A nomogram to predict lung cancer in pulmonary lesions for tuberculosis infection patients

Zhi Xia,[1-3] Xueyao Rong,[4,5] Qiong Chen,[4,5] Min Fang,[6,7] Jian Xiao[4,5]

[1]Department of Oncology, Hunan Provincial People's Hospital, Changsha; [2]Key Laboratory of Small Molecule Targeted Drug Research and Creation in Hunan Province, Changsha; [3]Hunan Provincial Clinical Medical Research Center for Hepatobiliary Pancreatic Tumors, Changsha; [4]Department of Geriatric Respiratory and Critical Care Medicine, Xiangya Hospital, Central South University, Changsha; [5]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha; [6]Hunan Provincial Key Laboratory of the Research and Development of Novel Pharmaceutical Preparations, the "Double-First Class" Application Characteristic Discipline of Hunan Province (Pharmaceutical Science), Changsha Medical University; [7]School of Pharmacy, Changsha Medical University, China

Correspondence: Jian Xiao, Department of Geriatric Respiratory and Critical Care Medicine, Xiangya Hospital, Central South University, Changsha, China.
E-mail: xiaojian_291@163.com

Key words: nomogram, lung cancer, tuberculosis infection, T-SPOT.TB.

Contributions: JX, QC, ZX, contributed to the study design; ZX, XR, contributed to analyzing the data and drafting the manuscript; ZX, XR, MF, contributed to the data acquisition; JX, QC, MF, contributed to the review and editing. All authors read and approved the final version of the manuscript.

Conflict of interest: the authors declare that they have no competing interests.

Ethics approval and consent to participate: this study was approved by Xiangya Hospital of Central South University (No. 202012230).

Informed consent: the requirement of informed consent was waived due to the retrospective nature of the study.

Patient consent for publication: not applicable.

Availability of data and materials: the datasets analyzed during the current study are not publicly available due to privacy and ethical restrictions, but are available from the corresponding author upon reasonable request.

Funding: this research was supported by the National Natural Science Foundation of China-Youth Science Fund (No. 82204453) and the Hunan Provincial People's Hospital Young Doctor Fund (No. BSJJ202207). The funding organizations played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Received: 14 November 2023.
Accepted: 1 March 2024.
Early view: 13 March 2024.

Monaldi Archives for Chest Disease 2025; 95:2847
doi: 10.4081/monaldi.2024.2847

Abstract

Similar clinical features make the differential diagnosis difficult, particularly between lung cancer and pulmonary tuberculosis (TB), without pathological evidence for patients with concomitant TB infection. Our study aimed to build a nomogram to predict malignant pulmonary lesions applicable to clinical practice. We retrospectively analyzed clinical characteristics, imaging features, and laboratory indicators of TB infection of patients diagnosed with lung cancer or active pulmonary TB at Xiangya Hospital of Central South University. A total of 158 cases from January 1, 2018, to May 30, 2019, were included in the training cohort. Predictive factors for lung cancer were screened by a multiple-stepwise logistic regression analysis. A nomogram model was established, and the discrimination, stability, and prediction performance of the model were analyzed. A total of 79 cases from June 1, 2019, to December 30, 2019, were used as the validation cohort to verify the predictive value of the model. Eight predictor variables, including age, pleural effusion, mediastinal lymph node, the number of positive tumor markers, the T cell spot test for TB, pulmonary lesion morphology, location, and distribution, were selected to construct the model. The corrected C-statistics and the Brier scores were 0.854 and 0.130 in the training cohort and 0.823 and 0.163 in the validation cohort. Calibration plots showed good performance, and decision curve analysis indicated a high net benefit. In conclusion, the nomogram model provides an effective method to calculate the probability of lung cancer in TB infection patients, and it has excellent discrimination, stability, and prediction performance in detecting a malignant diagnosis of undiagnosed pulmonary lesions.

Introduction

Lung cancer is a common and fatal disease with the highest incidence and mortality [1]. In China, the death rate of lung cancer has nearly quintupled over the past 30 years [2]. However, distinguishing lung cancer from pulmonary tuberculosis (TB) is a common challenge in clinical practice. Despite the different pathogenesis, biological markers, and radiological features of lung cancer and TB in many cases, a portion of atypical lung cancer cases can exhibit the presence of cavitary lesions, a tree-in-bud appearance, and

adjacent fibrocalcific foci, which probably raise suspicion for TB infection and highlight the importance of careful differential diagnosis. However, active TB is often difficult to discern from the concomitant presence of lung cancer, especially in the same lobe [3]. Similar clinical manifestations and imaging features, especially in patients with concomitant TB infection, result in misdiagnosed risk and additional medical costs. Given the above situation, it is necessary to improve the ability of differential diagnosis for this intractable situation.

TB infection mainly includes active TB and latent TB infection (LTBI). Active TB refers to the patient infected with TB, then reproducing in the body and resulting in related symptoms, while the latter means the patient infected with TB does not have any infectiousness, symptoms, etiological, or imaging evidence of active TB. Currently, there is no gold standard for the diagnosis of LTBI. The guideline for the management of LTBI provided by the World Health Organization recommends screening for TB infection using interferon-γ release assay (IGRA) and tuberculin skin test (TST) for the asymptomatic high-risk population [4]. However, active TB and LTBI usually manifest a positive result of the T cell spot test for TB (T-SPOT.TB), which is one of the most widely used IGRAs for diagnosing TB infection in decades [5]. Through incubating peripheral blood mononuclear cells with mycobacterium TB antigens, the assay can assess TB infection by counting the number of spot-forming cells (SFCs). Compared with IGRA-enzyme-linked immunosorbent assay (ELISA), which directly measures the concentration of interferon-γ after being stimulated by mycobacterium TB antigens, T-SPOT.TB had higher sensitivity and specificity for diagnostic active TB. TB tests were 82.9% and 78.6%, and those by IGRA-ELISA were 81.7% and 75.2% [6], suggesting almost perfect agreement between the IGRA-ELISA and the T-SPOT.TB. Compared to the low positive rate of microbiological examination and suboptimal specificity of purified protein derivative (PPD) test, which is one of the most common TSTs, the sensitivity of T-SPOT.TB for diagnosis of TB infection nears 90%, and specificity surpasses 95% [7-10]. In a guideline published by the Centers for Disease Control and Prevention, T-SPOT.TB has been recommended to detect TB infection and is more efficient than PPD in many situations [11]. However, T-SPOT.TB could not discriminate between active TB and LTBI, and more than 90% LTBI population will keep this status lifetime [12].

