Journal of Inflammation Research Dovepress

Open Access Full Text Article ORIGINAL RESEARCH

Development of a Nomogram Based on Transcriptional Signatures, IFN-γ Response and Neutrophils for Diagnosis of Tuberculosis

Yan-Hua Liu[1,*], Jin-Wen Su[2,*], Jing Jiang[3], Bing-Fen Yang[1], Zhi-Hong Cao[1], Fei Zhai[1], Wen-Na Sun[1], Ling-Xia Zhang[1], Xiao-Xing Cheng[1]

[1]Beijing Key Laboratory of New Techniques of Tuberculosis Diagnosis and Treatment, Institute of Tuberculosis Research, Senior Department of Tuberculosis, the Eighth Medical Center of PLA General Hospital, Beijing, 100091, People's Republic of China; [2]Division of Critical Care Medicine, Senior Department of Tuberculosis, the Eighth Medical Center of PLA General Hospital, Beijing, 100091, People's Republic of China; [3]Institute of Research, Beijing Key Laboratory of Organ Transplantation and Immune Regulation, Senior Department of Respiratory and Critical Care Medicine, the Eighth Medical Center of PLA General Hospital, Beijing, 100091, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xiao-Xing Cheng; Ling-Xia Zhang, Beijing Key Laboratory of New Techniques of Tuberculosis Diagnosis and Treatment, Institute of Tuberculosis Research, Senior Department of Tuberculosis, Eighth Medical Center of PLA General Hospital, 17 hei Shan Hu Road, Haidian, Beijing, 100091, People's Republic of China, Fax +86 10 51520496; +86 10 66775969, Email xcheng79@outlook.com; zhanglingxia7@163.com

Purpose: Tuberculosis (TB) is a major global health threat and its diagnosis remains challenging. This study aimed to develop a nomogram that incorporated peripheral blood transcriptional signatures and other blood tests for the diagnosis of tuberculosis.

Patients and Methods: Patients with TB, patients with other definite pulmonary diseases (OPD), individuals with latent tuberculosis infection (LTBI), and healthy controls (HC) were retrospectively enrolled between May 2017 and April 2018. The results of the interferon-γ release assay (IGRA) and blood counts were obtained from medical records, and the transcripts of 10 genes were detected using reverse transcription polymerase chain reaction (RT-PCR). Variable selection was performed using least absolute shrinkage and selection operator regression (LASSO) and multivariate logistic regression was performed for the optimal prediction model with backward direction. The model was displayed as a nomogram, and its performance was evaluated for discrimination ability, calibration ability, and clinical usefulness. Internal validation of the prediction model was conducted using bootstrap resampling.

Results: A total of 185 participants were enrolled, including 84 patients with TB and 101 controls. A prediction nomogram composed of IGRA, percentage of neutrophils, and expression levels of CD64, granzyme A (GZMA), and PR/SET domain 1 (PRDM1) was established. The nomogram demonstrated good discrimination, with an unadjusted area under the curve (AUC) of 0.914 (95% CI: 0.875–0.954) and a bootstrap-corrected AUC of 0.914 (95% CI: 0.874–0.947). With a cutoff value of 0.519, the sensitivity and specificity for discriminating PTB from controls were 0.81 and 0.871, respectively. The nomogram also showed good calibration with the Hosmer–Lemeshow test (P=0.58) and good clinical practicality displayed by the decision curve analysis.

Conclusion: A nomogram composed of IGRA, percentage of neutrophils, and expression of CD64, GZMA, and PRDM1 was established. The nomogram demonstrated a sensitivity and specificity of 81% and 87%, respectively, for differentiating TB from controls.

Keywords: tuberculosis, prediction model, peripheral blood, interferon gamma release assay, transcripts of gene

Introduction

Tuberculosis (TB), which is caused by *Mycobacterium tuberculosis* (*Mtb*), is a major global health concern. An estimated one-quarter of the world's population is infected with *Mtb*.[1] The vast majority of infected individuals remain clinically asymptomatic, known as latent TB infection (LTBI) in their lifetime. Although only 5–10% of infected individuals progress to active TB due to failed host immune control, approximately 10 million people are diagnosed with TB and

Received: 28 May 2024
Accepted: 31 October 2024
Published: 13 November 2024

1.2 million deaths are caused by TB annually.[2] Early diagnosis with high accuracy and sensitivity is vital for TB transmission control.

TB diagnosis remains challenging and largely depends on "bacteriologically confirmed" diagnostic tests, mainly based on culture, which are time-consuming and have low sensitivity. Xpert MTB/RIF assay is a sensitive method for rapid diagnosis of TB and rifampicin resistance. However, the diagnostic performance of Xpert MTB/RIF is reduced in patients with bacteria-negative TB, young children, people living with HIV, and individuals with extrapulmonary TB.[3] Therefore, rapid and nonpathogen-based tests are required to improve TB diagnosis.

Immunological tests, including the tuberculin skin test (TST) and interferon-γ (IFN-γ) release assay (IGRA), are based on the human immune response to TB antigens and are widely used to detect *Mtb* infections with high sensitivity and specificity.[4] However, they are not recommended for TB diagnosis because of their poor ability to discriminate between TB and LTBI. Monocytes and neutrophils in peripheral blood are abnormal in patients with TB and might be potential biomarkers for the diagnosis of TB.[5] Transcriptional signatures selected using high-throughput techniques can be used for the diagnosis of TB.[6–8] Some genes such as cluster of differentiation 64 (CD64), Fc fragment of IgG receptor 1b (FCGR1B), guanylate binding protein 1 (GBP1), and granzyme A (GZMA), which are more easily detectable than proteins, have been identified as potential biomarkers to differentiate TB from LTBI and benefit the diagnosis of TB.[9–11] The diagnostic efficacy may be improved when the transcriptional signature is combined with IGRA and the percentage of blood cell subsets. However, few studies have investigated the performance of a combination of transcriptional signatures, IGRA, and percentages of cell subsets in the diagnosis of TB.

Nomogram is a reliable tool for creating a simple intuitive graph for a statistical predictive model that quantifies the risk of a clinical event. In the present study, we developed a nomogram by integrating blood-based tests, and evaluated its potential for TB diagnosis.

Materials and Methods

Subjects

All participants were recruited from the Department of Tuberculosis, Eighth Medical Center of PLA General Hospital, between May 2, 2017, and April 31, 2018. The participants from this time frame were selected because we had available RNA samples from which cDNA was synthesized as part of a study specifically designed to investigate the expression of TB-specific genes. Patients with pulmonary tuberculosis (PTB) were considered confirmed cases if the results of *Mtb* culture and/or Xpert MTB/RIF (Cepheid, Carlsbad, California, USA) were positive and were enrolled in the patient group. The control group included patients with other pulmonary diseases (OPD), individuals with latent tuberculosis infection (LTBI), and healthy controls (HC). LTBI and HC, who were free of any clinical signs of TB, had normal radiography findings, and no TB history, were discriminated based on the results of IGRA. All participants were HIV-negative (Figure 1).

The study was conducted in accordance with the ethical standards of the Declaration of Helsinki. It was approved by the Ethics Committee of the Eighth Medical Center of PLA General Hospital (Approval No.309201702211512), and informed consent was obtained from all participants.

Laboratory Tests, Radiological and Pathological Examination

Laboratory tests, including sputum smear staining, culture, Xpert MTB/RIF, blood counts, and radiological and pathological examinations, were obtained from the electronic medical records of the participants.

