Multimodal machine learning-based model for differentiating nontuberculous mycobacteria from *mycobacterium tuberculosis*

Hong-ling Li*, Ri-zeng Zhi, Hua-sheng Liu, Mei Wang and Si-jie Yu

Department of Infectious Diseases, Zhoushan Hospital, Wenzhou Medical University, Zhoushan, Zhejiang, China

Objective: To develop and evaluate the effectiveness of multimodal machine learning approach for the differentiation of NTM from MTB.

Methods: The clinical data and CT images of 175 patients were retrospectively obtained. We established clinical data-based model, radiomics-based model, and multimodal (clinical plus radiomics) model gradually using 5 machine learning algorithms (Logistic, XGBoost, AdaBoost, RandomForest, and LightGBM). Optimal algorithm in each model was selected after evaluating the differentiation performance both in training and validation sets. The model performance was further verified using external new MTB and NTM patient data. Performance was also compared with the existing approaches and model.

Results: The clinical data-based model contained age, gender, and IL-6, and the RandomForest algorithm achieved the optimal learning model. Two key radiomics features of CT images were identified and then used to establish the radiomics model, finding that model from Logistic algorithm was the optimal. The multimodal model contained age, IL-6, and the 2 radiomics features, and the optimal model was from LightGBM algorithm. The optimal multimodal model had the highest AUC value, accuracy, sensitivity, and negative predictive value compared with the optimal clinical or radiomics models, and its' favorable performance was also verified in the external test dataset (accuracy = 0.745, sensitivity = 0.900). Additionally, the performance of multimodal model was better than that of the radiologist, NGS detection, and existing machine learning model, with an increased accuracy of 26, 4, and 6%, respectively.

Conclusion: This is the first study to establish multimodal model to distinguish NTM from MTB and it performs well in differentiating them, which has the potential to aid clinical decision-making for experienced radiologists.

1 Introduction

Tuberculosis is a chronic infectious disease caused by *mycobacterium tuberculosis* (MTB). MTB can invade all organs of the body, especially lung. Tuberculosis caused by MTB is one of the most serious public health problems in the world (1). The tuberculosis burden in China is only lower than in India and Indonesia, ranking third (2). With the increasing number of patients with human immunodeficiency virus (HIV) year by year and the massive use of immunosuppressants, the incidence of opportunistic infections and disease burden caused by non-tuberculous mycobacteria (NTM) are rising globally (3, 4). NTM is a major cause of morbidity and mortality in progressive lung diseases. However, the clinical manifestations of MTB and NTM are similar in symptoms such as low fever, cough, and reduced body weight (5), making it difficult to distinguish them. Therefore, choosing a fast, accurate, and clinically applicable method for distinguishing NTM and MTB is of great significance.

The bacteria culture is the "gold" reference standard and takes 2 to 6 weeks to produce diagnostic results. However, due to the different treatment plans between the two types of diseases, NTM patients will miss the best treatment opportunity due to the long time consumption, leading to disease progression (6). In addition, bacterial culture shows low sensitivity. Other methods (7) including GeneXpert MTB/RIF Ultra, chest X-ray, and tuberculosis loop-mediated isothermal amplification fail to distinguish between NTM and MTB. Recently, next-generation sequencing (NGS) based technology has been successfully applied for the routine characterization of tuberculosis (8, 9). Therefore, it is extremely necessary to combine multiple methods for the differentiation between NTM and MTB.

Although chest computed tomography (CT) examination also faced a similar predicament of difficulty in identification, there are still subtle differences in imaging performance between the two diseases (10). This indicates that CT signs may play an important role in disease differentiation and it is necessary to mine more useful information behind CT images. In recent years, radiomics has shown great potential in the diagnosis and differential diagnosis of lung diseases through high-throughput extraction and mining of data features (11, 12), which may provide a feasible method for distinguishing NTM from MTB. Radiomics is a non-invasive and objective image analysis tool, that uses computer algorithms to mine deep information in images such as CT, Magnetic Resonance Imaging (MRI), and positron emission tomography (PET), thereby reflecting the heterogeneity of the lesion area. However, the discriminating model based on these imaging pictures showed different differentiation performances. For example, the previous study quantifies bronchiectasis regions in CT images and explored a machine learning approach, finding that the model achieved an area under the curve (AUC) of 0.84 and an accuracy of 0.85 (13). However, the other research reported an accuracy of 0.74 (14). It followed that the performance difference between different machine learning models was the key point when applying these models for distinguishing lung diseases.

Machine learning (ML) is a technique, which can automatically extract useful models from large-scale heterogeneous datasets based on complex algorithms. These models can then be utilized for outcome prediction (15). ML has enhanced the integration of computer science and statistics with medical problems, and it is now extensively employed in disease diagnosis, cancer treatment, and other medical research areas (16–18). In tuberculosis research, ML is also widely applied for the diagnosis, treatment and differential diagnosis of tuberculosis. Yao et al. (19) combined plasma proteins with ML to establish seven models for the diagnosis of active tuberculosis. Among them, the support vector machine (SVM) model demonstrated the best performance, achieving an AUC exceeding 0.89. In addition, ML in conjunction with histological information can diagnose latent tuberculosis (20). Research has shown that the ML fusion model based on longitudinal CT scan image histology performs well in predicting the poor prognosis of TB treatment, with internally validated AUC and externally validated AUC of 0.767–0.802 and 0.831–0.857, respectively, enabling early preventive measures against unfavorable prognoses (21). Although some researchers have also built models to differentiate between NTM and MTB by combining urinary metabolomics or CT imaging information with ML (22, 23), these models are limited in their clinical interpretability due to the only utilization of single laboratory or CT imaging parameter. Currently, limited researches performed machine learning to distinguish NTM from MTB, especially from aspect of multidimensions (clinical characteristics, laboratory test, CT/MRI images, etc.).

In this study, we conducted five machine learning algorithms and established an optimal multimodal model containing clinical, laboratory test, and radiomics data of CT images for distinguishing NTM from MTB. We compared its differentiation performance with single clinical or radiomics based models. We also verified its differentiation performance in the external new dataset, and then compared its' differentiation ability model with the existing approaches and machine learning model. The contributions of our study were as follows: (1) at present, few study was performed for distinguishing them using machine learning models. Our study aimed to establish a machine learning model for differentiating NTM from MTB, which can promote the application of new technologies, thereby advancing research progress in this fields. (2) Currently, most of studies were based on the one dimension, such as only CT or X-ray images. This is the first study to consider multidimensions (clinical characteristics, laboratory test, CT images) and use multimodal model to distinguish them. Our research provides a new perspective and strategy for differentiating NTM from MTB, which is of great significance for doctors to choose appropriate treatment plans.

2 Methods

2.1 Data source

This study retrospectively enrolled patients with pulmonary infection who were admitted to our hospital between April 2020 and December 2023. All the patients underwent a CT examination.

The diagnostic criteria were presented as follows. NTM was diagnosed according to the Euro-American 2020 edition (24). MTB was diagnosed according to the rapid tuberculosis diagnostic criteria with 2023 World Health Organization edition (25). To differentiate the NTM and MTB, we conducted the T-SPOT test using a testing kit produced by DEAOU (Guangzhou) according to the kit instructions. Next-generation sequencing (NGS) of alveolar lavage fluid and bacteria culture were also performed to differentiate the NTM and

MTB. The collection of alveolar lavage fluid and next-generation sequencing information including sample processing and DNA extraction, and library generation and sequencing can be found in the previous study (26). Integrating the results of multiple tests, the NTM and MTB patients were classified.

The inclusion criteria included (1) age ranged from 18 to 80 years; (2) diagnosed with NTM or MTB infection; (3) had bacterial culture results; (4) had NGS test results; (5) had T-SPOT test results; (6) with at least 2 set of lung CT images available. We excluded these patients with lung cancer, fungal infection, pneumoconiosis, and mixed infections of TB and NTM. The process of patient screening and enrolling was presented in Figure 1. After screening, 99 MTB patients and 76 NTM patients were enrolled in the final analyses.

It should be noted that there was a data imbalance regarding the MTB and NTM patient number, which may cause the potential for producing overfitting problems. Due to the requirements of periodical for follow-up after diagnosis, at least 2 images of NTM/MTB patients were taken. Considering the slow rate of radiological abnormal changes in NTM, longer follow-up time was required. Hence, more CT images of one NTM patient were collected than MTB. Therefore,

although there was unbalance in the proportion of patients, CT images were as balanced as possible.

2.2 Machine learning (ML) approach

This study aimed to construct 3 models (clinical model, radiomics model, multimodal model) for differentiating NTM from MTB using 5 machine learning algorithms (Logistic, XGBoost, AdaBoost, RandomForest, and LightGBM), and to explore which algorithms was more suitable. The purpose of classified multi-model method was to select the best model, rather than directly modeling to get the final model. In this study, we used the training/verification mechanism of 5-fold cross-validation to summarize the performance of each model in many trainings, focusing on the overall performance of each model. The overall workflow of the models' development, validation, and comparison was presented in Figure 2.

