Journal of Infection and Public Health 18 (2025) 102740

Contents lists available at ScienceDirect

Journal of Infection and Public Health

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

Original Article

Early differential diagnosis models of Talaromycosis and Tuberculosis in HIV-negative hosts using clinical data and machine learning

Ye Qiu [a,b,1], Zheng-tu Li [a,1], Shi-xiong Yang [d,1], Wu-shu Chen [f], Yong Zhang [g], Qun-yu Kong [g], Ling-rui Chen [g], Jie Huang [d], Lü Lin [d], Kan Xie [d], Wen Zeng [e], Shao-qiang Li [a], Yang-qing Zhan [a], Yan Wang [a], Jian-quan Zhang [c,*], Feng Ye [a,**]

[a] State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong 510120, China
[b] Department of Respiratory and Critical Medicine, The Affiliated Tumour Hospital of Guangxi Medical University, Nanning, Guangxi 530021, China
[c] Department of Respiratory and Critical Medicine, The Eighth Affiliated Hospital, Sun Yat-Sen University, Shenzhen, Guangdong 518000, China
[d] Department of Tuberculosis Ward, Guangxi Nanning Fourth People's Hospital, Nanning, Guangxi 530021, China
[e] Department of Respiratory and Critical Medicine, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi 530021, China
[f] Nanshan School of Guangzhou Medical University, Guangzhou, Guangdong 510120, China
[g] Chongqing Nanpeng Artificial Intelligence Technology Research Institute Co., Ltd, Chongqing 401123, China

ARTICLE INFO

Article history:
Received 8 November 2024
Received in revised form 2 March 2025
Accepted 4 March 2025

Keywords:
Talaromyces marneffei
HIV-negative
Pulmonary tuberculosis
Clinical prediction models
Differential diagnosis

ABSTRACT

Background: *Talaromyces marneffei* is an emerging pathogen, and the number of infections in HIV-negative individuals is increasing. In HIV-negative individuals, talaromycosis is usually misdiagnosed as another disease, especially tuberculosis (TB).
Methods: We retrospectively extracted the clinical data of HIV-negative patients with *Talaromyces marneffei* infection from 2018 to 2023, analyzed the differences between TB patients and talaromycosis patients and attempted to establish differential diagnosis models utilizing clinical prediction models for these two diseases.
Results: Overall, 718 patients, including 137 patients with talaromycosis and 581 patients with pulmonary tuberculosis (PTB), were enrolled in this study. According to the multivariate analysis, age > 65 years, expectoration, and PLT count were independent predictors for TB. Fever, chest pain, gasping, rash, lymphadenectasis, osteolysis, Neu count, EOS count, and ALB were independent predictors for talaromycosis. Receiver operating characteristic (ROC) curve analysis of the training set showed that the area under the curve (AUC) (95 % CI) of the clinical differential model based on logistic regression analysis was 0.918 (0.884–0.953). The model was verified in the validation set. ROC curve analysis of the validation set showed that the AUC (95 % CI) was 0.900 (0.841–0.959).
Conclusion: These new differential diagnosis models can calculate the probability of either talaromycosis or tuberculosis.

Introduction

Talaromycosis is an emerging and critical invasive pulmonary mycosis caused by the thermally dimorphic fungus *Talaromyces marneffei* (TM) [1,2]. Immunocompromised individuals, especially those with human immunodeficiency virus (HIV) infection [2], are susceptible to this mycosis, which is endemic throughout southeast Asia and southern China, especially in Guangxi, Guangzhou and Hong Kong [1]. In the last two decades, talaromycosis has accounted for 16 % of HIVassociated hospital admissions [3–5] and is the second most common cause of HIVassociated bloodstream infections and death (mortality up to 28 %) [1]. Due to the highly active antiretroviral therapy used in patients with HIV infection, the epidemiology of talaromycosis is changing, with a decreasing trend among HIV-positive hosts but a rapidly increasing incidence in HIV-negative hosts with immunocompromising conditions such as anti-IFN-γ autoantibody syndrome, autoimmune diseases, primary immunodeficiency, glucocorticoids and/or immunosuppression [6,7].

* Corresponding author.
** Correspondence to: The First Affiliated Hospital of Guangzhou Medical University, 151 Yanjiang Xi Road, Guangzhou, Guangdong 510120, China.
E-mail addresses: jqzhang2002@126.com (J.-q. Zhang), yefeng@gird.cn (F. Ye).
[1] These authors contributed equally as co-first authors.

https://doi.org/10.1016/j.jiph.2025.102740

Approximately 17300 talaromycosis cases and 4900 associated deaths occur annually [4]. Even with active antifungal treatment, the mortality and recurrence rates of TM infection in HIV-negative individuals are still high (up to 31.4 % and 26.0 %, respectively) [5,6,8]. In 2021, the Lancet Global Health journal launched a global call for research to increase awareness of this neglected and lethal mycosis [4]. In the next year, TM was listed as a fungal priority pathogen by the World Health Organization [9].

Despite the calamitous toll of TM, the global focus on diagnostic and therapeutic strategies for this pathogen in HIV-negative individuals remains poor. Furthermore, due to the lack of specific manifestations and insufficient knowledge of this pathogen, talaromycosis in HIV-negative individuals is usually misdiagnosed as another disease, especially tuberculosis (TB); the rate of TB misdiagnosis can be as high as 80.7 %, and 38.1 % of patients receive antituberculosis treatment) [1,8]. The long delay in diagnosis of talaromycosis has become a critical determinant of prognosis. The diagnostic confusion arises from shared clinical-radiological patterns (e.g., chronic fever with cavitary lung lesions) and limitations in conventional microbiology. For instance, TM's small intracellular yeasts may be mistaken for granuloma fragments or atypical mycobacteria on rapid stains, delaying fungal-specific testing. Moreover, 62 % of misdiagnosed TM cases in our cohort showed false-positive T-SPOT. TB results, likely due to bystander T-cell activation [1,8]. These factors, compounded by empiric TB treatment practices in endemic areas, contribute to the observed 80.7 % misdiagnosis rate.

