European Radiology (2023) 33:391–400
https://doi.org/10.1007/s00330-022-08997-9

CHEST

# Machine learning and radiomics for the prediction of multidrug resistance in cavitary pulmonary tuberculosis: a multicentre study

Ye Li[1] · Bing Wang[2] · Limin Wen[3] · Hengxing Li[3] · Fang He[4] · Jian Wu[1] · Shan Gao[2] · Dailun Hou[1]

Received: 2 April 2022 / Revised: 29 May 2022 / Accepted: 29 June 2022 / Published online: 19 July 2022


## Abstract

**Objectives** Multidrug-resistant tuberculosis (MDR-TB) is a major challenge to global health security. Early identification of MDR-TB patients increases the likelihood of treatment success and interrupts transmission. We aimed to develop a predictive model for MDR to cavitary pulmonary TB using CT radiomics features.

**Methods** This retrospective study included 257 consecutive patients with proven active cavitary TB (training cohort: 187 patients from Beijing Chest Hospital; testing cohort: 70 patients from Infectious Disease Hospital of Heilongjiang Province). Radiomics features were extracted from the segmented cavitation. A radiomics model was constructed to predict MDR using a random forest classifier. Meaningful clinical characteristics and subjective CT findings comprised the clinical model. The radiomics and clinical models were combined to create a combined model. ROC curves were used to validate the capability of the models in the training and testing cohorts.

**Results** Twenty-one radiomics features were selected as optimal predictors to build the model for predicting MDR-TB. The AUCs of the radiomics model were significantly higher than those of the clinical model in either the training cohort (0.844 versus 0.589, $p < 0.05$) or the testing cohort (0.829 versus 0.500, $p < 0.05$). The AUCs of the radiomics model were slightly lower than those of the combined model in the training cohort (0.844 versus 0.881, $p > 0.05$) and testing cohort (0.829 versus 0.834, $p > 0.05$), but there was no significant difference.

**Conclusions** The radiomics model has the potential to predict MDR in cavitary TB patients and thus has the potential to be a diagnostic tool.

**Key Points**

- *This is the first study to build and validate models that distinguish MDR-TB from DS-TB with clinical and radiomics features based on cavitation.*
- *The radiomics model demonstrated good performance and might potentially aid in prior TB characterisation treatment.*
- *This noninvasive and convenient technique can be used as a diagnosis tool into routine clinical practice.*

**Keywords** Pulmonary tuberculosis · Cavitation · Radiomics · Machine learning · Drug resistance

**Abbreviations**

| | |
|---|---|
| AUC | Area under the ROC curve |
| CT | Computed tomography |
| DST | Drug-sensitive test |
| DS-TB | Drug-sensitive TB |
| GLCM | Grey-level co-occurrence matrix features |
| GLDM | Grey-level dependence matrix |
| GLRLM | Grey-level run-length matrix |
| GLSZM | Grey-level size-zone matrix |
| MDR-TB | Multidrug-resistant TB |

✉ Dailun Hou
hou.dl@mail.ccmu.edu.cn

[1] Department of Radiology, Beijing Chest Hospital, Capital Medical University, Beijing 101149, China

[2] Department of Radiology, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing 101149, China

[3] Department of Radiology, Infectious Disease Hospital of Heilongjiang Province, Harbin 150500, China

[4] Department of Radiology, Guangxi Zhuang Autonomous Region Chest Hospital, Liuzhou 545000, China

| | |
|---|---|
| MIC | Maximal information coefficient |
| MINE | Maximal information–based nonparametric exploration |
| NGTDM | Neighbouring grey-tone difference matrix |
| RFC | Random forest classifier |
| ROC | Receiver operator characteristic |
| ROIs | Regions of interest |
| TB | Tuberculosis |

## Introduction

Pulmonary tuberculosis (TB) is one of the leading causes of death worldwide and the top cause of death from a single infectious disease globally until the coronavirus (COVID-19) pandemic, ranking above HIV/AIDS [1]. The presence of multidrug-resistant TB (MDR-TB) is a major challenge to global health security, and it is the only major airborne drug-resistant epidemic. The number of confirmed MDR-TB cases over the past 5 years has almost doubled globally [2]. MDR-TB refers to TB infection resistant to at least two of the most powerful anti-TB drugs, isoniazid and rifampicin [3]. Microbiologic culture and sputum smear microscopy after Ziehl-Neelsen staining are necessary for diagnosing TB [4]. Currently, there are many genotypic and phenotypic tests to detect MDR-TB strains and determine whether they are resistant to drugs used [5]. However, using sputum instead of a culture isolate as a sample to detect MDR-TB has poor sensitivity. The results from the culture isolate are generally only available after 4–8 weeks [1]. Hence, the early identification of patients with MDR-TB increases the likelihood of treatment success and interrupts transmission.

Chest computed tomography (CT) has shown good performance as a triage tool, in monitoring imaging changes, and in evaluating the severity of pulmonary TB. Several studies have demonstrated that multiple cavities, masses, bronchiectasis, and lymph nodes are more commonly observed in MDR-TB [6]. Cavitation in the lung parenchyma is an important imaging feature for diagnosing TB, and it can be easily detected using CT [7]. A cavity is defined as a pathological, gas-filled space within a pulmonary consolidation in the lung parenchyma produced by the expulsion of the necrotic part of the lesion via the bronchial tree [8]. In addition, cavitation may also be seen in primary bronchogenic carcinoma, pulmonary metastasis and pneumonia. The variation in the cavity is an essential sign for measuring the response to TB treatment [9]. The existence of cavitation is common in TB, especially in MDR-TB. The presence of cavities is strongly associated with a high TB mycobacterial load, representing an advanced state of the disease [10]. However, most imaging features in prior studies were heterogeneous and nonquantitative, and few quantitative studies have focused on cavitary TB.