Logistic regression models are more traditional single-model classification algorithms that aim to identify the connection between features and the likelihood of a specific (binary) outcome. It is widely
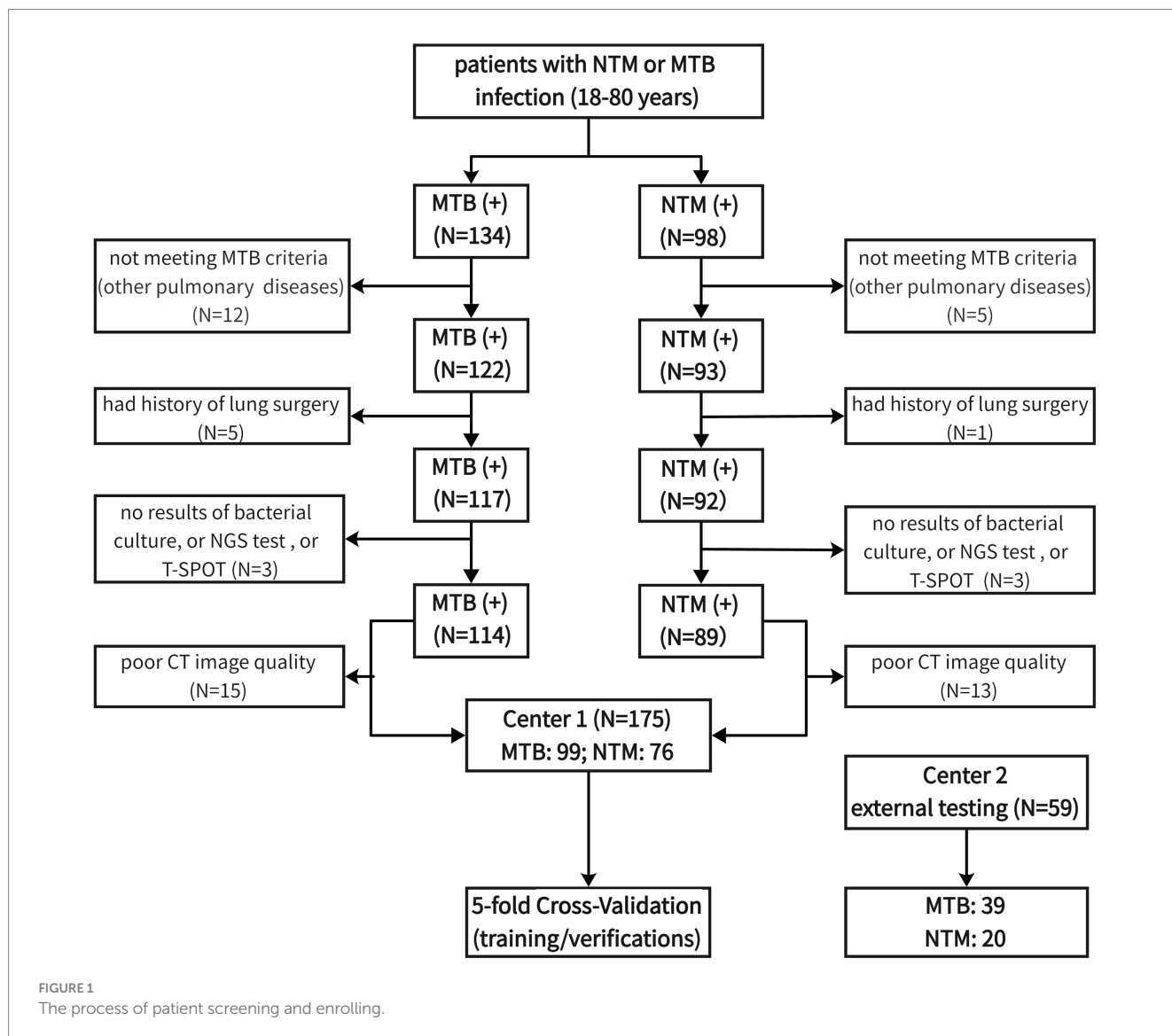

FIGURE 1
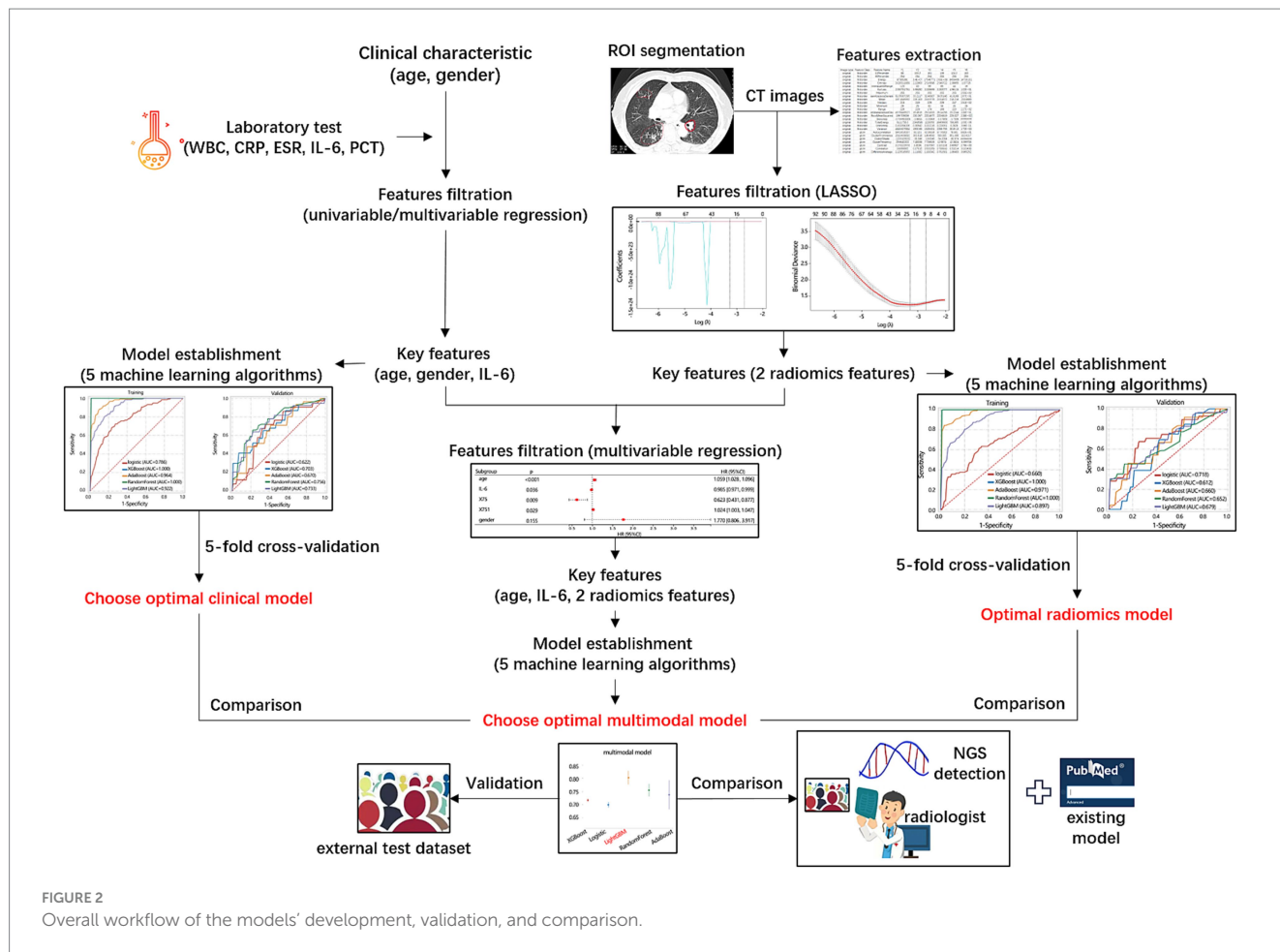The process of patient screening and enrolling.

FIGURE 2
Overall workflow of the models' development, validation, and comparison.

employed by medical professionals for its ability to calculate odds ratio (27).

Models such as RandomForest, XGBoost, LightGBM, and AdaBoost are ensemble models based on decision trees. These ensemble models can get a strong model through taking the strengths of all single models. This strong classifier achieves a relative best performance.

RandomForest algorithm is the embodiment of group intelligence, which creates different training sets by randomly sampling rows (bagging) and columns (feature bagging) from the dataset. The decision tree of RandomForest algorithm is grown by bootstrap. RandomForest reduces training variance and improves model generalization and integration. And it can be utilized without parameter tuning, offering variable importance information for classification and high predictive accuracy (28).

Conversely, XGBoost, LightGBM, and AdaBoost are based on the idea of gradient boosting. XGBoost performs a second-order Taylor expansion of the loss function and employs various techniques to minimize overfitting. It utilizes the structure score (gain) of the tree to determine split points, improving tree quality, and implements parallel and distributed computation for increased efficiency. For detailed algorithm information, consult the literature (29).

The LightGBM algorithm employs a histogram approximation algorithm to generate a histogram that discretises continuous features. It splits trees at leaf nodes with the largest lift. And using the GOSS

technique, the samples with larger gradient are preferred, so the optimal split can be found faster and the training efficiency is improved (30, 31).

AdaBoost is a relatively new nonlinear ML algorithm that builds a tree by adjusting sample weights and combines multiple trees to become a strong classifier. It does not require feature screening, can perform automatic feature selection, and has a low risk of overfitting. For specific steps, refer to the relevant literature (32).

2.3 Key clinical features identification

We collected the clinical data of patients including the age (years), gender, white blood cell (WBC, 10^9/L), erythrocyte sedimentation rate (ESR, mm/h), C-reactive protein (CRP, mg/L), Interleukin-6 (IL-6, pg./mL), and procalcitonin (PCT, ng/mL). The difference in these variables between NTM and MTB groups was first compared to identify the differential variables. These differential variables were then enrolled in Logistic regression analysis to determine the independent factor associated with the disease types. Then Receiver Operating Characteristic (ROC) analysis and area under curve (AUC) were used to assess the discriminating performance of these independent factors. Decision Curve Analysis (DCA) was used to evaluate their obtained clinical net benefit for discriminating disease.

2.4 Differentiation model construction based on the key clinical features

Based on the independent clinical factors, a clinical features-based model for discriminating NTM from MTB was first constructed and validated using 5-fold cross-validation. The cross-validation can partially resolve the overfitting problem. In addition, we added L2 penalty to control the complexity of the model to prevent it from being too complicated. We also applied early stop-ping, learning rate adjustments, and drop-out to prevent overfitting. The detailed parameters of the Logistic algorithm were as follows: $C = 1.0$; max-iter = 100; penalty = l2; tol = 0.0001. The parameters of the XGBoost algorithm were as follows: learning-rate = None; max-depth = None; min child-weight = None; reg-lambda = None. The parameters of the AdaBoost algorithm were as follows: learning-rate = 1.0; n-estimators = 50. The parameters of the RandomForest algorithm were as follows: criterion = gini; max-depth = None; min impurity-decrease = 0.0; n-estimators = 20. The parameters of the LightGBM algorithm were as follows: boosting type = gbdt; learning rate = 0.1; max depth = -1; n-estimators = 100; num-leaves = 31. The discriminating performance of models from 5 algorithms was evaluated by ROC and DCA analyses.

2.5 CT images and radiomics features extraction

The CT images of all patients were also obtained. CT scans were performed using 64-slice CT scanners with the following parameters: tube voltage 120 kV; automatic tube current modulation 300 mA; detector collimation 64 × 0.625 mm; thread pitch 0.993; section thickness 2 mm; section interval 2 mm. The scanning area ranged from the apex pulmonic to the bottom of the lung.

