Original Research

Whole-Lung Computed Tomography Radiomics-Based Machine Learning Classification of Nontuberculous Mycobacterial Lung Disease Across Geographically Distinct Cohorts

Anusha Kanagala, MS[1] Bryan Garcia, MD[2] Taru S. Dutt, PhD[3] Stéphanie M. Aguilera, BS[4,5] Abhiraj S. Pudhota, MS[2] Dhiral D. Panjwani, MS[1] Nikhil Dukkipati, MBBS[2] Amit Gaggar, MD[2] Threnesan Naidoo, MD[6,7] Leon Jololian, PhD[1] Surya P. Bhatt, MD[2,4] Camilla Margaroli, PhD[3*] Sandeep Bodduluri, PhD[1,2,4*]

[1]Department of Electrical and Computer Engineering, University of Alabama at Birmingham, Birmingham, Alabama, United States

[2]Division of Pulmonary, Allergy and Critical Care Medicine, University of Alabama at Birmingham, Birmingham, Alabama, United States

[3]Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, United States

[4]Center for Lung Analytics and Imaging Research (CLAIR), University of Alabama at Birmingham, Birmingham, Alabama, United States

[5]Department of Biomedical Engineering, University of Alabama at Birmingham, Birmingham, Alabama, United States

[6]School of Pathology, Faculty of Medicine and Health Sciences, Walter Sisulu University, Mthatha, Eastern Cape, South Africa

[7]Africa Health Research Institute, Durban, South Africa

**(Co-senior authors)*

Address correspondence to:
Sandeep Bodduluri, PhD
University of Alabama at Birmingham
Division of Pulmonary, Allergy and Critical Care Medicine
THT 422, 1720, 2nd Avenue South
Birmingham, AL 35294
Email: sbodduluri@uabmc.edu

***Running Head*: Whole-Lung CT Radiomics for NTM Classification**

***Keywords*:** computed tomography; nontuberculosis mycobacteria lung disease; radiomics; segmentation; linear discriminant analysis;

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

***Abbreviations*:**
NTM-LD – Nontuberculosis Mycobacteria Lung Disease; CT – Computed Tomography; LDA – Linear Discriminant Analysis; AUC – Area Under the Curve; PTB – Pulmonary Tuberculosis; CF – Cystic Fibrosis; GLCM – Gray-Level Co-occurrence Matrices ; GLRLM – Gray-Level Run-Length Matrices ; GLSZM – Gray-Level Size Zone Matrices ; NGTDM – Neighboring Gray-Tone Difference Matrices; ;SHAP – SHapley Additive exPlanations

***Funding Support*:** This work was supported by NHLBI K01HL163249 (SB), NHLBI R01 HL151421 (SPB), UAB Research and Development Program Pilot Program – ROWE19RO (DAVIS)(CM), NIH-NIAID R00AI173401 (TSD). TSD served as a Director of Bacterial Countermeasures in Solaris Vaccines. No tobacco company funded.

Date of Acceptance: October 20, 2026 | ***Published Online Date***: November 12, 2026

Citation: Kanagala A, Garcia B, Dutt TS, et al. Whole-lung computed tomography radiomics-based machine learning classification of nontuberculous mycobacterial lung disease across geographically distinct cohorts. *Chronic Obstr Pulm Dis.* 2026; Published online November 12, 2026.
https://doi.org/10.15326/jcopdf.2026.0867

This article has an online supplement.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

ABSTRACT

Nontuberculous mycobacterial lung disease (NTM-LD) is highly heterogeneous, geographically and etiologically, hindering effective timely identification. Prior CT radiomics studies require manual segmentation of pathology. We developed a whole-lung CT radiomics-based machine learning approach and identified common features across two geographically distinct NTM-LD cohorts. We included 1,300 chest CT scans from China (871 TB; 429 NTM, Dataset 1) and 173 independent NTM cohort from the University of Alabama at Birmingham (UAB), US. Whole-lung regions were automatically segmented on each scan, and 85 quantitative radiomic features were extracted using a standardized image-processing pipeline. We evaluated two frameworks to assess model performance and generalizability: (1) training on Dataset 1 with external validation on Dataset 2, and (2) training on the combined cohort. Linear discriminant analysis (LDA) was used as the primary classification method. Cross-cohort concordance analysis was performed to evaluate the reproducibility of radiomic features across datasets. In Scenario 1, the LDA classifier trained on Dataset 1 achieved an AUC of 0.79 (95% CI, 0.73–0.84) with high specificity (0.91). On the external UAB cohort, the model achieved an AUC of 0.94 (95% CI, 0.90–0.97). In Scenario 2, the combined cohort model achieved an AUC of 0.81 (95% CI, 0.76–0.85) with improved sensitivity (0.61) and precision (0.82). Feature importance analysis identified 16 features consistently ranked among the top 20 in both scenarios, predominantly texture-based descriptors reflecting distinct parenchymal patterns between mycobacterial species. In conclusion, whole-lung CT radiomics enables interpretable NTM-LD classification across geographically distinct populations without manual annotation. Suggesting population-independent parenchymal signatures of NTM-LD.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

INTRODUCTION

Nontuberculous mycobacterial lung disease (NTM-LD) and pulmonary tuberculosis (PTB) represent distinct clinical entities requiring fundamentally different treatment and management approaches, yet their differentiation poses substantial diagnostic challenges.[1,2] At present, distinguishing NTM-LD from PTB often relies on sputum-based microbiologic testing, including smear microscopy, mycobacterial culture, and nucleic acid amplification tests such as Xpert MTB/RIF. However, culture remains time-consuming because mycobacteria require weeks to grow, and molecular assays can detect DNA from nonviable bacilli, limiting their ability to distinguish active infection from residual bacterial material. The global incidence of NTM-LD has increased substantially over recent decades, with prevalence now exceeding that of TB in many developed nations including the United States, while TB remains a leading cause of infectious disease mortality worldwide.[3,4] Both conditions manifest with overlapping clinical presentations and radiologic findings on chest CT imaging, including nodular infiltrates, cavitary lesions, bronchiectasis, and tree-in-bud patterns.[5,6] Early differentiation is essential not only for appropriate antimicrobial selection but also for infection control measures, as faster identification can reduce community exposure to PTB and avoid unnecessary airborne isolation in patients with NTM-LD. Additional challenges in the timely clinical diagnosis of NTM-LD are posed by the microbiological diversity of NTM, which includes more than 150 species, and the etiological differences of underlying lung diseases acting as predisposing factors for NTM-LD (PMID: 35914904). These challenges result in a highly geographically heterogeneous disease affecting people worldwide. One of the most accessible and rapid approaches for the initial evaluation of mycobacterial lung disease is radiological imaging, which often provides the earliest clinically actionable information, particularly when definitive microbiologic diagnosis is

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

delayed by low bacterial burden and prolonged culture turnaround times. Traditional radiologic interpretation relies on subjective visual assessment, which demonstrates variable interobserver agreement.[8] Further, a formal scoring or staging system for NTM-LD and PTB does not currently exist, hindering the ability to uniformly monitor lung remodeling and disease progression. While certain imaging features have been associated with each condition, PTB characteristically presenting with upper lobe predominance and cavitation, and NTM-LD with middle lobe and lingular bronchiectasis, substantial overlap exists, and prior studies have shown that these features do not offer adequate markers for reliable discrimination.[5,9]

Radiomics, the high-throughput extraction of quantitative imaging features, has emerged as a promising approach to augment clinical decision-making.[10,11] Recent studies have applied radiomics to NTM-LD versus PTB differentiation with encouraging results. Yan and colleagues[12] demonstrated that radiomic features extracted from cavities achieved AUC values exceeding 0.84 for differentiating NTM-LD from PTB. Similarly, consolidation-based radiomics models have achieved AUC values above 0.90.[13] However, these lesion-specific approaches require manual delineation of individual cavities or consolidations by experienced radiologists, a labor-intensive process that limits scalability and introduces potential inter-reader variability.