Clinical prediction models (CPMs), previously referred to as prognosis models (PMs), strategically harness the clinical attributes of the subject under study and are designed to predict the likelihood of the clinical outcome or prognosis associated with a particular disease [10,11]. CPMs are currently a valuable resource for clinicians in clinical diagnosis and treatment and are widely used in clinical research. In differential diagnosis, the application of machine learning models rooted in clinical characteristics is extensive, exhibiting commendable diagnostic efficacy and robust performance [12,13]. Diagnostic models that can differentiate between talaromycosis and TB in HIV-negative hosts are lacking and are therefore urgently needed.

Thus, in this study, we examined multicenter data on talaromycosis in HIV-negative hosts, comprehensively investigated the fundamental characteristics of talaromycosis, elucidated the distinctions in clinical characteristics between talaromycosis and TB, and tried to establish differential diagnosis models leveraging CPMs for these two diseases. This study aimed to provide clinicians with a valuable tool for differentiating talaromycosis from TB infection in HIV-negative hosts and to enhance clinicians' understanding and awareness of talaromycosis, thereby improving the prognosis of patients with this disease.

Method

Study design

This national multicenter retrospective longitudinal study was conducted at China, between February 2012 and October 2023. The Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement was used as a reporting guideline [14]. Baseline data were retrospectively collected from medical records, with written informed consent obtained during patient enrollment to authorize both retrospective data use and prospective follow-up assessments.

Study population

Definition/diagnostic criteria of talaromycosis

The definition and diagnostic criteria of talaromycosis followed the Global Guidelines for the Diagnosis and Management of Endemic Mycoses: An Initiative of the European Confederation of Medical Mycology in cooperation with the International Society for Human and Animal Mycology from 2021 [3]. A presumptive diagnosis of talaromycosis is made based on smear, culture, microscopic or histopathological examination of clinical specimens (including blood, skin, lung, marrow, bronchoalveolar lavage fluid [BALF], stool, urine, lymph node, sputum, or secretions). The characteristics of TM include the identification of a transverse septum in a dividing yeast cell that is 3–6 μm in diameter, round-to-oval in shape, extracellular, and present within macrophages. Antigen detection (Mp1p), TM commercial antigen detection assays, qPCR assays and metagenomic next-generation sequencing (mNGS) are also used to diagnose TM.

Definition/diagnostic criteria of pulmonary tuberculosis (PTB)

The definition and diagnostic criteria for pulmonary tuberculosis followed the Diagnosis of Tuberculosis in Adults and Children: An Initiative of the American Thoracic Society (ATS) Clinical Practice Guideline: Summary for Clinicians (2016 version) and the WHO consolidated guidelines on tuberculosis: Module 3: diagnosis-rapid diagnostics for tuberculosis detection (2021) [15–17].

Inclusion and exclusion criteria for participants

Inclusion criteria for PTB patients

HIV-negative adult participants were included if they met the following criteria: age ≥ 18 years, a chest radiograph suggestive of pulmonary tuberculosis, and less than 7 days of previous tuberculosis treatment with one of the following conditions: ① a positive acid-fast bacilli smear of sputum or BALF (grade 1 or higher on the WHO and International Union Against Tuberculosis and Lung Disease Scale), ② a positive Xpert MTB/RIF test (Cepheid, Sunnyvale, CA, USA), ③ positive mNGS results, ④ a semiquantitative *M. tuberculosis* bacterial load of medium or high, or ⑤ a positive culture or histopathological examination for *Mycobacterium tuberculosis* using sputum, BALF or lung tissue. All participants provided written informed consent.

Inclusion criteria for talaromycosis patients

HIV-negative adult participants were included if they met the following criteria: age ≥ 18 years, a definitive talaromycosis diagnosis confirmed by the Global Guidelines for the Diagnosis and Management of Endemic Mycoses: An Initiative of the European Confederation of Medical Mycology in cooperation with the International Society for Human and Animal Mycology (2021 version), and no antifungal treatment. All participants provided written informed consent.

Exclusion criteria for PTB and talaromycosis patients

The main exclusion criteria were as follows: PTB patients who had 7 days or more of tuberculosis treatment before the baseline visit, HIV-positive patients, patients who had TM and TB coinfection, or patients who had TM and nontuberculous mycobacteria coinfection.

Statistical analysis

Logistic regression analysis was used to determine which variables were predictive of talaromycosis and tuberculosis. The variables that were significant at the 0.05 level were included in the final multiple logistic model. The samples were divided into two sets at a ratio of 7:3 using stratified sampling; 70 % of the total study samples were used for model development (the derivation cohort), and the remaining 30 % of the samples were used as the test cohort for validation. The discrimination ability was evaluated by the area under the curve (AUC) of the receiver operating characteristic (ROC) curve, and the degree of calibration was assessed using the Hosmer–Lemeshow goodness-of-fit test. Clinical applicability was elucidated via clinical decision curve analysis (DCA) and a clinical impact curve (CIC). All the statistical analyses were performed using R version 4.3.1. The data are shown as numbers and percentages. Data with missing rates below 20 % were interpolated using the K-nearest neighbor (KNN) algorithm and weighted average algorithm. Categorical comparisons were made by using chi-square or Fisher's exact tests where appropriate.

Results

Baseline characteristics

A total of 718 patients, including 137 patients with talaromycosis and 581 patients with PTB, were enrolled in this study. All patients in this study, including both TB and TM cohorts, were recruited from the Chinese population. We included all eligible TB and TM patients diagnosed within the predefined study period across participating centers. There were statistically significant differences in age, cough, expectoration, fever, chest pain, gasping, rash, lymphadenectasis, osteolysis, white blood cell (WBC) count, neutrophil (Neu) count, leukomonocyte (Leu) count, platelet (PLT) count, hemoglobin (HGB), eosinophil (EOS) count, $CD4^+$ T lymphocyte count, $CD8^+$ T lymphocyte count, globulin (GLB) concentration, and albumin (ALB) concentration between patients with talaromycosis and patients with PTB ($P < 0.05$) (Table 1). According to the analysis of baseline characteristics, age > 65 years, cough and expectoration were significantly more common in patients with TB than in patients with talaromycosis. However, age ≤ 65 years, fever, chest pain, gasping, rash, lymphadenectasis, osteolysis, elevated WBC, Neu, Leu, PLT, EOS, GLB, ALB, decreased $CD4^+$ T and $CD8^+$ T lymphocytes, and anemia were significantly more common in patients with talaromycosis.