IGRA

The T cell detection kit for TB infection (in vitro release ELISA) was purchased from Beijing Wantai (Beijing Wantai Biopharmaceutical Co. Ltd., Beijing, China). The test was performed strictly according to instructions of the kit. The detection range was from 2 to 400 pg/mL; results which were ≥14 pg/mL and ≥N/4 (N is background controlled detection value) are determined as positive.

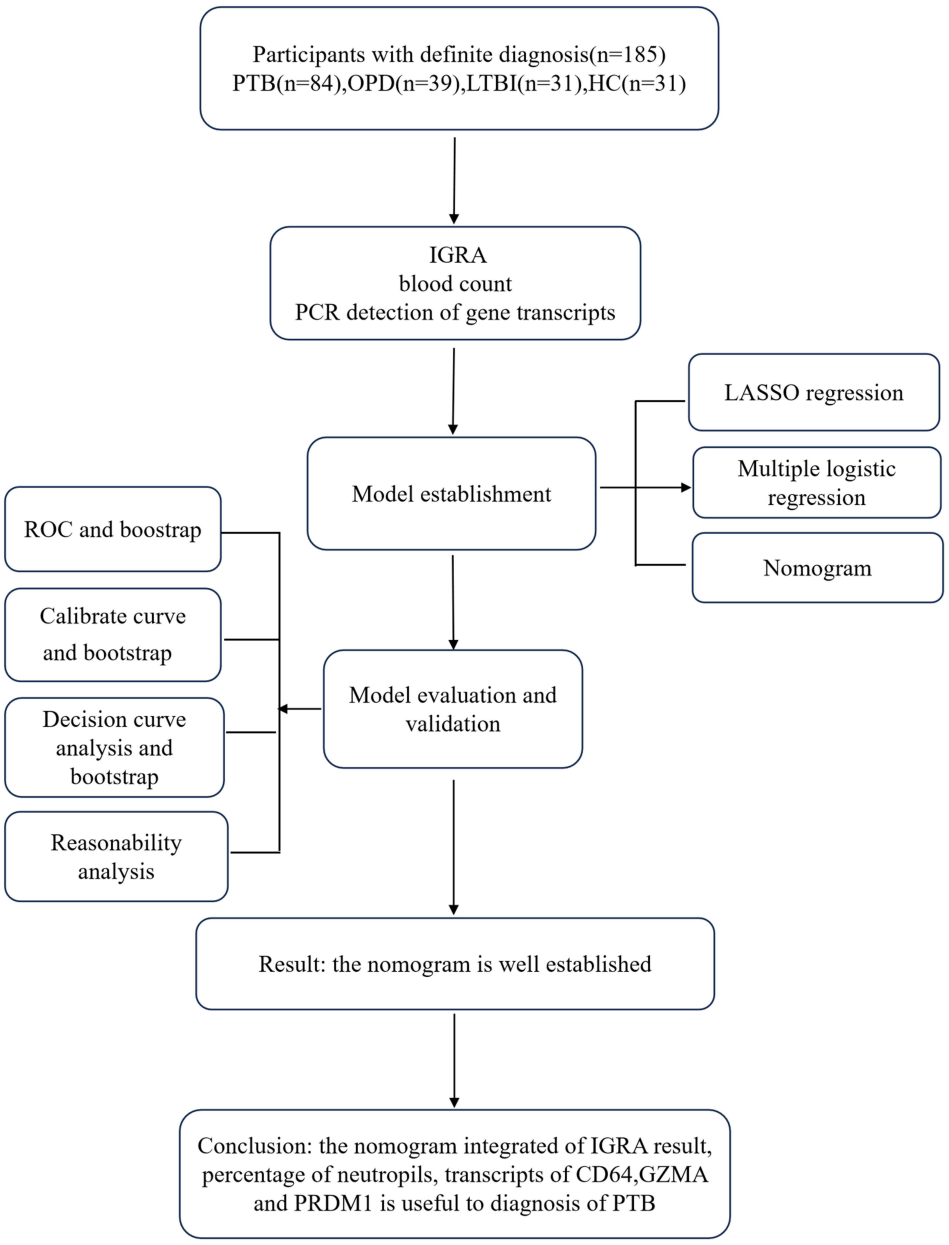

Figure 1 Enrollment of participants and statistical analysis. All patients with PTB were definite TB. OPD included patients with lung cancer, pneumonia, acute exacerbations of chronic obstructive pulmonary disease, and non-tuberculosis mycobacterium.
Abbreviations: PTB, pulmonary tuberculosis; TB, tuberculosis; OPD, other pulmonary disease.

RT-PCR Detection

Peripheral blood mononuclear cells (PBMCs) were isolated from peripheral whole blood by density gradient centrifugation using Ficoll-Paque™ Plus (Cytiva, Uppsala, Sweden), according to the manufacturer's protocol. PBMCs were lysed with TRIzol® Reagent (Invitrogen, Carlsbad, California, USA) for RNA extraction. RNAs were reverse-transcribed into cDNA using a PrimeScript™ RT-PCR Kit (Takara Bio Inc., Kusatsu, Shiga, Japan) with oligo-dT primers according to the manufacturer's protocol. Real-time PCR was performed using a KAPA SYBR® FAST qPCR Kit (KAPA Biosystems, Wilmington, MA, USA) on a LightCycler 480 II (Roche Diagnostics, Basel, Switzerland). PCR was performed under the following conditions: thermal denaturation at 95°C for 3 min, followed by 40 cycles of amplification and quantification (95°C for 5s and 60°C for 30s). Melting curve analysis was performed to verify the primer specificity. Glyceraldehyde-

3-phosphate dehydrogenase (GAPDH) was used as the internal control. Relative gene expression levels were calculated using the cycle threshold of GAPDH minus the cycle threshold of the target gene (− ΔCt=Ct(GAPDH) − Ct (target gene)). Primers were designed using the Primer-BLAST tool, available at the National Center for Biotechnology Information (NCBI). The primer sequences used were listed in Table 1.

Statistical Analysis

All statistical analyses were performed using the R statistical software (version 4.3.2, https://www.r-project.org/). The data in the characteristic baseline were analyzed using the "compareGroups" package. The least absolute shrinkage and selection operator (LASSO) was used to select variables through the "glmnet" package. Multivariate logistic regression was performed for optimal prediction model by the "glm" function with backward direction. Two kinds of nomogram construction were conducted by the "rms" and "regplot" package. Receiver operator characteristic curve (ROC) and area under the curve (AUC) calculations were performed using the "pROC" package. The calibration curves were plotted using the "rms" package. The Hosmer-Lemeshow test was performed with the "HLtest.r" code. The decision curve (DCA) was conducted with the "rmda" package. The internal validation of the prediction model was performed using the bootstrap method. All statistical tests were 2- tailed, and statistical significance was set at $P < 0.05$ (Figure 1).

Results

Characteristics of the Participants

A total of 185 participants, including 84 patients with confirmed PTB and 101 controls, were enrolled in this study (Table 2). Among the 39 patients with OPD, 12 had lung cancer, 24 had pneumonia, 2 had chronic obstructive pulmonary disease, and 1 had non-tuberculosis mycobacterial infection. 31 individuals with LTBI and 31 HC were also enrolled as controls. The characteristics and results of the IGRA, blood counts, and relative mRNA expression levels of the 10 genes were summarized in Table 2. The constituent ratio of sex in the two groups was well balanced (67.3% and 63.1% in subjects without TB and subjects with TB, respectively; $P = 0.655$). There was also no significant difference in age between the two groups (51% and 49% in subjects without TB and subjects with TB, respectively; $P = 0.384$).