All CT images were assessed by two radiologists with 5 years of diagnostic experience in CT, respectively. They were blinded to the histopathological and clinical data of patients. The 2 radiologists manually segmented the region of interest (ROI) of CT images by using 3D slicer software, and the intraclass correlation coefficient (ICC) between the 2 radiologists indicated the consistency of extracted features. ICC > 8 suggested a good consistency between observers.

For feature extraction, all images and ROIs were batched into 3D slicer software. The extracted radiomics features included 162 first-order features, 216 gray-level co-occurrence matrix (GLCM) features, 126 gray-level dependence matrix (GLDM) features, 144 gray-level run-length matrix (GLRLM) features, 144 gray-level size zone matrix (GLSZM) features, and 45 neighboring gray-tone difference matrix (NGTDM) features.

2.6 Differentiation model construction based on radiomics features of CT images

A total of 837 radiomics features were extracted. We then used the Least absolute shrinkage and selection operator (LASSO) analysis to remove the redundant features. By introducing a penalty coefficient (λ), the coefficients of most features will be compressed to 0. The retained features in the final optimal LASSO model were selected for further analysis. LASSO analysis is a regularization method for regression analysis. By introducing L1 regularization into the regression model, the coefficients of some features are reduced to zero, thus realizing variable

selection. LAASO can reduce the complexity of the model, and improve the prediction performance of the model. Therefore, application of LASSO analysis can partially resolve the overfitting problem. Then, the differences of the retained features after LASSO analysis between NTM and MTB groups were compared. The differential features between the 2 groups were enrolled in the logistic regression analysis to explore the independent factors. The clinical value of independent factors in disease discrimination was assessed by ROC and DCA analysis.

Based on the independent features, a radiomics features-based differentiation model was constructed in a training set using 5 machine learning algorithms including Logistic, XGBoost, AdaBoost, RandomForest, and LightGBM. Their performance was also verified in the validation set by ROC and DCA analyses. Detailed information for the model construction and validation can be found in the Methods 2.3 section.

2.7 Multimodal differentiation model construction

The key clinical and radiomics features, that are independently related to the disease types, have been identified in the above analyses. Then these various indicators were enrolled in logistic regression analysis to further determine the independent variables. Based on the key clinical and radiomics features, a multimodal differentiation model was constructed using 5 machine learning algorithms. Their performance was also verified in the validation set by ROC and DCA analyses. Detailed information for the model construction and validation can be found in the Methods 2.3 section. Especially, the importance ranking of features within the multimodal model was explored using 3 optimal algorithms.

2.8 Differentiation performance comparison and verification

Then, the differentiation performance of the 3 models (clinical-based, radiomics-based, multimodal-based) was compared from aspects of AUC, cutoff, accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Kappa values. Although we have collected data for 4 years to establish and evaluate our models, there was limit to accurately assessing their performance due to the lack of test data, especially for NTM. Therefore, it was necessary to obtain a new external dataset from public data or data from other institutions. Unfortunately, we failed to obtain the test data from other institutions due to some reasons. Currently, no publica data on NTM was provided.

Therefore, another external dataset containing NTM and MTB cases who met the diagnostic criteria were retrospectively found. From January 2024 to December 2024, totally 59 patients (20 NTM and 39 MTB patients) were collected. These data differed from the patient data used when constructing the learning model, and were the newly generated data after performing learning model. Only one image was collected per patient in the new external dataset. We also validated the performance of our models for differentiating NTM and MTB using the new external test dataset. The differentiation performances of 3 models were also compared in the new external test dataset.

In addition, we compared the differentiation performance of multimodal model in the new external test dataset with other measures including NGS detection method, radiologist assessment

TABLE 1 The baseline characteristics of patients grouped by infection types.

		MTB (*n* = 99)	NTM (*n* = 76)	*p*
Age (years)		57 [38, 69]	66 [59, 72]	<0.001
WBC (10⁹/L)		5.9 [4.8, 7.3]	5.4 [4.5, 6.8]	0.109
CRP (mg/L)		5.0 [1.7, 24.7]	4.4 [1.3, 15.9]	0.323
ESR (mm/h)		32 [15, 62]	27 [12, 57]	0.779
IL-6 (pg/mL)		39.5 [16.0, 60.8]	15.6 [2.5, 45.9]	<0.001
Gender	Male	64 (64.646)	31 (40.789)	0.002
	Female	35 (35.354)	45 (59.211)	
PCT (ng/mL)	<0.01	7 (7.071)	10 (13.158)	0.178
	≥0.01	92 (92.929)	66 (86.842)	

WBC, white blood cell; ESR, erythrocyte sedimentation rate; CRP, C-reactive protein; PCT, procalcitonin.

with 5-year experiences, and existing machine learning model. In terms of radiologist assessment, one radiologist identified the pathogen that infected the patients in the testing dataset by only scanning CT images without any other reminding. The radiologist was informed that the patients were only infected by one type of pathogen, either NTM or MTB. The existing machine learning model was searched from the published articles. Currently, few studies reported the machine learning model for the differentiation of NTM from MTB. Finally, a deep learning model based on CT images by Wang et al. was selected for the performance comparison (33). For the differentiation performance comparison among different methods, AUC, accuracy, sensitivity, specificity, MTB-precision, and NTM-precision were set as the indicators.

2.9 Statistical analysis

The continuous variables were presented as median and quartiles as they did not confirm to the normal distribution, and their differences between the 2 groups were compared with the Mann–Whitney U test. The categorical variables were expressed as frequency and percent, and their distribution difference between the 2 groups was analyzed by χ^2 test. Univariable and multivariable logistic regression analyses were used to explore the association between variables and disease types. ROC and DCA were used to assess the discriminating performance. $p < 0.05$ was considered statistical significance.

3 Results

3.1 Characteristics of patients

In this study, we enrolled 175 patients infected by NMT or MTB to our analysis the median age was 63 [51, 70] years old, with average age of 58 years old. The median white blood cell was 5.6 [4.7, 7.2] (10⁹/L). The median CRP was 4.4 [1.3, 21.0] (mg/L). The median ESR was 32.0 [12.5, 59.5] (mg/L). The median IL-6 was 26.65 [4.38, 56.22] (pg/mL). Male patients accounted for 54% of all patients. Of the 76 patients diagnosed with NTM, 35 patients (46.1%), 23 patients (30.2%), 15 patients (19.7%), and 3 patients (4.0%) were infected by *M. avium*, *M. intracellulare*, *M. abscessus*, and *M. kansasii* strains of NTM, respectively.

The baseline characteristics of MTB (*N* = 99) and NTM (*N* = 76) patients are presented in Table 1. It followed that age ($p < 0.001$), gender ($p = 0.002$), and IL-6 level ($p < 0.001$) showed significant differences between MTB and NTM groups. NTM group had higher median age (66 vs. 57), but lower WBC (5.4 vs. 5.9), CRP (4.4 vs. 5.0), ESR (27 vs. 32), and IL-6 (15.6 vs. 39.5) levels than MTB group.

3.2 Clinical feature-based differentiation model

Based on 3 significant baseline characteristics, we next explored their association with the disease types. The univariable logistic regression presented their significant association (Figure 3A, all $p < 0.05$), and the multivariable regression analysis further displayed their independent association (Figure 3B, all $p < 0.05$). Their differential performance of TB and NTM was then assessed by ROC analysis (Figure 3C), finding that age had the highest AUC value (0.672), followed by IL-6 (0.667). DCA analysis showed that the 3 clinical features obtained similar clinical net benefits for discriminating TB and NTM (Figure 3D).

Subsequently, we aimed to construct a clinical feature-based differentiation model. Before model construction, all patients were assigned to training and validation sets. The 3 clinical feature-based differentiation models were first constructed in the training set using 5 deep learning algorithms, and then the differentiation performance of the model was verified in the validation set. The results showed that XGBoost and RandomForest had the most favorable differentiation performance in the training set (Figure 4A). In the validation set, the RandomForest algorithm achieved the highest AUC value (Figure 4B). Integrating the results of the training set and validation set, XGBoost algorithm may show the overfitting and RandomForest algorithm had more stable differentiation performance. Hence, 3 clinical feature-based models by the RandomForest algorithm were regarded as the optimal deep learning model.