The dataset was randomly divided into a training set and a testing set at a 7:3 ratio. The training set totaling 503 patients (96 patients with talaromycosis and 407 patients with PTB) was used to build the model, and the test set totaling 215 patients (41 patients with talaromycosis and 174 patients with PTB), was used to evaluate the accuracy of the model in predicting unknown samples. The baseline characteristics of the training set and the validation set were similar ($P > 0.05$) (Table 2).

The baseline characteristics of patients with talaromycosis and TB patients in the training set and validation set were also compared (Supplementary tables).

Univariate and multivariate logistic regression analyses

Univariate logistic regression analyses revealed that age > 65 years, expectoration, fever, chest pain, gasping, rash, lymphadenectasis, osteolysis, WBC count, Neu count, PLT count, EOS count, $CD8^+$ T lymphocyte count and ALB concentration were significant ($P < 0.05$). Multivariate logistic regression analysis revealed that age > 65 years, expectoration, fever, chest pain, gasping, rash, lymphadenectasis, osteolysis, Neu count, PLT count, EOS count, $CD8^+$ T lymphocyte count and ALB concentration were independent predictors for distinguishing between talaromycosis and TB ($P < 0.05$). In multivariate analysis, age > 65 years (OR: 0.095, 95 % CI: 0.031–0.250, $p < 0.001$), expectoration (OR: 0.359, 95 % CI: 0.154–0.840, $p = 0.017$), and PLT count (OR: 0.390, 95 % CI: 0.178–0.817, $p = 0.015$) were independent predictors for TB. However, fever (OR: 2.502, 95 % CI: 1.258–5.036, $p = 0.009$), chest pain (OR: 3.390, 95 % CI: 1.365–1.365, $p = 0.008$), gasping (OR: 2.664, 95 % CI: 1.268–5.671, $p = 0.010$), rash (OR: 13.267, 95 % CI: 3.309–70.179, $p = 0.001$), lymphadenectasis (OR: 9.559, 95 % CI: 4.206–22.624, $p = 0.000$), osteolysis (OR: 47.114, 95 % CI: 6.563–362.410, $p = 0.000$), Neu count (OR: 5.794, 95 % CI: 2.792–12.686, $p = 0.000$), EOS count (OR: 2.813, 95 % CI: 1.338–5.916, $p = 0.006$), and ALB concentration (OR: 2.251, 95 % CI: 1.043–5.078, $p = 0.043$) were independent predictors for talaromycosis (Table 3, Fig. 1A).

We used these independent variables to establish models for differential diagnosis between talaromycosis and TB using CPMs. Age > 65 y, expectoration, fever, chest pain, gasping, rash, lymphadenectasis, osteolysis, Neu count, PLT count, EOS count, and ALB concentration were found to be significant, and the Akaike information criterion (AIC) was 266.27 ($P < 0.05$).

Table 1
Comparison of demographic and clinical characteristics between patients with talaromycosis and patients with tuberculosis.

Variable	PTB (N = 581)	TM (N = 137)	P value
Sex			0.126
-Male	412 (71 %)	88 (64 %)	
-Female	169 (29 %)	49 (36 %)	
Age			< 0.001
-≤ 65 y	406 (70 %)	122 (89 %)	
- > 65 y	175 (30 %)	15 (11 %)	
Height	163 (158, 168)	162 (158, 165)	0.068
Weight	52 (46, 60)	52 (47, 56)	0.116
Cough	528 (91 %)	115 (84 %)	0.017
Expectoration	499 (86 %)	96 (70 %)	< 0.001
Fever	186 (32 %)	92 (67 %)	< 0.001
Chest pain	48 (8.3 %)	29 (21 %)	< 0.001
Gasping	142 (24 %)	50 (36 %)	0.004
Rash	5 (0.9 %)	35 (26 %)	< 0.001
Lymphadenectasis	32 (5.5 %)	70 (51 %)	< 0.001
Osteolysis	6 (1.0 %)	23 (17 %)	< 0.001
WBC			< 0.001
$3.5–9.5 \times 10^9$/L	442 (76 %)	42 (31 %)	
$> 9.5 \times 10^9$/L	139 (24 %)	95 (69 %)	
Neu			< 0.001
$1.8–6.3 \times 10^9$/L	380 (65 %)	33 (24 %)	
$> 6.3 \times 10^9$/L	201 (35 %)	104 (76 %)	
Leu			< 0.001
$0.8–3.5 \times 10^9$/L	576 (99 %)	90 (66 %)	
$> 3.5 \times 10^9$/L	5 (0.9 %)	47 (34 %)	
PLT			0.077
$125–350 \times 10^9$/L	338 (58 %)	91 (66 %)	
$> 350 \times 10^9$/L	243 (42 %)	46 (34 %)	
HGB			0.480
130–175 g/L	569 (98 %)	136 (99 %)	
< 130 g/L	12 (2.1 %)	1 (0.7 %)	
EOS			< 0.001
$0.02–0.5 \times 10^9$/L	466 (80 %)	72 (53 %)	
$> 0.5 \times 10^9$/L	115 (20 %)	65 (47 %)	
$CD4^+$ T lymphocyte			0.042
561–1137 cells/uL	561 (97 %)	127 (93 %)	
< 561 cells/uL	20 (3.4 %)	10 (7.3 %)	
$CD8^+$ T lymphocyte			0.003
220–1130 cells/uL	576 (99 %)	130 (95 %)	
< 220 cells/uL	5 (0.9 %)	7 (5.1 %)	
GLO			0.046
20–40 g/L	498 (86 %)	108 (79 %)	
> 40 g/L	83 (14 %)	29 (21 %)	
ALB			< 0.001
40–55 g/L	264 (45 %)	32 (23 %)	
< 40 g/L	317 (55 %)	105 (77 %)	

The data are expressed as numbers (percentages). WBC=white blood cell, Neu=neutrophil, Leu=leukomonocyte, PLT=platelet, HGB=hemoglobin, EOS=eosinophil, GLO=globulin, ALB=albumin

Table 2
Comparison of demographic and clinical characteristics between the training and validation sets.