Variables Selected by LASSO Regression Analysis

Sixteen parameters were used as initial independent variables: IGRA results (IGRA), percentage of monocytes (monocytes), percentage of neutrophils (neutrophils), ratio of neutrophils to lymphocytes (NLR), ratio of monocytes to lymphocytes (MLR), and mRNA expression levels of genes encoding ATP-binding cassette transporter 2 (ABCA2), B-cell lymphoma 6 (BCL6), cluster of differentiation 64 (CD64), Fc fragment of IgG receptor 1b (FCGR1B), GATA-

Table 1 Sequences of Primers Used in the Quantitative Real-Time PCR

Genes	Forward primers (5' to 3')	Reverse primers (5' to 3')
ABCA2	CTCGTGCACCTCATGACCAG	GAGTCACGTTGCCCACAAAA
BCL6	TGGAGCATGTTGTGGACACT	TTGTTCTCCACCACCTCACG
CD64	CGACCCCCAGCTACAGAATC	TTCCACGCATGACACCTCAA
FCGR1B	TCTCAATGGCACAGCCACTC	ACCAGCTTATCCTTCCACGC
GATA2	CCTGGCGCACAACTACATGG	ACATCTGGCCTCCGGTCA
GBP1	CGACTGATGGCGAATCCAGA	GCAGAACTAGGATGTGGCCT
GZMA	CAGTTGTCGTTTCTCTCCTGC	TGCAGTCAACACCCAGTCTTT
HP	GCATGTCTAAGTACCAAGA	GTCACCTTCACATACACA
LTF	TGCAGGCAAATGTGGTTTGG	AGGATCAGGGTCACTGCTTTG
PRDM1	TTAAGCCCATCCCTGCCAAC	GCTCGGTTGCTTTAGACTGC
GAPDH	TGTTGCCATCAATGACCCCT	TCGCCCCACTTGATTTTGGA

Abbreviations: ABCA2, ATP-binding cassette transporter 2; BCL6, B-cell lymphoma 6; CD64, cluster of differentiation 64; FCGR1B, Fc fragment of IgG receptor 1b; GATA2, GATA-binding protein 2; GBP1, guanylate-binding protein 1; GZMA, granzyme A; HP, haptoglobin; LTF, lactotransferrin; PRDM1, PR/SET domain 1; GAPDH, Glyceraldehyde-3-phosphate dehydrogenase.

Table 2 Baseline Characteristics of the Enrolled Participants

	All subjects	Subjects without TB	Subjects with TB	P value
	n=185	n=101	n=84	
Definition, n(%)				<0.001
TB	84 (45.4)	0 (0.00)	84 (100)	
OPD	39 (21.1)	39 (38.6)	0 (0.00)	
LTBI	31 (16.8)	31 (30.7)	0 (0.00)	
HC	31 (16.8)	31 (30.7)	0 (0.00)	
Sex, n(%)				0.655
Male	121 (65.4)	68 (67.3)	53 (63.1)	
Female	64 (34.6)	33 (32.7)	31 (36.9)	
Age, year				
Mean(IQR)	50.0 [41.0;60.0]	51.0 [44.0;60.0]	49.0 [38.0;62.2]	0.384
IGRA, n(%)				<0.001
Negative	82 (44.3)	64 (63.4)	18 (21.4)	
Positive	103 (55.7)	37 (36.6)	66 (78.6)	
Neutrophils,%				
Mean(IQR)	62.0 [54.6;67.8]	62.5 [55.0;67.2]	61.5 [53.8;67.9]	0.669
Lymphocytes,%				
Mean(IQR)	27.7 [21.8;33.3]	28.0 [24.0;33.3]	27.3 [19.2;33.4]	0.19
Monocytes,%				
Mean(IQR)	7.36 [5.89;8.90]	6.65 [5.16;7.95]	8.20 [6.96;9.50]	<0.001
NLR				
Mean(IQR)	2.27 [1.66;3.03]	2.25 [1.69;2.79]	2.29 [1.66;3.57]	0.413
MLR				
Mean(IQR)	0.25 [0.19;0.37]	0.23 [0.18;0.31]	0.32 [0.21;0.46]	<0.001
ABCA2,-ΔCT				
Mean(IQR)	−6.03 [−6.64;-5.34]	−5.52 [−6.28;-5.04]	−6.36 [−6.79;-5.93]	<0.001
BCL6,-ΔCT				
Mean(IQR)	−3.39 [−3.70;-2.88]	−3.48 [−3.79;-3.04]	−3.13 [−3.58;-2.80]	0.013
CD64,-ΔCT				
Mean(IQR)	−3.84 [−4.78;-3.06]	−4.46 [−5.02;-3.64]	−3.20 [−3.90;-2.56]	<0.001
FCGR1B,-ΔCT				
Mean(IQR)	−5.62 [−6.49;-4.55]	−6.18 [−6.89;-5.14]	−4.76 [−5.81;-4.18]	<0.001
GATA2,-ΔCT				
Mean(IQR)	−7.67 [−8.40;-6.90]	−7.24 [−8.00;-6.66]	−7.94 [−8.53;-7.40]	<0.001
GBP1,-ΔCT				
Mean(IQR)	−3.96 [−4.62;-3.17]	−4.28 [−4.78;-3.62]	−3.43 [−4.16;-2.74]	<0.001
GZMA,-ΔCT				
Mean(IQR)	−3.07 [−3.80;-2.40]	−2.46 [−3.25;-1.94]	−3.50 [−4.25;-3.08]	<0.001
HP,-ΔCT				
Mean(IQR)	−6.92 [−7.71;-6.11]	−6.82 [−7.71;-6.02]	−6.95 [−7.67;-6.16]	0.66
LTF,-ΔCT				
Mean(IQR)	−4.62 [−5.90;-3.36]	−4.58 [−6.08;-3.52]	−4.65 [−5.84;-3.35]	0.583
PRDM1,-ΔCT				
Mean(IQR)	−3.92 [−4.52;-3.22]	−3.36 [−4.08;-3.00]	−4.32 [−4.85;-3.85]	<0.001

Notes: P values were obtained from univariate analysis between subjects with TB and those without TB. Categorical variables were compared using Pearson's $\chi 2$ test. Continuous variables were compared using the Mann–Whitney *U*-test.
Abbreviations: TB, tuberculosis; OPD, other pulmonary diseases; LTBI, latent tuberculosis infection; HC, healthy controls; IGRA, IFN-γ release assay; NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; IQR, interquartile range; ABCA2, ATP-binding cassette transporter 2; BCL6, B-cell lymphoma 6; CD64, cluster of differentiation 64; FCGR1B, Fc fragment of IgG receptor 1b; GATA2, GATA binding protein 2; GBP1, guanylate binding protein 1; GZMA, granzyme A; HP, haptoglobin; LTF, lactotransferrin; PRDM1, PR/SET domain 1; -ΔCT=CT (housekeeping gene)-CT (target gene).