3.3 Radiomics features-based differentiation model

Besides the key clinical features, the CT images can also provide some valuable information for disease differentiation. To reveal the potential value of CT images, we explored significant radiomics
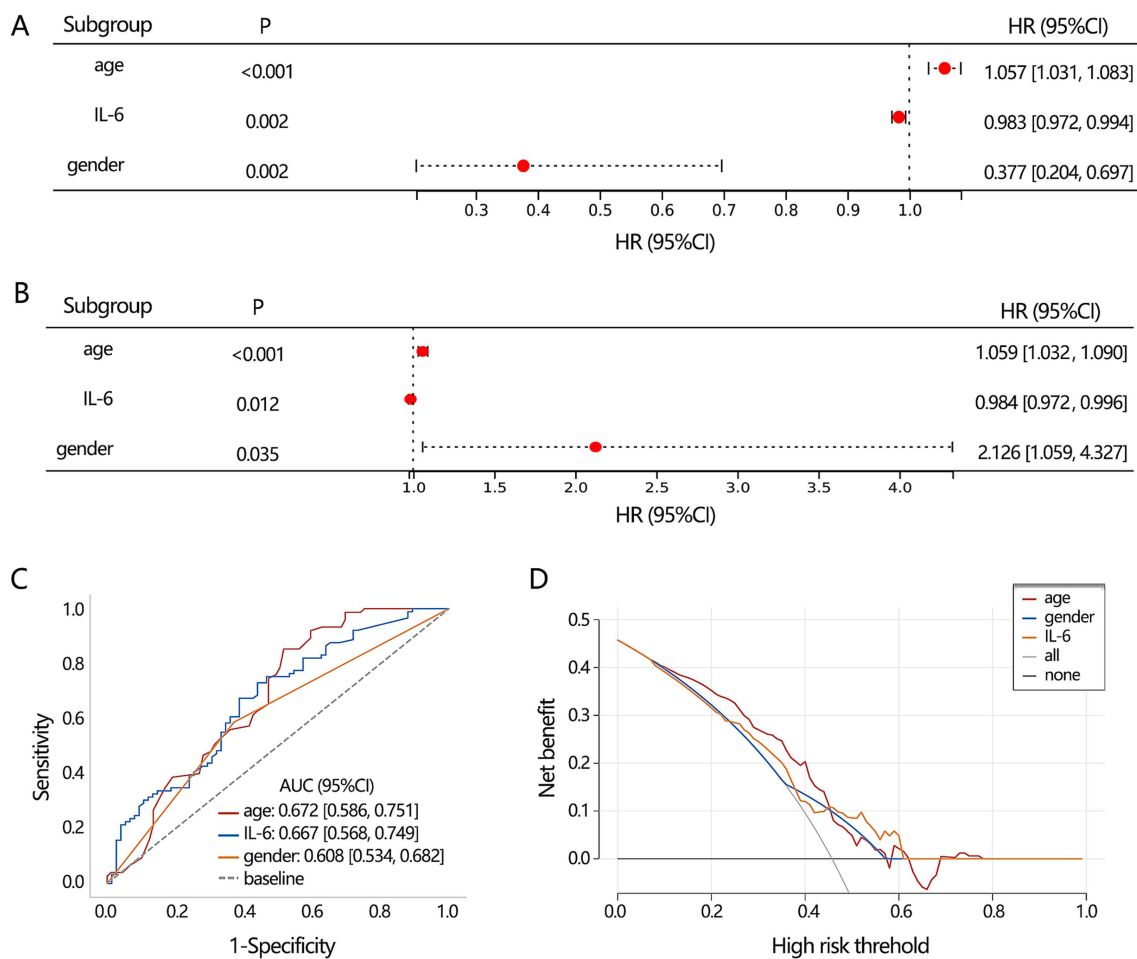
FIGURE 3
The differentiation value assessment on the significant clinical indicators. **(A)** Univariable and **(B)** multivariable logistic regression analyses. **(C)** ROC analysis was conducted to assess the discriminating performance of TB and NTM. **(D)** DCA analysis was used to assess the clinical net benefit.
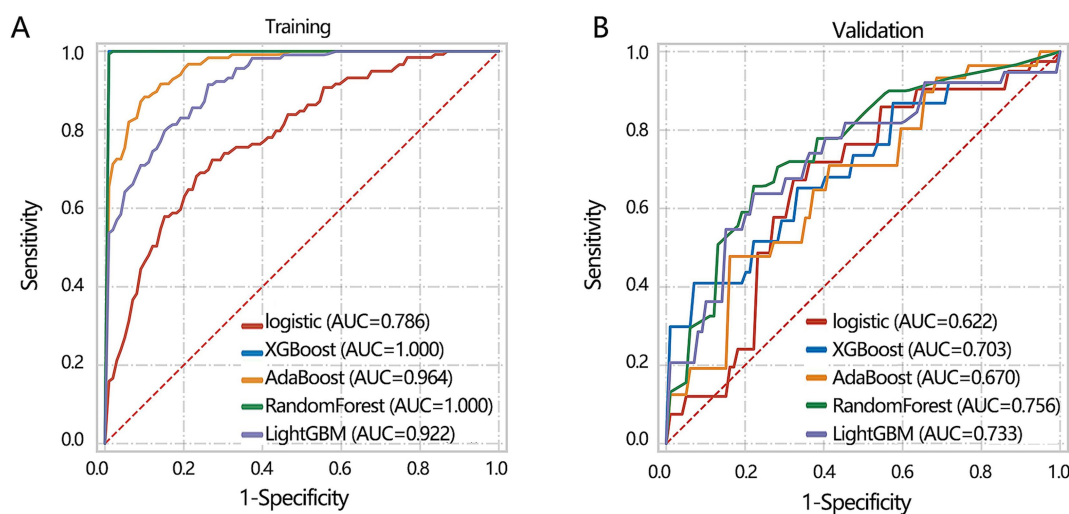
FIGURE 4
The machine learning model construction and validation based on 3 clinical indicators including age, IL-6, and gender. **(A)** Training set. **(B)** Validation set.
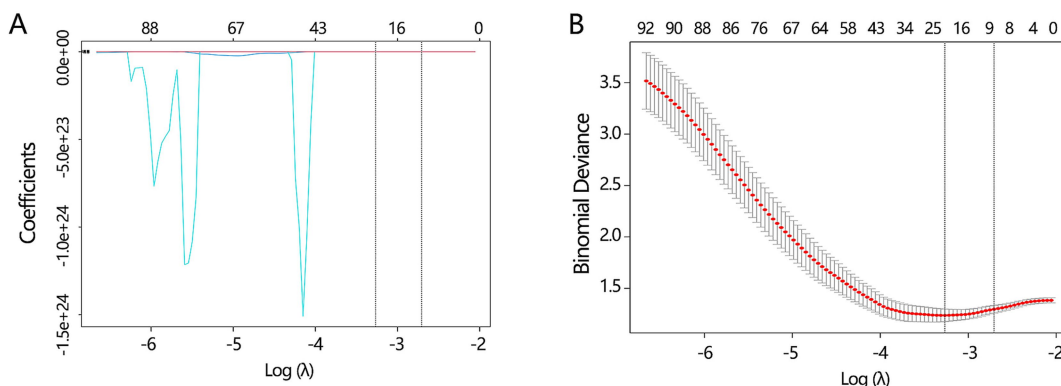
FIGURE 5
The key radiomics features selection. **(A)** The LASSO analysis was conducted to filter the redundant features among all radiomics features. **(B)** The optimal LASSO model was obtained when the standard error of minimum distance presented the $\lambda = 0.067$.

TABLE 2 The differences of 7 radiomics features between 2 groups.

Features	Image type	Feature class	Feature name	TB ($n = 99$)	NTM ($n = 76$)	p
X75	Original	GLSZM	Gray Level Variance	5.617 ± 1.260	5.047 ± 1.300	0.007
X210	wavelet-HLL	GLCM	Correlation	0.121 [0.066, 0.195]	0.200 [0.122, 0.282]	<0.001
X387	wavelet-HLH	FIRSTORDER	Skewness	−0.182 [−0.390, −0.014]	−0.303 [−0.466, −0.124]	0.014
X539	wavelet-HLH	GLSZM	GrayLevelNonUniformityNormalized	0.516 [0.504, 0.541]	0.523 [0.510, 0.565]	0.046
X640	wavelet-HLH	GLSZM	SizeZoneNonUniformityNormalized	0.460 [0.424, 0.561]	0.430 [0.394, 0.505]	0.006
X709	wavelet-HLH	GLSZM	GrayLevelNonUniformityNormalized	0.500 [0.500, 0.500]	0.500 [0.500, 0.500]	0.097
X751	wavelet-LLL	FIRSTORDER	Maximum	745.900 [735.390, 765.304]	755.041 [744.533, 767.961]	0.022

TABLE 3 The association of related radiomics features with disease type.

	Univariable		Multivariable	
Features	OR (95%CI)	p	OR (95%CI)	p
X75	0.704 [0.542, 0.915]	0.009	0.636 [0.446, 0.886]	0.009
X210	168.781 [6.789, 4195.777]	0.002	114.373 [2.273, 7595.81]	0.021
X387	0.357 [0.142, 0.899]	0.029	0.499 [0.180, 1.169]	0.137
X539	11.532 [9.205, 14.447]	0.015	6.157 [0.138, 13.969]	0.113
X640	0.013 [0.001, 0.250]	0.004	0.048 [0.001, 1.670]	0.110
X751	1.019 [1.003, 1.035]	0.020	1.035 [1.014, 1.060]	0.003

The features information can be found in Table 2.

features contributing to disease differentiation behind the CT images. A total of 837 radiomics features were extracted. To identify the most valuable features among 837 features, we performed the LASSO analysis to remove the redundant features by compressing their coefficients to 0. The optimal LASSO model (Figures 5A,B) was obtained when the standard error of minimum distance presented the $\lambda = 0.067$, containing 7 non-zero radiomics features. The detailed information on these 7 radiomics features is shown in Table 2.

Among 7 features, X709 (GLSZM-Gray Level Non-Uniformity Normalized) showed no difference between the 2 groups (Table 2). The remaining 6 features were further found to correlate with the

disease type in univariable logistic regression analysis (Table 3, all $p < 0.05$), and multivariable regression analysis showed that X75 (GLSZM, Gray Level Variance), X210 (GLCM, Correlation), and X751 (FIRSTORDER, Maximum) were independent factors of disease type (all $p < 0.05$). It should be stated that the confidence interval of X210 was significantly abnormal, hence X210 was removed from our analysis. Finally, only X75 and X751 features were entered into our further analyses.