Radiomics is based on the conversion of medical images to high-dimensional quantitative information, which provides important insights into pathophysiology. Radiomics can extract multiple types of quantitative data from radiological images, such as shape, distribution of attenuation or intensity, and spatial information [11]. Radiomics has been primarily applied to a variety of tumours, including brain, nasopharynx, and lung cancers [12]. However, to the best of our knowledge, few studies have published radiomics analyses of cavitary TB for MDR-TB prediction.

We hypothesised that the radiomics features of cavitary TB patients may be useful for predicting MDR-TB because of the detailed quantitative information that can be obtained from a segmented cavity. The aim of this study was to evaluate whether CT-based radiomics features from cavitation could distinguish MDR-TB from drug-sensitive TB (DS-TB) and predict MDR-TB successfully.

## Materials and methods

### Participants

This retrospective study was approved by the ethics committees of Beijing Chest Hospital, Capital Medical University (No. 36, 2021) and Infectious Disease Hospital of Heilongjiang Province, and the requirement for informed consent was waived.

In the study, all patients with cavitary TB were enrolled according to the inclusion and exclusion criteria. The inclusion criteria were as follows: (a) patients were confirmed to have active TB with the use of sputum microscopy, culture, or polymerase chain reaction test; (b) the drug-sensitive test (DST) of *Mycobacterium tuberculosis* isolates was performed to distinguish DS-TB from MDR-TB and quality control was routinely performed during the DST; (c) there was a clearly visible cavity on lung CT; and (d) all CT scans were performed within 30 days of TB diagnosis to exclude the confounding effects of medications on the measurements. The exclusion criteria included the following: (a) a history of thoracoabdominal trauma or other pulmonary diseases such as lung cancer, silicosis, or COPD; (b) diabetes or HIV seropositivity; and (c) image artefacts or incomplete clinical information.

According to the above criteria, we recruited a total of 257 patients who were diagnosed with cavitary TB. Finally, a total of 187 patients, which included 115 DS-TB and 72 MDR-TB patients from Beijing Chest Hospital, were enrolled as the training cohort from December 2016 to July 2019. In addition, a testing cohort from Infectious Disease Hospital of Heilongjiang Province was subsequently enrolled, including 35 patients with DS-TB and 35 patients with MDR-TB, from August 2018 to May 2020. The patient recruitment flowchart is shown in Fig. 1.

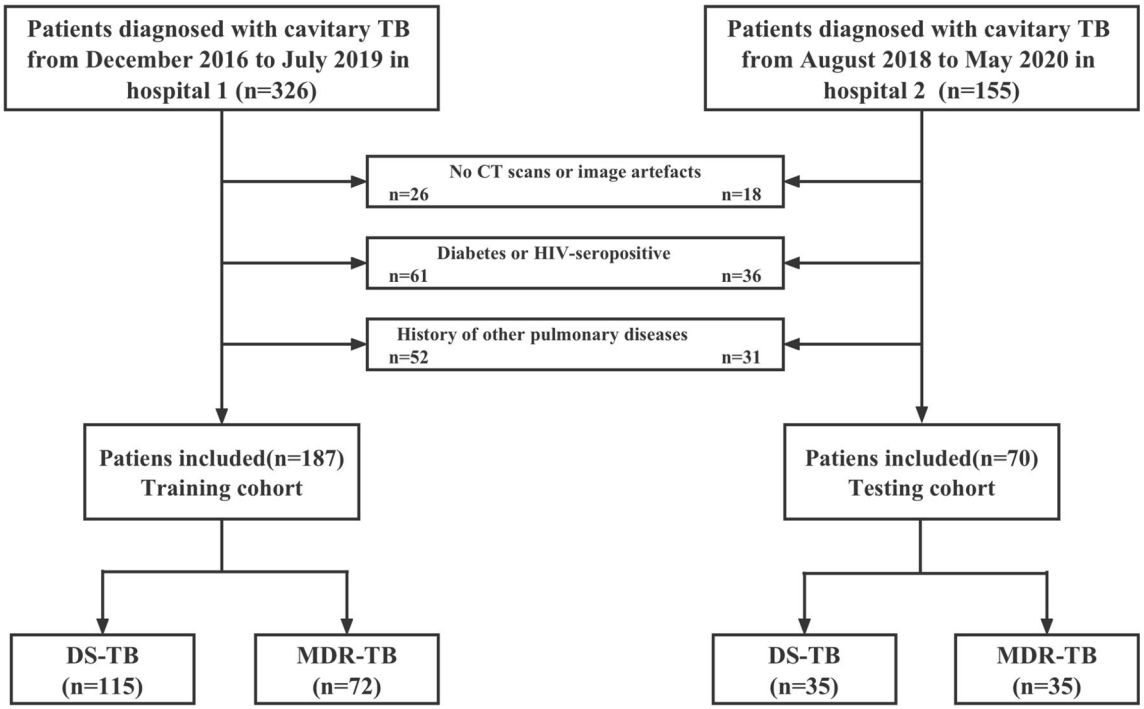


**Fig. 1** Flowchart of patient selection

## CT image acquisition

All CT scans were performed with an Optima CT 680 system and Light Speed VCT (GE Healthcare). The scanning parameters were as follows: tube voltage, 120 kV; automatic tube current modulation; detector collimation, 64 × 0.625 mm; rotation time, 500 ms; and pitch, 1.375. The image reconstruction parameters were as follows: slice thickness, 1.25 mm; increment, 1.25 mm; field of view, 15 cm; and matrix, 512 × 512. Reconstructed images were transferred from the hospital's picture archiving and communication system to 3D Slicer (http://www.slicer.org) for radiomics analysis.

## Evaluation of subjective CT findings

CT images were independently reviewed by two radiologists with at least 10 years of experience in analysing chest images who were blinded to the final clinical diagnosis based on the DST. The primary signs assessed in the CT images (Fig. 2) were as follows: (a) tree-in-bud and small centrilobular nodules [13]; (b) single large nodule and surrounding satellite lesions; (c) consolidation (lobular or subsegmental, segmental or lobar); (d) fibro stripe; and (e) calcified nodules. A consensus was reached through discussion in cases of disagreement.