binding protein 2 (GATA2), guanylate-binding protein 1 (GBP1), granzyme A (GZMA), haptoglobin (HP), lactotransferrin (LTF), and PR/SET domain 1 (PRDM1). LASSO regression analysis, a shrinkage and variable selection method for logistic regression models, was used to select a subset of the predictors. LASSO regression analysis can impose a constraint on the model parameters that causes regression coefficients for some variables to shrink toward zero. Variables with zero regression coefficients were excluded from the model, whereas variables with non-zero regression coefficients were most strongly associated with the response variable. The distribution of the coefficients for the 16 features was shown in Figure 2A, where the coefficient profile was plotted against the log (λ) sequence. Ten times

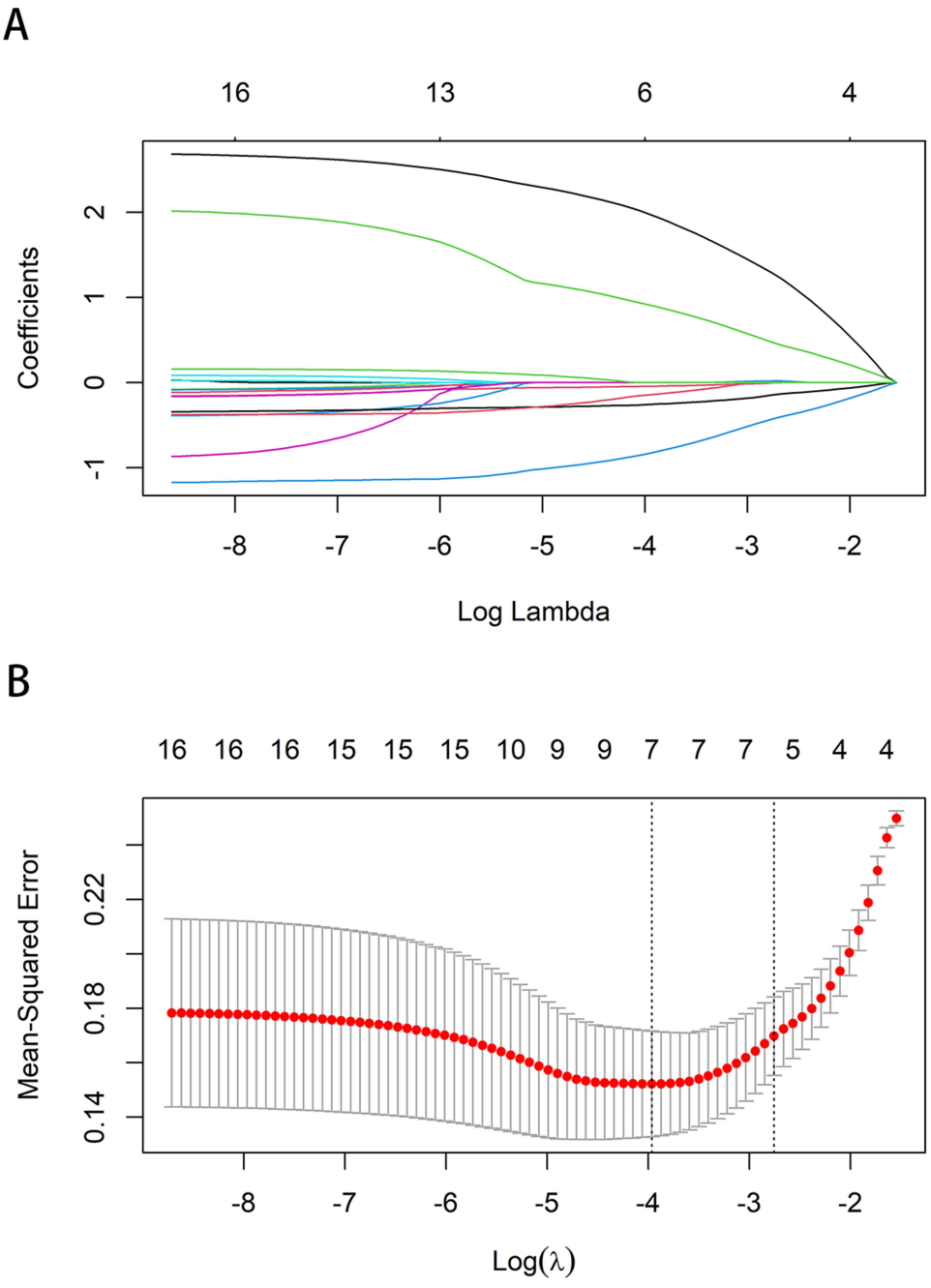

Figure 2 Variable selection using LASSO. (**A**) Each colored curve represents the LASSO coefficient profile of a variable against the log (lambda) sequence. (**B**) Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria(Lambda.1se), which gave a model with good performance and the least number of independent variables.
Abbreviation: LASSO, least absolute shrinkage and selection operator.

K cross-validations were used for the centralization and normalization of the included variables, and the best lambda value was selected. As shown in Figure 2B, dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria (Lambda.1se), which gave a model with good performance and the least number of independent variables. Six predictors were selected using LASSO: IGRA, neutrophils, monocytes, CD64, GZMA, and PRDM1.

Predictors Model Developed by Multivariable Logistic Regression Analysis

After the LASSO analysis, six variables were selected using multivariable logistic regression analysis. Multivariable logistic analysis based on the backward stepwise approach demonstrated that the prediction model containing IGRA, percentage of neutrophils, CD64, GZMA and PRDM1, had miminum Akaike information criterion (AIC) value. The occurrence of TB was positively correlated with IGRA (P<0.001) and expression of CD64 (P<0.001), and negatively correlated with the percentage of neutrophils (P=0.001) and expression of PRDM1 (P=0.001). Although GZMA showed no significant relationship with TB (P=0.139), it was included in the final prediction model (Table 3).

A nomogram that included five TB predictors was established. The score for each predictor was obtained according to the following criteria. The positive and negative IGRA results yielded scores of 27 and 0, respectively. When the percentage of neutrophils was 30%, the score was 59; with every 5 percentage increase, the score decreased by five points. When -ΔCt (CD64) was −7.0, the score was 0; with every 0.5 Ct value increase, the score increased by six or seven points. When -ΔCt (GZMA) was −11.0, the score was 45, and with every 1 Ct value increase, the score decreased by four points. When -ΔCt (PRDM1) was −7.0, the score was 64, and with every 0.5 Ct value increase, the score decreased by 6 points. The total score for each subject was the sum of the five predictor scores (Figure 3).

Performance Evaluation and Validation of the Nomogram

AUC was calculated to evaluate the diagnostic potential of the nomogram. The nomogram had an AUC of 0.914 (95% CI: 0.875–0.954) for differentiating TB from controls. With a cutoff value of 0.519, the sensitivity and specificity were 81% (95% CI: 73–89%) and 87% (95% CI: 81–94%), respectively (Table 4). An AUC of 0.914 (95% CI: 0.874–0.947) with 500 bootstrap repetitions indicated that this nomogram might be a potential biomarker for differentiating TB from controls (Figure 4). The stratified analysis of the nomogram's diagnostic efficacy in TB patients, based on bacteriological status (Table 5). Patients with bacteria-positive TB exhibited superior accuracy, sensitivity, and positive predictive value (PPV), with respective measures of 0.923 (95% CI: 0.882–0.964), 0.861 (95% CI: 0.778–0.946), and 0.812 (95% CI: 0.719–0.904). In contrast, patients with bacteria-negative TB had relatively lower metrics: accuracy at 0.895 (95% CI: 0.84–0.95), sensitivity at 0.632 (95% CI: 0.415–0.848), and a notably reduced PPV of 0.48 (95% CI: 0.284–0.676). Notably, specificity was maintained at a high level of 0.871 (95% CI: 0.806–0.937) across both groups, demonstrating the nomogram's consistency in ruling out TB. The negative predictive value (NPV) was also high in both subsets, suggesting a reliable exclusion of disease.