As a TB-endemic country, the population with TB infection is around 40 % in China, and this epidemiological situation can reduce the diagnostic efficiency of T-SPOT.TB [13]. A study confirmed that LTBI makes T-SPOT.TB unreliable in China, and an increasing proportion of T-SPOT.TB-positive patients have LTBI rather than active TB [14]. The positive result often imposes a great difficulty on the differential diagnosis between lung cancer and TB in pulmonary lesion cases without pathological evidence and typical symptomatology, and even results in missed diagnosis or misdiagnosis, leading to treatment delay and inappropriate medication. The literature has reported that nearly 40% of pulmonary nodules are benign, and atypical TB is the main disease misdiagnosed as lung cancer [15,16]. A single auxiliary diagnosis method is difficult to provide enough information in these difficult cases, and clinicians usually comprehensively take multiple clinical indicators into consideration before decision-making. Previous studies provided diagnostic evidence from blood transcriptional profiles, but it is not a simple and practical approach nowadays [17]. The literature provides a radiomics model to differentiate TB and lung cancer, adopting parameters of lung computerized tomography (CT), but that model suits radiologists rather than clinicians [18]. Otherwise, these methods ignored the coexistence of lung cancer and TB infection. A clinical model to predict lung cancer in undiagnosed pulmonary lesions in TB infection patients is necessary.

Nomograms, simple and effective prediction tools in clinical application, show a good performance in predicting outcomes [19]. In the present study, we constructed a nomogram model to quantify the possibility of lung cancer in pulmonary lesions cases with concomitant TB infection, which could provide a direction for clinical diagnosis.

Materials and Methods

Study patients and data collection

Patients diagnosed as pulmonary TB or lung cancer between January 1, 2018, and December 30, 2019, at the Xiangya Hospital of Central South University were retrospectively collected. The study was conducted in December 2020. According to the literature, we regard a case with a positive T-SPOT.TB as a TB infection case [20,21]. All patients with solitary or multiple pulmonary nodules or masses combined with TB infection were enrolled. The study was approved by the Institutional Ethics Committee of the Xiangya Hospital of Central South University. Participant consent for patients was abandoned due to the retrospective study design, and patients' information follows the data protection and privacy regulations strictly.

The inclusion criteria: i) physicians were unable to determine morphologically whether it was lung cancer or TB during the patient's imaging evaluation on admission; ii) the patient with diagnosed pulmonary TB or pathology-proved diagnosed lung cancer after admitting to hospital; iii) TB infection is diagnosed by TB bacteria founding in sputum or pleural effusion specimens, or effectiveness of diagnostic anti-TB treatment, or with caseous necrosis in pathological reports for focal biopsy specimen [4]; iv) with complete evaluation of TB-related test, lung tumor markers and lung CT scan in first hospitalization.

Cases with the following conditions were excluded: i) patients with diagnosed active pulmonary TB or outer-pulmonary TB before admission; ii) history of non-pulmonary tumors; iii) history of anti-TB treatment prior to the diagnosis of lung cancer; iv) usage of immunosuppressant medications; v) immunosuppression; vi) critical missing clinical data.

Eligible cases between January 1, 2018, and May 30, 2019, were incorporated into the training cohort for development of the nomogram, and cases between June 1, 2019, and December 30, 2019, were entered into the validation cohort. After the model development in the training cohort, the validation cohort is used to test the predictive accuracy of the model in unknown data and thus evaluate its generalization ability.

Demographical and predictor variables

Clinical information and outcomes of lung CT and laboratory tests were collected from electronic medical records. The following data were obtained: i) demographics – age, gender and smoking status; ii) imaging features from the report of lung CT scan – pleural effusion (none, small, moderate, or large according to according to CT feature) [22], lesions' location (unilateral or bilateral lung) and distribution in lung lobes (single or multiple lobes covered by the lesion in single side lung), morphology, status of mediastinal lymph nodes; iii) laboratory indicators – PPD, mycobacterium TB antibody (TBAB) test, T-SPOT.TB, erythrocyte sedimentation rate (ESR), peripheral blood monocyte counts, and seven tumor markers test,

including cancer antigen 125 (CA125), cancer antigen 242 (CA242), carcinoma embryonic antigen (CEA), neuron-specific enolase (NSE), cytokeratin 19 fragment 21-1, cytokeratin-19 and squamous cell carcinoma. The T-SPOT.TB (an interferon (IFN)-γ release assay) is based on detecting secreted IFN-γ in M. TB-specific T-cells were stimulated by mycobacterium-specific antigens: early secreted antigenic target 6 or culture filtrate protein 10, which have been successfully utilized in T-cell effect tests to determine whether M. tuberculosis infection exists. The assay records the number of SFCs, with higher values indicating stronger T-cell responses to these antigens. An induration ≥10 mm is considered a suitable cutoff for a positive PPD test in China [23]. ESR>15 mm/h in males or ESR≥20 mm/h in females is regarded as a positive result.

Development of the nomogram

A nomogram model was constructed using selected risk variables according to the outcome of univariate analysis and multiple stepwise regression. Through transforming regression coefficients of each predictive variable, the nomogram presents an appropriate point scale which can quantify probabilities of outcome. The R package 'rms' was used in the entire process.