ROC analysis showed that X75 had a higher AUC value for discriminating the disease than the X751 (Figure 6A, P for Delong test<0.001), with a sensitivity of 0.713 and specificity of 0.591. The
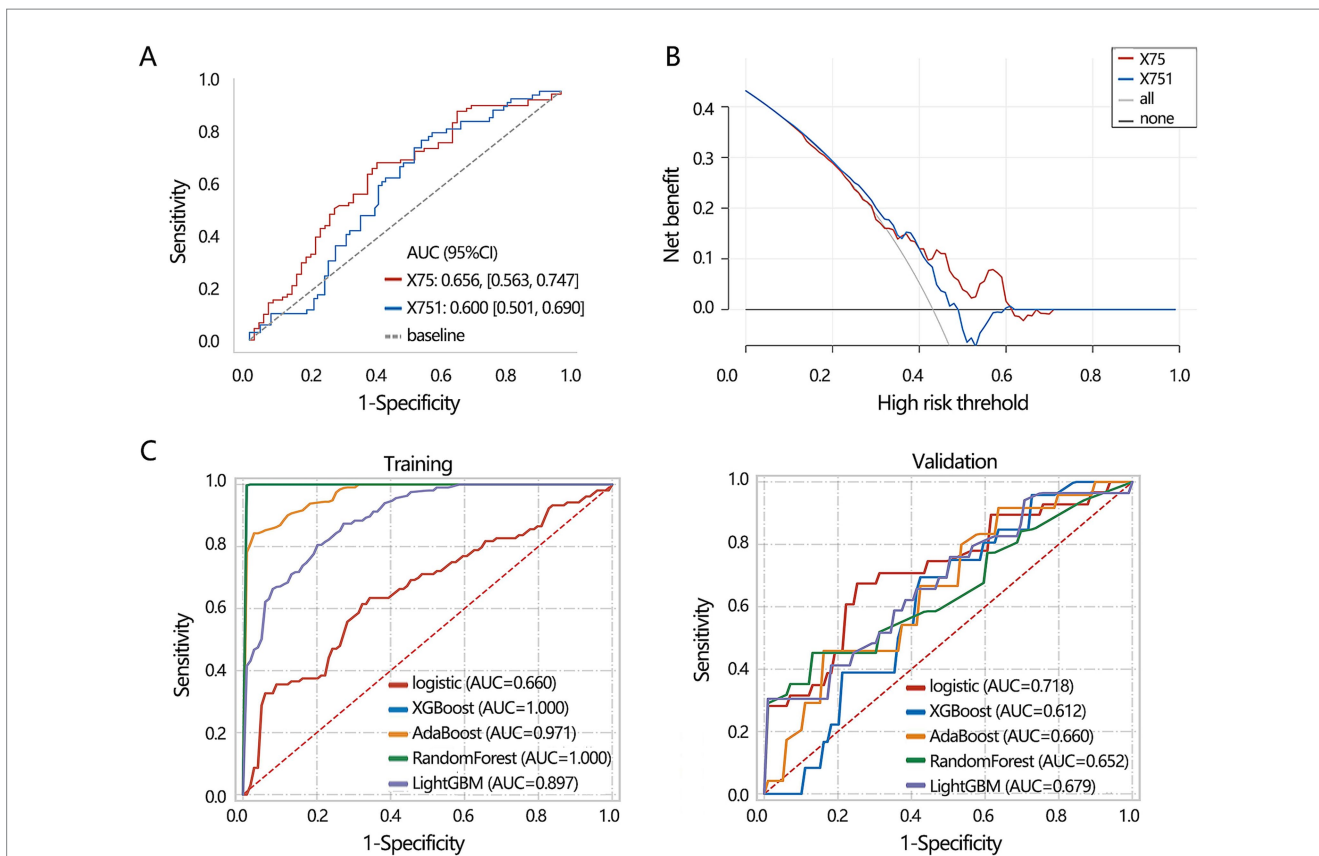
FIGURE 6
Machine learning model construction based on 2 radiomics features including X75 and X751. **(A)** ROC analysis was conducted to assess the discriminating performance of single radiomics features. **(B)** DCA analysis was used to assess the clinical net benefit of 2 radiomics features for discriminating disease. **(C)** Machine learning model construction and validation based on 2 radiomics features. X75: Gray Level Variance (GLSZM); X751: Maximum (FIRSTORDER).

obtained clinical net benefit of X75 was also superior to that of X751 (Figure 6B). Based on X75 and X751, a 2 radiomics features-based differentiation model was constructed using 5 deep learning algorithms in our training set, and the differentiation performance of the model was verified in the validation set. In the training set, XGBoost and RandomForest algorithms achieved the highest AUC values, while they achieved the lowest differentiation performance in the validation set (Figure 6C). It followed that the Logistic algorithm achieved a relatively stable performance both in training and validation set. Hence, 2 radiomics features-based model from the Logistic algorithm was regarded as the optimal deep learning model.

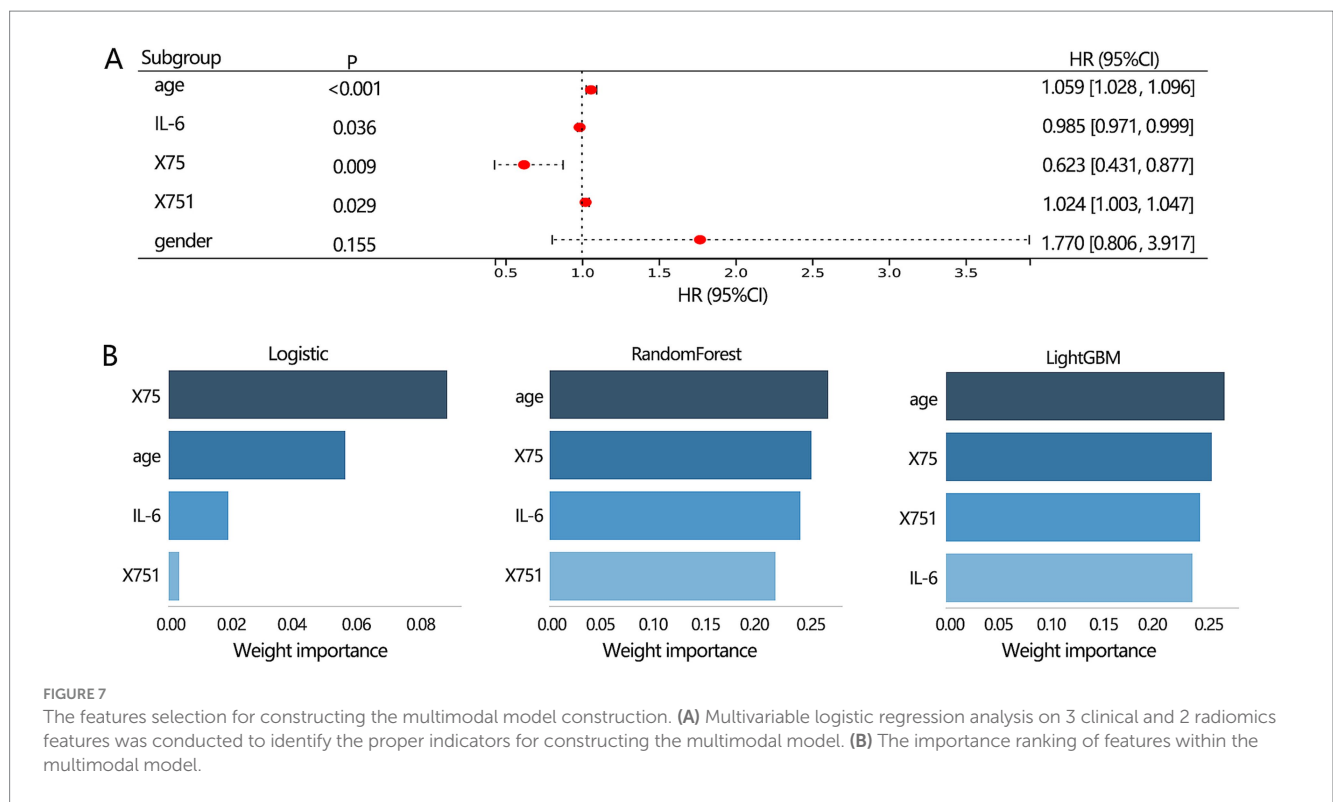## 3.4 Construction of multimodal differentiation model

The above results have demonstrated the importance of 3 clinical features (age, gender, IL-6 level) and 2 radiomics features (X75, X751) in the differentiation of lung disease, respectively. We further combined these 5 features and confirmed their independent role in the disease types. The multivariable regression analysis showed that age, IL-6, X75, and X751 were all independently related to the disease types (Figure 7A). Further, we constructed a multimodal

differentiation model based on the 4 features in the training set, followed by verification in the validation set. Considering the model performance in training and validation sets, the LightGBM algorithm achieved a relatively favorable and stable differentiation performance (Table 4). Therefore, the multimodal model constructed by the LightGBM algorithm was regarded as the optimal model.

Our results showed that the best 3 clinical features-based models, 2 radiomics features-based models, and the multimodal model were from RandomForest algorithm, Logistic algorithm, and LightGBM algorithm, respectively. Therefore, we next used the 3 algorithms to rank the importance of 4 features within the multimodal model. The results showed (Figure 7B) that age and X75 were the top 2 features among the 3 algorithms.

3.5 The performance comparisons of different models for differentiating NTM from MTB

Finally, we compared the differentiation performance of 3 optimal models constructed by corresponding algorithms only in the validation set due to the stability consideration. The optimal multimodal model from the LightGBM algorithm (Figure 8A) had the highest AUC value (0.804) compared with the optimal clinical model

FIGURE 7
The features selection for constructing the multimodal model construction. **(A)** Multivariable logistic regression analysis on 3 clinical and 2 radiomics features was conducted to identify the proper indicators for constructing the multimodal model. **(B)** The importance ranking of features within the multimodal model.

(0.756) and radiomic model (0.718). In addition, the optimal multimodal model (Table 5) had the highest accuracy (0.724), sensitivity (0.875), and NPV (0.738). However, the multimodal model had the lowest specificity (0.693). The PPV values of the optimal multimodal model was in between the clinical model and the radiomics model. The clinical net benefit of 3 models seemed to be similar (Figure 8B). To better visualize the net benefit, we next performed the DCA analysis only based on the Logistic regression (Figure 8C) among the whole population, finding that the multimodal model had better clinical net benefit.

3.6 Performance verification and comparison using the external testing dataset

We further verified the prediction performance among 3 clinical data based-model, 2 radiomics based-model, and multimodal model for differentiating NTM from MTB using a new external testing dataset. There were 36 male and 23 female patients in the new external testing dataset. The median age, IL-6, X75, and X751 were 61 [52, 67] years old, 10.9 [6.4, 33.2] (pg/mL), 4.77 [3.68, 5.42], and 766 [751, 771], respectively. The comparison results showed that the multimodal model had the highest AUC, accuracy, sensitivity, and NPV, but had the lowest specificity (Table 6). These results were similar to those from the validation set, which suggested the stability of our findings.

We also compared the differentiation performance of our multimodal model with the only NGS detection method, only radiologist assessment, and existing machine learning model constructed by Wang et al. (33) for differentiating NTM from

MTB. The results (Table 7; Figure 9) showed that our multimodal model had the highest AUC, accuracy, and sensitivity. The NGS detection also had favorable sensitivity. Radiologist assessment had the highest specificity and favorable accuracy. The published model by Wang et al. using CT images only had the highest precision for identifying NTM (yes vs. no). Our multimodal model improved accuracy than NGS, radiologist, and existing machine learning model, with an increased accuracy of 26, 4, and 6%, respectively. It also significantly improved sensitivity than radiologist and existing machine learning model, with an increased sensitivity of 4.5 and 15%, respectively, but was similar with that of NGS detection. These results highlighted the superiority of our multimodal model for differentiating NTM from MTB compared with existing approaches or existing machine learning model. Our model significantly improved the differentiation performance and accuracy, and provided favorable sensitivity at the same time.

4 Discussion

In this study, we initially established an optimal 3 clinical features-based differentiation model. We also constructed an optimal 2 radiomics features-based differentiation model. The differentiation performance of the clinical model was superior to the radiomics model. Finally, we developed an optimal multimodal differentiation model containing clinical and radiomics data. After analysis, our multimodal differentiation model showed more favorable differentiation performance compared with the single clinical or radiomics model. Our study suggested the necessity of the combination of clinical data and radiomics data in disease differentiation.

