



Original Article

## Post-treatment Lung Tuberculosis Sequelae: an Inexpensive Clinical-Laboratory Nomogram to Predict Tissue Destruction

Min Xu<sup>1</sup>, Jiahui He<sup>2</sup> and Jianju Feng<sup>3</sup>.

<sup>1</sup> Department of Respiratory and Critical Care Medicine, Zhuji People's Hospital of Zhejiang Province, Zhuji Affiliated Hospital of Wenzhou Medical University, Zhuji, Zhejiang 311800, China.

<sup>2</sup> Department of Traditional Chinese Medicine, Zhuji People's Hospital of Zhejiang Province, Zhuji Affiliated Hospital of Wenzhou Medical University, Zhuji, Zhejiang 311800, China.

<sup>3</sup> Department of Radiology, Zhuji People's Hospital of Zhejiang Province, Zhuji Affiliated Hospital of Wenzhou Medical University, Zhuji, Zhejiang 311800, China.

**Competing interests:** The authors declare no competing interest.

**Abstract. Background:** Post-treatment lung destruction (LD) impairs quality of life in pulmonary tuberculosis (TB) survivors, yet early risk-stratification tools are lacking. We aimed to develop and internally validate a clinical-laboratory nomogram to predict LD at completion of standard anti-TB therapy.

**Methods:** In this retrospective cohort, we enrolled 205 treatment-naïve adults with pulmonary TB from April 2021 to April 2025. LD was defined on follow-up chest CT as extensive fibrosis, bronchiectasis with volume loss, or parenchymal destruction. Twenty-two baseline demographic, clinical, laboratory, and imaging variables were screened. Least absolute shrinkage and selection operator (LASSO; 10-fold cross-validation) was used for variable selection, followed by Akaike information criterion (AIC)-guided stepwise multivariable logistic regression. Model performance was compared with random forest (RF) and support vector machine (SVM) classifiers. Discrimination (area under the receiver-operating characteristic curve, AUC), calibration (bootstrap-corrected curve; Brier score), and clinical utility (decision-curve analysis, DCA) were assessed; internal validation used 1,000-sample bootstrap resampling.

**Results:** LD occurred in 61/205 patients (29.8%). Nine predictors—silicosis, drug resistance, symptom-to-treatment delay, lymphocyte count, C-reactive protein, aspartate aminotransferase,  $\gamma$ -glutamyl transferase, albumin, and baseline atelectasis/cavity—composed the final model. The nomogram showed excellent discrimination (AUC = 0.93, 95% CI 0.897-0.971; optimism-corrected AUC = 0.93) and good calibration (Brier = 0.13). Across 10-40% risk thresholds, DCA indicated a higher net benefit than treat-all or treat-none strategies. Logistic regression slightly outperformed RF (AUC = 0.91) and SVM (AUC = 0.92) while retaining interpretability.

**Conclusions:** An inexpensive, easily applicable nomogram integrating routine clinical and laboratory indices accurately predicts post-treatment LD in TB patients. The tool can support personalized follow-up and timely interventions, warranting external validation in multicenter prospective cohorts.

**Keywords:** Lung tuberculosis sequelae; Pulmonary tuberculosis; Anti-TB therapy.

**Citation:** Xu M., He J., Feng J. Post-treatment lung tuberculosis sequelae: an inexpensive clinical-laboratory nomogram to predict tissue destruction. *Mediterr J Hematol Infect Dis* 2025, 17(1): e2025075, DOI: <http://dx.doi.org/10.4084/MJHID.2025.075>

**Published:** November 01, 2025

**Received:** July 30, 2025

**Accepted:** October 21, 2025

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by-nc/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Introduction.** Tuberculosis (TB) remains a major global health challenge.<sup>1,2</sup> As treatment outcomes improve, the number of survivors is increasing.<sup>3</sup> Post-tuberculosis lung disease (PTLD) - including fibrotic volume loss, bronchiectasis, and cavitation - has emerged as a key source of morbidity beyond microbiological cure.<sup>4,5</sup> These sequelae worsen respiratory symptoms, impair lung function, and diminish quality of life.<sup>6</sup> However, PTLD is under-recognized, follow-up is inconsistent, and comprehensive management guidance is lacking.<sup>6,7</sup> Limited clinician awareness and the absence of standardized post-treatment protocols contribute to suboptimal care.