Evaluation and validation of the nomogram

The Brier score is known as a popular measure for evaluating the overall prediction accuracy of a binary outcome. It is defined as the mean square error between the observed value of a binary outcome and its predicted probability. In the present study, Brier scores were used to calculate the performance of the nomogram model in the training and validation cohorts, and lower scores indicated higher predictive accuracy. The predictive power was measured by the area under the receiver-operating characteristic curve (AUC), also called the concordance index (namely C-statistics), which indicates the probability that the predicted result will agree with what is actually observed, and bootstrapping validation with 100 resamples was conducted to calculate the corrected value [24]. The calibration curve provided a comparison between the expected and observed conversion probabilities. Decision curve analysis (DCA) is a method for evaluating and comparing prediction models that incorporates clinical consequences, requires only the data set on which the models are tested, and can be applied to models that have either continuous or dichotomous results [25]. DCA was conducted to assess the clinical utility of the nomogram developed in present study, and the DCA plot can show the net benefit of nomogram-based decisions at different threshold probabilities, and three curves on DCA respectively present cases with the model predicting outcome, all cases with the outcome, and no cases with the outcome. The 'rms', 'pROC' and 'dca' packages of R were used in the process.

Statistical analysis

R statistical software (v.3.6.1) was used for statistical analyses and graphical visualization. The null hypotheses were rejected at p-values lower than 0.05. Univariate logistic regression analysis was used to find variables related to the final diagnosis of lung cancer in the training cohort. All the significant variables were included in a stepwise multivariate analysis. Continuous variables are expressed as mean (standard deviation) and compared using an unpaired t-test or Mann-Whitney test.

Categorical variables were compared using the χ^2 test or the Fisher exact test. Odds ratio and correspondence 95% confidence interval (CI) were used to present the strength of the correlations. Sensitivity, a classifier represents the positive correctly classified samples to the total number of positive samples, whereas specificity is expressed as the ratio of the correctly classified negative samples to the total number of negative samples. These two classifiers are used for evaluating the classification performance in the diagnosis of lung cancer or TB infection.

Results

Clinical characteristics of patients

A total of 237 patients with concomitant TB infection had been diagnosed as active pulmonary TB or lung cancer, and these cases were incorporated in our study. Meanwhile, a total of 158 patients were assigned into the training cohort, and the remaining 79 were incorporated into the validation cohort. The clinicopathologic characteristics of cases are listed in Table 1. There is no statistic difference between the baseline clinicopathologic data of training and validation cohorts.

Independent predictive factors for lung cancer in cases with concomitant tuberculosis infection

A total of 67 (42.41%) and 40 (50.63%) patients were diagnosed with lung cancer in the training and validation cohorts, showing a near rate of LTBI in pulmonary lesion cases with concomitant TB infection. Between TB and lung cancer cases in both the training and validation cohorts, we found no significant difference in the positive rate of TB-related indicators, including PPD and TBAB, and laboratory indicators such as ESR and monocyte counts (as shown in Table 2). Although there was a significant difference in the positive rate of tumor markers, the sensitivity and specificity were suboptimal: 68.42% and 72.28% in the training cohort and 71.43% and 65.91% in the validation cohort, respectively.

Almost all indicators of imaging features displayed a significant difference between TB and lung cancer cases in both data sets. To identify the variables predicting lung cancer in cases with TB infection, univariate logistic analysis was used to analyze all variables listed in Table 3. The result reveals ten variables related to lung cancer in cases with TB infection, such as TB-related indicators, including PPD and TBAB, and laboratory indicators, including ESR and monocyte counts.

Then, the multivariate logistic regression analysis shows age, pleural effusion, status of mediastinal lymph nodes, the number of positive lung tumor markers, T-SPOT.TB, lesions' morphology, location, and distribution, which were suitable variables for the construction of the nomogram model (Table 3). Among these variables, pleural effusion, the number of positive lung tumor markers, T-SPOT.TB, lesions' morphology and distribution were independent predictive factors for lung cancer in pulmonary lesions combined with TB infection.

Building and validating a predictive nomogram model

Based on variables screened by multiple stepwise regression, a predictive nomogram is established for the risk assessment of lung cancer in pulmonary lesions combined with TB infection (Figure 1). Each variable is assigned a score according to the clinical characteristics of each individual, and the total score, which can reflect the probability of lung cancer, is computed by summing individual scores. The nomogram showed that the number of positive tumor

markers is a potent predictor for lung cancer, and the risk rises with the number of positive tumor markers increasing. In contrast to tumor markers, the strength of T-SPOT.TB was inversely related to the risk of lung cancer. Otherwise, imaging features of pulmonary lesions are important indicators for differentiating lung cancer from TB. In the training cohort, the C-index of the nomogram was 0.881(95% CI, 0.825-0.938) and 0.854 by bootstrapping analysis, and the Brier score was 0.130, suggesting that the model had good discriminative ability (Figure 2a). The calibration plots of the nomogram showed the agreement between predicted and observed situation was optimal (Figure 2b), and DCA shows that the predictive model can bring significant net benefits to predicting lung cancer in pulmonary lesions combined TB infection, demonstrating the potential application value of the predictive model in clinical practice (Figure 2c). In the validation cohort, the C-index was 0.851(95% CI, 0.768-0.933), and 0.823 by bootstrapping analysis, and the brier score was and 0.163, and the calibration plots and DCA also have a good perform, confirming this predictive nomogram can serve as an excellent diagnostic tool for lung cancer in cases with concomitant TB infection (Figure 3).

Discussion

We developed and validated a predictive nomogram based on clinical features to help distinguish lung cancer in patients with concomitant TB infection. The nomogram including case history, imaging features, and laboratory indicators, which is easily obtained in clinical practice, shows a good discrimination and calibration.

The increasing number of lung cancer patients has made cases comorbid with LTBI more common in recent years, especially

Table 1. Clinical and pathological features of cases.