## Cavity segmentation

Regions of interest (ROIs) were manually contoured along the whole surface of the cavity layer by layer, including circumambient satellite lesions in reference to images in lung windows (Fig. 3). Segmentation was manually contoured in 3D Slicer software, strictly performed by an experienced chest radiologist and confirmed by another chest radiologist with 10 years of experience in lung CT to exclude irrelevant components. Both radiologists were blinded to the diagnosis of TB.

## Radiomics feature extraction

The radiomics features were extracted from each ROI. Detailed information about the radiomics features is provided in Supplementary Material 1 and PyRadiomics official documentation (https://pyradiomics.readthedocs.io/en/latest/features.html).

## Feature selection

Not all features contribute to the positive performance of classification, and some features might add noise to it. The maximal information coefficient (MIC) is a measure of two-variable dependence designed specifically for the rapid exploration of many-dimensional data sets [14]. MIC is part of a larger family of maximal information-based nonparametric exploration (MINE) statistics, which can be used not only to identify important relationships in data sets but also to characterise them. According to the MIC values, we selected the top 300 relevant features in the training cohort after data splitting. The random forest classifier (RFC) method was used to

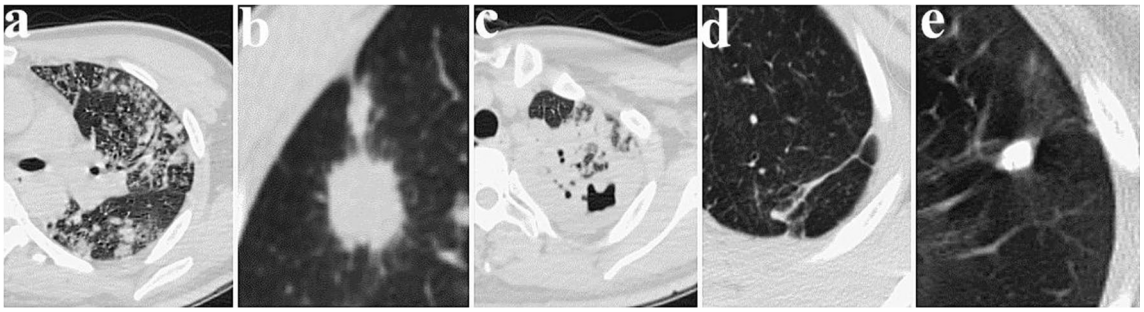


**Fig. 2** Axial lung CT images of subjective CT findings for TB patients. **a** Tree-in-bud and small centrilobular nodules; **b** single large nodule and surrounding satellite lesions; **c** consolidation (lobular or subsegmental, segmental or lobar); **d** fibro stripe; **e** nodules with calcification

construct the prediction model because of its high variance-bias trade-off capability [15]. RFC can give estimates of what variables are important in the classification. Then we used the embedded feature selection to choose predictable features based on the top 300 relevant features (Supplementary Material 2). Tenfold cross validation was implemented to avoid overfitting. The two steps of feature selection were carried out by the Python scikit-learn package (version 3.8, Scikit-learn Version 0. 21, http://scikit-learn.org/).

## Model construction

In the training cohort, the clinical characteristics and subjective CT findings were selected to build a clinical model by multivariate logistic regression. The radiomics model to predict MDR-TB based on the selected radiomics features was trained using the RFC method. Additionally, the combined model was the combination of the radiomics model and clinical model by the RFC method, which included the selected radiomics features and clinical variables (clinical characteristics and subjective CT findings) with significant differences. These three models were all trained and validated on the training cohort which were randomly divided into two parts by the ratio of 7:3 using tenfold cross validation. Finally, the best model which was separately selected from all three models was tested on the external testing cohort.

## Statistics

Statistical analysis was performed with SPSS software (version 21) and the Python Scikit-learn package. Cohen's kappa test was used to analyse interreader agreements of subjective findings ($\kappa$ values of poor, fair, moderate, substantial, and near-perfect agreement were < 0.00, 0.21–0.40, 0.41–0.60, 0.61–0.80, and 0.81–1.00, respectively). Qualitative variables (sex and subjective CT findings) were presented as frequencies. Differences between qualitative variables were compared with the chi-square test. Then the age was compared by the Mann-Whitney *U* test. Variables with significant differences were used to develop the clinical model by logistic regression. The performance of all three models in both cohorts was evaluated with receiver operator characteristic (ROC) curves, which were used to calculate the area under the ROC curve (AUC). Thereafter, we used the DeLong test to separately analyse the AUCs of the clinical model, radiomics model and combined model. We individually calculated accuracy, precision (positive predictive value), recall (sensitivity) and F1 score. F1 score is the harmonic average of the precision

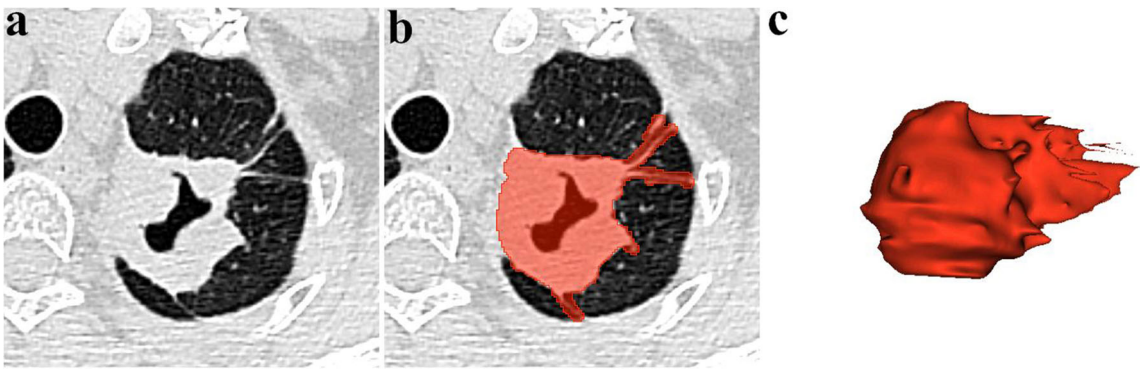


**Fig. 3** Work flow of cavity segmentation. **a** Axial lung CT image shows the cavity. **b** Cavity segmentation and ROI delineation. **c** 3D volume construction based on the ROI

and recall, ranging from 0 to 1. Statistical tests were conducted with $p < 0.05$ as an indicator of statistical significance.