In pulmonary medicine, early risk stratification is challenging because radiological manifestations often appear late, and baseline predictors are incompletely characterized. This challenge is well illustrated by the management of indeterminate pulmonary nodules (IPNs), whose detection has risen with widespread CT screening;<sup>8</sup> complex prediction tools can be difficult to interpret or apply at the bedside.<sup>9,10</sup> By analogy, there is a strong rationale for parsimonious clinical-laboratory nomograms that rely on routinely collected variables at treatment initiation. Such tools could individualize follow-up, prioritize pulmonary rehabilitation referrals, and inform preventive strategies, thereby optimizing care for TB survivors.<sup>11,12</sup>

Despite growing recognition that TB survivors may develop disabling structural sequelae long after microbiological cure, no validated multivariable model integrating demographic, biochemical, and baseline imaging data predicts post-treatment lung destruction (LD). We therefore aimed to develop and internally validate a concise nomogram - grounded in routinely collected clinical-laboratory indices and chest CT features - to estimate an individual's probability of LD at completion of standard anti-TB therapy, and to benchmark its performance against random forest and support vector machine classifiers while assessing real-world clinical utility using decision-curve analysis.

## Methods.

**Study design and setting.** We conducted a retrospective, single-center cohort study at Zhuji People's Hospital (Zhejiang Province, China). The study protocol was approved by the hospital's Institutional Review Board, and all procedures were conducted in accordance with the Declaration of Helsinki. Owing to the retrospective design, the requirement for written informed consent was waived.

**Participants.** Consecutive adults ( $\geq 18$  years) with bacteriologically confirmed, drug-naïve pulmonary TB who completed standard first-line therapy and follow-up

between April 2021 and April 2025 were screened. We excluded patients who (i) received  $<6$  months of anti-TB therapy; (ii) had pre-existing extensive fibrocavitary lung disease due to non-TB etiologies; (iii) lacked a baseline chest CT or the protocol-mandated follow-up CT obtained within 4 weeks of treatment completion; or (iv) had incomplete electronic records for key covariates.

**Outcome definition.** The primary endpoint - post-treatment lung destruction (LD) - was prespecified as the presence on follow-up CT of any of the following in  $\geq 1$  lobe: (i) volume loss with architectural distortion; (ii) confluent fibrocystic bronchiectasis occupying  $\geq 50\%$  of the lobe; or (iii) parenchymal replacement with dense fibrosis or atelectatic tissue. The definition followed national imaging consensus statements and was adjudicated by two thoracic radiologists ( $\geq 5$  years' experience) who were blinded to clinical data. Disagreements were resolved by consensus; interobserver reliability was quantified in a randomly selected subset.

**Candidate predictors.** Twenty-two baseline variables were extracted from electronic medical and radiology records:

**Demographics and comorbidities:** age, sex, diabetes, chronic obstructive pulmonary disease, bronchiectasis, silicosis, and inhaled corticosteroid use.

**Disease-specific factors:** drug-resistant TB (yes/no) defined at baseline as rifampicin resistance detected by GeneXpert MTB/RIF and/or phenotypic culture-based DST indicating resistance to rifampicin and/or isoniazid; symptom-to-treatment delay (days).

**Laboratory indices:** white blood cell and lymphocyte counts; lymphocyte percentage; C-reactive protein (CRP); alanine aminotransferase (ALT) and aspartate aminotransferase (AST);  $\gamma$ -glutamyl transferase (GGT); serum albumin; and prealbumin.

**Baseline CT descriptors:** number of affected lobes (0-5), presence of pleural effusion, atelectasis, and pulmonary cavity.

Variables were chosen for clinical plausibility and  $\geq 90\%$  availability at most TB centers.