	Training (n=158) (No. %)	Validation (n=79) (No. %)	p
Diagnosis			
Tuberculosis	91 (57.59)	39 (49.37)	0.2302
Lung cancer	67 (42.41)	40 (50.63)	
Age, mean (SD)	59.80 (12.95)	62.14 (11.95)	0.169
Sex			
Male	115 (72.78)	55 (69.62)	0.61
Female	43 (27.22)	24 (30.38)	
Smoking			
No	63 (39.87)	36 (45.57)	0.4019
Yes	95 (60.13)	43 (54.43)	
Pleural effusion			
None or small volume	140 (88.61)	68 (86.08)	0.6746
Moderate or large volume	18 (11.39)	11 (13.92)	
PPD test			
Negative (<10 mm)	76 (48.1)	32 (40.5)	0.2684
Positive (≥10 mm)	82 (51.9)	47 (59.5)	
TBAB			
Negative	124 (88.57)	70 (95.89)	0.0824
Positive	16 (11.43)	3 (4.11)	
ESR (mm/h), mean (SD)	61.17 (34.1)	64.33 (35.97)	0.529
Monocyte counts			
Normal	114 (72.15)	50 (63.29)	0.1637
Abnormal	44 (27.85)	29 (36.71)	
Number of markers			
0	101 (63.92)	44 (55.69)	0.4557
1	31 (19.62)	20 (25.32)	
>1	26 (16.46)	15 (18.99)	
T-SPOT.TB, mean (SD)	36.22 (16.34)	35.96 (16.74)	0.91
Mediastinal lymph nodes			
Normal	68 (43.04)	31 (39.24)	0.5763
Enlarged	90 (56.96)	48 (60.76)	
Morphology			
Nodule	87 (55.06)	35 (44.30)	0.1182
Mass	71 (44.94)	44 (55.70)	
Location			
Unilateral lung	83 (52.53)	40 (50.63)	0.7827
Bilateral lung	75 (47.47)	39 (49.37)	
Lung lobe			
Single	122 (77.22)	65 (82.28)	0.3678
Multiple	36 (22.78)	14 (17.72)	

SD, standard deviation; PPD, purified protein derivative test; TBAB, tuberculosis antibody; ESR, erythrocyte sedimentation rate; T-SPOT.TB, T cell spot test for tuberculosis.

among older individuals in a high-prevalence setting [2,26]. In this study, we found a significant difference in smoking between lung cancer and TB cases, but it does not have a significant contribution to malignant pulmonary lesions in TB infection patients, suggesting that smoking has little predictive value for this situation. Although smoking is a proven risk factor for lung cancer, the impact of smoking related to infection is probably leading to the increase in the risk of TB. Lung nodules are mostly caused by long-term stimulation of chronic inflammation, and cigarette smoking can provoke inflammation and aggravate the growth of lung cancer. However, smoking cessation contributes to a reduction in the size and number of benign lung nodules, indicating the partly reversible effect of smoking, while TB infection could cause persistent inflammation, and it is regarded as a predisposing risk for lung cancer [3,27]. Gender difference was observed in the training cohorts, but it was not significant in the validation cohorts. Gender difference is more associated with the smoking rate: men generally smoke more than women in the world. It might lead to a high prevalence of lung cancer or TB in men, but no evidence supports that sex is a predictive factor for lung cancer in TB infection.

Studies reported that more than 20% of lung cancer cases had LTBI in Japan and Italy [21,28]. 28.2% of cases with newly diagnosed lung cancer had concomitant LTBI in Taiwan [20]. Similarly, in mainland China, researchers found that the positive rate of T-SPOT.TB was 23.8% in lung cancer patients [29]. This situation makes T-SPOT.TB unreliable in differentiating lung cancer from active TB [14]. A study reported that the performance of using T-SPOT.TB in distinguishing tuberculoma from lung cancer was not satisfactory, but the specificity improved with the positive cutoff value increase, thereby achieving a better efficiency for diagnosing tuberculoma [29]. A later study confirmed the difference in spot number between LTBI and active TB [30]. Taking a cue from these studies, we found a significant difference in the number of SFCs between lung cancer and TB cases, and then took it into model development, and confirmed it as a valuable factor to predict lung cancer in cases with concomitant TB infection.

Table 2. The statistical analysis of variables in training and validation cohorts.

	Training (n=158) (No. %)		p	Validation (n=79) (No. %)		p
Diagnosis	TB (n=91)	LC (n=67)		TB (n=39)	LC (n=40)	
Sex						
Male	59 (64.84)	56 (83.58)	0.015	23 (5897)	32 (80)	0.074
Female	32 (35.16)	11 (16.42)		16 (4103)	8 (20)	
Age, mean (SD)	57.99 (14.94)	62.25 (9.16)	0.094	60.03 (13.92)	64.2 (9.36)	0.124
Smoking						
No	45 (49.45)	18 (26.87)	0.006	25 (64.10)	11 (27.5)	0.002
Yes	46 (50.55)	49 (73.13)		14 (35.89)	29 (72.5)	
Pleural effusion						
None or small volume	75 (82.42)	65 (97.01)	0.009	31 (79.49)	37 (92.5)	0.179
Moderate or large volume	16 (17.58)	2 (2.99)		8 (20.51)	3 (7.5)	
PPD test						
Negative (<10 mm)	44 (48.35)	32 (47.76)	1	16 (41.03)	16 (40)	1
Positive (≥10 mm)	47 (51.65)	35 (52.24)		23 (58.97)	24 (60)	
TBAB						
Negative	68 (87.18)	56 (90.32)	0.754	35 (94.59)	35 (97.22)	1
Positive	10 (12.82)	6 (9.68)		2 (5.41)	1 (2.78)	
ESR						
Negative	8 (9.10)	8 (12.31)	1	2 (5.26)	4 (10.00)	1
Positive	80 (90.90)	57 (87.69)		36 (94.74)	36 (90.00)	
Monocyte counts						
Normal	68 (74.73)	46 (68.66)	0.508	26 (66.67)	24 (60)	0.703
Abnormal	23 (25.27)	21 (31.34)		13 (33.33)	16 (40)	
Tumor marker						
Negative	73 (80.22)	28 (41.79)	0.000	29 (74.36)	15 (37.5)	0.002
Positive	18 (19.78)	39 (58.21)		10 (25.64)	25 (62.5)	
Mediastinal lymph nodes						
Normal	49 (53.85)	19 (28.36)	0.002	18 (46.15)	13 (32.5)	0.312
Enlarged	42 (46.15)	48 (71.64)		21 (53.86)	27 (67.5)	
Morphology						
Nodule	66 (72.53)	21 (31.34)	0.000	27 (69.23)	8 (20)	0.000
Mass	25 (27.47)	46 (68.66)		12 (30.77)	32 (80)	
Location						
Unilateral lung	38 (41.76)	45 (67.16)	0.003	16 (41.03)	24 (60)	0.144
Bilateral lung	53 (58.24)	22 (32.84)		23 (58.97)	16 (40)	
Lung lobe						
Single	62 (68.13)	60 (89.55)	0.003	27 (69.23)	38 (95)	0.007
Multiple	29 (31.87)	7 (10.45)		12 (30.77)	2 (5)	