Variable	Training set (N = 503)	Validation set (N = 215)	P value
Disease			0.996
Tuberculosis	407 (81 %)	174 (81 %)	
Talaromycosis	96 (19 %)	41 (19 %)	
Sex			0.821
Male	349 (69 %)	151 (70 %)	
Female	154 (31 %)	64 (30 %)	
Age			0.593
≤ 65 y	367 (73 %)	161 (75 %)	
> 65 y	136 (27 %)	54 (25 %)	
Height	164 (158, 168)	162 (158, 168)	0.410
Weight	52 (46, 59)	52 (47, 60)	0.482
Cough	454 (90 %)	189 (88 %)	0.345
Expectoration	414 (82 %)	181 (84 %)	0.540
Fever	193 (38 %)	85 (40 %)	0.769
Chest pain	52 (10 %)	25 (12 %)	0.609
Gasping	138 (27 %)	54 (25 %)	0.520
Rash	26 (5.2 %)	14 (6.5 %)	0.472
Lymphadenectasis	72 (14 %)	30 (14 %)	0.899
Osteolysis	20 (4.0 %)	9 (4.2 %)	0.896
WBC			0.990
$3.5–9.5 \times 10^9$/L	339 (67 %)	145 (67 %)	
$> 9.5 \times 10^9$/L	164 (33 %)	70 (33 %)	
Neu			0.912
$1.8–6.3 \times 10^9$/L	290 (58 %)	123 (57 %)	
$> 6.3 \times 10^9$/L	213 (42 %)	92 (43 %)	
Leu			0.419
$0.8–3.5 \times 10^9$/L	464 (92 %)	202 (94 %)	
$> 3.5 \times 10^9$/L	39 (7.8 %)	13 (6.0 %)	
PLT			0.160
$125–350 \times 10^9$/L	309 (61 %)	120 (56 %)	
$> 350 \times 10^9$/L	194 (39 %)	95 (44 %)	
HGB			0.363
130–175 g/L	492 (98 %)	213 (99 %)	
< 130 g/L	11 (2.2 %)	2 (0.9 %)	
EOS			0.182
$0.02–0.5 \times 10^9$/L	384 (76 %)	154 (72 %)	
$> 0.5 \times 10^9$/L	119 (24 %)	61 (28 %)	
$CD4^+$ T lymphocyte			0.689
561–1137 cells/uL	481 (96 %)	207 (96 %)	
< 561 cells/uL	22 (4.4 %)	8 (3.7 %)	
$CD8^+$ T lymphocyte			> 0.999
220–1130 cells/uL	494 (98 %)	212 (99 %)	
< 220 cells/uL	9 (1.8 %)	3 (1.4 %)	
GLO			0.917
20–40 g/L	425 (84 %)	181 (84 %)	
> 40 g/L	78 (16 %)	34 (16 %)	
ALB			0.547
40–55 g/L	211 (42 %)	85 (40 %)	
< 40 g/L	292 (58 %)	130 (60 %)	

The data are expressed as numbers (percentages). WBC=white blood cell, Neu=neutrophil, Leu=leukomonocyte, PLT=platelet, HGB=hemoglobin, EOS=eosinophil, GLO=globulin, ALB=albumin

Construction and validation of the models

The nomogram for PTB and talaromycosis was developed by using multivariate logistic regression analysis models as the final diagnostic models after factor selection (Fig. 1B). The nomogram was internally validated in the training cohort. The indices for the training set were as follows: X-squared = 15.586, df = 18, p value = 0.6214. The indices for the test set were as follows: X-squared = 24.72, df = 18, p value = 0.1328. The Hosmer–Lemeshow (H-L) test indicated that there was no significant difference between the predicted value and the true value. The calibration curves are shown in Fig. 1C and D.

ROC curve analysis of the training set showed that the AUC (95 % CI) of the clinical differential model constructed based on logistic regression analysis was 0.918 (0.884–0.953), suggesting that the model had good predictive value.

The model was verified in the validation set. ROC curve analysis of the validation set showed that (95 % CI) was 0.900 (0.841–0.959), indicating that the model performed well in the validation dataset (Fig. 2A and B).

Clinical efficacy analysis

DCA was used to test the clinical efficacy of the model. The results are shown in Figs. 2C, D and E. The curve of the constructed model was significantly higher than baseline in both the test set and the validation set, which indicated that the model could provide better net benefits in most threshold ranges. A clinical impact curve was generated. The red curve indicates the number of people who are diagnosed with talaromycosis according to the simple model under each threshold probability. The blue curve shows the number of true talaromycosis cases at each threshold probability.

Discussion

Talaromycosis shares similar clinical and pathologic presentations with pulmonary tuberculosis [8]. The definitive diagnostic benchmarks for talaromycosis were determined by fungal microscopy, histology, and fungal cultures [3]. TB diagnosis has analogous criteria. However, microbiological evaluations, such as acid-fast bacilli (AFB), periodic acid Schiff (PAS) staining and positive culture for TM or Mtb, can assist in differential diagnosis, but the findings have low sensitivity, and the process is time consuming. In addition, the substantial overlap in the epidemiological spectra of talaromycosis and TB often prompts clinicians to initiate empirical antituberculosis treatment for patients with talaromycosis at an early stage. Differential diagnosis between the 2 diseases is very important in TB-endemic Asian countries such as Southeast Asia and southern China, where the incidence of talaromycosis is increasing [1–7]. The resemblance in clinical presentation between talaromycosis and TB extends beyond pulmonary manifestations and encompasses systemic effects, contributing to a myriad of diagnostic challenges in clinical settings [18–20]. The natural history and treatment options of these 2 diseases are distinct; therefore, misdiagnosis or delayed diagnosis can affect not only treatment outcomes but also health care costs and economic losses. This intricacy underscores the critical need for refined diagnostic approaches in addressing these clinical scenarios. The model identifies high-risk patients using routine clinical data, enabling prioritized testing or early empiric antifungals while awaiting results, and the model complements (not replaces) lab diagnostics.