## Results

### Clinical characteristics of the patients

The clinical characteristics and subjective CT findings of the 257 patients are listed in Table 1. As shown in Table 1, the age of the MDR-TB patients and DS-TB patients had significant difference in both the training and testing cohorts ($p < 0.05$). There was no significant difference in gender between patients with MDR-TB and DS-TB in either cohort. There was excellent interobserver agreement with regard to tree-in-bud and small centrilobular nodules, single large nodules and surrounding satellite lesions, consolidation (lobular or subsegmental, segmental or lobar), fibro stripes and calcified nodules ($k = 0.726$, 0.743, 0.797, 0.783 and 0.843, respectively). In terms of these CT findings, the MDR-TB and DS-TB groups had significant differences in single large nodules and calcified nodules ($p < 0.05$). The other remaining findings showed no significant difference in either the training cohort or the testing cohort. As a result, age, single large nodule and calcified nodules were selected to establish a clinical model.

### Radiomics feature selection

There were 300 radiomics features extracted from the MIC values. Then, we identified the 21 most important MDR predictive biomarkers for model construction based on the optimal threshold, which was set to 0.007 by the learning curve (details shown in Supplementary Materials 3 and Fig. S1). The normalised importance of the 21 features is shown in Fig. 4. Detailed values and distributions of these features are shown by the violin plot in Fig. 5.

### Performance outcomes of the clinical, radiomics and combined models

The ROC curves and AUCs of these three models in the training and testing cohorts are shown in Fig. 6. The radiomics model showed a favourable discriminatory ability in the training cohort, with an AUC of 0.844 (95% CI, 0.725 to 0.949)

**Table 1** Clinical characteristics and subjective CT findings from DS-TB and MDR-TB in the training cohort and testing cohort

| Characteristic | Training cohort ($n$ = 187) | | *p value* | Testing cohort ($n$ = 70) | | *p value* |
|---|---|---|---|---|---|---|
| | DS-TB ($n$ = 115) | MDR-TB ($n$ = 72) | | DS-TB ($n$ = 35) | MDR-TB ($n$ = 35) | |
| Gender, *n* (%) | | | | | | |
| Male | 76 (66.09) | 51 (70.83) | 0.499 | 25 (71.43) | 23 (65.71) | 0.797 |
| Female | 39 (33.91) | 21 (29.17) | | 10 (28.57) | 12 (34.29) | |
| Age (mean ± SD years) | 39.59 ± 15.1 | 34.87 ± 11.46 | 0.003* | 36.27 ± 13.12 | 30.16 ± 7.49 | 0.044* |
| Tree-in-bud and small centrilobular nodules, *n* (%) | | | | | | |
| Presence | 84 (73.04) | 47 (65.28) | 0.259 | 22 (62.86) | 20 (57.14) | 0.808 |
| Absence | 31 (26.96) | 25 (34.72) | | 13 (37.14) | 15 (42.86) | |
| Single large nodule and surrounding satellite lesions, *n* (%) | | | | | | |
| Presence | 69 (60.00) | 27 (37.50) | 0.003* | 25 (71.43) | 12 (34.29) | 0.004* |
| Absence | 46 (40.00) | 45 (62.50) | | 10 (28.57) | 23 (65.71) | |
| Consolidation (lobular or subsegmental, segmental or lobar), *n* (%) | | | | | | |
| Presence | 56 (48.70) | 38 (52.78) | 0.587 | 19 (54.29) | 22 (62.86) | 0.628 |
| Absence | 59 (51.30) | 34 (47.22) | | 16 (45.71) | 13 (37.14) | |
| Fibro stripe, *n* (%) | | | | | | |
| Presence | 31 (26.96) | 25 (34.72) | 0.259 | 14 (40.00) | 10 (28.57) | 0.450 |
| Absence | 84 (73.04) | 47 (65.28) | | 21 (60.00) | 25 (71.43) | |
| Calcified nodules, *n* (%) | | | | | | |
| Presence | 25 (21.74) | 5 (6.94) | 0.007* | 11 (31.43) | 3 (8.57) | 0.034* |
| Absence | 90 (78.26) | 67 (93.06) | | 24 (68.57) | 32 (91.43) | |