A calibration curve was used to estimate the consistency between the nomogram-predicted probability of TB and actual outcomes. As shown in Figure 5A, the calibration curve with 500 bootstrap repetitions suggested good agreement

Table 3 Multivariate Logistic Regression Analysis of the Variables Selected in the Nomogram

Variables	β	OR (95% CI)	P value
IGRA	2.535	12.62(4.71–33.82)	<0.001
Percentage of neutrophils	−0.083	0.92(0.87–0.97)	0.001
CD64	1.234	3.43(2.12–5.56)	<0.001
GZMA	−0.38	0.68(0.41–1.13)	0.139
PRDM1	−1.075	0.34(0.18–0.65)	0.001

Abbreviations: IGRA, interferon gamma release assay; CD64, cluster of differentiation 64; GZMA, granzyme A; PRDM1, PR/SET domain 1; OR, Odds ratio; CI, confidence interval.

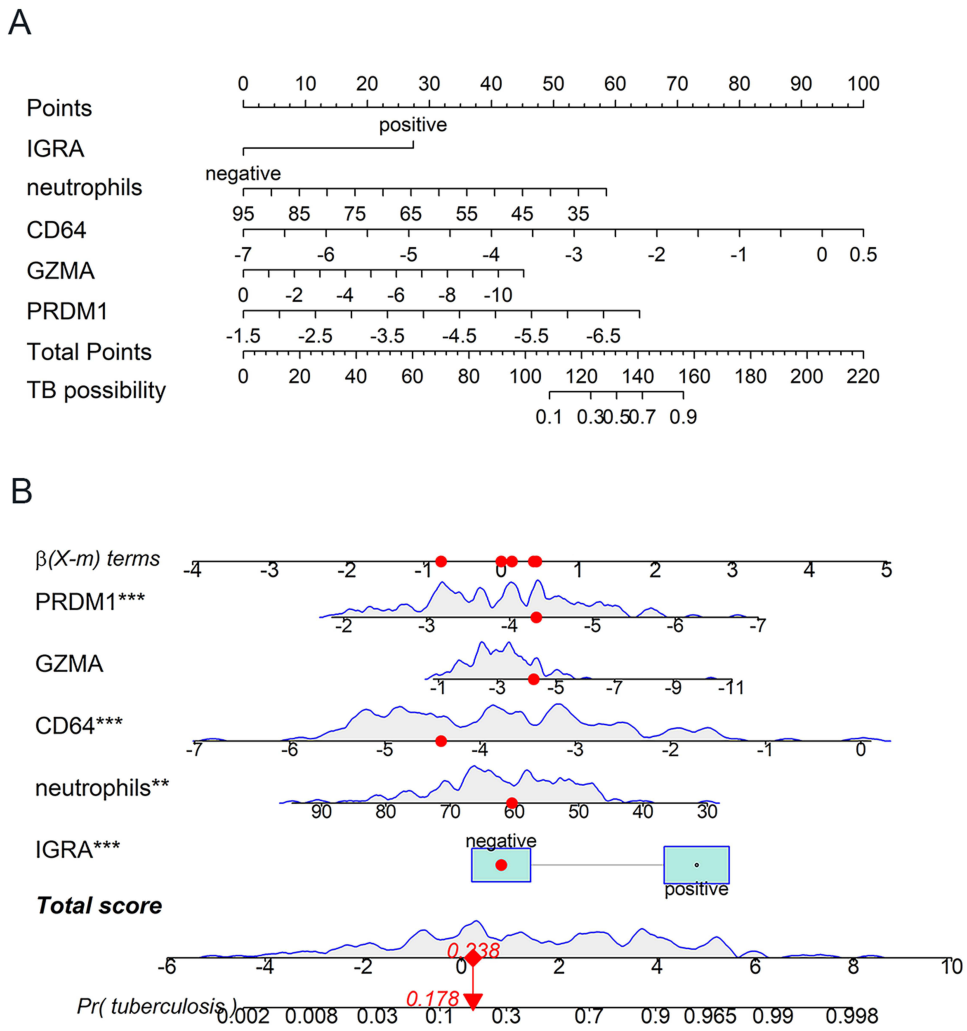

Figure 3 The Nomogram composed of 5 variables to predict the probability of tuberculosis. (**A**) The standard nomogram was presented, utilizing a fixed scoring algorithm to predict the likelihood of tuberculosis based on established clinical and laboratory parameters. "Points" referred to points for the individual risk factor and were added together to the "Total Points". "TB possibility" was calculated according to the "Total Points". (**B**) The dynamic nomogram was displayed, enabling the computation of an individual's risk for tuberculosis through an interactive and personalized scoring approach.This panel illustrated the β (X-m) terms for each predictor, where β represented the regression coefficient, X was the variable value, and m was the mean of the variable. The significance levels were denoted by asterisks (***p<0.001, **p<0.01). The "Total score" (equivalent to "Total Points" in panel (**A**) was calculated by summing the individual predictor scores, which corresponded to the Pr(tuberculosis)(equivalent to "TB possibility" in panel (**A**), ranging from 0.002 to 0.998. The red dots and arrows indicated an example case with a Total score of 0.238, corresponding to a tuberculosis probability of 0.178 (17.8%).

between the model prediction and the actual probabilities. The Hosmer-Lemeshow test yielded a non-significant P-value of 0.581, indicating good calibration power.

DCA, which reflects the clinical usefulness of a nomogram, was used to evaluate its clinical diagnostic performance with 500 bootstrap replicates. As shown in Figure 5B, if the threshold probabilities for patients ranged between 0 and 0.7, using this model to predict TB would be more beneficial than using either predict-none or predict-all. The nomogram might be a promising tool for clinical decision-making.

AUC of each variable was calculated. The AUC of IGRA, percentage of neutrophils, CD64, GZMA and PRDM1 was 0.703 (95% CI: 0.639 to 0.769), 0.518 (95% CI: 0.434 to 0.603), 0.759 (95% CI: 0.690 to 0.828), 0.801 (95% CI: 0.738 to 0.865) and 0.76 (95% CI: 0.691 to 0.829), respectively (Figure 6A). The DCA for each variable was shown in Figure 6B.