**Comparator machine-learning models.** To benchmark performance, we trained a random forest (RF) classifier (randomForest package; 500 trees;  $\sqrt{p}$  variables tried at each split) and a support vector machine (SVM) with a radial basis function kernel (e1071 package; cost and  $\gamma$  tuned via grid search). Hyperparameters for both algorithms were optimized using a five-fold cross-validated grid search.

*Internal validation and performance metrics.* Discrimination was assessed by the area under the receiver-operating characteristic (ROC) curve (AUC) with 95% confidence intervals estimated by DeLong's method. Calibration was evaluated using calibration plots, the Hosmer-Lemeshow test, and the Brier score. Clinical usefulness was appraised by decision-curve analysis (DCA; dca.R) across risk thresholds of 10-40%. To correct for optimism, 1,000-sample bootstrap resampling was used to estimate the bias-adjusted AUC and calibration slope. Additional metrics included accuracy, sensitivity, specificity, positive and negative predictive values, and F1 score.

*Inter-observer reliability.* To ensure endpoint robustness, 25 follow-up CT scans (stratified 1:1 by LD status) were randomly selected for blinded duplicate reading. Agreement on LD presence (binary) was summarized using Cohen's  $\kappa$  (irr package) and raw concordance. In 18 scans, the percentage of lung destruction was scored on a 0-100% scale and analyzed with the two-way random-effects intraclass correlation coefficient

[ICC(2,1)] and Bland-Altman limits of agreement (Bland-AltmanLeh package).

*Sample-size and events-per-variable considerations.* With 61 LD events, the final nine-predictor model satisfied a conservative events-per-variable (EPV) ratio of 6.8. Bootstrap internal validation and penalized shrinkage helped mitigate overfitting given the modest sample size.

*Statistical analysis.* All analyses were conducted in R (version 4.3) using the following packages: glmnet (4.1-8), MASS (7.3-60), rms (6.7-3), pROC (1.18.4), randomForest (4.7-1.1), e1071 (1.7-13), mice (3.16.0), irr (0.84.1), and BlandAltmanLeh (0.5.2). Two-sided  $p < 0.05$  was considered statistically significant.

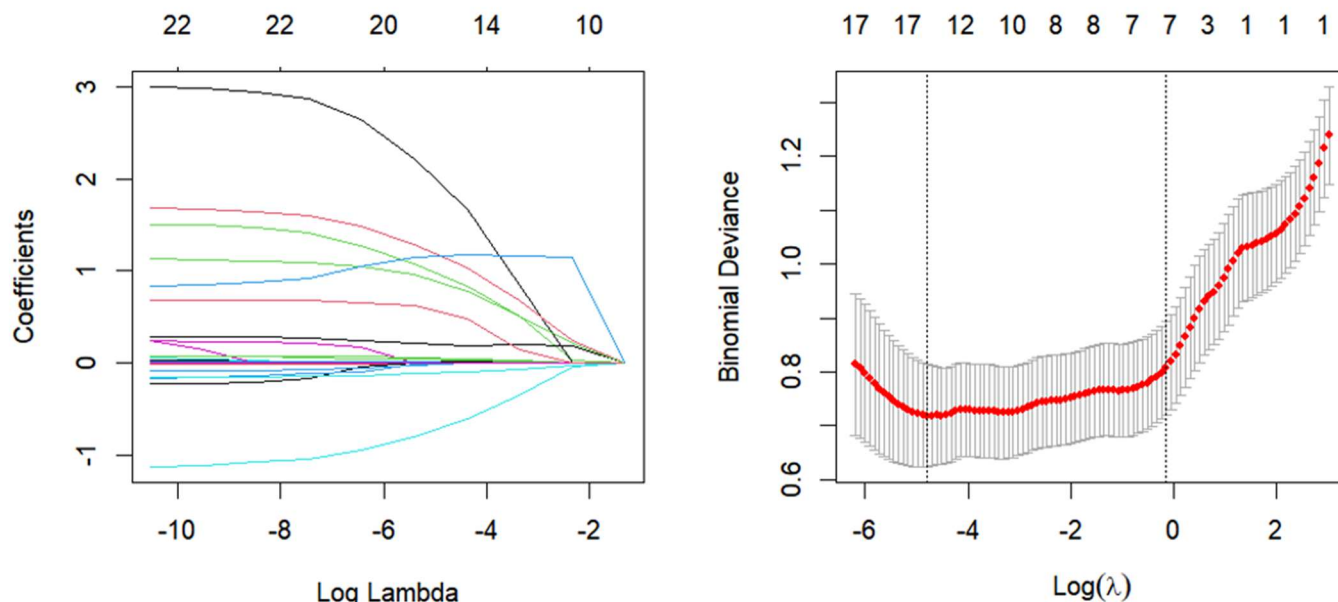
## Results.