Note: Differences were assessed by Mann-Whitney *U* test or chi-square test

*SD* standard deviation

*$p < 0.05$

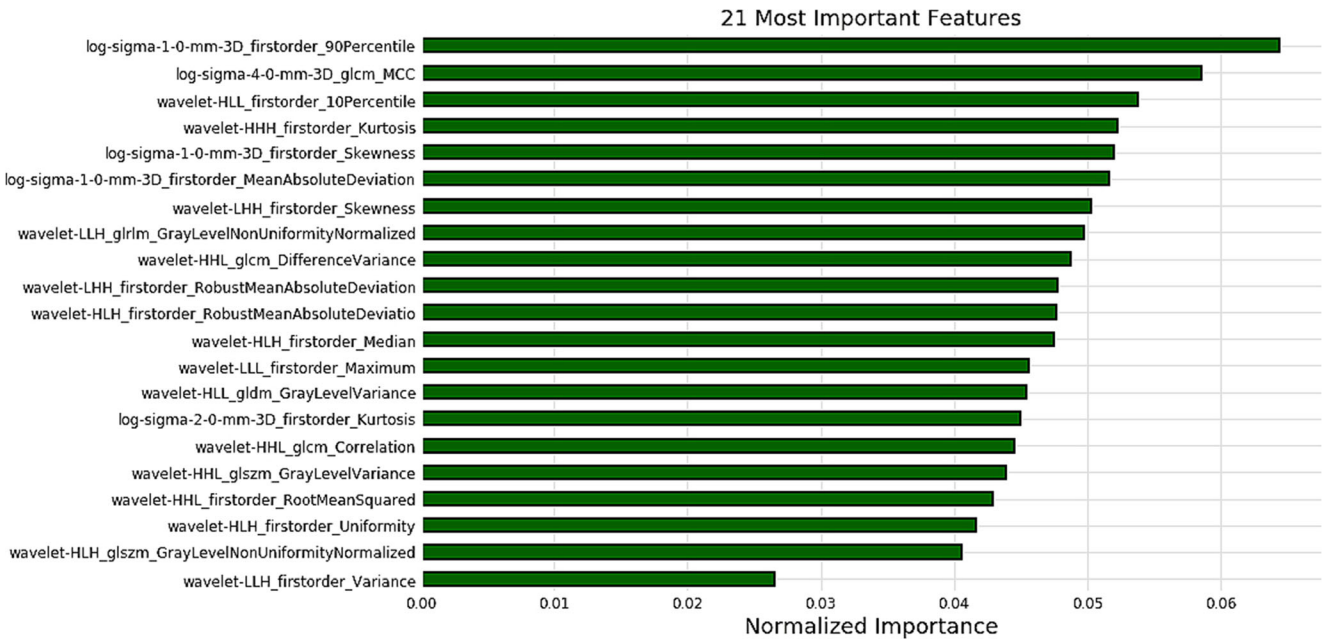


**Fig. 4** The 21 features with the highest normalised importance were selected and included

and F1 score of 0.827, which was confirmed in the testing cohort AUC of 0.829 (95% CI, 0.724 to 0.916) and F1 score of 0.809. The AUCs and F1 score of the clinical model were 0.589 (95% CI, 0.429 to 0.753) and 0.500 (95% CI, 0.366 to 0.631) and 0.483 and 0.465 in the training cohort and testing cohort. Finally, the combined model showed the best performance in the training cohort, with an AUC of 0.881 (95% CI, 0.780 to 0.969) and F1 score of 0.852, which was confirmed in the testing cohort AUC of 0.834 (95% CI, 0.730 to 0.917) and F1 score of 0.824. In addition, the AUCs of the radiomics model (0.844 and 0.829) were much higher than those of the clinical model (0.589 and 0.500) in both the training cohort ($p < 0.0001$) and testing cohort ($p = 0.003$). The AUCs of the combined model achieved the highest values (0.881 and 0.834) among all of the models in both the training and testing cohorts. More specifically, there were significant differences between the combined model and clinical model in both the training ($p < 0.0001$) and testing ($p < 0.0001$) cohorts. Meanwhile, there was no significant difference between the combined model and the radiomics model in the training ($p = 0.165$) and testing ($p = 0.861$) cohorts. The accuracy, precision, recall and F1 score of all three models in the training and testing cohorts are summarised in Table 2.

## Discussion

The main finding of this study is that we developed a machine learning model for differentiating DS-TB from MDR-TB based on CT radiomics features, which achieved good accuracy in an independent external testing cohort. In this retrospective study, radiomics was first used for predicting MDR in cavitary TB. The constructed radiomics model showed good predictive performance (AUC = 0.844 and 0.829), and the performance was much higher than that of the clinical model and was similar to that of the combined model.

An analysis of clinical characteristics showed that the age of MDR-TB patients had significant differences with those of DS-TB patients, whereas sex did not show a significant difference, similar to the results of previous studies [13]. The year of effective anti-TB drug therapies may relate to the difference. Sedentary lifestyle and decreased immune response were other potential factors [16]. An analysis of subjective CT findings revealed that there were significant differences in single large nodule and calcified nodules between the two groups. There have been different viewpoints in this subjective finding. A previous study reported that calcification, large nodules and calcified lymph nodes are more frequently seen in DS-TB than in MDR-TB [13, 17], whereas other reports indicated that calcified granulomas are more common in patients with MDR-TB [6, 18]. This difference may be attributed to the aetiology of MDR-TB (primary or acquired MDR-TB). Tree-in-bud and small centrilobular nodules, consolidation and fibro stripes were the most common CT signs in active TB [19]. Hence, there were no evident differences in these findings between MDR-TB patients and DS-TB patients.

Cavitation is not only an essential imaging characteristic but also a key pathological feature of TB [20]. The presence of cavitation always indicates poor treatment outcomes and risk of drug resistance [21]. In addition, there is a great risk of person-to-person transmission among patients with cavitary