Deep learning methods have also been applied to this classification task. Wang and colleagues[14] reported an AUC of 0.71 using the BoTNet50 convolutional neural network, while 3D-ResNet models have achieved AUC values of 0.78-0.86 in a cohort from China.[15] Although promising, these deep learning approaches function as "black boxes" with limited interpretability and require substantial computational resources, potentially hindering clinical adoption. A critical gap exists in the literature: no study has systematically evaluated whole-lung radiomic features for

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

NTM-LD classification across geographically and demographically distinct populations. Whole-lung analysis offers various potential advantages: automated lung segmentation eliminates the need for manual lesion annotation, captures global parenchymal patterns that may not be apparent in individual lesions, and facilitates deployment in settings where expert radiologic annotation is unavailable. In this study, we evaluated whether whole-lung radiomic features could achieve robust NTM-LD classification performance across cohorts from China and the United States while offering substantially greater scalability and interpretability than existing approaches.

METHODS

Study Population

This study utilized two geographically distinct cohorts. Dataset 1 comprised a publicly available collection of chest CT scans from patients diagnosed with pulmonary mycobacterial disease from China.[15] This dataset represents the first comprehensive CT imaging resource combining NTM-LD and PTB cases with multispecies mycobacterial information. The dataset includes anonymized imaging data and clinical descriptors obtained prior to treatment initiation. Cases were selected with microbiologically confirmed diagnoses and CT scans of adequate quality for quantitative analysis. The final study cohort comprised 1,300 patients: 871 with culture-confirmed PTB and 429 with culture-confirmed NTM-LD, representing the largest radiomics cohort assembled for this classification task to date. Dataset 2 comprised an independent clinical NTM-LD cohort (N=173) acquired at the University of Alabama at Birmingham (UAB; IRB-300010437). (PMID: 41046132)[18] NTM-LD patients were identified from the UAB

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Bronchiectasis Research Registry as having pulmonary infections with NTM but not with TB based on prior sputum cultures in the year prior. Patients with cystic fibrosis (CF) were excluded because the underlying structural lung abnormalities characteristic of CF, including diffuse bronchiectasis, mucus plugging, and progressive parenchymal destruction, can confound radiomic feature extraction and introduce disease-related imaging signatures unrelated to mycobacterial infection. Patients with extrapulmonary manifestations of NTM (i.e., skin and soft tissue) were also excluded from the study. CT scans were obtained retrospectively from the most recent visit. Clinical data was collected in October 2024. Notably, Dataset 2 did not include any TB cases, reflecting the low prevalence of active PTB in the United States; the role of this cohort in our analysis is detailed below under the experimental design. Bronchiectasis severity in the UAB cohort was assessed using the Reiff score (Supplementary Table S1), a validated CT-based scoring system (range 0-18). The distribution was as follows: mild (Reiff 0-3), 47 patients (27.2%); moderate (Reiff 4-7), 92 patients (53.1%); severe (Reiff >7), 28 patients (16.2%); no score available, 6 patients (3.5%). The mean Reiff score was 5.2 ± 2.6 (median 5.0), indicating predominantly mild-to-moderate bronchiectasis. Additionally, 35 patients (20.2%) were asymptomatic at the time of CT acquisition, with a lower mean Reiff score (3.6 ± 2.2), and 41 patients (23.7%) had cavitary disease.

Ethics Statement:

This study was approved by the University of Alabama at Birmingham Institutional Review Board (IRB-300010437) for retrospective use of the UAB Bronchiectasis Research Registry cohort (Dataset 2).

CT Imaging Protocol

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

CT imaging was performed using standardized protocols across participating centers.[15] The scan range encompassed the entire thorax from lung apex to base, with patients positioned supine during end-inspiratory breath-hold. Dataset 1 (China) scans were acquired at 512 x 512 matrix resolution on GE BrightSpeed and Canon Aquilion Prime 128 machines, 120 kVp, with thin slices (1.0–1.25 mm) to capture fine details of the lung tissue. Dataset 2 (UAB) scans were acquired at 512 x 512 matrix resolution across multiple clinical scanners reflecting real-world imaging heterogeneity: GE Medical Systems (1.25–5.0 mm slice thickness, 70 – 140 kVp), Siemens (0.6 - 5.0 mm, 100 – 150 kVp), Philips (1.5 – 5.0 mm, 100 – 120 kVp), Canon/Toshiba (1.0–5.0 mm, 120 kVp), and Hitachi (1.25 mm, 120 kVp). This multi-vendor, multi-model imaging heterogeneity reflects typical real-world clinical practice.

Automated Lung Segmentation

In contrast to prior radiomics studies requiring manual lesion delineation,[12,13] we employed fully automated deep learning-based lung segmentation.[16] This algorithm employs a U-Net architecture trained on diverse CT datasets to achieve robust segmentation across varying imaging protocols and disease presentations. Segmentation quality was visually assessed for each case.

Radiomic Feature Extraction

Radiomic feature extraction was performed using the PyRadiomics (version 3.1.1) library.[10,11] A total of 85 quantitative features was extracted from the segmented whole-lung parenchyma, capturing global disease burden rather than individual lesion characteristics. First-order statistics captured intensity distribution characteristics, including mean, median, standard deviation, skewness, kurtosis, and energy. Shape-based features quantified three-dimensional lung

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

morphology, including volume, surface area, sphericity, and elongation. Texture features were derived from gray-level co-occurrence matrices (GLCM), gray-level run-length matrices (GLRLM), gray-level size zone matrices (GLSZM), and neighboring gray-tone difference matrices (NGTDM). Images were processed at native voxel spacing without resampling, preserving the original spatial resolution (China cohort: ~0.8 × 0.8 × 1.25 mm; UAB cohort: ~0.6 × 0.6 × 1.25 mm). Intensity discretization used a fixed bin width of 25 HU. No image-level intensity normalization was applied, as both datasets utilized calibrated Hounsfield Unit values from clinical CT scanners. All extracted radiomic features underwent z-score normalization (zero mean, unit variance) computed on the training set and applied to the test set, harmonizing inter-scanner differences in feature magnitude while preserving discriminative information.

Statistical Analysis and Model Development

Two experimental scenarios were designed to evaluate the generalizability of whole-lung radiomic features for NTM-LD classification. In Scenario 1 (Cross-Cohort External Validation), the LDA classifier was trained on Dataset 1 (China cohort; N=1,300; NTM=429, TB=871) using a 70/30 stratified train-test split (train=909, test=391). The trained model was then applied to Dataset 2 (UAB cohort; N=173, all NTM-LD) as an external validation set to evaluate whether NTM-LD radiomic signatures learned from the Chinese cohort generalize to an American NTM-LD population. In Scenario 2 (Combined Cohort Analysis), Datasets 1 and 2 were combined into a single cohort (N=1,473; NTM=602, TB=871) with a group-aware 70/30 stratified split maintaining proportional representation from each source dataset and disease class (Dataset 1 TB: train=610, test=261; Dataset 1 NTM: train=300, test=129; Dataset 2 NTM: train=121, test=52). Linear discriminant analysis (LDA) was selected as the classification algorithm owing to its

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

computational efficiency with high-dimensional data, interpretability, and established performance in medical imaging applications.[17] LDA offers explicit feature weights that enhance clinical interpretability.

Model performance was evaluated on the independent test set using Area under the Curve (AUC) and precision.[19,20] Additional metrics included sensitivity, specificity, and overall accuracy. The F1 score (harmonic mean of precision and sensitivity, range 0-1) was also reported, as it is particularly informative for imbalanced discrimination tasks. Balanced accuracy and Matthews correlation coefficient (MCC) were computed alongside AUC, sensitivity, specificity, and F1 score to account for class imbalance in Dataset 1. Confidence intervals for AUC were computed using bootstrap resampling with 1,000 iterations. Feature importance was assessed using SHapley Additive exPlanations (SHAP) values to identify radiomic features contributing most substantially to classification decisions.[21] Performance was benchmarked against published results from the BoTNet50 deep learning model evaluated on the same dataset.[14] All statistical analyses were performed using Python 3.8 with scikit-learn.[17] Statistical significance was defined as $P < 0.05$ (two-sided).