*Patient characteristics.* A total of 205 treatment-naïve adults with pulmonary tuberculosis were included (Table 1). The mean age was  $49 \pm 19$  years, and approximately three-quarters were male. Lung

**Table 1.** Baseline demographic, clinical, and imaging characteristics.

Characteristic	No LD (n = 144)	LD (n = 61)	p-value†
Age (year), mean±SD	45.9 ± 18.4	56.6 ± 16.9	<0.001
Male sex, n (%)	99 (68.8%)	50 (82.0%)	0.077
Bronchiectasis, n (%)	5 (3.5%)	16 (26.2%)	<0.001
COPD, n (%)	18 (12.5%)	18 (29.5%)	0.006
Inhaled corticosteroids, n (%)	13 (9.0%)	16 (26.2%)	0.003
Diabetes, n (%)	10 (6.9%)	13 (21.3%)	0.006
Silicosis, n (%)	2 (1.4%)	9 (14.8%)	<0.001
Drug-resistant TB, n (%)	6 (4.2%)	22 (36.1%)	<0.001
Symptom-to-treatment delay (day), mean±SD	12.6 ± 9.3	24.7 ± 13.8	<0.001
White-blood-cell count ( $\times 10^9/L$ ), mean±SD	7.8 ± 2.4	7.5 ± 3.3	0.620
Lymphocyte count ( $\times 10^9/L$ ), mean±SD	1.7 ± 0.6	1.2 ± 0.6	<0.001
Lymphocyte fraction (%), mean±SD	23.2 ± 9.6	18.9 ± 9.0	<0.001
C-reactive protein (mg/L), mean±SD	11.1 ± 17.7	37.3 ± 42.1	<0.001
Alanine aminotransferase (U/L), mean±SD	27.2 ± 14.9	37.1 ± 28.9	0.013
Aspartate aminotransferase (U/L), mean±SD	28.0 ± 14.1	37.5 ± 22.3	0.003
$\gamma$ -Glutamyl-transferase (U/L), mean±SD	34.0 ± 20.5	52.6 ± 38.8	<0.001
Albumin (g/L), mean±SD	40.6 ± 5.0	35.4 ± 4.7	<0.001
Pre-albumin (mg/L), mean±SD	268.2 ± 98.9	169.9 ± 76.7	<0.001
Number of affected lobes, n (%)			<0.001
0	7 (4.9%)	0 (0.0%)	–
1	41 (28.5%)	4 (6.6%)	–
2	63 (43.8%)	6 (9.8%)	–
3	19 (13.2%)	21 (34.4%)	–
4	9 (6.3%)	18 (29.5%)	–
5	5 (3.5%)	12 (19.7%)	–
Pleural effusion, n (%)	11 (7.6%)	16 (26.2%)	<0.001
Atelectasis, n (%)	4 (2.8%)	18 (29.5%)	<0.001
Pulmonary cavity, n (%)	42 (29.2%)	57 (93.4%)	<0.001

† Welch two-sample t-test for continuous variables; Pearson  $\chi^2$  test for categorical variables. Values are mean ± SD or number (percentage).



**Figure 1.** LASSO variable-selection path for candidate predictors of post-treatment lung destruction. LASSO variable-selection path for candidate predictors of post-treatment LD. Coefficients shrink toward zero with increasing penalty ( $\lambda$ ). The dashed lines mark  $\lambda$  minimizing ten-fold cross-validated deviance (left) and the one-standard-error  $\lambda$  (right).