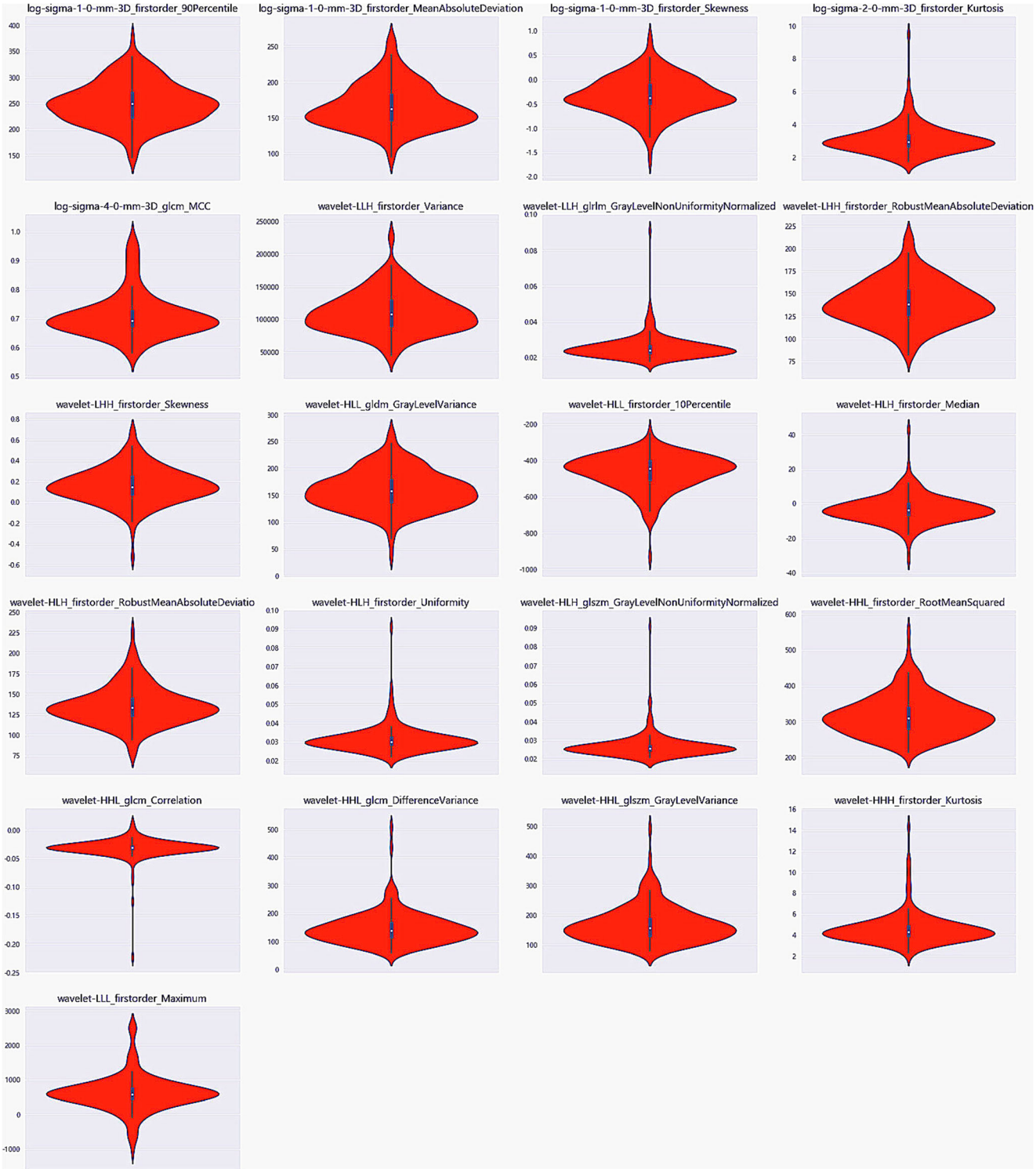


**Fig. 5** These violin plots show the detailed values and distributions of 21 features

TB [22]. The cavities of TB patients are heterogeneous in size, morphology and wall composition [23]. Cavities are commonly seen in the lung upper lobe, especially in immunocompetent adults [24]. The walls of the cavity are associated with the efficacy of therapy. Thinner walls are usually observed after effective treatment, whereas thicker cavity walls present higher concentrations of bacilli in the sputum [25]. Satellite lesions (small nodules distant to the cavity) are also evident in 10–20% of cases [26]. The radiological manifestations of TB cavities are heterogeneous, with some patients having single or multiple cavities surrounded by consolidated, fibronodular or mixed patterns [27]. It has been noted that cavities were

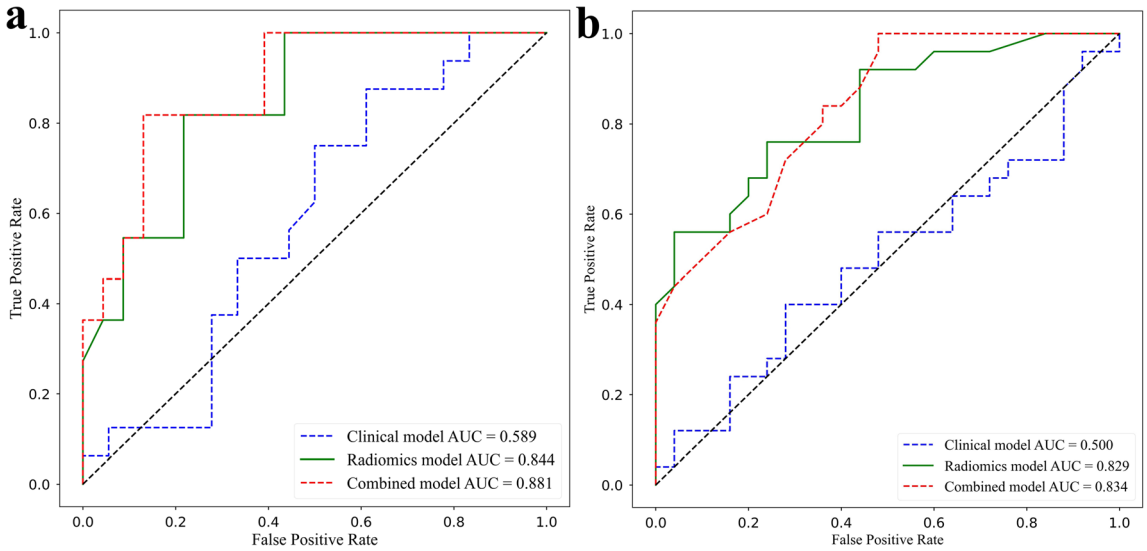


**Fig. 6** ROC curves of the clinical, radiomics, and combined model. **a** Training cohort. **b** Testing cohort

more common for MDR-TB than for DS-TB [28]. Large numbers of mycobacteria harboured in cavities and limited drug penetration may contribute to MDR-TB [6]. It has been reported that larger cavities and more cavities can be easily observed in MDR-TB than in DS-TB [16]. In summary, previous studies have focused on subjective CT findings, and it is difficult to make a diagnosis of MDR-TB based on these findings, especially in cavitary TB. Thus, it is necessary to explore the specific differences in cavitation in both MDR-TB and DS-TB by using quantitative analysis, and radiomics seems promising in this regard.

