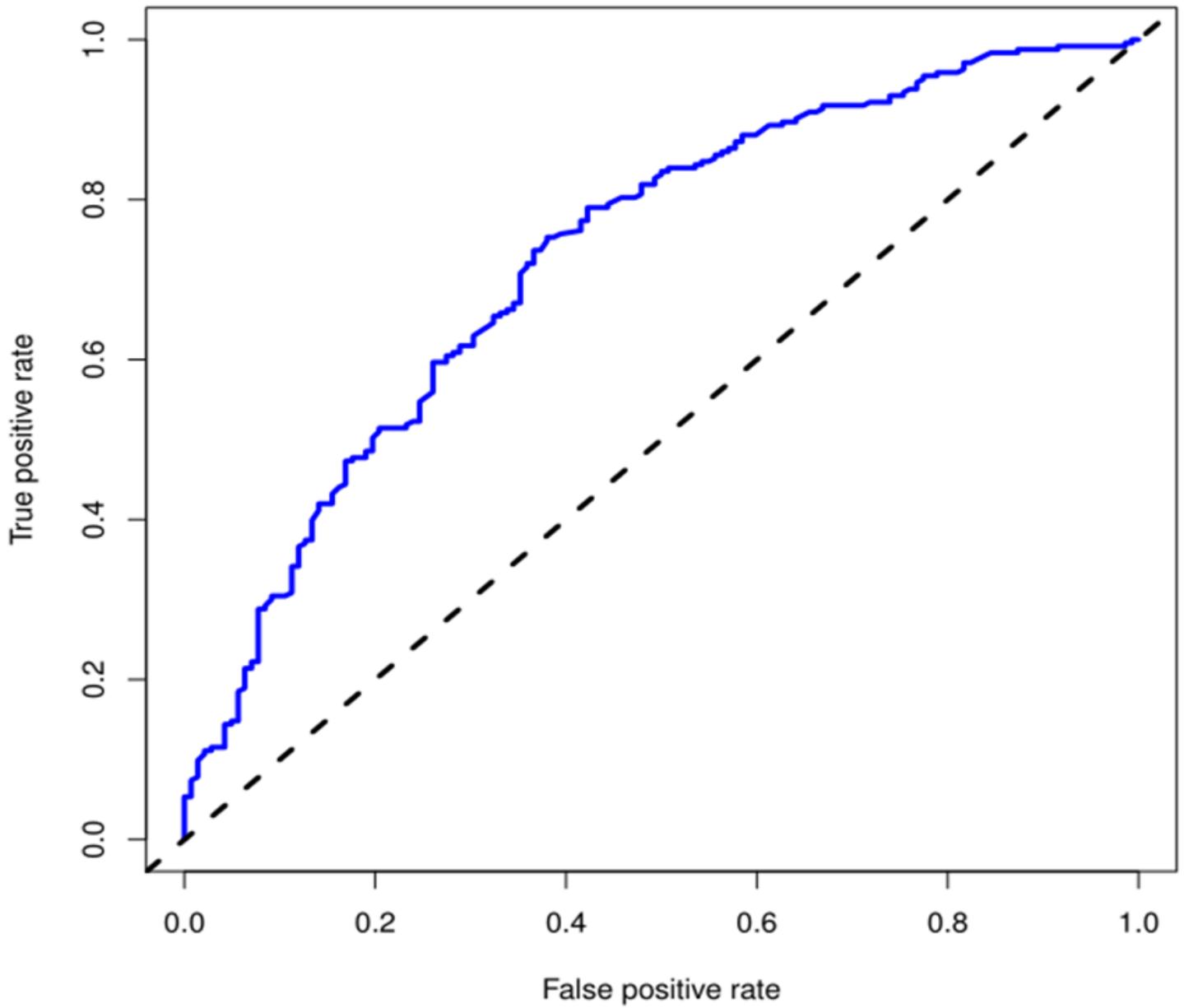


Figure 4

The AUC (representative the discrimination of the model) of the model. Note: The dotted vertical lines indicate the 95% confidence interval. AUC, area under the curve.