For timely and accurate differential diagnosis, we combined clinical symptoms and laboratory, radiological, and pathologic features and established an accurate prediction model to differentiate between talaromycosis and PTB. This model could be useful in clinical practice because it can mitigate the occurrence of misdiagnoses in patients with talaromycosis and TB. Leveraging data sourced from four distinct centers specializing in talaromycosis and TB and drawing upon readily available clinical factors, our model exhibited commendable predictive performance, as indicated by a robust C-statistic of 0.918 (95 % CI 0.884–0.953), attesting to its strong internal validity. The findings of our study suggested a noteworthy association between specific clinical factors and the likelihood of talaromycosis as opposed to TB in patients younger than 65. Notably, the presence of chest pain, gasping, rash, osteolysis, abnormal lymph nodes, and aberrations in Neu count, EOS count, and ALB levels increased the probability of talaromycosis. These results underscore the potential clinical utility of our diagnostic model in enhancing accuracy and aiding in targeted diagnostic decision-making.

The prediction model can also fit the actual clinical features and guide clinical practice. Osteolysis is common in patients with

Table 3
Results of the univariable and multivariable logistic regression analyses in the training set for the diagnosis of talaromycosis.

Variable	Univariable Analysis		Multivariable Analysis	
	OR (95 % CI)	P value	OR (95 % CI)	P value
Sex	1.04 (0.64–1.68)	0.881		
Age > 65 y	0.29 (0.15–0.57)	< 0.001	0.095 (0.031–0.250)	< 0.001
Cough	0.80 (0.39–1.62)	0.529		
Expectoration	0.49 (0.29–0.83)	0.008	0.359 (0.154–0.840)	0.017
Fever	4.57 (2.84–7.36)	< 0.001	2.502 (1.258–5.036)	0.009
Chest pain	3.40 (1.85–6.23)	< 0.001	3.390 (1.365–8.335)	0.008
Gasping	1.90 (1.19–3.04)	0.007	2.664 (1.268–5.671)	0.010
Rash	42.43 (12.42–144.97)	< 0.001	13.267 (3.309–70.179)	0.001
Lymphadenectasis	13.48 (7.67–23.70)	< 0.001	9.559 (4.206–22.642)	< 0.001
Osteolysis	20.15 (6.56–61.86)	< 0.001	47.114 (6.563–362.410)	< 0.001
WBC > 9.5×10^9/L	6.52 (4.02–10.58)	< 0.001	1.77 (0.65–4.83)	0.266
N > 6.3×10^9/L	5.66 (3.42–9.38)	< 0.001	5.794 (2.792–12.686)	< 0.001
PLT > 350×10^9/L	0.60 (0.37–0.97)	0.037	0.390 (0.178–0.817)	0.015
HGB < 130 g/L	0.42 (0.05–3.30)	0.408		
EOS > 0.5×10^9/L	3.97 (2.47–6.37)	< 0.001	2.813 (1.338–5.916)	0.006
$CD4^+$ T lymphocyte < 561 cells/uL	1.63 (0.62–4.28)	0.322		
$CD8^+$ T lymphocyte < 561 cells/uL	8.98 (2.20–36.57)	0.002	5.28 (0.43–64.14)	0.191
ALB < 40 g/L	3.62 (2.11–6.20)	< 0.001	2.251 (1.043–5.078)	0.000

After excluding covariates from the baseline factors in the training set, the remaining variables were subjected to univariable logistic regression analyses with the outcome variables (subgroups), and those with statistically significant (P < 0.05) results in the univariable analyses were subjected to logistic multivariable analyses, with a backward rule for the removal of redundant variables to identify influences affecting the outcome of the two diseases. WBC=white blood cell, Neu=neutrophil, Leu=leukomonocyte, PLT=platelet, HGB=hemoglobin, EOS=eosinophil, GLO=globulin, ALB=albumin.