The AUC of the other two models was also calculated: one contained IGRA and percentage of neutrophils, with an AUC of 0.702 (95% CI: 0.628–0.777), and the other contained CD64, GZMA, and PRDM1, with an AUC of 0.814 (95%

Table 4 Diagnostic Efficiency of the Three Models for Tuberculosis Diagnosis

Parameter	5 Predictors' Nomogram (95% CI)	IGRA+Neutrophils (95% CI)	3 Transcripts (95% CI)
AUC	0.914(0.875 to 0.954)	0.702(0.628 to 0.777)	0.814(0.753 to 0.875)
Accuracy	0.843(0.842 to 0.845)	0.697(0.695 to 0.700)	0.762(0.760 to 0.764)
Sensitivity	0.810(0.726 to 0.893)	0.774(0.684 to 0.863)	0.821(0.740 to 0.903)
Specificity	0.871(0.806 to 0.937)	0.634(0.540 to 0.728)	0.713(0.625 to 0.801)
PPV	0.840(0.760 to 0.919)	0.637(0.544 to 0.731)	0.704(0.614 to 0.794)
NPV	0.846(0.777 to 0.915)	0.771(0.681 to 0.861)	0.828(0.748 to 0.907)
PLR	6.289(3.747 to 10.557)	2.112(1.594 to 2.799)	2.861(2.071 to 3.952)
NLR	0.219(0.140 to 0.342)	0.357(0.234 to 0.545)	0.250(0.156 to 0.403)

Notes: 5 predictors nomogram, model including IGRA, percentage of neutrophils; CD64, GZMA, and PRDM1; IGRA +neutrophils, model including IGRA and percentage of neutrophils; 3 transcripts, model including CD64, GZMA, and PRDM1.
Abbreviations: AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; PLR, positive likelihood ratio; NLR, negative likelihood ratio; CI, confidence interval.

CI: 0.753–0.875) (Figure 6C). DCA also showed that the threshold range of the two models was smaller than that of the nomogram (Figure 6D).

These results suggested that the nomogram, including the five predictor parameters, showed good performance for TB diagnosis.

Discussion

This study recruited 185 participants to examine the feasibility and efficacy of a nomogram based on transcriptional signatures, IFN-γ responses, and neutrophils for the diagnosis of TB. Following the selection of variables using statistical methods and subsequent modelling, we established a nomogram that included five parameters: IGRA, percentage of neutrophils, and transcripts of CD64, GZMA, and PRDM1. The prediction model discriminated PTB from OPD, LTBI, and HC. The model showed good diagnostic potential with an AUC of 0.914 (95% CI: 0.875–0.954). The sensitivity and specificity were 81% (95% CI: 73–89%) and 87% (95% CI: 81–94%), respectively. The model also demonstrated good calibration power and good clinical decisions based on the calibration plot and DCA. We achieved similar results in discrimination, calibration, and clinical decisions of the model in the internal validation using the "bootstrap" method.

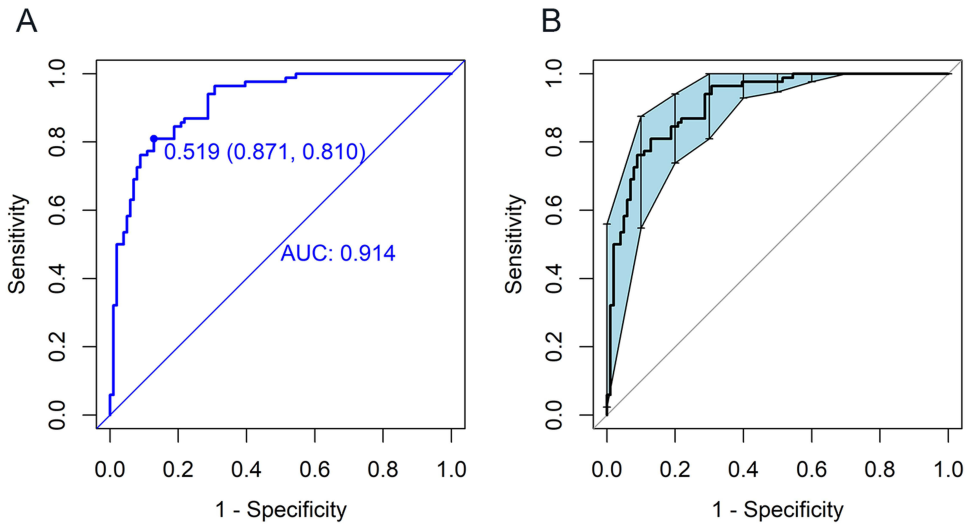

Figure 4 Discrimination of the nomogram for discriminating PTB from non-TB patients. (**A**) An AUC of 0.914 showed good discrimination of the nomogram with a cutoff value of 0.519 (sensitivity and specificity were 81% and 87%, respectively); (**B**) ROC curve plotted with 500 bootstrap replicates.
Abbreviations: PTB, pulmonary tuberculosis; TB, tuberculosis; AUC, area under the curve; ROC, Receiver operator characteristic curve.

Table 5 Nomogram Diagnostic Performance in Bacteria-Positive and Bacteria-Negative Tuberculosis

Parameter	Patients with Bacteria-positive TB (n=65)	Patients with Bacteria-negative TB (n=19)
Accuracy	0.923(0.882–0.964)	0.895(0.84–0.95)
Sensitivity	0.861(0.778–0.946)	0.632(0.415–0.848)
Specificity	0.871(0.806–0.937)	0.871(0.806–0.937)
PPV	0.812(0.719–0.904)	0.48(0.284–0.676)
NPV	0.907(0.849–0.965)	0.926(0.874–0.979)
PLR	6.693(3.993–11.221)	4.907(2.659–9.055)
NLR	0.159(0.086–0.293)	0.423(0.234–0.765)

Notes: Numbers in parentheses are 95% confidence intervals. "bacteria-positive TB" refers to tuberculosis with confirmed bacterial growth or identification. "bacteria-negative TB" refers to tuberculosis without confirmed bacterial growth or identification.

Abbreviations: TB, tuberculosis; PPV, positive predictive value; NPV, negative predictive value; PLR, positive likelihood ratio; NLR, negative likelihood ratio.

These findings indicate that the combination of IGRA, neutrophil percentage, and expression of CD64, GZMA, and PRDM1 is a potential biomarker for TB diagnosis.

IGRA, a test for IFN-γ response to *Mtb* antigens Early Secretory Antigenic Target 6-kDa (ESAT-6), culture filtrate protein 10 (CFP10) or/and TB7.7 (Rv2654c)), has good sensitivity and specificity for the detection of *Mtb* infection. This blood-based test in TB-endemic countries has been limited by their poor ability to discriminate active TB disease from asymptomatic LTBI.[4,12] In this study, the AUC of IGRA was only 0.703 (95% CI: 0.639–0.769) when distinguishing PTB cases from non-TB cases. The combined detection of IGRA and other biomarkers has the potential to improve the specificity and sensitivity of TB diagnosis. The performance of IGRA is improved when combined with tumor necrosis factor-alpha (TNF-α).[13,14]

Some genes showed potential for discriminating TB from non-TB.[15–17] The performance of these genes may be improved when combined with the IGRA. Based on our previous results and those of other studies, we selected ten genes, including FCGR1B, ABCA2, PRDM1, GATA2, CD64, GZMA, BCL6, HP, LTF, and GBP1, for modelling in the present study.[9–11,15,18–21] In our previous studies, FCGR1B, ABCA2, PRDM1, and GATA2 were identified as potential biomarkers for discriminating TB from LTBI by PCR assay.[11,18,19] CD64 has been associated with TB and has shown the potential to assist in the diagnosis of TB in some studies.[6,9,10,21] GZMA, BCL6, HP, LTF and GBP1 were also shown to