RESULTS

Cohort Characteristics

Table 1 summarizes the characteristics of the two study cohorts. Dataset 1 (China cohort) comprised 1,300 patients with culture-confirmed pulmonary mycobacterial disease, including 871 PTB cases (67.0%) and 429 NTM-LD cases (33.0%), collected from multiple centers in China. This publicly available dataset contains multispecies mycobacterial CT imaging data acquired

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

across several Chinese institutions, providing broad geographic and demographic representation within the East Asian population. Patients with NTM-LD were older than those with PTB (mean age 55.3 ± 17.0 vs. 47.3 ± 18.3 years, $p < 0.001$) and more often female (38.2% vs. 28.2%, $p < 0.001$), although males predominated in both groups. Race/ethnicity, BMI, pulmonary function, smoking status, and symptom data were not available except strain types for NTM (Supplementary Table S2) for this cohort. All CT examinations in Dataset 1 were acquired on GE BrightSpeed scanners with a slice thickness of 1.0–1.25 mm. Dataset 2 (UAB cohort) comprised 173 patients with culture-confirmed NTM-LD exclusively, identified from the UAB Bronchiectasis Research Registry in Birmingham, Alabama, United States. The UAB cohort was predominantly female (71.1%), non-Hispanic White (86.1%), with a mean age of 69.8 ± 11.8 years and mean BMI of 22.8 ± 5.8 kg/m². Mean percent predicted FEV_1 was 71.3% ± 23.3%, reflecting moderate airflow limitation consistent with underlying bronchiectasis. Among UAB participants, 43.4% were current or former smokers. The majority (79.8%) were symptomatic at the time of CT acquisition.

Classification Performance

The whole-lung radiomics-based LDA classifier demonstrated strong discriminative performance on the Dataset 1 internal test set (Table 2). In Scenario 1, the model achieved an AUC of 0.79 (95% CI, 0.73–0.84), sensitivity of 0.47 (95% CI, 0.39–0.56), specificity of 0.91 (95% CI, 0.88–0.94), accuracy of 0.767 (95% CI, 0.73–0.81), and an F1 score of 0.57. When applied to the external UAB cohort (Dataset 2, N=173), the model achieved an accuracy of 0.94 (95% CI, 0.90–0.97), correctly detecting 162 of 173 NTM-LD cases (sensitivity 93.6%). This cross-cohort performance demonstrates that NTM-LD radiomic signatures learned from the Chinese cohort

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

generalize effectively to an American NTM-LD population despite differences in demographics, mycobacterial species distributions, and imaging protocols.

In Scenario 2 (Combined Cohort; Table 3), the model achieved an AUC of 0.81 (95% CI, 0.76–0.85), sensitivity of 0.61 (95% CI, 0.54–0.69), specificity of 0.90 (95% CI, 0.87–0.94), accuracy of 0.79 (95% CI, 0.75–0.82), and an F1 score of 0.70. The improved AUC (0.81 vs. 0.79), sensitivity (0.61 vs. 0.47), and F1 score (0.70 vs. 0.57) relative to Scenario 1 indicate that incorporating NTM-LD cases from a second geographic population enriches the training data and improves the model's ability to capture heterogeneous NTM-LD presentations.

Receiver operating characteristic curve analysis demonstrated stable performance across threshold ranges, with particularly reliable behavior at high specificity levels (>0.90), where the model preserved its discriminative strength (Figure 2). Precision-recall analysis revealed consistently high precision values, remaining above 0.70 at elevated recall thresholds. Model robustness was further supported by bootstrap analysis, which produced narrow confidence intervals and demonstrated stable generalization across resampled datasets.

Comparison with Prior Approaches

Table 4 presents a comparison of our whole-lung approach with published methods for NTM-TB differentiation. Cavity-specific radiomics models have achieved AUC values of 0.84,[12] and consolidation-based radiomics achieved AUC values of 0.90, but both require manual lesion-specific annotation.[13] Our whole-lung approach achieves competitive performance (AUC = 0.79 in Scenario 1; AUC = 0.81 in Scenario 2) while requiring only automated lung segmentation, substantially reducing annotation burden and enabling application to all patients regardless of specific lesion morphology.

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Pre-Proof

Feature Importance Analysis

SHAP-based feature importance analysis identified texture heterogeneity measures and morphologic descriptors as the most influential parameters for distinguishing NTM-LD from TB (Figure 3). The top-ranked features organized by radiomic class were: (1) Shape features: Voxel Volume and Mesh Volume, reflecting three-dimensional lung volume and disease-related volumetric changes; (2) GLSZM texture features: Large Area Emphasis and Zone Variance, quantifying the prevalence of large homogeneous zones and spatial heterogeneity of parenchymal texture patterns; (3) GLCM texture features: Inverse Difference Moment (Idm), Inverse Difference (Id), Autocorrelation, Difference Average, Inverse Variance, Joint Entropy, and Entropy, collectively capturing local textural uniformity, spatial intensity correlations, and gray-level complexity; (4) GLRLM texture features: Run Percentage, High Gray Level Run Emphasis, Run Length NonUniformity Normalized, Long Run Emphasis, and Run Variance, encoding the distribution of consecutive voxel runs and reflecting fine-to-coarse granularity of parenchymal patterns; and (5) First-order features: Short Run High Gray Level Emphasis, indicating small bright foci such as nodules or calcifications. Critically, 16 of these features were consistently ranked among the top 20 in both experimental scenarios, providing strong evidence for population-independent radiomic signatures of NTM-LD.

Cross-Cohort Validation with Shared Radiomic Features

We evaluated the generalizability of discriminative radiomic features across geographically distinct cohorts. We performed cross-cohort validation using a shared feature set derived from both datasets. Each dataset was partitioned into training (70%) and testing (30%) subsets with

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

stratified sampling. Feature selection was performed independently within each training partition, and the intersection yielded 16 shared radiomic features common to both geospatial cohorts (Table 5). These 16 shared features spanned four radiomic classes. Shape descriptors (VoxelVolume, MeshVolume) captured three-dimensional lung morphology. GLCM texture features (Idm, Id, Autocorrelation, JointEntropy, DifferenceAverage, InverseVariance, Entropy) encoded local gray-level spatial dependencies and textural complexity. GLRLM features (HighGrayLevelRunEmphasis, RunLengthNonUniformityNormalized, LongRunEmphasis, RunVariance, RunPercentage) characterized the distribution and heterogeneity of consecutive voxel intensity runs. GLSZM features (LargeAreaEmphasis, ZoneVariance) quantified the size and variability of connected homogeneous zones. The predominance of texture-based features across both scenarios suggests that NTM-LD and TB produce fundamentally different global patterns of parenchymal spatial heterogeneity detectable through whole-lung analysis.

An LDA classifier trained and evaluated on Dataset 1 using only these 16 features retained 94.4% of the full model's discriminative performance (AUC 0.74 [95% CI, 0.69–0.80] vs. 0.79 [95% CI, 0.73–0.84]). Notably, the reduced model surpassed the full model in specificity (0.95 [95% CI, 0.93–0.98] vs. 0.91 [95% CI, 0.88–0.94]), with accuracy of 0.78 (95% CI, 0.74–0.82). The sensitivity trade-off (0.43 [95% CI, 0.35–0.52] vs. 0.65 [95% CI, 0.58–0.72]) yielded an F1 of 0.57.

The 16 texture and shape features were predominantly texture features, GLCM and GLRLM descriptors encoding spatial gray-level dependencies and run-length distributions, retaining 94.4% of discriminative capacity with overlapping confidence intervals (AUC 0.74 [95% CI, 0.69–0.80] vs. 0.79 [95% CI, 0.73–0.84]). The concurrent improvement in specificity suggests that feature reduction functions as implicit regularization, isolating high-confidence TB-specific

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

structural patterns while attenuating overfitting to cohort-specific variance. The reduction in sensitivity likely reflects exclusion of features that captured heterogeneous NTM presentations in individual cohorts but did not replicate across populations, consistent with the greater radiographic variability characteristic of NTM disease.