TABLE 4 Multimodal model construction and validation using 5 machine learning methods.

| | Training set | | | | | Validation set | | | | |
	XGBoost	Logistic	LightGBM	RandomForest	AdaBoost	XGBoost	Logistic	LightGBM	RandomForest	AdaBoost
AUC	1.000	0.808	0.954	1.000	0.999	0.717	0.699	0.804	0.755	0.738
Cutoff	0.741	0.452	0.522	0.575	0.500	0.741	0.452	0.522	0.575	0.500
Accuracy	0.991	0.750	0.895	0.987	0.978	0.690	0.672	0.724	0.655	0.655
Sensitivity	1.000	0.792	0.857	0.990	0.990	0.750	0.811	0.875	0.893	0.844
Specificity	1.000	0.726	0.943	0.992	0.985	0.700	0.643	0.698	0.600	0.577
PPV	1.000	0.716	0.926	1.000	0.979	0.762	0.608	0.673	0.833	0.708
NPV	0.984	0.785	0.872	0.977	0.977	0.675	0.752	0.738	0.609	0.604
Kappa	0.982	0.498	0.786	0.973	0.955	0.374	0.351	0.409	0.296	0.307

NPV, Negative predictive value; PPV, positive predictive value.

In our multimodal differentiation model, 4 key features were contained including age, IL-6, and 2 radiomics features. Spatial epidemiologic analysis showed that a higher risk for NTM infection was associated with older age, rurality, and more flooding (34). The previous study showed that females were 1.4 times more likely to infect NTM than males, clustering persons with age ≥ 65 years (35). We found that the NTM patients had a larger age compared with MTB patients and the median age of NTM patients was 66 years old, which was consistent with the previous findings. NTM infections have become a neglected and emerging problem in geriatric patients, and the older adult population is more susceptible to NTM and experiences increased morbidities (36). Our study and other findings (33, 37, 38) all suggested that NTM was more common in older adult people, which may be because NTM was a type of opportunistic pathogen, and patients infected with NTM often developed symptoms due to their aging and weakened immune system. However, we found no difference in gender distribution between the NTM and MTB groups, although our results were consistent with the previous study (33). It follows that patient characteristics alone are insufficient in differentiating between NTM and MTB.

It should be stated that NTM-infected old mice had significant dysrhythmia, cardiac hypertrophy, cardiac fibrosis, and elevated CD45[+] leukocyte levels and expression of inflammatory genes in heart tissue (39). It follows that NTM infections may contribute to cardiac dysfunction in the older adult population, which is a cause for concern. Besides the involvement of elevated age in the NTM infection, this study also found a decreased IL-6 level in the NTM group. The previous study (40) also reported that the production of IL-6 in the NTM infection group occurred to a significantly lesser extent, and p38 and extracellular regulated protein kinases (ERK1/2) played essential roles in the production of IL-6 during NTM infection. The impaired induction of p38 and ERK1/2 expression in response to NTM may contribute to host susceptibility to NTM lung disease. In addition, the level of IL-6 in macrophages infected with NTM can also be regulated by XLOC_002383/miR-146a-5p/TRAF6 axis (TRAF6, TNF receptor associated factor 6) (41). IL-6 has been regarded as a biomarker for discriminating NTM and other lung disease.

In addition, we also obtained 2 important radiomics features for discriminating lung disease. Especially, the importance of GLSZM [Gray Level Variance] (X75) was highlighted. Currently, there is no study reporting the role of GLSZM [Gray Level Variance] in the discrimination of NTM lung disease. The texture features represented by GLSZM mainly reflected the variability (heterogeneity of image texture) in the measurement area, and lower values indicated that the regions in the image are more homogeneous (42). Gray Level Variance is a regional-scale heterogeneity indices derived from GLSZM. In this study, NTM group had higher Gray Level Variance value, suggesting the higher heterogeneity in NTM than in MTB. In addition, Gray Level Variance value had favorable performance for differentiating NTM from NTM. The results suggested that distinctive textural features between NTM and MTB could be better captured by regional-scale lesion heterogeneity. Imaged heterogeneity may be due to the regional differences in cellularity, proliferation, hypoxia, angiogenesis, and necrosis (43). The previous study indicated that decreased lung tissue oxygenation may contribute to the development of NTM disease (44), and the level of hypoxia in NTM lesions in mice was more severe than that observed in the setting of tuberculosis. It followed that the hypoxia level presented regional differences between NTM and TB. In
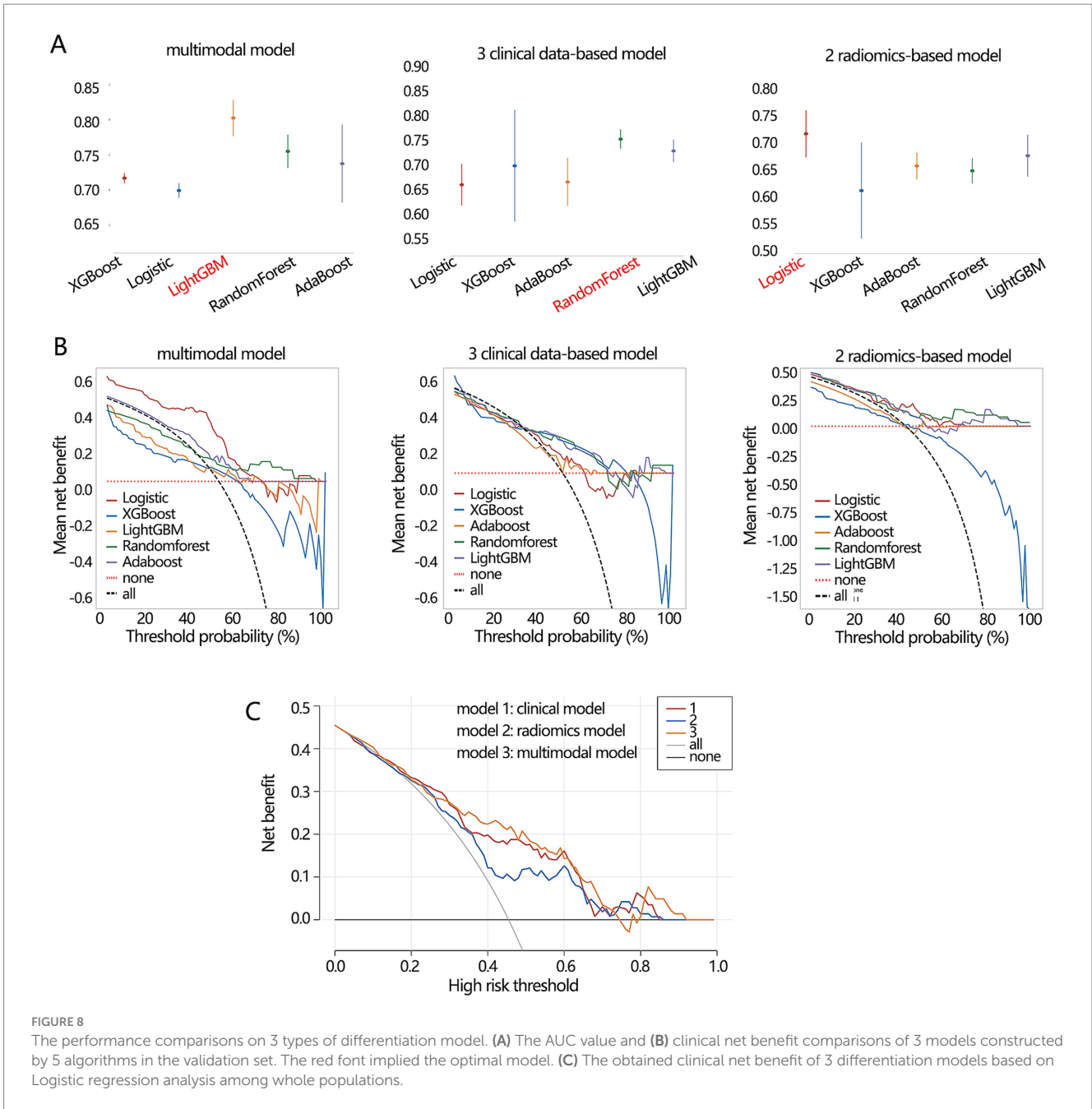
FIGURE 8
The performance comparisons on 3 types of differentiation model. **(A)** The AUC value and **(B)** clinical net benefit comparisons of 3 models constructed by 5 algorithms in the validation set. The red font implied the optimal model. **(C)** The obtained clinical net benefit of 3 differentiation models based on Logistic regression analysis among whole populations.

TABLE 5 The comparisons among different differentiation model.

	3 Clinical data model (RandomForest)		2 Radiomics model (Logistic)		Multimodal model (LightGBM)	
	Training	Validation	Training	Validation	Training	Validation
AUC	1.000	0.756	0.660	0.718	0.954	0.804
Cutoff	0.500	0.500	0.442	0.442	0.522	0.522
Accuracy	0.981	0.712	0.652	0.677	0.895	0.724
Sensitivity	1.000	0.690	0.615	0.708	0.857	0.875
Specificity	0.993	0.767	0.693	0.733	0.943	0.698
PPV	1.000	0.728	0.589	0.633	0.926	0.673
NPV	0.966	0.707	0.699	0.719	0.872	0.738
kappa	0.961	0.425	0.288	0.353	0.786	0.409

NPV, negative predictive value; PPV, positive predictive value.