TB, tuberculosis; LC, lung cancer; SD, standard deviation; PPD, purified protein derivative test; TBAB, tuberculosis antibody; ESR, erythrocyte sedimentation rate.

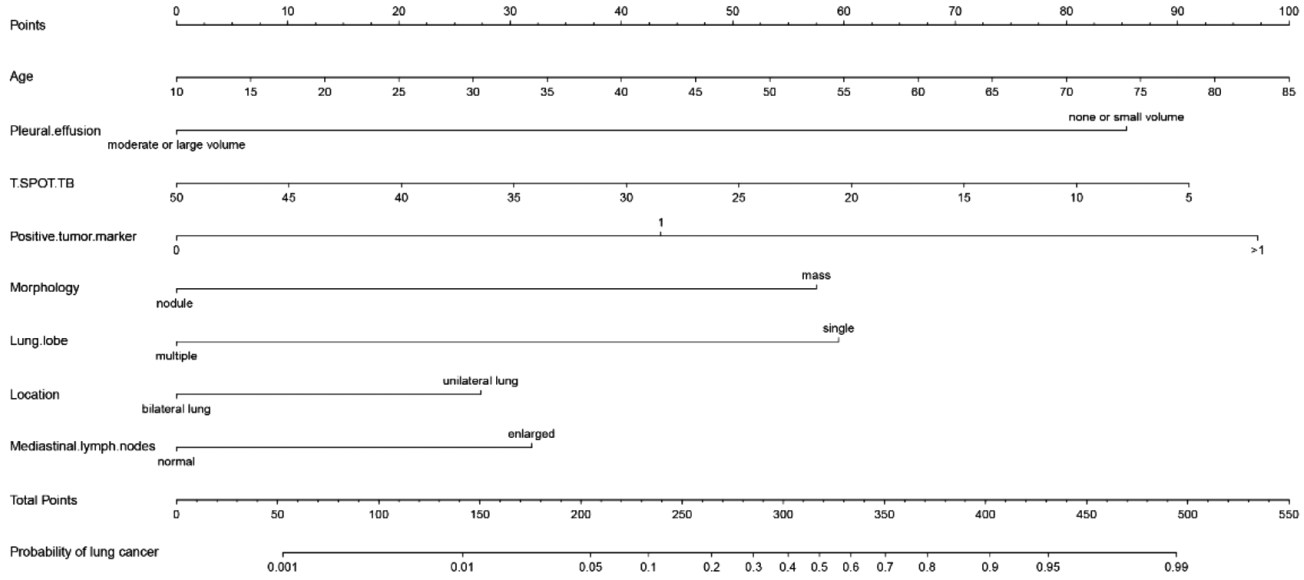

Figure 1. Development of a nomogram for predicting lung cancer cases with TB infection. The nomogram included age, pleural effusion, status of mediastinal lymph nodes, T-SPOT.TB, the number of positive tumor markers, lesions' morphology, location, and distribution. The nomogram summed the scores for each scale and variable. The total score on each scale indicated the risk of lung cancer.

Table 3. Univariate and multivariate logistic regression analysis of variables to predict the risk of lung cancer.

	Univariate analysis			Multivariate analysis		
	OR	95% CI	p	OR	95% CI	p
Age	1.028	1.002-1.056	0.0436	1.035	0.999-1.077	0.065
Sex						
Female *vs.* male	0.362	0.161-0.768	0.0104			
Smoking						
Yes *vs.* no	2.663	1.367-5.339	0.005			
Pleural effusion						
None or small *vs.* moderate or large volume	0.1442	0.022-0.532	0.012	0.108	0.0096-0.699	0.038
PPD test						
<10 *vs.* >10mm	1.024	0.544-1.929	0.941			
TBAB						
Positive *vs.* negative	0.729	0.235-2.087	0.563			
ESR	0.997	0.987-1.006	0.477			
Monocyte counts						
Positive *vs.* negative	1.35	0.668-2.724	0.401			
Number of tumor marker						
1 *vs.* 0	4.128	1.798-9.83	0.000	3.107	1.047- 9.687	0.043
>2 *vs.* 0	8.69	3.327-25.861	0.000	12.558	2.949-73.785	0.002
T-SPOT.TB	0.964	0.944- 0.983	0.000	0.949	0.92-0.976	0.000
Lung lobe						
Multiple *vs.* single lung lobe	0.249	0.095-0.584	0.002	0.212	0.056-0.687	0.014
Typical TB areas *vs.* other	1.354	0.718-2.562	0.3497			
Morphology						
Mass *vs.* nodule	5.783	2.938-11.76	0.000	4.4695	1.849- 11.448	0.001
Mediastinal lymph nodes						
Enlarged *vs.* normal	2.947	1.522-5.864	0.0016	2.295	0.9267-5.867	0.075
Location						
Bilateral *vs.* unilateral lung	0.351	0.179-0.671	0.0018	0.491	0.195-1.206	0.123

OR, odds ratio; CI, confidence interval; PPD, purified protein derivative test; TBAB, tuberculosis antibody; ESR, erythrocyte sedimentation rate; T-SPOT.TB, T cell spot test for tuberculosis.