DISCUSSION

In this study, we demonstrated that whole-lung radiomic features extracted from chest CT scans enable robust classification of NTM-LD across geographically and demographically distinct populations using an interpretable machine learning approach that requires no manual lesion annotation. The radiomics-based LDA classifier achieved strong discriminative performance both in a single-cohort setting (AUC = 0.79) and when applied cross-nationally (external validation accuracy = 0.94), while a combined cohort model further improved classification (AUC = 0.81). These results exceeded published deep learning methods[14] while approaching performance of lesion-specific radiomics models[12,13] all without requiring the labor-intensive manual delineation of individual cavities or consolidations.

A key finding is the strong generalizability of NTM-LD radiomic signatures across populations. When the model trained exclusively on the Chinese cohort (Dataset 1) was applied to the American UAB cohort (Dataset 2), it achieved an accuracy of 0.94, correctly identifying 162 of 173 NTM-LD cases. This cross-cohort performance is notable because the two populations differ substantially in demographics, geographic origin, likely mycobacterial species distributions, and imaging protocols. The fact that whole-lung radiomic features captured NTM-LD patterns that transcend these sources of variability suggests that NTM-LD produces

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

fundamental, population-independent alterations. However, this finding should be interpreted cautiously, as the high external AUC may also reflect characteristics of the UAB cohort and imaging acquisition. Specifically, the bronchiectasis-registry origin of the UAB cohort may preferentially capture patients with more extensive or severe structural disease, while differences in CT vendors and acquisition protocols may contribute additional variation in imaging features. These factors are plausible and non-exclusive contributors to the observed external performance and therefore should be considered alongside, rather than excluded by, the possibility of a population-independent NTM-LD signal.

The identification of 16 radiomic features consistently ranked among the top 20 in both experimental scenarios further reinforces this conclusion. The discriminative features identified through SHAP capture global parenchymal patterns consistent with known pathophysiologic differences between TB and NTM.[5,6] TB characteristically produces more extensive tissue destruction with coarse, heterogeneous parenchymal changes, whereas NTM lung disease, particularly that caused by Mycobacterium avium complex, may demonstrate more diffuse or nodular involvement with relatively preserved parenchymal architecture.[1,23] The predominance of texture-driven and intensity-driven features in our classification model quantitatively captures these radiologic distinctions at a whole-lung level. The importance of shape-based metrics further suggests that the overall geometric configuration of affected lung contributes meaningfully to differentiation between these disease entities.

Deep learning methods, while achieving moderate performance (AUC = 0.71-0.78),[14,15] function as "black boxes" that provide limited insight into classification rationale. Our feature-based radiomics model with explicit SHAP-based feature importance quantification identifies specific imaging characteristics driving predictions. The predominance of texture heterogeneity

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

features (zone variance, inverse difference moment) aligns with known differences in parenchymal destruction patterns between TB and NTM,[5,23] providing clinical face validity that may enhance physician trust and adoption. Importantly, cross-cohort validation demonstrated that a reduced set of 16 shared features retained 94.4% of the full model's performance while improving specificity. These shared features were predominantly texture-based (GLCM and GLRLM), capturing spatial gray-level heterogeneity across the lung parenchyma. This supports that global texture patterns are robust and reproducible markers of TB versus NTM across populations. The improved specificity with fewer features suggests implicit regularization, reducing cohort-specific noise, whereas the decrease in sensitivity likely reflects exclusion of features associated with the more heterogeneous presentation of NTM disease.

Recent work by Biciusca and colleagues[24] highlighted the challenge of differentiating radiologic features among slow-growing NTM species, finding that *M. intracellulare* exhibited a tendency toward higher overall radiologic severity scores compared with other species. Such species-level heterogeneity within NTM infections underscores the complexity of this diagnostic challenge. Likewise, heterogeneity in the etiology of NTM-LD may lead to distinct patterns of disease progression that are currently unknown due to the lack a formal scoring/staging system and standardized analytic tools. Our model's ability to distinguish NTM as a group from TB, despite this intra-group variability, suggests that quantitative radiomic features capture fundamental differences in parenchymal involvement patterns between these mycobacterial categories. The high specificity (0.91) of our classifier warrants particular attention, as it indicates robust ability to correctly classify non-NTM cases and supports confident clinical decision-making.[4] However, given the class imbalance in Dataset 1 (871 PTB [67.0%] vs. 429 NTM-LD [33.0%]), a trivial classifier always predicting TB would achieve an accuracy of 0.67 and a

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

specificity of 1.00; we therefore also report balanced accuracy and Matthews correlation coefficient (Table X) to reflect discriminative gain over this baseline. The combined-cohort design of Scenario 2 partly addressed this imbalance, improving sensitivity from 0.47 to 0.61 while maintaining specificity at 0.90. A highly specific classifier minimizes misclassification, ensuring that patients with NTM-LD are correctly identified and can receive timely, appropriate therapeutic regimens.[2] The precision-recall characteristics of our model support its utility in clinical workflows where confident NTM-LD identification is needed alongside microbiologic confirmation.

Our study has several limitations. First, our analysis was retrospective, utilizing a publicly available dataset and a clinical registry cohort, both with inherent selection biases given the cohort's location. Prospective and longitudinal validation in diverse clinical settings is necessary before clinical implementation. Second, Dataset 2 did not include TB cases, precluding traditional sensitivity/specificity evaluation on this cohort for NTM vs. TB discrimination; its role was to validate the generalizability of NTM-LD radiomic signatures rather than to serve as a full NTM vs. TB discrimination test. Moreover, models were trained/validated only for binary NTM-LD vs. PTB discrimination; performance against other radiographically overlapping entities (bacterial, fungal, viral pneumonia, other bronchiectasis etiologies) is untested. Third, we analyzed whole-lung radiomic features without lesion-specific analysis; a combined approach incorporating both whole-lung and lesion-level features might further improve classification accuracy. Fourth, while our model outperformed deep learning approaches evaluated on the same dataset, more advanced neural network architectures may achieve superior performance. Fifth, clinical integration would require validation of workflow feasibility, comparison with expert radiologist interpretations, and assessment of the model's impact on clinical decision-making.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Finally, reconstruction kernels and radiation dose varied across scanners in Dataset 2 (Supplementary Table S3), and these differences in image noise may have affected radiomic feature values. These parameters were not adjusted for in our analysis, and CTDIvol was unavailable for more than half of the scans. Future work should evaluate kernel- and dose-specific harmonization.

CONCLUSION

Whole-lung CT radiomics enables interpretable machine learning-based classification of NTM-LD across geographically and demographically distinct populations without requiring manual lesion annotation. This automated approach achieves performance exceeding deep learning methods and approaching that of lesion-specific radiomics models while offering substantially greater scalability and interpretability. The identification of 16 shared radiomic features that generalize across Chinese and American cohorts suggests that NTM-LD produces distinct, reproducible, population-independent parenchymal signatures detectable through quantitative imaging analysis. This lightweight, interpretable approach may facilitate NTM-LD detection when microbiologic confirmation is pending, particularly in diverse clinical settings. Future directions include prospective multicenter validation, severity-stratified evaluation to determine whether whole-lung radiomics retains discriminative value in early or mild disease, using the Reiff score or similar measures, where improved clinical triage may be particularly valuable, integration with clinical and laboratory data including biomarkers, comparison with expert radiologist interpretations, and combination with lesion-specific features to potentially enhance discriminative performance. [25]

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Acknowledgement:

Kanagala, Aguilera, Garcia, Dutt, Margaroli, and Bodduluri contributed to the conceptualization of the study. Kanagala, Aguilera, and Bodduluri performed the data analysis. All authors contributed to the drafting and critical review of the manuscript, and all authors approved the final version for submission. The authors thank Han and colleagues for making the mycobacterial CT imaging dataset publicly available, enabling this research.

Declaration of interest:

Anusha Kanagala, Taru S. Dutt, Abhiraj S. Pudhota, Dhiral D. Panjwani, Nikhil Dukkipati, Amit Gaggar, Threnesan Naidoo, Leon Jololian, Surya P. Bhatt, and Camilla Margaroli: declare that they have no conflicts of interest.