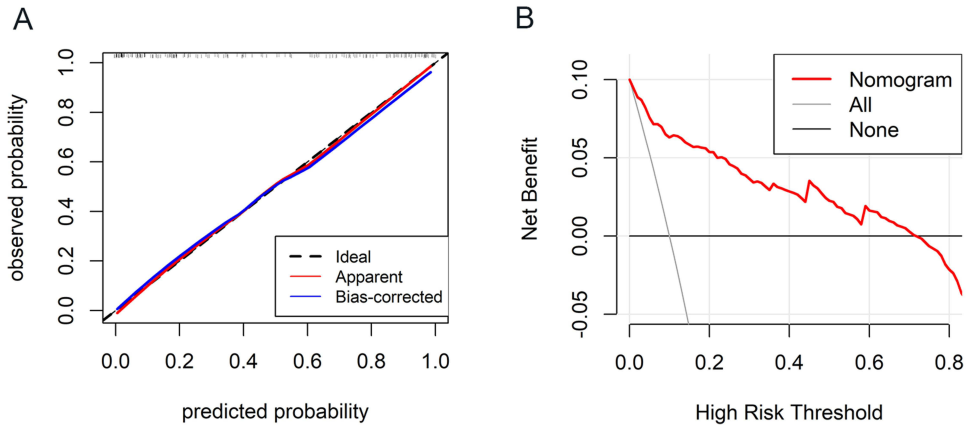

Figure 5 Evaluation of the accuracy and clinical usefulness of the nomogram. (**A**) the calibration curve was presented with 500 repetitions of bootstrap; "Apparent" referred to the calibration curve of the nomogram without correction; "Bias-corrected" referred to the calibration curve of the nomogram with 500 repetitions of bootstrap; (**B**) DCA was used to evaluate the performance in clinical diagnosis of the nomogram with 500 repetitions of bootstrap.
Abbreviation: DCA, the decision curve.

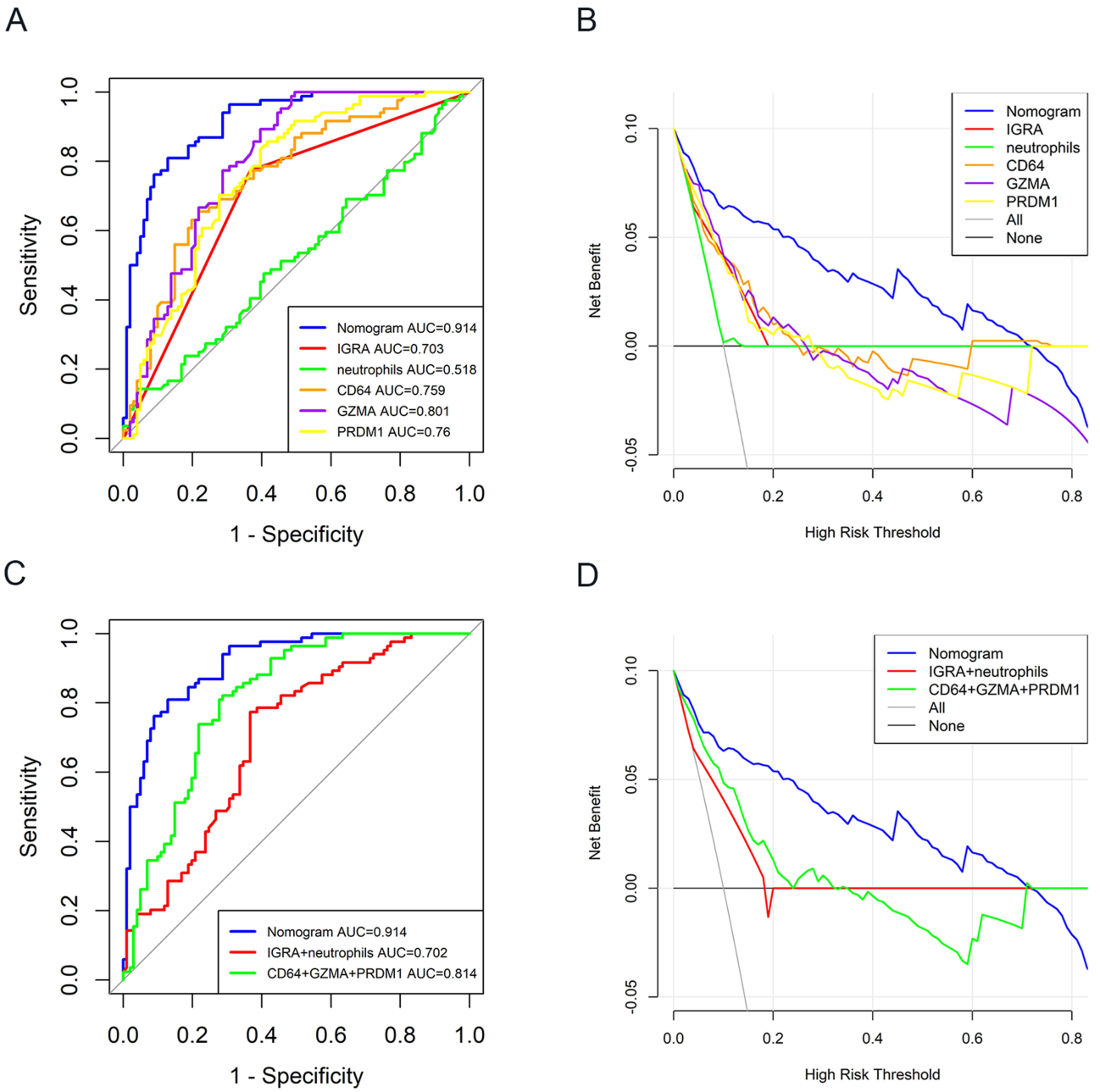

Figure 6 Comparison of the nomogram with individual predictors and the other two models. (**A and C**) ROCs were presented to compare the discriminatory accuracy of the nomogram with individual predictors and other models; (**B** and **D**) DCAs comparing the net benefit of the nomogram versus individual predictors and other models were shown.
Abbreviation: DCA, the decision curve.

be potential biomarkers for TB.[15,21] These genes, except HP and LTF, were significantly differentially expressed in patients with TB compared with those without TB. CD64 and GZMA had AUC of 0.759 (95% CI: 0.690 to 0.828) and 0.801 (95% CI: 0.738 to 0.865), and PRDM1 had an AUC of 0.76 (95% CI: 0.691 to 0.829) in this study. The 3 genes were chosen for the final model, and the discrimination power was slightly improved when combined (AUC, 0.814,95% CI, 0.753–0.875).

Abnormal percentages of peripheral blood cell subsets have been associated with TB. For example, the percentages of monocytes and neutrophils, MLR, and NLR were higher in patients with TB than in HC and might be auxiliary diagnostic biomarkers that are easy to obtain and low-cost. In this study, we selected monocytes, neutrophils,

lymphocytes, NLR, and MLR. The percentage of neutrophils was chosen by LASSO and multivariate logistic regression in the final model, even though the percentage of neutrophils showed poor discrimination with an AUC of 0.518 (95% CI: 0.434–0.603).

When IGRA, CD64, GZMA, PRDM1, and neutrophils were combined, the nomogram showed good prediction ability, with an AUC of 0.914, which was better than that of any individual variable used in the model.