For another TB-related auxiliary diagnosis method, we found no significant difference in the positive rate of PPD and TBAB in both training and validation cohorts. PPD is a skin test based on the principle of type IV allergy, used to detect whether the body has been infected with TB, but it cannot differentiate LTBI from active TB [31], and many factors might arise the variability in its result: the false-positive results due to prior Bacillus Calmette-Guerin vaccination or exposure to other non-tuberculous mycobacteria, as well as the operator-bias inherent to the assay; false negative results due to reasons such as anergy, recent live virus vaccination or overwhelming active TB infection and improper administration [32]. TBAB is a well-used method for detecting the TB antibody in clinical practice, but the diagnostic efficiency cannot meet the requirements for accurate TB diagnosis. TB antibody production generally needs 2-3 weeks after infection, and it only exists in the early stage and then disappears in the later stage of the TB infection. Moreover, the individual differences in antigen recognition are inescapable characteristics of the human TB humoral immune response, which also leads to the suboptimal outcome of TBAB [33,34]. Studies reported that patients at different stages of TB infection may induce an immune response to different antigens, and their sera can contain unrecognized antibodies against varied TB antigens [33,35]. Moreover, antigens and extracellular proteins derived from dead bacteria can lead to a false-positive outcome [35]. Therefore, the sensitivity and specificity of TBAB for TB diagnosis vary greatly [36], and the World Health Organization has not recommended TBAB as a diagnostic tool [37]. In a word, these TB-related auxiliary diagnosis experiments could hardly be applied to differentially diagnose for lung cancer in TB-endemic regions and provide valuable evidence for clinical decision-making.

Tumor markers are potent indicators for early screening or monitoring the recurrence of lung cancer. However, elevated values of these markers can also be detected in pulmonary TB. In the present study, we found a significant statistical difference in the positive rate

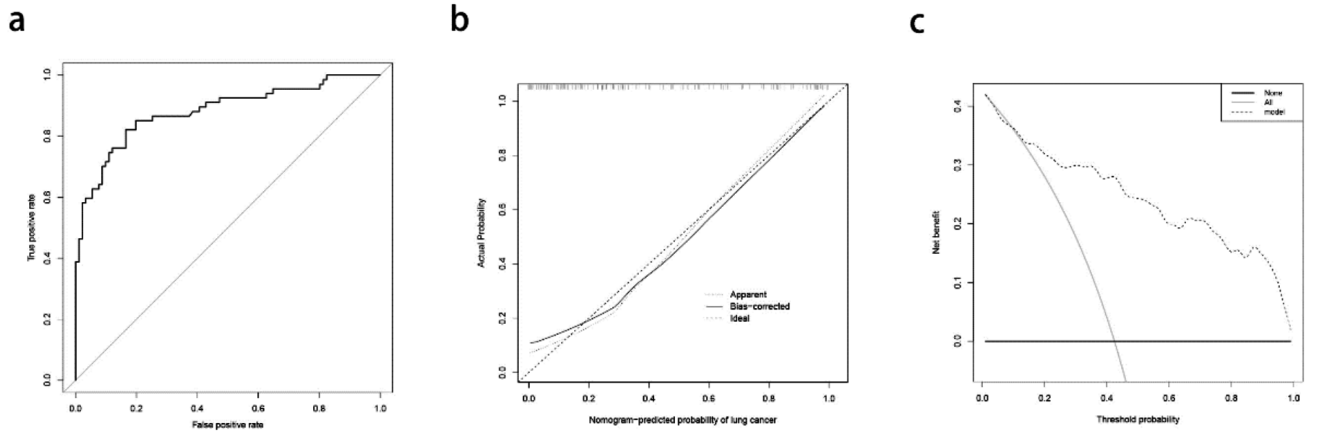

Figure 2. Evaluation of the nomogram model in the training cohort. a) The receiver operating characteristic curve indicates the good discriminative ability of lung cancer predicted by the nomogram model; b) the calibration curve shows the optimal agreement between predicted and observed situations; c) decision curve analysis demonstrates the potential application value of the model for predicting lung cancer.

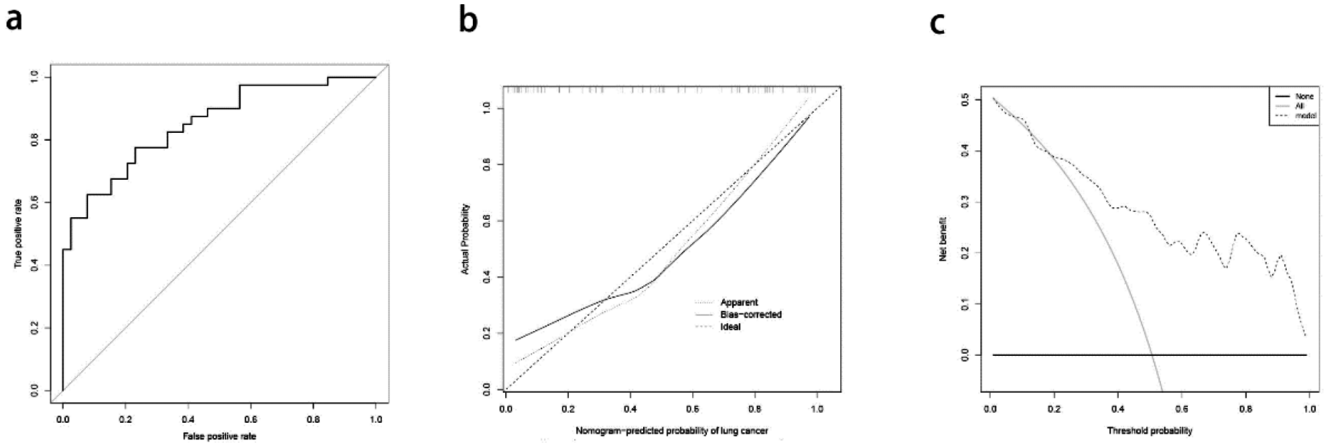

Figure 3. Evaluation of the nomogram model in the validation cohort. a) The receiver operating characteristic curve indicates the good discriminative ability of lung cancer predicted by the nomogram model; b) the calibration curve shows the optimal agreement between predicted and observed situations; c) decision curve analysis demonstrates the potential application value of the model for predicting lung cancer.