Bryan Garcia: Reports a relationship with Insmed Incorporated that includes: speaking and lecture fees. Reports a relationship with Vertex Pharmaceuticals Incorporated that includes: consulting or advisory. Reports a relationship with Gilead Sciences Inc that includes: equity or stocks. Reports a relationship with Johnson and Johnson that includes: equity or stocks. Reports a relationship with AbbVie Inc that includes: equity or stocks. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Stéphanie M. Aguilera: Reports research support provided by The University of Alabama at Birmingham and has no other competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Sandeep Bodduluri: reports financial support was provided by National Institutes of Health and

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

a relationship with Spiromatics, Inc. that includes: equity or stocks.

Data Availability: The data (Dataset 1) that support the findings of this study are openly available at https://doi.org/10.34740/kaggle/dsv/10818141.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

REFERENCES

1. 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(1):74. doi:10.1186/s12929-020-00667-6

2. Daley CL, Iaccarino JM, Lange C, et al. Treatment of nontuberculous mycobacterial pulmonary disease: an official ATS/ERS/ESCMID/IDSA clinical practice guideline. *Clin Infect Dis.* 2020;71(4):e1-e36. doi:10.1093/cid/ciaa241

3. Chung C, Lee KN, Han K, et al. The effect of smoking on nontuberculous mycobacterial pulmonary disease and tuberculosis: a nationwide retrospective cohort study. *Sci Rep.* 2024;14(1):22653. doi:10.1038/s41598-024-72438-4

4. World Health Organization**.** *Global Tuberculosis Report 2025***.** World Health Organization; 2025.

5. Haworth CS, Banks J, Capstick T, et al. British Thoracic Society guidelines for the management of non-tuberculous mycobacterial pulmonary disease (NTM-PD). *Thorax.* 2017;72(suppl 2):ii1-ii64. doi:10.1136/thoraxjnl-2017-210927

6. Kwak N, Lee JH, Kim HJ, Kim SA, Yim JJ. New-onset nontuberculous mycobacterial pulmonary disease in bronchiectasis: tracking the clinical and radiographic changes. *BMC Pulm Med.* 2020;20(1):293. doi:10.1186/s12890-020-01331-3

7. Forbes BA, Hall GS, Miller MB, et al. Practical guidance for clinical microbiology laboratories: mycobacteria. *Clin Microbiol Rev.* 2018;31(2):e00038-17. doi:10.1128/CMR.00038-17

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

8. Ko JM, Park HJ, Kim CH. Comparative chest computed tomography findings of non-tuberculous mycobacterial lung diseases and pulmonary tuberculosis in patients with acid fast bacilli smear-positive sputum. *J Thorac Dis.* 2022;14(12):4893-4904. doi:10.21037/jtd-22-1108

9. Zheng L, Wang Y, Zhang W, 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(1):2401613. doi:10.1080/07853890.2024.2401613

10. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77(21):e104-e107. doi:10.1158/0008-5472.CAN-17-0339

11. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology.* 2020;295(2):328-338. doi:10.1148/radiol.2020191145

12. Yan Q, Wang W, Zhao W, et al. Differentiating nontuberculous mycobacterium pulmonary disease from pulmonary tuberculosis through the analysis of the cavity features in CT images using radiomics. *BMC Pulm Med.* 2022;22(1):4. doi:10.1186/s12890-021-01766-2

13. Yan Q, Zhao W, Kong H, et al. CT-based radiomics analysis of consolidation characteristics in differentiating pulmonary disease of non-tuberculous mycobacterium from pulmonary tuberculosis. *Exp Ther Med.* 2024;27(3):112. doi:10.3892/etm.2024.12400

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

14. Wang L, Ding W, Mo Y, 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(13):4293-4306. doi:10.1007/s00259-021-05432-x

15. Han Z, Zhang Y, Ding W, et al. An integrated mycobacterial CT imaging dataset with multispecies information. *Sci Data.* 2025;12(1):533. doi:10.1038/s41597-025-04621-x

16. Hofmanninger J, Prayer F, Pan J, Röhrich S, Prosch H, Langs G. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. *Eur Radiol Exp.* 2020;4(1):50. doi:10.1186/s41747-020-00173-2

17. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825-2830.

18. Garcia, B., Mullins, M., Lim, L., Henostroza, G., & Margaroli, C. (2025). Sex-Associated Radiographic and Clinical Differences in Nontuberculous Mycobacteria Pulmonary Disease. Chronic obstructive pulmonary diseases (Miami, Fla.), 12(6), 455–465. https://doi.org/10.15326/jcopdf.2025.0622

19. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29-36. doi:10.1148/radiology.143.1.7063747

20. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One.* 2015;10(3):e0118432. doi:10.1371/journal.pone.0118432

21. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems 30.* 2017:4765-4774.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

22. Ji S, Xu W, Sun J, et al. Retrospective analysis of patients with non-tuberculous mycobacteria from a primary hospital in Southeast China. *Sci Rep.* 2020;10(1):1060. doi:10.1038/s41598-020-58105-4

23. Griffith DE, Aksamit T, Brown-Elliott BA, et al. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *Am J Respir Crit Care Med.* 2007;175(4):367-416. doi:10.1164/rccm.200604-571ST

24. Biciusca T, Zielbauer AS, Anton T, et al. Differential radiological features of patients infected or colonised with slow-growing non-tuberculous mycobacteria. *Sci Rep.* 2024;14(1):13295. doi:10.1038/s41598-024-64029-0

25. Chen X, Yang X, Li T, et al. Multimodal machine learning-based model for differentiating nontuberculous mycobacteria from Mycobacterium tuberculosis. *Front Public Health.* 2025;13:1470072. doi:10.3389/fpubh.2025.1470072

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Table 1. Cohort Characteristics

Characteristic	Dataset 1 (China)		Dataset 2 (UAB)
	TB (N=871)	NTM (N=429)	NTM (N=173)
Age, years, mean (SD)	47.3(18.3)	55.3(17.0)	69.8 (11.8)
Sex, n (%)			
Female	246 (28.2)	164 (38.2)	123 (71.1)
Male	625 (71.8)	265 (61.8)	50 (28.9)
Race/ethnicity, n (%)			
Non-Hispanic White	NR	NR	149 (86.1)
Asian	NR	NR	3 (1.7)
African American	NR	NR	13 (7.5)
Other	NR	NR	8 (4.6)
BMI, kg/m², mean (SD)	NR	NR	22.8 (5.8)
% Predicted FEV_1, mean (SD)	NR	NR	71.3 (23.3)
Smoking status, n (%)			
Patient/person who smokes	NR	NR	75 (43.4)
Non-smoker	NR	NR	95 (54.9)
Symptomatic, n (%)			
Yes	NR	NR	138 (79.8)
No	NR	NR	35 (20.2)
Geographic origin	China	China	Birmingham, AL, USA
CT scanner	GE BrightSpeed	GE BrightSpeed	GE BrightSpeed / Canon Aquilion Prime 128 / Toshiba / Philips / Siemens / Hitachi
Slice thickness, mm	1.0-1.25	1.0 -1.25	0.6 - 0.8

Continuous variables presented as mean (SD); categorical variables as n (%). NR = not reported.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Table 2. Classification Performance: Scenario 1 (Train Dataset 1, Internal Test=391) and External Validation on Dataset 2 (N=173)

Performance Metric	Scenario 1: Internal Test (n=391)	95% CI	External Validation D2 (n=173)	95% CI
AUC	0.79	0.73–0.84	—	—
Sensitivity	0.47	0.39–0.56	0.94	0.90–0.97
Specificity	0.91	0.88–0.94	—	—
Accuracy	0.77	0.73–0.81	0.94	0.90–0.97
F1 Score	0.57	0.51–0.60	0.97	0.95-0.99
NTM-LD Detected	61/129	—	162/173	—

AUC = area under the receiver operating characteristic curve; CI = confidence interval; LDA = linear discriminant analysis; External validation on Dataset 2 reports detection rate for NTM-LD; traditional specificity metrics are not applicable as Dataset 2 contains no TB cases. Dash (—) indicates metric not applicable.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Table 3. Classification Performance: Scenario 2 (Combined Cohort D1+D2, N=1,473)