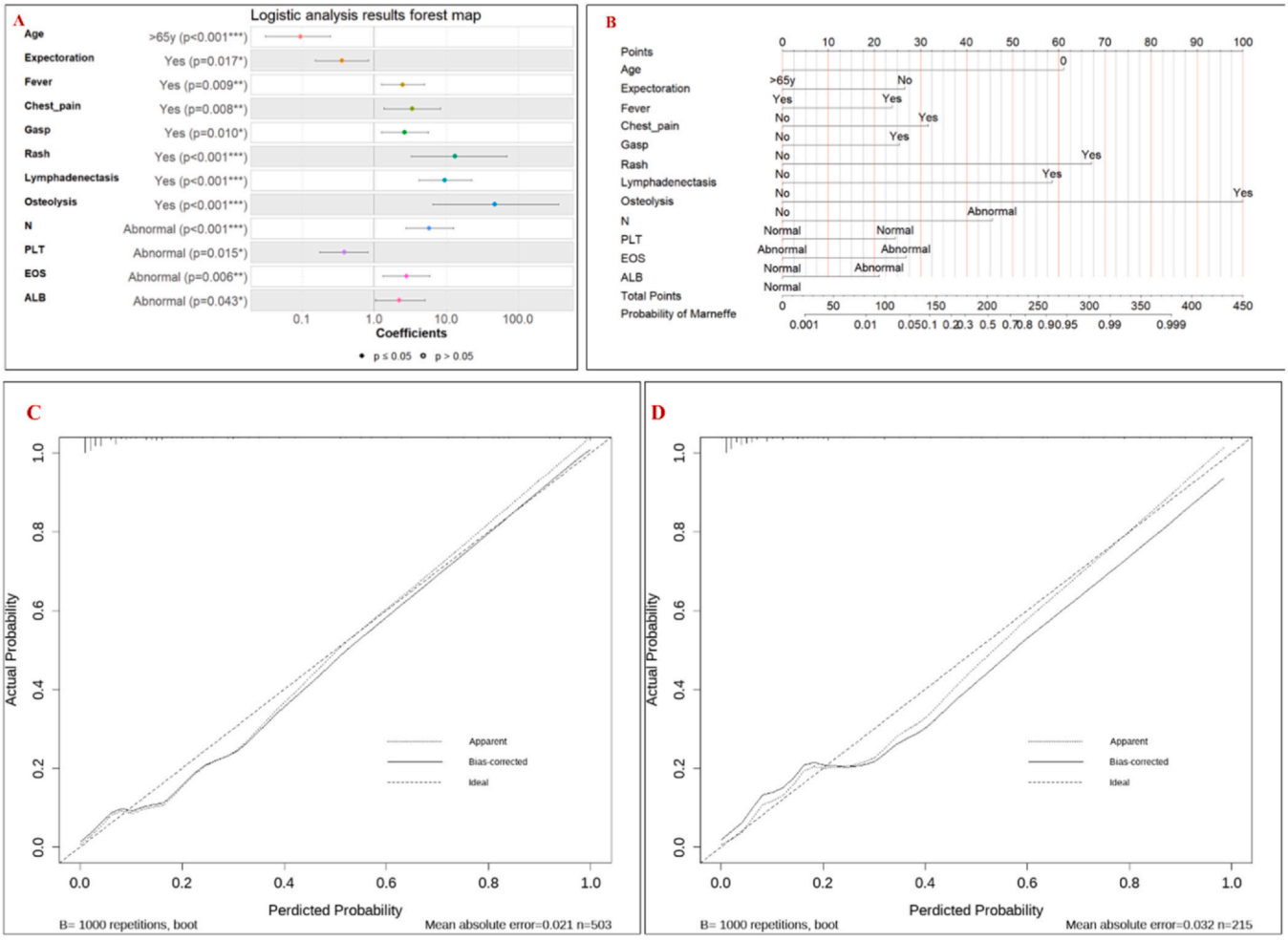

Fig. 1. (A) The variables with statistical significance in the stepwise backward analysis were screened, and the forest map of logistic regression analysis results was constructed. Age > 65 years, expectoration, and PLT were independent predictors for TB. Fever, chest pain, gasping, rash, lymphadenectasis, osteolysis, Neu count, EOS count, and ALB concentration were independent predictors for talaromycosis. (B) The nomogram for PTB and TSM was developed by using the multivariate logistic regression analysis models as the final diagnosis models after factor selection. The calibration curves and Hosmer−Lemeshow(H-L) test indicated that there was no significant difference between the predicted value (C) and the true value (D).

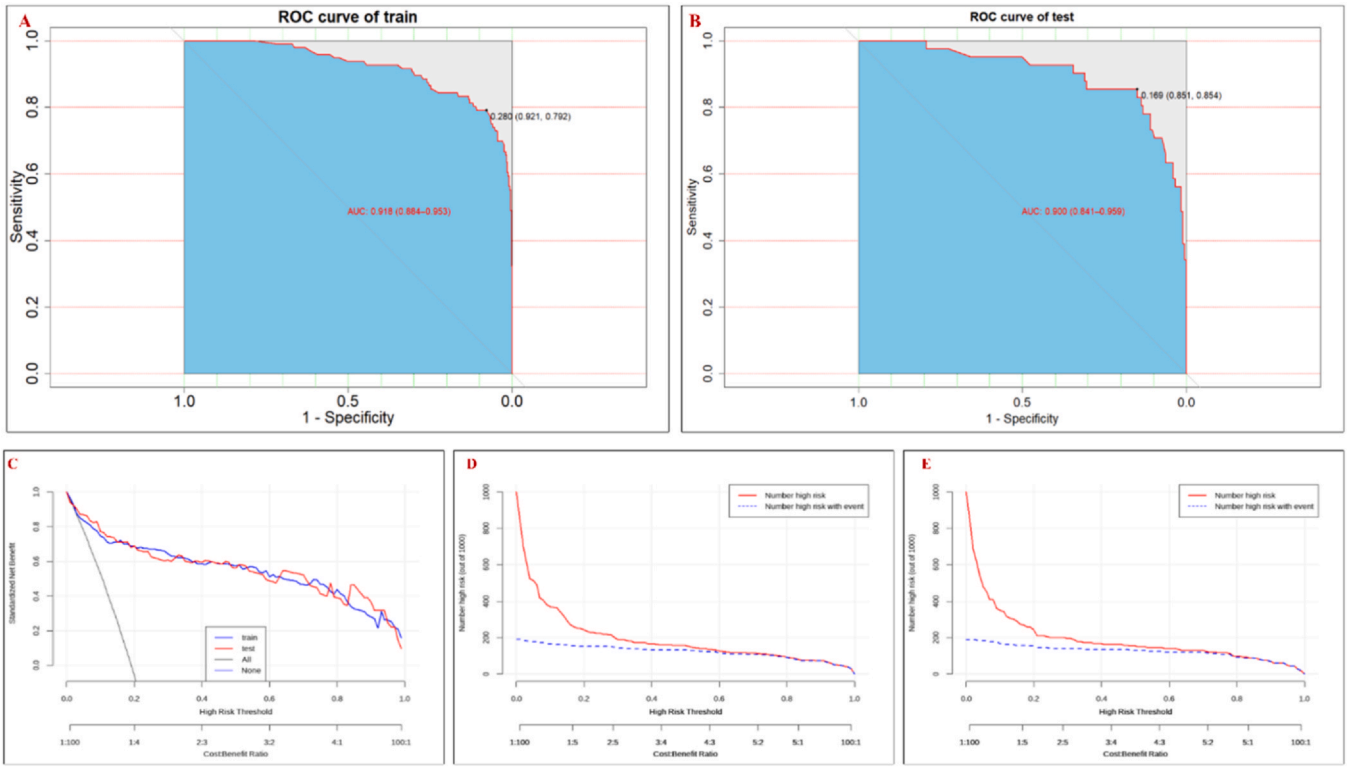

Fig. 2. ROC curve analysis of the training set showed that the AUC (95 % CI) of the clinical differential model constructed based on logistic regression analysis was 0.918 (0.884–0.953). (B) ROC curve analysis of the validation set showed that the AUC (95 % CI) was 0.900 (0.841–0.959). Decision curve analysis (DCA) was used to test the clinical efficacy of the model, and a clinical impact curve was generated. The red curve indicates the number of people who are diagnosed with talaromycosis by the simple model under each threshold probability (C). The blue curve shows the number of true talaromycosis patients at each threshold probability. The curve of the constructed model is significantly higher than the baseline in both the test set (D) and the validation set (E), which means that the model can provide better net benefits in most threshold ranges.

talaromycosis, especially in HIV-negative patients [21,22], and chest pain, gasping, and rash are also common [23]. The calibration curve, serving as a graphical representation of the outcomes from the H-L goodness-of-fit test, offers valuable insights into the evaluation of the multivariate logistic regression analysis. This visualization allows for an intuitive examination of the correlation between the predicted probability and the true probability. Notably, the observed outcome suggests a lack of significant disparity between the predicted values and the actual values, as discerned from the graphical representation. The insights from the DCA underscored the model's capacity to accrue superior net benefits across a spectrum of threshold ranges [24].