Our study tried to explore more detailed information on cavitation and quantify not only the common features, such as intensity, shape, size or volume, but also the internal texture features that are unable to be obtained by macroscopic observation but can be reflected by radiomics analysis. Currently, radiomics is widely used due to its potential to build predictive or prognostic models [29]. Many previous studies have confirmed that combining CT- or PET-CT-based radiomics features and clinical characteristics could effectively differentiate solitary TB nodules from lung adenocarcinoma [30, 31]. To our knowledge, no published study has focused on machine learning and the CT radiomics of cavitary TB, as was done in this study. We attempted to differentiate MDR-TB from DS-TB, as this is the most relevant clinical question and is oriented towards a very different kind of medication.

The performance of radiomics models might be substantially influenced by the selection of image features. These features are crucial to the model performance, and too many features could easily lead to model overfitting and reduce the performance of the model. To overcome this crucial problem, we took many steps. First, we used the MIC to select 300 features according to the important relationships between features and different diagnoses of TB. Then, we used the RFC to measure the Gini importance of each feature and identified 21 features. Finally, the RFC was selected to build the radiomics model and combined model. The RFC is a model of integrated learning that consists of a number of decision tree classification machines. Thus, overfitting is avoided, and the prediction accuracy is improved without significantly increasing the computation.

**Table 2** Predictive performance of three models in the training and testing cohorts

| Index | Training cohort | | | Testing cohort | | |
|---|---|---|---|---|---|---|
| | Clinical model | Radiomics model | Combined model | Clinical model | Radiomics model | Combined model |
| AUC | 0.589 | 0.844 | 0.881 | 0.500 | 0.829 | 0.834 |
| Accuracy | 0.559 | 0.765 | 0.794 | 0.440 | 0.720 | 0.767 |
| Precision | 0.538 | 0.800 | 0.829 | 0.366 | 0.823 | 0.838 |
| Recall | 0.438 | 0.856 | 0.877 | 0.639 | 0.796 | 0.811 |
| F1 score | 0.483 | 0.827 | 0.852 | 0.465 | 0.809 | 0.824 |

The clinical model established in our study showed some classification power, whereas it was not precise enough for clinical use. In contrast, our radiomics model showed good performance in the training cohort (AUC = 0.844) and similar performance in the independent testing cohort. The F1 score provides greater focus on the classification of interest, as the score varies more when the dataset is imbalanced (the ratio of positive and negative classifications is different). According to the F1 score, the radiomics and combined model also showed good predictive performance in both training cohort and testing cohort. This implies that the radiomics model will show good stability and generalizability in clinical practice. There were no significant difference between the radiomics model and the combined model. This indicates that radiomics features of cavitation may be dominant factors of MDR. According to the radiomics model, the 21 most important radiomics features were selected and composed of 14 first-order features, 3 GLCM features, 2 GLSZM features, 1 GLRLM feature and 1 GLDM feature. Five features were based on LoG-sigma–transformed images, which emphasises areas of grey-level change, with a low sigma value emphasizing fine textures and a high sigma value emphasising coarse textures. Sixteen features based on wavelet-transformed images were obtained from all possible combinations of applying either a high- or a low-pass filter in each of the three dimensions. The first-order features mainly describe the distribution of voxel intensities within the image region defined by the mask through commonly used and basic metrics. The GLCM feature describes the image's second-order joint probability function. The GLRLM feature quantifies grey-level runs, which are defined as the length (in number of pixels) of consecutive pixels that have the same grey-level value. The GLDM feature quantifies the grey-level dependencies in an image. The GLSZM feature quantifies the grey-level zones in an image. The results indicate that cavitation in DS-TB and MDR-TB is heterogeneous, where many texture features are relevant to MDR.

However, several limitations in this study still exist. First, this study was a retrospective analysis, and the number of patients was not very large. As a result, inherent selection bias may exist. Second, considering the smaller sample size of the testing cohort, future research should focus on prospective studies to increase the generalizability. Third, our study only discussed predictive models for MDR of cavitary TB, but models that can be applied to other types of TB should be further studied.

In conclusion, radiomics features based on cavitation have significant values in distinguishing MDR-TB from DS-TB. Our study may potentially aid in early TB characterisation by integrating the multidisciplinary approach currently based on clinical and subjective CT findings and radiomics. Future large-scale multicentre studies should be carried out to further confirm the preliminary results so that this noninvasive and convenient technique can be used as a diagnostic tool in routine clinical practice.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00330-022-08997-9.

**Acknowledgements** This research was funded by the Beijing Hospitals Authority Clinical Medicine Development of Special Funding. The funding source provided financial support without any influence on the study design and interpretation of data. We would like to thank our groups from many hospitals for the data collection and interpretation.

**Funding** This study has received funding from the Beijing Hospitals Authority Clinical Medicine Development of Special Funding (XMLX202146). The funders had a role in the study design and submitted the paper for publication.