Performance Metric	Value	95% CI
AUC	0.81	0.76–0.85
Sensitivity	0.61	0.54–0.69
Specificity	0.90	0.87–0.94
Accuracy	0.79	0.75–0.82
F1 Score	0.70	0.64-0.75

Combined cohort with group-aware stratified 70/30 split (D1 TB: train=610, test=261; D1 NTM: train=300, test=129; D2 NTM: train=121, test=52). Dash (—) indicates metric not applicable.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Table 4. Comparison of Classification Approaches for NTM-TB Differentiation

Study/ Approach	Method	AUC	Manual Annotation	Interpretable	Cohort
Yan et al[12]	Cavity radiomics + LR	0.84	Yes	Yes	Shandong Public Health Clinical Center Jinan Infectious Disease Hospital
Yan et al[13]	Consolidation radiomics	0.90	Yes	Yes	Shandong Public Health Clinical Center Jinan Infectious Disease Hospital
Han et al[15]	BoTNet50 CNN	0.71	No	No	Mycobacterial CT Imaging Dataset (Kaggle)
Wang et al[14]	3D-ResNet	0.78	No	No	Tianjin Haihe Hospital, China
Current study	Whole-lung radiomics	0.79 / 0.81	No	Yes	D1: China (multicenter); D2: UAB NTM-LD Registry (USA)

AUC = area under the receiver operating characteristic curve; CNN = convolutional neural network; LDA = linear discriminant analysis; LR = logistic regression; UAB = University of Alabama at Birmingham. Current study results shown for Scenario 1 (Dataset 1 only) and Scenario 2 (combined D1+D2 cohort).

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Table 5. Sixteen Shared Radiomic Features Identified Across Both Experimental Scenarios

Radiomic Class	Feature	Description	Importance Rank (16-Feature Model)
Shape	VoxelVolume	Total volume of segmented lung voxels	3
Shape	MeshVolume	Volume computed from triangular mesh of lung surface	4
GLCM (Texture)	Autocorrelation	Spatial correlation of gray-level co-occurrences	11
GLCM (Texture)	Idm (Inverse Difference Moment)	Local homogeneity of gray-level distribution	13
GLCM (Texture)	Id (Inverse Difference)	Smoothness of gray-level transitions between voxels	8
GLCM (Texture)	JointEntropy	Randomness in joint gray-level distribution of voxel pairs	14
GLCM (Texture)	DifferenceAverage	Mean absolute difference between paired gray levels	16
GLCM (Texture)	InverseVariance	Inverse-weighted variance of gray-level differences	10
GLCM (Texture)	Entropy	Uncertainty/complexity in gray-level spatial distribution	15
GLRLM (Texture)	HighGrayLevelRunEmphasis	Concentration of high-intensity consecutive voxel runs	12
GLRLM (Texture)	RunLengthNonUniformityNormalized	Heterogeneity of run lengths across the image	6
GLRLM (Texture)	LongRunEmphasis	Predominance of long consecutive voxel runs	7
GLRLM (Texture)	RunVariance	Dispersion of run length values	9
GLRLM (Texture)	RunPercentage	Fraction of voxels that form runs relative to total	5
GLSZM (Texture)	LargeAreaEmphasis	Predominance of large, connected zones of similar intensity	1
GLSZM (Texture)	ZoneVariance	Dispersion of zone sizes across the image	2

GLCM = gray-level co-occurrence matrix; GLRLM = gray-level run-length matrix; GLSZM = gray-level size-zone matrix. Importance rank derived from the 16-feature reduced LDA model (Scenario 1). Features are grouped by radiomic class and ordered by importance within each class.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

FIGURES

Figure 1. Radiomics analysis pipeline for distinguishing NTM from TB. The workflow comprises: input chest CT images from NTM and TB patient cohorts, automated lung parenchyma segmentation, quantitative radiomics feature extraction including density, texture, and shape metrics, feature selection and dimensionality reduction, Linear Discriminant Analysis (LDA) classification, and model performance evaluation with feature importance.

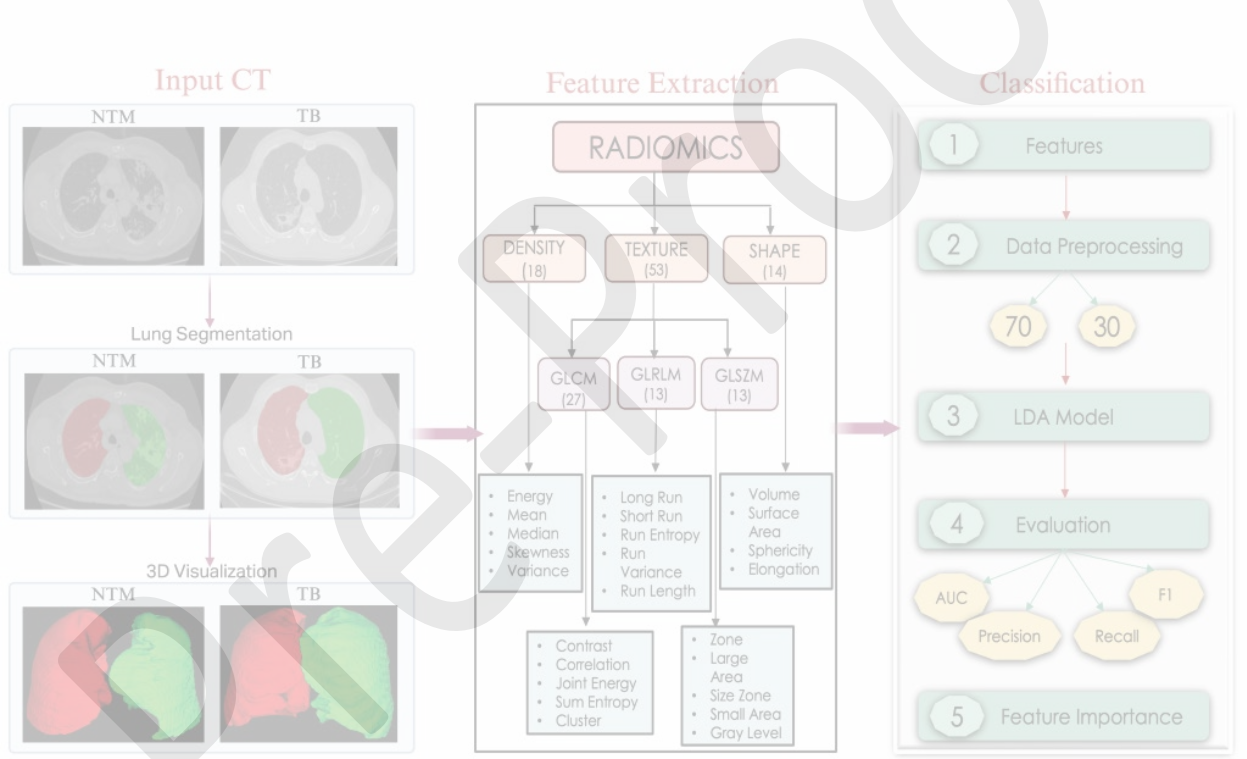
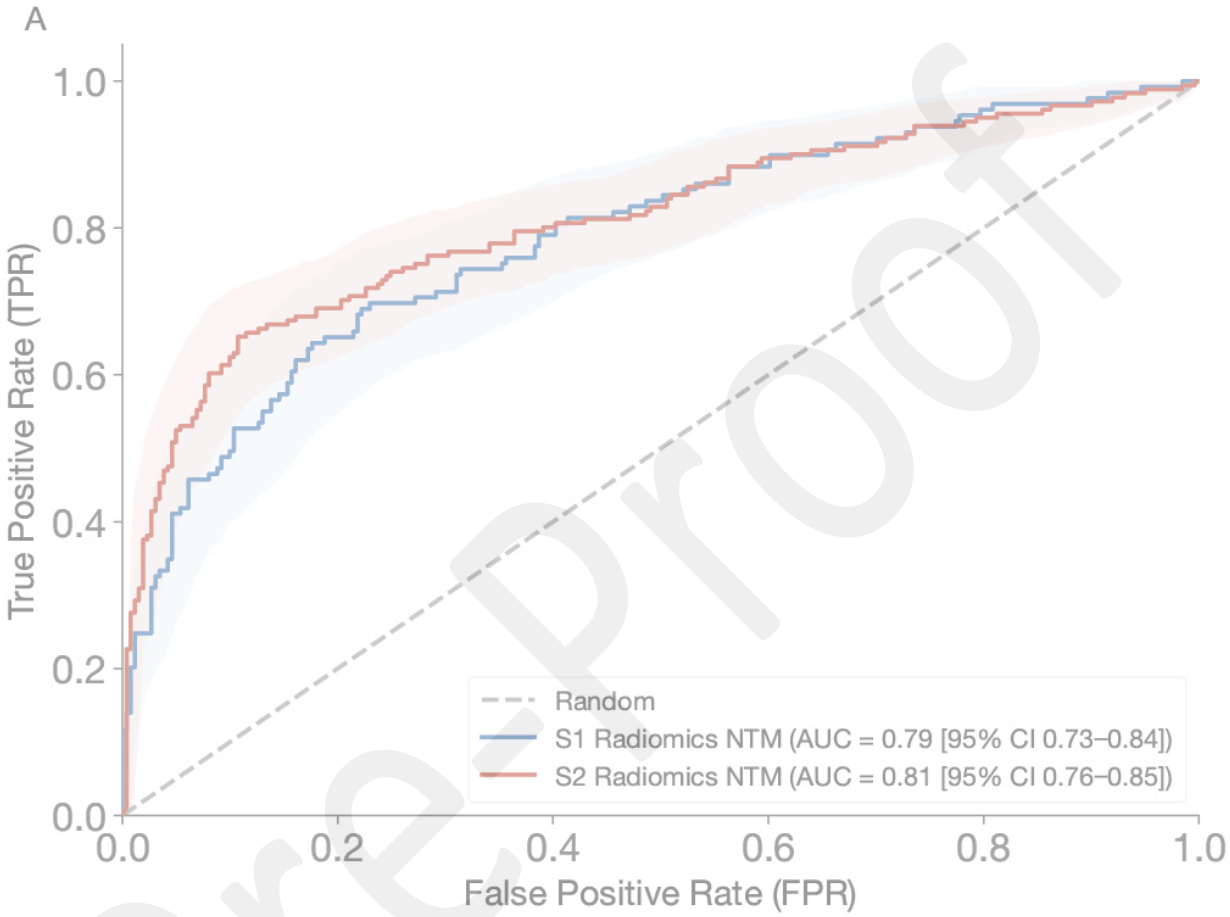