The meticulous validation of clinical prediction models is a pivotal step in their development. Evaluation within the internal validation cohort revealed no compromise in model calibration, suggesting a lack of overfitting. However, a critical step in this process involves external validation across diverse patient populations, a crucial precursor to widespread model adoption. Remarkably, these models demonstrated commendable performance in the geographically distinct Ontario cohort, even in the face of an anticipated decline in discriminative ability (C statistic, 0.918; 95 % CI 0.884–0.953). Traditionally, decision-making models with a C statistic surpassing 0.70 and strength exceeding 0.80, have clinical usefulness. In this context, the model exhibits promise in supporting clinical decision-making, given its robust C statistic [24,25]. Furthermore, the net reclassification improvement observed in the full model, when contrasted with simpler models, underscores the enhanced accuracy achievable in decision-making for follow-up. This finding emphasizes the potential of employing this comprehensive model to refine decision accuracy, surpassing outcomes based solely on fewer variables.

The host immune response and TM interaction is framed by the Damage-Response Framework (DRF). The DRF states that talaromycosis results from the host-fungus interplay, with the host's immune status dictating the outcome. The fungus's virulence, including its phase transition and melanin production, combined with the host's immune response, particularly macrophage polarization (M1 vs. M2), determines whether the infection is controlled or progresses. The DRF underscores the need for a balanced immune response to manage infection and minimize host damage [26].

The Mp1p antigen assay enables rapid detection within 6 hours, particularly in urine samples from immunocompromised hosts [27,28]. The 4D1-GNA monoclonal antibody offers species-specific identification via immunochromatography, avoiding cross-reactivity with Aspergillus and outperforming traditional GM testing [29]. While host surrogate markers like 1,3-β-D-glucan (BDG) lack TM specificity, their elevation in disseminated cases supports early empiric therapy when combined with clinical suspicion [30]. These methods complement rather than replace culture/molecular diagnostics, and we now emphasize their role in resource-limited settings where delays in fungal identification worsen mortality.

A notable strength of this investigation lies in the deliberate incorporation of the predominant symptoms associated with talaromycosis and TB in routine clinical scenarios, where their similarities often lead to a heightened probability of misdiagnosis. By amalgamating these symptoms, we conducted a thorough exploration to discern the distinctive indicators for talaromycosis and TB. Given the relatively common occurrence of TB, this comprehensive approach aims to increase the precision of initial clinical diagnoses, thereby mitigating the likelihood of unnecessary treatments.

While our study contributes valuable insights, it is imperative to acknowledge certain limitations. First, the scope of patient

enrollment was confined to individuals from Guangdong and Guangxi in China, representing a regional focus that may not be fully representative of the broader geographic landscape. Additionally, our study was limited by its narrow racial perspective, as it exclusively included Asian individuals. Furthermore, the presence of some missing values within the dataset potentially influenced the accuracy and comprehensiveness of our results. Acknowledging these constraints is essential for contextualizing the generalizability and robustness of our findings.

Conclusions

A multivariate model employing routine laboratory data and clinical features demonstrated proficiency in differentiating between talaromycosis and TB. However, the integration of this model into routine clinical care necessitates additional investigation to ascertain its utility and robustness. Further research is imperative to gauge its effectiveness and potential impact in diverse clinical settings, fostering a more comprehensive understanding of its applicability in enhancing diagnostic precision and patient care.

Ethics statement

This multicenter retrospective study (Chinese Clinical Trial Registry, ChiCTR2000029306) was conducted at the First Affiliated Hospital of Guangzhou Medical University, the First Affiliated Hospital of Guangxi Medical University, the Guangxi Medical University Cancer Hospital, and the Fourth People's Hospital of Nanning, China, between February 2012 and October 2023. The Ethics Review Boards of the First Affiliated Hospital of Guangxi Medical University (reference number 2018. KY-E-094) and the First Affiliated Hospital of Guangzhou Medical University (reference number 2019026) approved the study, which was conducted in accordance with Good Clinical Practice standards and the Declaration of Helsinki.

Authors' contributions

All authors fulfilled the contribution requirements as per the International Committee of Medical Journal Editors guidelines regarding the roles of authors and contributors. Qiu Y, Li ZT and Yang SX analyzed all the data and wrote the first draft of the manuscript. Zhang JQ and Ye F conceived and designed the study, and modified the finally manuscript. Chen WX, Zhang Y, Kong QY, Chen LR, Huang J, Lin L, Xie K, Zeng Wen, Li SQ, Zhan YQ and Wang Y contributed to the collection of data from the electronic medical records. All authors contributed to the data acquisition, data analysis, or data interpretation and reviewed and approved the final version of the manuscript.

Authors' contributions

All authors fulfilled the contribution requirements as per the International Committee of Medical Journal Editors guidelines regarding the roles of authors and contributors. Qiu Y, Li ZT and Yang SX analyzed all the data and wrote the first draft of the manuscript. Zhang JQ and Ye F conceived and designed the study, and modified the finally manuscript. Chen WX, Zhang Y, Kong QY, Chen LR, Huang J, Lin L, Xie K, Zeng Wen, Li SQ, Zhan YQ and Wang Y contributed to the collection of data from the electronic medical records. All authors contributed to the data acquisition, data analysis, or data interpretation and reviewed and approved the final version of the manuscript.

Consent for publication

Written informed consent for publication was obtained from all participants.

Clinical trial number

Not applicable.