TABLE 6 The performance comparison of different models for differentiating NTM and MTB in the external testing dataset.
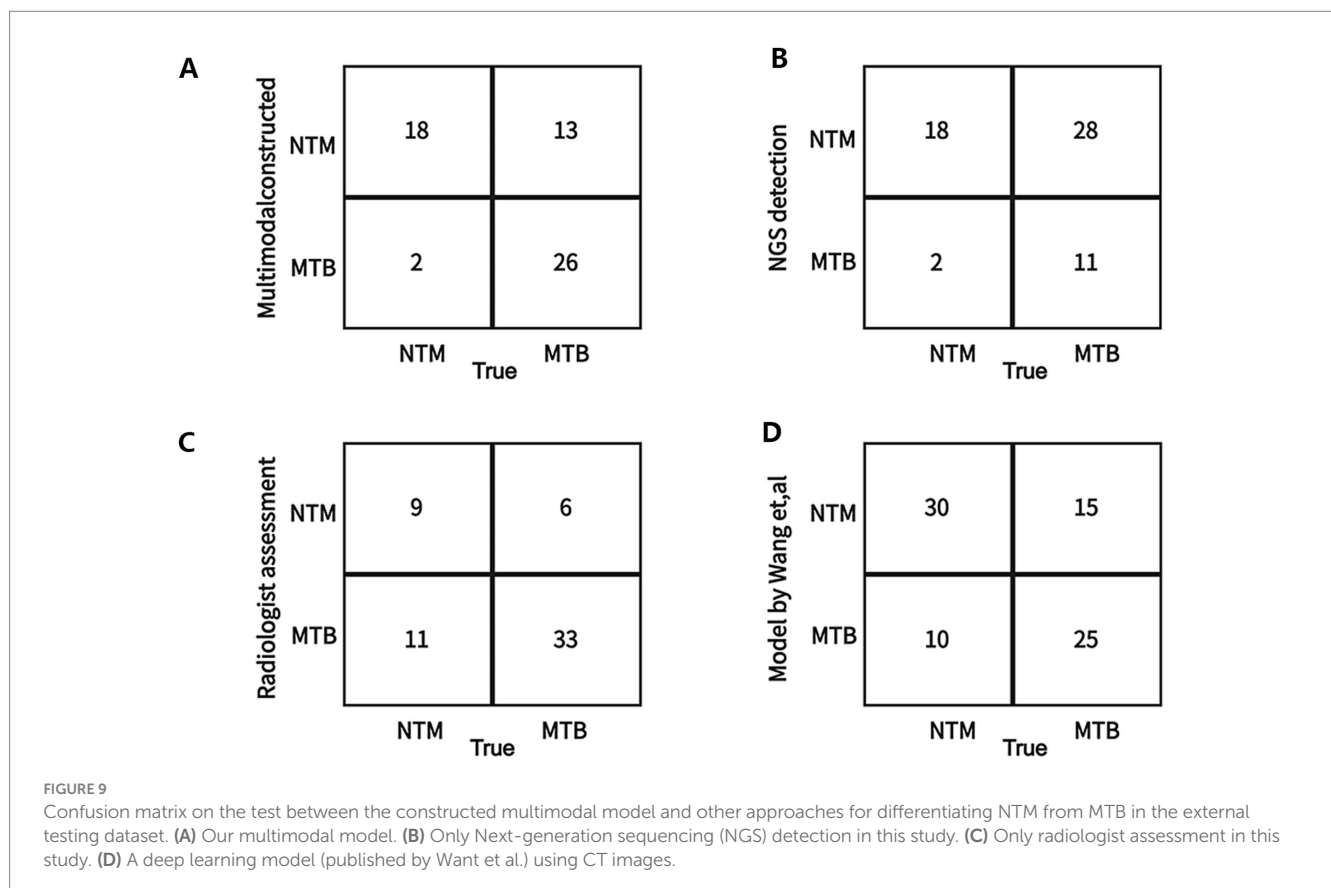
	3 Clinical data model (RandomForest)	2 Radiomics model (Logistic)	Multimodal model (LightGBM)
AUC	0.657	0.766	0.858
cutoff	0.550	0.310	0.204
Accuracy	0.583	0.744	0.745
Sensitivity	0.200	0.833	0.900
Specificity	0.857	0.710	0.667
PPV	0.500	0.526	0.581
NPV	0.600	0.917	0.929

AUC, area under curve; NPV, negative predictive value; PPV, positive predictive value.

TABLE 7 The performance comparison between the multimodal model and other approaches for differentiating NTM from MTB in the external testing dataset.

	Our model	Only NGS detection	Only radiologist	Published model by Wang et al.
Details	Multimodal constructed	DNA sequencing	5-year experience	Only CT images
AUC	0.86	0.59	0.65	0.78
Accuracy	0.75	0.49	0.71	0.69
Sensitivity	0.90	0.90	0.45	0.75
Specificity	0.67	0.28	0.85	0.63
MTB-Precision	0.93	0.85	0.75	0.71
NTM-Precision	0.58	0.39	0.60	0.67

AUC, area under curve; NGS, Next-generation sequencing.

FIGURE 9
Confusion matrix on the test between the constructed multimodal model and other approaches for differentiating NTM from MTB in the external testing dataset. (A) Our multimodal model. (B) Only Next-generation sequencing (NGS) detection in this study. (C) Only radiologist assessment in this study. (D) A deep learning model (published by Want et al.) using CT images.

addition, NTM and TB also showed the difference in hematological profiles. A pilot study found that TB patients had higher basophils and platelets levels, but lower eosinophils level than NTM patients (45). Moreover, MTB-infected cells had lower level of phagosome-lysosome fusion and apoptosis than NTM-infected cells (46). We speculated that the final selection of GLSZM [Gray Level Variance] as key features may be related to the heterogeneity in hypoxia, blood, etc. between NTM and MTB. NTM infections are common but are often cofounded with TB because of the similarity of symptoms, therefore, it is necessary to find more heterogeneity between them.

The previous study (13) also established a machine learning-based differentiation of NTM and MTB using CT images, finding that feature extracting from bronchiectasis was relatively more informative than that from a cavity or the combination (bronchiectasis+ cavity). This study highlighted the effectiveness difference of different regions (cavities, bronchiectasis, and their combination). In addition, chest X-rays from suspects of mycobacterial lung disease were also used to distinguish between TB or NTM patients by artificial intelligence, finding that deep neural networks had a better performance than pulmonologists on classifying patients (47). It follows that the deep learning model may achieve favorable differentiation performance, and a more effective model needs to be investigated.

Finally, several limitations should be stated. In this study, the construction of the multimodal model depended on high-quality CT imaging, and the sample size was not sufficiently large, which may influence the generalizability and application of the models. The DCA was used to evaluate the clinical net benefit of 3 models, but it was only based on Logistic regression and we failed to use more methods to validate the clinical net benefit difference due to different algorithms. In addition, we just verified the model performance by an external testing dataset from our hospital and the external dataset lacks diversity, which may have a selection bias and limit the generalizability of our findings. Another limitation is the absence of comparison with existing clinical methods or diagnostic tools. Without this comparison, it is difficult to assess the added value of the proposed approach. The model validation and comparison are the other key research topic in the future.

5 Conclusion

This study developed a multimodal learning model to classify NTM from MTB, with greater accuracy, sensitivity, and negative predictive value than the single clinical or radiomics based models. Our multimodal model improved accuracy than NGS, radiologist, and existing machine learning model. It also significantly improved sensitivity than radiologist and existing machine learning model. These results highlighted the superiority of our multimodal model for differentiating NTM from MTB compared with existing approaches. Our study can promote the application of new technologies, thereby advancing research progress in this fields. This is the first study to consider multidimensions and use multimodal model to distinguish diseases, which provides a new perspective and strategy for differentiating NTM from MTB, and help doctors to choose appropriate treatment plans.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Zhoushan Hospital (Ethical Approval Number: 2024–090). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

H-lL: Conceptualization, Data curation, Formal analysis, Supervision, Writing – original draft, Writing – review & editing. R-zZ: Data curation, Writing – original draft. H-sL: Methodology, Writing – original draft. MW: Formal analysis, Writing – original draft. S-jY: Data curation, Writing – original draft.

Funding

Conflict of interest

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

Publisher's note

References

1. Furin J, Cox H, Pai M. Tuberculosis. *Lancet*. (2019) 393:1642–56. doi: 10.1016/S0140-6736(19)30308-3

2. Bagcchi S. WHO's global tuberculosis report 2022. *Lancet Microbe*. (2023) 4:e20. doi: 10.1016/S2666-5247(22)00359-7

3. Dahl VN, Molhave M, Floe A, van Ingen J, Schon T, Lillebaek T, et al. Global trends of pulmonary infections with nontuberculous mycobacteria: a systematic review. *Int J Infect Dis*. (2022) 125:120–31. doi: 10.1016/j.ijid.2022.10.013

4. Ratnatunga CN, Lutzky VP, Kupz A, Doolan DL, Reid DW, Field M, et al. The rise of non-tuberculosis mycobacterial lung disease. *Front Immunol*. (2020) 11:303. doi: 10.3389/fimmu.2020.00303

5. Gopalaswamy R, Shanmugam S, Mondal R, Subbian S. Of tuberculosis and non-tuberculous mycobacterial infections - a comparative analysis of epidemiology, diagnosis and treatment. *J Biomed Sci*. (2020) 27:74. doi: 10.1186/s12929-020-00667-6

6. Liu CF, Song YM, He WC, Liu DX, He P, Bao JJ, et al. Nontuberculous mycobacteria in China: incidence and antimicrobial resistance spectrum from a nationwide survey. *Infect Dis Poverty*. (2021) 10:59. doi: 10.1186/s40249-021-00844-1

7. Acharya B, Acharya A, Gautam S, Ghimire SP, Mishra G, Parajuli N, et al. Advances in diagnosis of tuberculosis: an update into molecular diagnosis of *Mycobacterium tuberculosis*. *Mol Biol Rep*. (2020) 47:4065–75. doi: 10.1007/s11033-020-05413-7

8. Gautam SS, Mac Aogain M, Cooley LA, Haug G, Fyfe JA, Globan M, et al. Molecular epidemiology of tuberculosis in Tasmania and genomic characterisation of its first known multi-drug resistant case. *PLoS One*. (2018) 13:e0192351. doi: 10.1371/journal.pone.0192351

9. Yang W, Jiang J, Zhao Q, Ren HQ, Yao XX, Sun SY, et al. A case of tuberculosis misdiagnosed as sarcoidosis and then confirmed by NGS testing. *Clin Lab*. (2024) 70. doi: 10.7754/Clin.Lab.2023.230823

10. Chu HQ, Li B, Zhao L, Huang DD, Zhang ZM, Xu JF, et al. Chest imaging comparison between non-tuberculous and tuberculosis mycobacteria in sputum acid fast bacilli smear-positive patients. *Eur Rev Med Pharmacol Sci*. (2015) 19:2429–39. doi: 10.1183/13993003.congress-2015.pa2674

11. Ma J, Zhou Z, Ren Y, Xiong J, Fu L, Wang Q, et al. Computerized detection of lung nodules through radiomics. *Med Phys*. (2017) 44:4148–58. doi: 10.1002/mp.12331