Figure 2. A, Receiver operating characteristic (ROC) curves for Scenario 1 (S1) (Dataset 1 only) with AUC 0.79 (95% CI, 0.73–0.84) and for Scenario 2 (S2) (combined D1+D2 cohort) with AUC 0.81 (95% CI, 0.76–0.85). and B, Precision-recall (PR) curves for Scenario 1 (S1) (Dataset

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

1 only) with AP 0.70 (95% CI, 0.73–0.84) and for Scenario 2 (S2) (combined D1+D2 cohort) with AP 0.79 (95% CI, 0.76–0.85) for the whole-lung radiomics-based linear discriminant analysis classifier.



Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

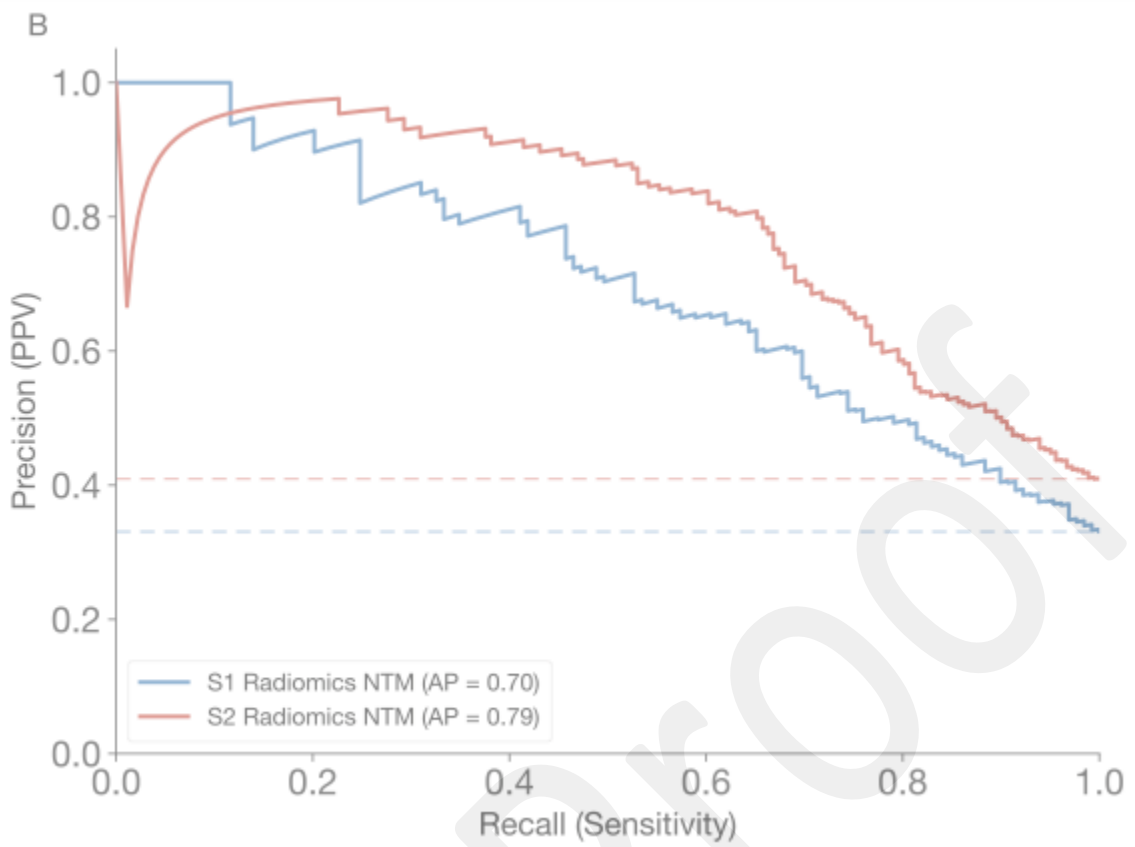
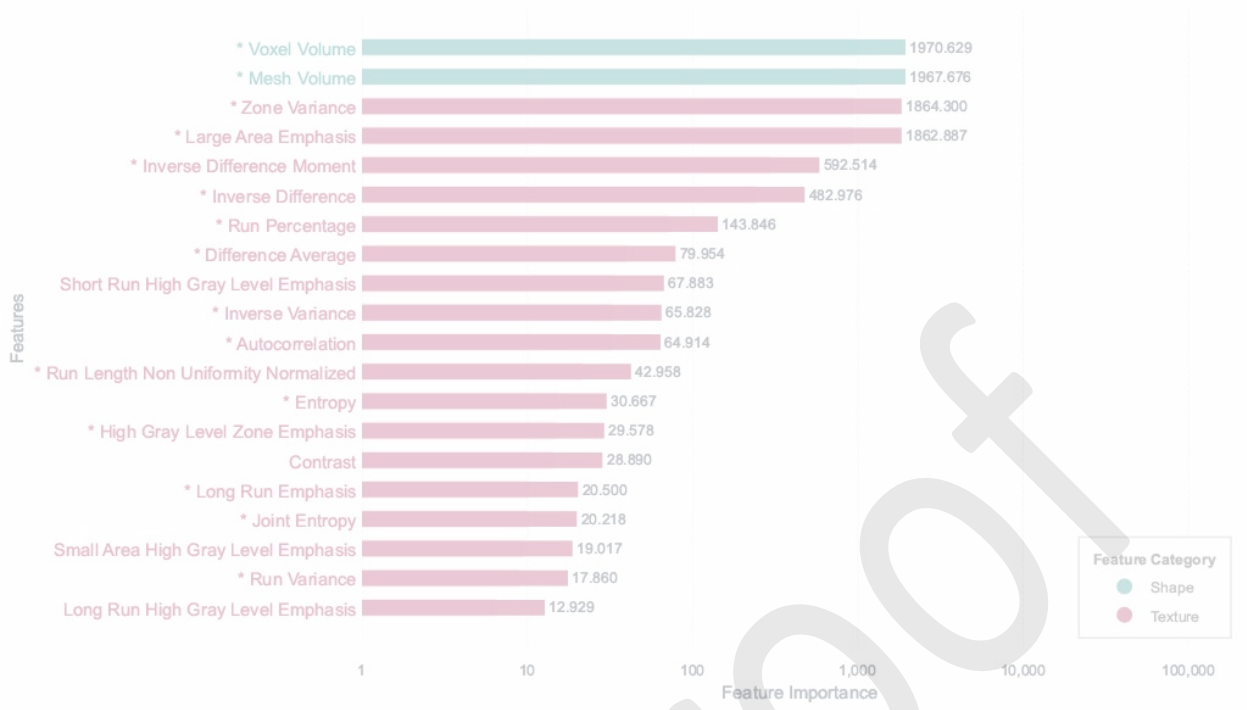
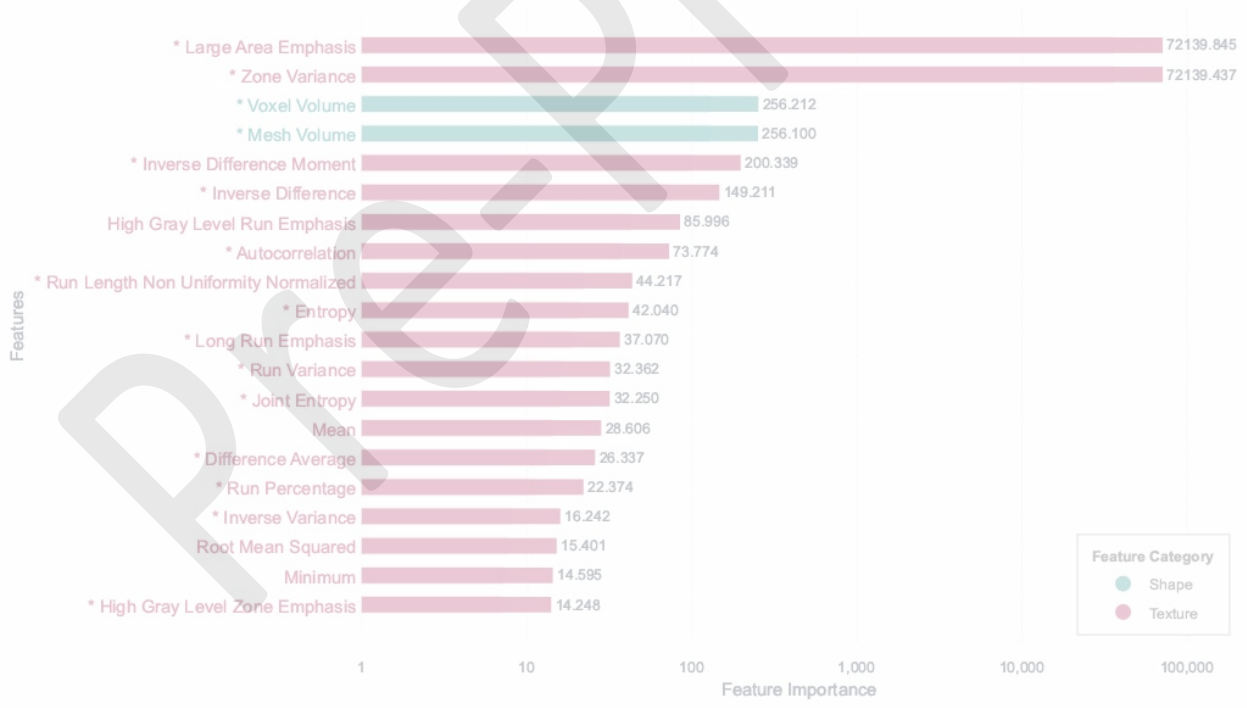