of lung tumor markers between the TB and lung cancer groups, but widely used markers such as CA125, NSE and CEA are positive in a proportion of TB cases, suggesting their sensitivity is less than optimal for diagnosis of lung cancer. CA125 is commonly used as a specific tumor marker for ovarian cancer, but its elevation is also seen in some non-gynecological diseases. A recent study reported that 75% of active pulmonary TB cases showed an increased concentration of serum CA125 [38]. A previous study has shown that several tumor markers are suboptimal in distinguishing non-small cell lung cancer from TB: AUC of CA125, NSE, and CEA were 0.626, 0.716, and 0.589, respectively [39]. In the present study, we indicated that the number of positive tumor markers is a more potent predictor than the positive status, and if the number of tumor markers increases, pulmonary lesions would more likely be diagnosed as lung cancer.

The radiological similarities between lung cancer and TB are the main reasons contributing to misdiagnosis or missed diagnosis for an indeterminate pulmonary nodule in asymptomatic individuals [40]. Even the positron emission tomography-CT could not well discriminate them, and TB infection can lead to a high false-positive rate and low specificity in the detection of lung cancer [41,42]. In the present study, the proportion of lung cancer and TB is nearly equal in TB infection patients, and the final model indicated that the mass in the unilateral lung and limited to a single lung lobe is prone to be malignant, which tends to accompany enlarged mediastinal lymph nodes. The result provided a comprehensive evaluation of the imaging features of lung cancer. It is interesting that the number of lung lobes covered by lesions is an independent predictor rather than lesion located in typical TB areas (dummy variables of lung lobe in Table 3), which is accordance to previous study that reported lung cancer tended to occur on an upper lobe location or the same side as previous TB infection, and lesion's location is improper indicator for the differential diagnosis between lung cancer and TB [43].

Pleural effusion can be used to identify the nature of pulmonary lesions through detecting the content of adenosine deaminase, CEA, and other tumor markers. However, almost one-third of lung cancer cases develop a pleural effusion [44], while less than 20% cases of TB have a TB pleural effusion [45], which means analysis for these markers in pleural effusion only apply to a small portion of cases, and thus we use the volume of pleural effusion rather than biochemical indicators as a potential variable. We found that pleural effusion tend to be none or only small in lung cancer cases, and of moderate or large volume in TB cases, and this result is partially consistent to that of Wang *et al.*, who found that lung cancer accounts for 27.8% in difficult cases with undiagnosed pleural effusions, while the proportion of TB is 40%, which means that TB derived pleural effusion could be even more common in China, but there is no open data about the difference of its volume between TB and lung cancer cases [46].

Some limitations exist in the present study. First, our study was based on documents from a single institution; second, it is a retrospective study with a limited number of cases, in which selection bias existed inevitably; third, coexistence of active TB and lung cancer did not include into our data because of its scarcity, and this situation is against monism in diagnostic principle. A prospective research study is needed to validate the feasibility and efficiency of the nomogram model.

Conclusions

We provided a model to predict lung cancer in TB infection patients, which is simple to use in clinical practice and provides an estimation for undiagnosed pulmonary lesions.

References

1. Lovly CM. Expanding horizons for treatment of early-stage lung cancer. N Engl J Med 2022;386:2050-1.
2. Cao M, Chen W. Epidemiology of lung cancer in China. Thorac Cancer 2019;10:3-7.
3. Ho JC, Leung CC. Management of co-existent tuberculosis and lung cancer. Lung Cancer 2018;122:83-7.
4. WHO. Guidelines on the management of latent tuberculosis infection. Available from: https://iris.who.int/bitstream/handle/10665/136471/9789241548908_eng.pdf?sequence=1.
5. Targeted tuberculin testing and treatment of latent tuberculosis infection. American Thoracic Society. MMWR Recomm Rep 2000;49:1-51.
6. Ortiz-Brizuela E, Apriani L, Mukherjee T, et al. Assessing the diagnostic performance of new commercial interferon-γ release assays for mycobacterium tuberculosis infection: a systematic review and meta-analysis. Clin Infect Dis 2023;76:1989-99.
7. Liu H, Huang Y, Wang J, et al. Clinical evaluation of T-SPOT.TB assay in 1 084 tuberculosis suspects. Zhonghua Jie He He Hu Xi Za Zhi 2014;37:192-6. [Article in Chinese].
8. Santín Cerezales M, Domínguez Benítez J. Diagnosis of tuberculosis infection using interferon-γ-based assays. Enferm Infecc Microbiol Clin 2011;29 Suppl 1:26-33.
9. Gilani B, Sergent SR. Interferon test. Treasure Island (FL): StatPearls Publishing; 2021.
10. Zhu M, Zhu Z, Yang J, Hu K. Performance evaluation of IGRA-ELISA and T-SPOT.TB for diagnosing tuberculosis infection. Clin Lab 2019;65.
11. Mazurek GH, Jereb J, Vernon A, et al. Updated guidelines for using Interferon Gamma release assays to detect mycobacterium tuberculosis infection - United States, 2010. MMWR Recomm Rep 2010;59:1-25.
12. Vynnycky E, Fine PE. Lifetime risks, incubation period, and serial interval of tuberculosis. Am J Epidemiol 2000;152:247-63.
13. Guo Z, Xiao D, Wang X, et al. Epidemiological characteristics of pulmonary tuberculosis in mainland China from 2004 to 2015: a model-based analysis. BMC Public Health 2019;19:219.
14. Zhu C, Liu Z, Li Z, et al. The performance and limitation of T-SPOT.TB for the diagnosis of TB in a high prevalence setting. J Thorac Dis 2014;6:713-9.
15. Patz EF, Jr., Pinsky P, Gatsonis C, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. JAMA Intern Med 2014;174:269-74.
16. Takiguchi Y, Sekine I, Iwasawa S. Overdiagnosis in lung cancer screening with low-dose computed tomography. J Thorac Oncol 2013;8:e101-2.
17. Bloom CI, Graham CM, Berry MP, et al. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. PLoS One 2013;8:e70630.
18. Cui EN, Yu T, Shang SJ, et al. Radiomics model for distinguishing tuberculosis and lung cancer on computed tomography scans. World J Clin Cases 2020;8:5203-12.
19. Shariat SF, Capitanio U, Jeldres C, Karakiewicz PI. Can nomograms be superior to other prediction tools?. BJU Int 2009;103:492-5; discussion 495-7.
20. Fan WC, Ting WY, Lee MC, et al. Latent TB infection in newly diagnosed lung cancer patients - a multicenter prospective observational study. Lung Cancer 2014;85:472-8.
21. Tamura A, Fukami T, Hebisawa A, Takahashi F. Recent trends in the incidence of latent tuberculosis infection in Japanese patients