12. Coroller TP, Agrawal V, Huynh E, Narayan V, Lee SW, Mak RH, et al. Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC. *J Thorac Oncol*. (2017) 12:467–76. doi: 10.1016/j.jtho.2016.11.2226

13. Xing Z, Ding W, Zhang S, Zhong L, Wang L, Wang J, et al. Machine learning-based differentiation of nontuberculous mycobacteria lung disease and pulmonary tuberculosis using CT images. *Biomed Res Int*. (2020) 2020:6287545. doi: 10.1155/2020/6287545

14. Ying C, Li X, Lv S, Du P, Chen Y, Fu H, et al. T-SPOT with CT image analysis based on deep learning for early differential diagnosis of nontuberculous mycobacteria pulmonary disease and pulmonary tuberculosis. *Int J Infect Dis*. (2022) 125:42–50. doi: 10.1016/j.ijid.2022.09.031

15. Goecks J, Jalili V, Heiser LM, Gray JW. How machine learning will transform biomedicine. *Cell*. (2020) 181:92–101. doi: 10.1016/j.cell.2020.03.022

16. Faujdar J, Gupta P, Natrajan M, Das R, Chauhan DS, Katoch VM, et al. Mycobacterium indicus pranii as stand-alone or adjunct immunotherapeutic in treatment of experimental animal tuberculosis. *Indian J Med Res*. (2011) 134:696–703. doi: 10.4103/0971-5916.90999

17. Chaki J, Deshpande G. Brain disorder detection and diagnosis using machine learning and deep learning - a bibliometric analysis. *Curr Neuropharmacol*. (2024) 22:2191–216. doi: 10.2174/1570159X22999240531160344

18. Zhang C, Xu J, Tang R, Yang J, Wang W, Yu X, et al. Novel research and future prospects of artificial intelligence in cancer diagnosis and treatment. *J Hematol Oncol*. (2023) 16:114. doi: 10.1186/s13045-023-01514-5

19. Yao F, Zhang R, Lin Q, Xu H, Li W, Ou M, et al. Plasma immune profiling combined with machine learning contributes to diagnosis and prognosis of active pulmonary tuberculosis. *Emerg Microbes Infect*. (2024) 13:2370399. doi: 10.1080/22221751.2024.2370399

20. Li LS, Yang L, Zhuang L, Ye ZY, Zhao WG, Gong WP. From immunology to artificial intelligence: revolutionizing latent tuberculosis infection diagnosis with machine learning. *Mil Med Res*. (2023) 10:58. doi: 10.1186/s40779-023-00490-8

21. Nijiati M, Guo L, Abulizi A, Fan S, Wubuli A, Tuersun A, et al. Deep learning and radiomics of longitudinal CT scans for early prediction of tuberculosis treatment outcomes. *Eur J Radiol*. (2023) 169:111180. doi: 10.1016/j.ejrad.2023.111180

22. Anh NK, Phat NK, Thu NQ, Tien NTN, Eunsu C, Kim HS, et al. Discovery of urinary biosignatures for tuberculosis and nontuberculous mycobacteria classification using metabolomics and machine learning. *Sci Rep*. (2024) 14:15312. doi: 10.1038/s41598-024-66113-x

23. Zhou L, Wang Y, Zhu W, Zhao Y, Yu Y, Hu Q, et al. A retrospective study differentiating nontuberculous mycobacterial pulmonary disease from pulmonary tuberculosis on computed tomography using radiomics and machine learning algorithms. *Ann Med*. (2024) 56:2401613. doi: 10.1080/07853890.2024.2401613

24. Horne D, Skerrett S. Recent advances in nontuberculous mycobacterial lung infections. *F1000Res*. (2019) 8:1710. doi: 10.12688/f1000research.20096.1

25. Dahiya B, Mehta N, Soni A, Mehta PK. Diagnosis of extrapulmonary tuberculosis by gene Xpert MTB/RIF ultra assay. *Expert Rev Mol Diagn*. (2023) 23:561–82. doi: 10.1080/14737159.2023.2223980

26. Wu W, Han X, Zhao H, Sun H, Sun Q. Application value of next-generation sequencing of bronchial alveolar lavage fluid in emergency patients with infection. *Cell Mol biol (Noisy-le-grand)*. (2023) 69:45–9. doi: 10.14715/cmb/2023.69.8.7

27. Schober P, Vetter TR. Logistic regression in medical research. *Anesth Analg*. (2021) 132:365–6. doi: 10.1213/ANE.0000000000005247

28. Touw WG, Bayjanov JR, Overmars L, Backus L, Boekhorst J, Wels M, et al. Data mining in the life sciences with random Forest: a walk in the park or lost in the jungle? *Brief Bioinform*. (2013) 14:315–26. doi: 10.1093/bib/bbs034

29. Zhao Z, Yang W, Zhai Y, Liang Y, Zhao Y. Identify DNA-binding proteins through the extreme gradient boosting algorithm. *Front Genet*. (2021) 12:821996. doi: 10.3389/fgene.2021.821996

30. Rufo DD, Debelee TG, Ibenthal A, Negera WG. Diagnosis of diabetes mellitus using gradient boosting machine (light GBM). *Diagnostics (Basel)*. (2021) 11:1714. doi: 10.3390/diagnostics11091714

31. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Light GBM: a highly efficient gradient boosting decision tree. *Neural Inform Proces Syst*. (2017) 31:3149-3157.

32. Li S, Zeng Y, Chapman WC Jr, Erfanzadeh M, Nandy S, Mutch M, et al. Adaptive boosting (Ada boost)-based multiwavelength spatial frequency domain imaging and characterization for ex vivo human colorectal tissue assessment. *J Biophotonics*. (2020) 13:e201960241. doi: 10.1002/jbio.201960241

33. Wang L, Ding W, Mo Y, Shi D, Zhang S, Zhong L, et al. Distinguishing nontuberculous mycobacteria from *Mycobacterium tuberculosis* lung disease from CT images using a deep learning framework. *Eur J Nucl Med Mol Imaging*. (2021) 48:4293–306. doi: 10.1007/s00259-021-05432-x

34. Mejia-Chew C, Chavez MA, Lian M, McKee A, Garrett L, Bailey TC, et al. Spatial epidemiologic analysis and risk factors for nontuberculous mycobacteria infections, Missouri, USA, 2008-2019. *Emerg Infect Dis*. (2023) 29:1540–6. doi: 10.3201/eid2908.230378

35. Adjemian J, Olivier KN, Seitz AE, Holland SM, Prevots DR. Prevalence of nontuberculous mycobacterial lung disease in U.S. Medicare beneficiaries. *Am J Respir Crit Care Med*. (2012) 185:881–6. doi: 10.1164/rccm.201111-2016OC

36. Mirsaeidi M, Farshidpour M, Ebrahimi G, Aliberti S, Falkinham JO 3rd. Management of nontuberculous mycobacterial infection in the elderly. *Eur J Intern Med*. (2014) 25:356–63. doi: 10.1016/j.ejim.2014.03.008

37. Winthrop KL, McNelley E, Kendall B, Marshall-Olson A, Morris C, Cassidy M, et al. Pulmonary nontuberculous mycobacterial disease prevalence and clinical features: an emerging public health disease. *Am J Respir Crit Care Med*. (2010) 182:977–82. doi: 10.1164/rccm.201003-0503OC

38. Santos A, Carneiro S, Silva A, Gomes JP, Macedo R. Nontuberculous mycobacteria in Portugal: trends from the last decade. *Pulmonology*. (2022) 30:337–43. doi: 10.1016/j.pulmoe.2022.01.011

39. Headley CA, Gerberick A, Mehta S, Wu Q, Yu L, Fadda P, et al. Nontuberculous mycobacterium *M. avium* infection predisposes aged mice to cardiac abnormalities and inflammation. *Aging Cell*. (2019) 18:e12926. doi: 10.1111/acel.12926

40. Sim YS, Kim SY, Kim EJ, Shin SJ, Koh WJ. Impaired expression of MAPK is associated with the downregulation of TNF-alpha, IL-6, and IL-10 in *Mycobacterium abscessus* lung disease. *Tuberc Respir Dis (Seoul)*. (2012) 72:275–83. doi: 10.4046/trd.2012.72.3.275

41. Hu R, Molibeli KM, Zhu L, Li H, Chen C, Wang Y, et al. Long non-coding RNA-XLOC_002383 enhances the inhibitory effects of THP-1 macrophages on Mycobacterium avium and functions as a competing endogenous RNA by sponging mi R-146a-5p to target TRAF6. *Microbes Infect*. (2023) 25:105175. doi: 10.1016/j.micinf.2023.105175

42. Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative Radiomics for high-throughput image-based phenotyping. *Radiology*. (2020) 295:328–38. doi: 10.1148/radiol.2020191145

43. Lee HS, Oh JS, Park YS, Jang SJ, Choi IS, Ryu JS. Differentiating the grades of thymic epithelial tumor malignancy using textural features of intratumoral heterogeneity via (18) F-FDG PET/CT. *Ann Nucl Med*. (2016) 30:309–19. doi: 10.1007/s12149-016-1062-2

44. Kuroda F, Tanabe N, Igari H, Sakurai T, Sakao S, Tada Y, et al. Nontuberculous mycobacterium diseases and chronic thromboembolic pulmonary hypertension. *Intern Med*. (2014) 53:2273–9. doi: 10.2169/internalmedicine.53.2558

45. Sanogo F, Kodio O, Sarro YS, Diarra B, Coulibaly G, Tolofoudie M, et al. Hematological profiles of patients with tuberculosis and nontuberculous mycobacteria infections in Bamako, Mali. *Int J Mycobacteriol*. (2023) 12:235–40. doi: 10.4103/ijmy.ijmy_208_22

46. Feng Z, Bai X, Wang T, Garcia C, Bai A, Li L, et al. Differential responses by human macrophages to infection with Mycobacterium tuberculosis and non-tuberculous mycobacteria. *Front Microbiol*. (2020) 11:116. doi: 10.3389/fmicb.2020.00116

47. Liu CJ, Tsai CC, Kuo LC, Kuo PC, Lee MR, Wang JY, et al. A deep learning model using chest X-ray for identifying TB and NTM-LD patients: a cross-sectional study. *Insights Imaging*. (2023) 14:67. doi: 10.1186/s13244-023-01395-9