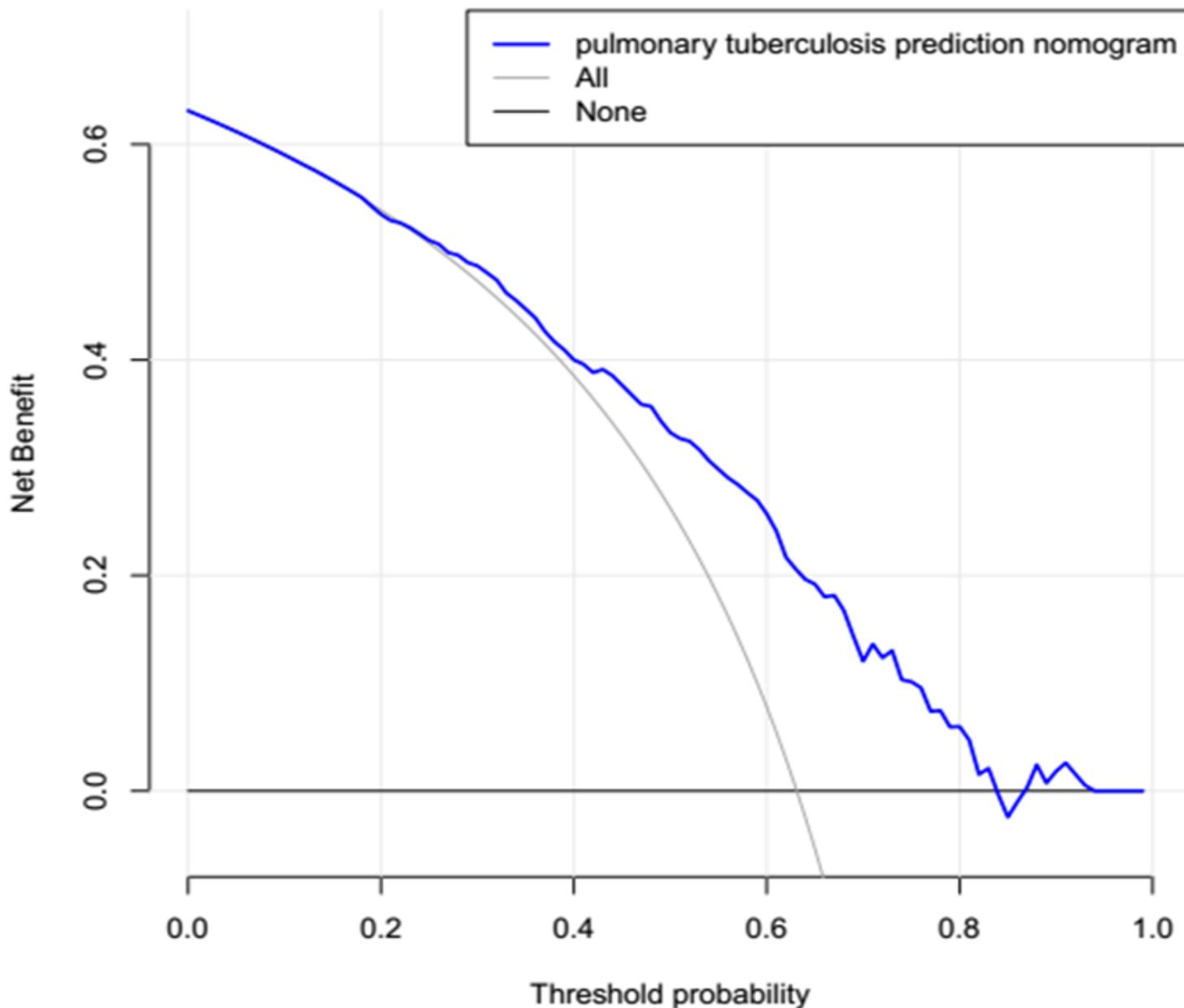


Figure 5

Decision curve analysis for the TB nomogram. Notes: The y-axis measures the net benefit. The dotted line represents the TB risk nomogram. The solid line represents the assumption that all patients are TB infection. Thin solid line represents the assumption that no infection. The decision curve showed that if the threshold probability of a patient and a doctor is 13 and 83%, respectively, using this TB nomogram in the current study to predict TB risk adds more benefit than the intervention-all-patients scheme or the intervention-none scheme.