Funding

This study was supported by grants from the National Natural Science Foundation of China (grant numbers NSFC 82202544 and NSFC 82270007), the Guangxi Natural Science Foundation (No. 2021GXNSFBA220064), the China Postdoctoral Science Foundation (No. 2023M730801), the State Key Laboratory of Respiratory Diseases (SKLRD-Z-202019), The Health Commission of Guangxi Zhuang Autonomous Region Natural Science Project (Z20211324) and Noncommunicable Chronic Diseases-National Science and Technology Major Project Fund Code: 2023ZD0506200; 2023ZD0506204.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Acknowledgment

We would like to extend sincere gratitude to the nurses and clinical staff who provided care for the patients, which provided us with a platform to conduct research into the respiratory system. Furthermore, we would also like to thank the AJE team for English language editing.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jiph.2025.102740.

References

[1] He L, et al. Talaromyces marneffei infection in non-HIV-infected patients in mainland China. Mycoses 2021;64(10):1170–6.
[2] Wang F, Han R, Chen S. An overlooked and underrated endemic mycosis-talaromycosis and the pathogenic fungus talaromyces marneffei. Clin Microbiol Rev 2023;36(1):e0005122.
[3] Thompson GR, 3rd, et al. Global guideline for the diagnosis and management of the endemic mycoses: an initiative of the European Confederation of Medical Mycology in cooperation with the International Society for Human and Animal Mycology. Lancet Infect Dis 2021;21(12):e364–74.
[4] Narayanasamy S, et al. A global call for talaromycosis to be recognised as a neglected tropical disease. Lancet Glob Health 2021;9(11):e1618–22.
[5] Jiang J, et al. Effects of Talaromyces marneffei infection on mortality of HIV/AIDS patients in southern China: a retrospective cohort study. Clin Microbiol Infect 2019;25(2):233–41.
[6] Cao C, Xi L, Chaturvedi V. Talaromycosis (Penicilliosis) due to talaromyces (Penicillium) marneffei: insights into the clinical trends of a major fungal disease 60 years after the discovery of the pathogen. Mycopathologia 2019;184(6):709–20.
[7] Qiu Y, et al. Immunodeficiency disease spectrum in HIV-negative individuals with talaromycosis. J Clin Immunol 2021;41(1):221–3.
[8] Qiu Y, et al. Determinants of prognosis in Talaromyces marneffei infections with respiratory system lesions. Chin Med J (Engl) 2019;132(16):1909–18.
[9] Liu W, Li RY. [Enlightenment of World Health Organization fungal priority pathogens list]. Zhonghua Liu Xing Bing Xue Za Zhi 2023;44(12):1984–7.
[10] van Smeden M, et al. Clinical prediction models: diagnosis versus prognosis. J Clin Epidemiol 2021;132:142–5.
[11] Ranstam J, Cook JA, Collins GS. Clinical prediction models. Br J Surg 2016;103(13):1886.

[12] Qi W, et al. Predictive models for predicting the risk of maternal postpartum depression: a systematic review and evaluation. J Affect Disord 2023;333:107–20.
[13] Liu J, et al. Nomogram predicting overall prognosis for invasive micropapillary carcinoma of the breast: a SEER-based population study. BMJ Open 2023;13(8):e072632.
[14] Collins GS, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Bmj 2015;350:g7594.
[15] WHO Guidelines Approved by the Guidelines Review Committee, in WHO consolidated guidelines on tuberculosis: Module 3: diagnosis – rapid diagnostics for tuberculosis detection. 2021, World Health Organization © World Health Organization 2021.: Geneva.
[16] Cross GB, et al. Rosuvastatin adjunctive therapy for rifampicin-susceptible pulmonary tuberculosis: a phase 2b, randomised, open-label, multicentre trial. Lancet Infect Dis 2023;23(7):847–55.
[17] Sterling TR, et al. Guidelines for the treatment of latent Tuberculosis infection: recommendations from the National Tuberculosis controllers association and CDC, 2020. MMWR Recomm Rep 2020;69(1):1–11.
[18] Yi S, Wang L. Clinical features of tuberculous pseudoaneurysm and risk factors for mortality. J Vasc Surg 2022;75(5):1729–1738.e2.
[19] Peng L, et al. Clinical features of patients with talaromycosis marneffei and microbiological characteristics of the causative strains. J Clin Lab Anal 2022;36(11):e24737.
[20] Pan M, et al. Assessment of talaromyces marneffei infection of the intestine in three patients and a systematic review of case reports. Open Forum Infect Dis 2020;7(6):ofaa128.
[21] Qiu Y, et al. Retrospective analysis of 14 cases of disseminated Penicillium marneffei infection with osteolytic lesions. BMC Infect Dis 2015;15:47.
[22] Wei HY, et al. Clinical characteristics and risk factors of Talaromyces marneffei infection in human immunodeficiency virus-negative patients: a retrospective observational study. World J Emerg Med 2021;12(4):281–6.
[23] Qiu Y, et al. Comparison of the clinical features of HIV-positive and HIV-negative hosts infected with Talaromyces marneffei: a multicenter, retrospective study. Int J Infect Dis 2023;132:93–8.
[24] Vickers AJ, Holland F. Decision curve analysis to evaluate the clinical benefit of prediction models. Spine J 2021;21(10):1643–8.
[25] Janssens A, Martens FK. Reflection on modern methods: revisiting the area under the ROC Curve. Int J Epidemiol 2020;49(4):1397–403.
[26] Pruksaphon K, et al. The microbial damage and host response framework: lesson learned from pathogenic survival trajectories and immunoinflammatory responses of Talaromyces marneffei infection. Front Immunol 2024;15:1448729.
[27] Lau SKP, Tsang CC, Woo PCY. Talaromyces marneffei genomic, transcriptomic, proteomic and metabolomic studies reveal mechanisms for environmental adaptations and virulence. Toxins 2017;9(6).
[28] Sze KH, et al. Talaromyces marneffei Mp1p Is a virulence factor that binds and sequesters a key proinflammatory lipid to dampen host innate immune response. Cell Chem Biol 2017;24(2):182–94.
[29] Shu F, et al. Evaluation of the yeast phase-specific monoclonal antibody 4D1 and Galanthus nivalis agglutinin sandwich ELISA to detect Talaromyces marneffei antigen in human urine. Front Cell Infect Microbiol 2023;13:1163868.
[30] Sun J, et al. Clinical characteristics and risk factors for poor prognosis among HIV patients with Talaromyces marneffei bloodstream infection. BMC Infect Dis 2021;21(1):514.