## Declarations

**Guarantor** The scientific guarantor of this publication is Dailun Hou.

**Conflict of interest** The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

**Statistics and biometry** No complex statistical methods were necessary for this paper.

**Informed consent** Written informed consent was waived by the Institutional Review Board.

**Ethical approval** Institutional Review Board approval was obtained.

**Methodology**

- retrospective
- diagnostic study
- multicentre study

## References

1. Organization WH, Programme GT (2021) Global tuberculosis report 2021[M]. Available online: https://www.who.int/teams/global-tuberculosis-programme/tbreports
2. Dheda K, Gumbo T, Maartens G et al (2019) The Lancet Respiratory Medicine Commission: 2019 update: epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant and incurable tuberculosis. Lancet Respir Med 7(9):820–826
3. Ahmad N, Ahuja SD, Akkerman OW et al (2018) Treatment correlates of successful outcomes in pulmonary multidrug-resistant tuberculosis: an individual patient data meta-analysis. Lancet 392(10150):821–834
4. Friedrich SO, Rachow A, Saathoff E et al (2013) Assessment of the sensitivity and specificity of Xpert MTB/RIF assay as an early sputum biomarker of response to tuberculosis treatment. Lancet Respir Med. 1(6):462–470
5. Xie YL, Chakravorty S, Armstrong DT et al (2017) Evaluation of a rapid molecular drug-susceptibility test for tuberculosis. N Engl J Med 377(11):1043–1054

6. Chung MJ, Lee KS, Koh WJ et al (2006) Drug-sensitive tuberculosis, multidrug-resistant tuberculosis, and nontuberculous mycobacterial pulmonary disease in non AIDS adults: comparisons of thin-section CT findings. Eur Radiol 16(9):1934–1941
7. Maduskar P, Hogeweg L, Jong PD et al (2014) Cavity contour segmentation in chest radiographs using supervised learning and dynamic programming. Medical Physics 41(7):071912. https://doi.org/10.1118/1.4881096
8. Gill RR, Matsusoka S, Hatabu H (2010) Cavities in the lung in oncology patients: imaging overview and differential diagnoses. Applied Radiology 39(6):10–21
9. Ralph AP, Ardian M, Wiguna A et al (2010) A simple, valid, numerical score for grading chest x-ray severity in adult smear-positive pulmonary tuberculosis. Thorax 65(10):863–869
10. Perrin FMR, Woodward N, Phillips PPJ et al (2010) Radiological cavitation, sputum mycobacterial load and treatment response in pulmonary tuberculosis. Int J Tuberc Lung Dis 14(12):1596–1602
11. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278(2):563–577
12. Kickingereder P, Burth S, Wick A et al (2016) Radiomic profiling of glioblastoma: identifying an imaging predictor of patient survival with improved performance over established clinical and radiologic risk models. Radiology. 280(3):880–889
13. Li D, He W, Chen B, Lv P (2017) Primary multidrug-resistant tuberculosis versus drug-sensitive tuberculosis in non-HIV-infected patients: comparisons of CT findings. PLoS One. 12(6): e0176354. https://doi.org/10.1371/journal.pone.0176354
14. Reshef DN, Reshef YA, Finucane HK et al (2011) Detecting novel associations in large data sets. Science. 334(6062):1518–1524
15. Breiman L (2001) Random forests, machine learning. Springer 45(1):5–32
16. Jihoon C, Yun LH, Soo LK et al (2009) Radiological findings of extensively drug-resistant pulmonary tuberculosis in non-AIDS adults: comparisons with findings of multidrug-resistant and drug-sensitive tuberculosis. Korean J Radiol 10(3):207–216
17. Yeom JA, Jeong YJ, Jeon D et al (2009) Imaging findings of primary multidrug-resistant tuberculosis: a comparison with findings of drug-sensitive tuberculosis. J Comput Assist Tomogr 33(6):956–960
18. Kim HC, Jin MG, Lee HJ et al (2004) Multidrug-resistant tuberculosis versus drug-sensitive tuberculosis in human immunodeficiency virus-negative patients: computed tomography features. J Comput Assist Tomogr 28(3):366–371
19. Ko JM, Park HJ, Kim CH, Song SW (2015) The relation between CT findings and sputum microbiology studies inactive pulmonary tuberculosis. Eur J Radiol. 84(11):2339–2344
20. Rozenshtein A, Hao F, Starc MT et al (2015) Radiographic appearance of pulmonary tuberculosis: dogma disproved. AJR Am J Roengenol 204(5):974–978
21. Geng E, Kreiswirth B, Burzynski J et al (2005) Clinical and radiographic correlates of primary and reactivation tuberculosis. JAMA 293(22):2740–2745
22. Turner RD, Chiu C et al (2017) Tuberculosis infectiousness and host susceptibility. J Infect Dis. 216(suppl_6):S636–S643
23. Ors F, Deniz O, Bozlar U et al (2007) High-resolution CT findings in patients with pulmonary tuberculosis: correlation with the degree of smear positivity. J Thorac Imaging 22(2):154–159
24. Won-Jung K, Joo JY, Jung KO et al (2010) Chest radiographic findings in primary pulmonary tuberculosis: observations from high school outbreaks. Korean J Radiol 11(6):612–617
25. Rosenthal A, Gabrielian A, Engle E et al (2017) The TB portals: an open-access, web-based platform for global drug-resistant-tuberculosis data sharing and analysis. J Clin Microbiol. 55(11): 3267–3282
26. Krysl J, Korzeniewska-Kosela M, Müller NL, FitzGerald JM (1994) Radiologic features of pulmonary tuberculosis: an assessment of 188 cases. Can Assoc Radiol J. 45(2):101–107
27. Leung AN (1999) Pulmonary tuberculosis: the essentials. Radiology. 210(2):307–322
28. Kim W, Lee KS, Kim HS et al (2016) CT and microbiologic follow-up in primary multidrug-resistant pulmonary tuberculosis. Acta Radiol. 57(2):197–204
29. Mayerhoefer ME, Materka A, Langs G et al (2020) Introduction to radiomics. J Nucl Med. 61(4):488–495
30. Feng B, Chen X, Chen Y et al (2020) Radiomics nomogram for preoperative differentiation of lung tuberculoma from adenocarcinoma in solitary pulmonary solid nodule. Eur J Radiol. 128: 109022. https://doi.org/10.1016/j.ejrad.2020.109022
31. Hu Y, Zhao X, Zhang J, Han J, Dai M (2021) Value of 18F-FDG PET/CT radiomic features to distinguish solitary lung adenocarcinoma from tuberculosis. Eur J Nucl Med Mol Imaging. 48(1):231–240