Our study introduces a novel diagnostic nomogram for TB, yet it is accompanied by several limitations. The relatively small sample size, gathered between May 2, 2017, and April 30, 2018, was strategically selected to ensure a detailed molecular analysis; however, it limits the generalizability of our findings. This concentrated recruitment period provided rich data but also highlights the need for broader and more diverse sampling in future studies to enhance the model's predictive accuracy and validate its effectiveness across various populations and settings. Additionally, our analysis, though focused for clarity, examined only a subset of genes within PBMCs, underscoring the necessity for a more comprehensive genomic analysis in subsequent research. An additional limitation is the observed reduced sensitivity of the nomogram in diagnosing bacteria-negative TB, which is a common challenge in the field of TB diagnosis. However, this limitation is counterbalanced by its high specificity. This attribute is particularly beneficial in settings where bacteriological testing is either constrained or inconclusive, effectively ruling out TB in the absence of microbiological confirmation. The nomogram's high NPV thus serves to bolster clinical decision-making, especially when advanced diagnostic tools are inaccessible or provide ambiguous results, underscoring its potential as a valuable diagnostic adjunct in the management of suspected TB cases.

In light of these limitations, future research should aim to expand the cohort size and diversity, thereby enhancing the model's predictive accuracy and generalizability. Particular attention should be given to refining the model's performance in bacteria-negative TB, an area of significant unmet need. By addressing these limitations, we can work towards developing a more robust and universally applicable diagnostic tool for TB.

Conclusion

In conclusion, a nomogram based on the integration of IGRA, the percentage of neutrophils, and three transcripts was developed. This model might be a potential biomarker for differentiating TB from LTBI and other diseases.

Data Sharing Statement

The relevant data are available from the corresponding author Xiao-Xing Cheng (E-mail: xcheng79@outlook.com) upon reasonable request.

Ethics Statement

This study was approved by the Ethics Committee of the Eighth Medical Center of the PLA General Hospital (Approval No.309201702211512), and informed consent was obtained from all participants. This study was conducted in accordance with the principles of the Declaration of Helsinki.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by grants from the Thirteenth Fifth Mega-Scientific Project on the Prevention and Treatment of AIDS, Viral Hepatitis and Other Infectious Diseases (No. 2017ZX10201301-007-002).

Disclosure

The authors declare that they have no conflict of interest.

References

1. Jonas DE, Riley SR, Lee LC, et al. Screening for latent tuberculosis infection in adults: updated evidence report and systematic review for the US preventive services task force. *JAMA*. 2023;329(17):1495–1509. doi:10.1001/jama.2023.3954
2. Bagcchi S. WHO's Global Tuberculosis Report 2022. *Lancet Microbe*. 2023;4(1):e20. doi:10.1016/S2666-5247(22)00359-7
3. Lewinsohn DM, Leonard MK, LoBue PA, et al. Official American thoracic society/infectious diseases society of America/centers for disease control and prevention clinical practice guidelines: diagnosis of tuberculosis in adults and children. *Clin Infect Dis*. 2017;64(2):e1–e33. doi:10.1093/cid/ciw694
4. Hamada Y, Gupta RK, Quartagno M, et al. Predictive performance of interferon-gamma release assays and the tuberculin skin test for incident tuberculosis: an individual participant data meta-analysis. *EClinicalMedicine*. 2023;56:101815. doi:10.1016/j.eclinm.2022.101815
5. Kissling M, Fritschi N, Baumann P, et al. Monocyte, lymphocyte and neutrophil ratios - easy-to-use biomarkers for the diagnosis of pediatric tuberculosis. *Pediatr Infect Dis J*. 2023;42(6):520–527. doi:10.1097/INF.0000000000003901
6. Gliddon HD, Kaforou M, Alikian M, et al. Identification of reduced host transcriptomic signatures for tuberculosis disease and digital PCR-based validation and quantification. *Front Immunol*. 2021;12:637164. doi:10.3389/fimmu.2021.637164
7. Gong Z, Gu Y, Xiong K, et al. Evaluation and validation of novel blood-derived biomarkers for precise and rapid diagnosis of tuberculosis in areas with a high-TB burden. *Front Microbiol*. 2021;12:650567. doi:10.3389/fmicb.2021.650567
8. Gupta RK, Turner CT, Venturini C, et al. Concise whole-blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. *Lancet Respir Med*. 2020;8(4):395–406. doi:10.1016/S2213-2600(19)30282-6
9. da Costa L L, Delcroix M, Dalla Costa ER, et al. Real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases. *Tuberculosis*. 2015;95(4):421–425. doi:10.1016/j.tube.2015.04.008
10. Liu Q, Gao Y, Ou Q, et al. Differential expression of CD64 in patients with *Mycobacterium tuberculosis* infection: a potential biomarker for clinical diagnosis and prognosis. *J Cell Mol Med*. 2020;24(23):13961–13972. doi:10.1111/jcmm.16004
11. Liu YH, Wang R, An HJ, Cheng XX. Diagnostic value of FCGR1B transcription levels in patients with active tuberculosis. *Zhonghua Jie He He Hu Xi Za Zhi*. 2022;45(4):373–378. doi:10.3760/cma.j.cn112147-20211213-00878
12. Kaul S, Nair V, Birla S, et al. Latent tuberculosis infection diagnosis among household contacts in a high tuberculosis burden area: a comparison between transcript signature and interferon gamma release assay. *Microbiol Spectr*. 2022;10(2):e0244521. doi:10.1128/spectrum.02445-21
13. Kim JY, Kang YA, Park JH, et al. Dual-release fluorospot assays for IFN-γ and TNF-α for diagnosing active tuberculosis. *Clin Microbiol Infect*. 2020;26(7):928–934. doi:10.1016/j.cmi.2019.11.003
14. Kim JY, Park JH, Kim MC, et al. Combined IFN-γ and TNF-α release assays to differentiate active tuberculosis from latent tuberculosis infection. *J Infect*. 2018;77(4):314–320. doi:10.1016/j.jinf.2018.04.011
15. Shao M, Wu F, Zhang J, et al. Screening potential biomarkers for distinguishing between latent and active tuberculosis in children using bioinformatic analysis. *Medicine*. 2021;100(5):e23207. doi:10.1097/MD.0000000000023207
16. Singhania A, Wilkinson RJ, Rodrigue M, Haldar P, O'Garra A. The value of transcriptomics in advancing the knowledge of the immune response and diagnosis of tuberculosis. *Nat Immunol*. 2018;19(11):1159–1168. doi:10.1038/s41590-018-0225-9
17. Turner CT, Gupta RK, Tsaliki E, et al. Blood transcriptional biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, observational, diagnostic accuracy study. *Lancet Respir Med*. 2020;8(4):407–419. doi:10.1016/S2213-2600(19)30469-2
18. Liu Y, Wang R, Su J, Yu S, Sun W, Cheng X. The mRNA expression of ABCA2 from individuals at different infection stages of *Mycobacterium tuberculosis* and its potential as a diagnostic biomarker. *Prog Mod Biomed*. 2023;23(2):211–215.
19. Liu Y, Wang R, Yang B, et al. The mRNA levels of PRDM1 and GATA2 were used as biomarkers to distinguish patients with active tuberculosis from those with latent TB infection. *J Pra Med*. 2022;38(13):1595–1598,1603.
20. Natarajan S, Ranganathan M, Hanna LE, Tripathy S. Transcriptional profiling and deriving a seven-gene signature that discriminates active and latent tuberculosis: an integrative bioinformatics approach. *Genes*. 2022;13(4):616. doi:10.3390/genes13040616
21. Sambarey A, Devaprasad A, Mohan A, et al. Unbiased identification of blood-based biomarkers for pulmonary tuberculosis by modeling and mining molecular interaction networks. *EBioMedicine*. 2017;15:112–126. doi:10.1016/j.ebiom.2016.12.009