with lung cancer: a small retrospective study. J Infect Chemother 2020;26:315-7.
22. Moy MP, Levsky JM, Berko NS, et al. A new, simple method for estimating pleural effusion size on CT scans. Chest 2013;143: 1054-9.
23. Pahal P, Sharma S. PPD skin test. Treasure Island (FL): StatPearls Publishing; 2021.
24. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-87.
25. Van Calster B, Wynants L, Verbeek JFM et al. Reporting and Interpreting decision curve analysis: a guide for investigators. Eur Urol 2018;74:796-804.
26. Molina-Romero C, Arrieta O, Hernández-Pando R. Tuberculosis and lung cancer. Salud Publica Mex 2019;61:286-91.
27. Maci E, Comito F, Frezza AM, et al. Lung nodule and functional changes in smokers after smoking cessation short-term treatment. Cancer Invest 2014;32:388-93.
28. Bordignon V, Bultrini S, Prignano G, et al. High prevalence of latent tuberculosis infection in autoimmune disorders such as psoriasis and in chronic respiratory diseases, including lung cancer. J Biol Regul Homeost Agents 2011;25:213-20.
29. Feng B, Li Y, Guo D, et al. Research on the value of the T cell spot test for tuberculosis for the diagnosis of lung cancer combined with pulmonary tuberculosis. Thorac Cancer 2018; 9:1231-4.
30. Ma Y, Li R, Shen J et al. Clinical effect of T-SPOT.TB test for the diagnosis of tuberculosis. BMC Infect Dis 2019;19:993.
31. Wu X, Li Q, Liang Y et al. Clinical evaluation of a homemade enzyme-linked immunospot assay for the diagnosis of active tuberculosis in China. Mol Biotechnol 2011;47:18-25.
32. de Lima Corvino DF, Shrestha S, Kosmin AR. Tuberculosis screening. Treasure Island (FL): StatPearls Publishing; 2021.
33. Wu X, Yang Y, Zhang J, et al. Comparison of antibody responses to seventeen antigens from Mycobacterium tuberculosis. Clin Chim Acta 2010;411:1520-8.
34. Samanich K, Belisle JT, Laal S. Homogeneity of antibody responses in tuberculosis patients. Infect Immun 2001;69: 4600-9.
35. Kunnath-Velayudhan S, Salamon H, Wang HY, et al. Dynamic antibody responses to the Mycobacterium tuberculosis proteome. Proc Natl Acad Sci U S A 2010;107:14703-8.
36. Bai XJ, Yang YR, Liang JQ, et al. Diagnostic performance and problem analysis of commercial tuberculosis antibody detection kits in China. Mil Med Res 2018;5:10.
37. WHO. Commercial serodiagnostic tests for diagnosis of tuberculosis: policy statement. Available from: https://iris.who.int/bitstream/handle/10665/44652/9789241502054_eng.pdf?sequence =1; 2011.
38. Mikačić M, Vasilj I, Vasilj M, et al. Tumor marker CA 125 in the diagnosis of active pulmonary tuberculosis - a study of adults in mostar, B&H. Psychiatr Danub 2017;29 Suppl 4:841-4.
39. Leng S, Zheng J, Jin Y, et al. Plasma cell-free DNA level and its integrity as biomarkers to distinguish non-small cell lung cancer from tuberculosis. Clin Chim Acta 2018;477:160-5.
40. Lang S, Sun J, Wang X, et al. Asymptomatic pulmonary tuberculosis mimicking lung cancer on imaging: a retrospective study. Exp Ther Med 2017;14:2180-8.
41. Niyonkuru A, Chen X, Bakari KH et al. Evaluation of the diagnostic efficacy of (18) F-Fluorine-2-Deoxy-D-Glucose PET/CT for lung cancer and pulmonary tuberculosis in a Tuberculosis-endemic Country. Cancer Med 2020;9:931-42.
42. Kim H, Kim HY, Goo JM, Kim Y. Lung cancer CT screening and lung-RADS in a tuberculosis-endemic country: the Korean lung cancer screening project (K-LUCAS). Radiology 2020;296:181-8.
43. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for Management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. Radiology 2017;284:228-43.
44. Jany B, Welte T. Pleural effusion in adults-etiology, diagnosis, and treatment. Dtsch Arztebl Int 2019;116:377-86.
45. Valdés L, Ferreiro L, Cruz-Ferro E, et al. Recent epidemiological trends in tuberculous pleural effusion in Galicia, Spain. Eur J Intern Med 2012;23:727-32.
46. Wang XJ, Yang Y, Wang Z, et al. Efficacy and safety of diagnostic thoracoscopy in undiagnosed pleural effusions. Respiration 2015;90:251-5.