Figure 3. Radiomic feature contributions to differentiation between TB and NTM-LD. A, Top 20 radiomic features driving model performance in Scenario 1 (Dataset 1 only), including texture heterogeneity measures and intensity-based features; asterisks (*) denote features shared with Scenario 2 (combined D1+D2 cohort). B, Top 20 radiomic features in Scenario 2 (combined D1+D2 cohort); asterisks (*) denote features shared with Scenario 1 (Dataset 1 only).

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867





Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Online Supplement

This supplementary material accompanies the manuscript. It contains three supplementary tables (Tables S1–S3) that provide additional data not included in the main text due to space limitations. These tables offer further detail on Reiff score, CTDIvol and Strain types provided for NTM in Dataset 1 and should be read alongside the main manuscript.

Table S1. Distribution of bronchiectasis severity by modified Reiff score (n = 173)

Severity category	Reiff score	n (%)
Mild	0–3	47 (27.2)
Moderate	4–7	92 (53.2)
Severe	>7	28 (16.2)
No score	NA	6 (3.5)

Mean Reiff score 5.2 ± 2.6; median 5.0 (IQR 3.0–7.0). NA, not available.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Table S2: Nontuberculous mycobacterial (NTM) species by dataset

NTM species	Dataset 1 (China) n (%)	Dataset 2(UAB) n (%)	Total n (%)
M. avium complex (MAC)	140 (32.6)	141 (81.5)	281 (46.7)
M. avium	41 (9.6)	133 (76.9)	174 (28.9)
M. intracellulare	99 (23.1)	0 (0.0)	99 (16.4)
MAC, species not specified	0 (0.0)	8 (4.6)	8 (1.3)
Other slow-growing NTM	85 (19.8)	7 (4.0)	92 (15.3)
M. kansasii	67 (15.6)	3 (1.7)	70 (11.6)
M. gordonae	12 (2.8)	4 (2.3)	16 (2.7)
M. szulgai	2 (0.5)	0 (0.0)	2 (0.3)
M. scrofulaceum	2 (0.5)	0 (0.0)	2 (0.3)
M. terrae	1 (0.2)	0 (0.0)	1 (0.2)
M. xenopi	1 (0.2)	0 (0.0)	1 (0.2)
Rapidly growing NTM	82 (19.1)	9 (5.2)	91 (15.1)
M. abscessus	70 (16.3)	7 (4.0)	77 (12.8)
M. fortuitum	11 (2.6)	0 (0.0)	11 (1.8)
M. chelonae	0 (0.0)	2 (1.2)	2 (0.3)
M. vaccae	1 (0.2)	0 (0.0)	1 (0.2)
Polymicrobial	0 (0.0)	12 (6.9)	12 (2.0)
Not identified / not available	122 (28.4)	4 (2.3)	126 (20.9)
Total	429	173	602

Percentages are calculated within each dataset. Mixed infection refers to isolation of more than one NTM species. NTM, nontuberculous mycobacteria; MAC, Mycobacterium avium complex.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867

Table S3. Dataset 2 CT scanner manufacturers, reconstruction kernels, and CTDIvol (n = 173)

Manufacturer	n (%)	Main kernels (n)	CTDIvol, mGy, median (range)
GE Medical Systems	97 (56.1)	STANDARD (48), BONE (36), other (13)	NA
Siemens	39 (22.5)	Br49d (17), Br40d/f (11), other (11)	4.19 (1.15–11.10)
Philips	30 (17.3)	C (14), L (12), other (4)	2.80 (1.30–21.70)
Canon/Toshiba	6 (3.5)	FC18 (2), FC52 (2), other (2)	4.05 (2.70–15.40)
Hitachi	1 (0.6)	(24)	NA
Total	173		4.00 (1.15–21.70)

CTDIvol was not available for Dataset 1; CTDIvol, volume computed tomography dose index; NA= not available.

Pre-Proof

Copyright *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* ©2026
Published online November 12, 2026 https://doi.org/10.15326/jcopdf.2026.0867