**Table 2.** Multivariable logistic regression predicting post-treatment lung destruction.

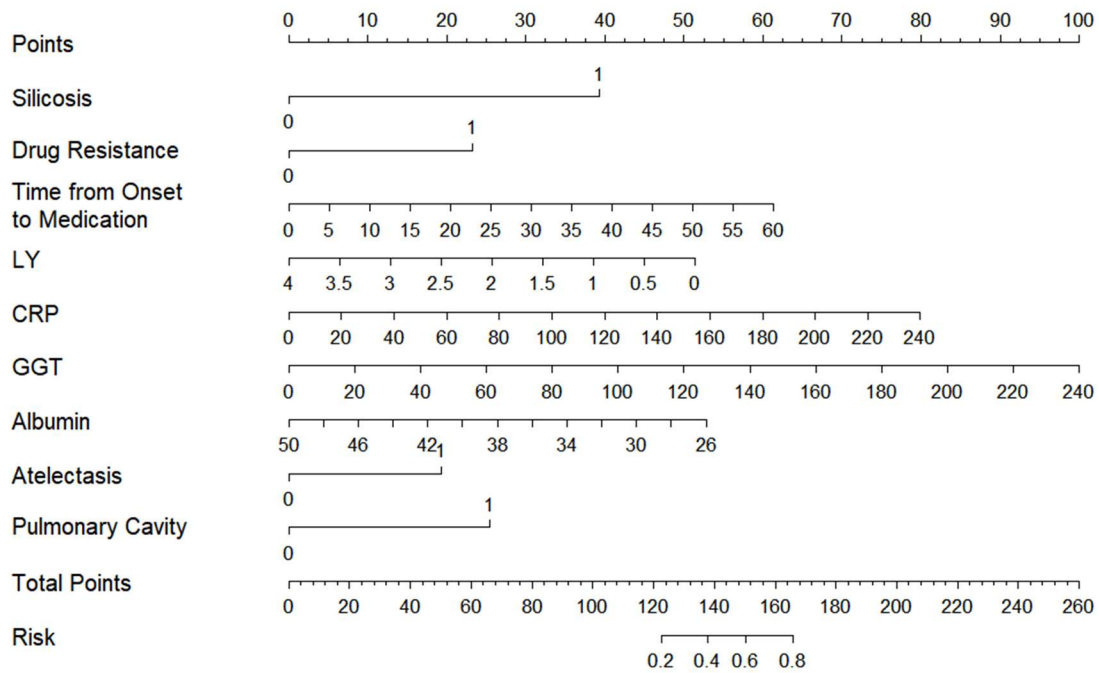
Predictor	$\beta$ -coefficient	SE	Adjusted OR (95% CI)	<i>p</i> -value
Intercept	0.96	2.51	–	0.701
Symptom-to-treatment delay (per 5 days)	0.08	0.02	1.08 (1.04 – 1.14)	0.001
Lymphocyte count (per $0.5 \times 10^9/L$ )	-1.15	0.49	0.32 (0.11 – 0.79)	0.019
C-reactive protein (per 10 mg/L)	0.03	0.01	1.03 (1.01 – 1.05)	0.011
Aspartate aminotransferase (per 5 U/L)	0.05	0.02	1.05 (1.01 – 1.10)	0.016
$\gamma$ -Glutamyl-transferase (per 5 U/L)	0.02	0.01	1.02 (1.00 – 1.04)	0.126
Albumin (per 1 g/L)	-0.17	0.06	0.85 (0.75 – 0.95)	0.006
Silicosis (yes vs. no)	2.82	1.05	16.79 (2.39 – 166.14)	0.007
Drug-resistant TB (yes vs. no)	1.90	0.72	6.68 (1.75 – 30.33)	0.008
Atelectasis/cavity (yes vs. no)	1.24	0.80	3.45 (0.75 – 18.01)	0.120

destruction (LD) was identified on follow-up chest CT in 61 participants (29.8%). At baseline, individuals who subsequently developed LD were, on average, about a decade older; had a longer symptom-to-treatment delay; exhibited higher systemic inflammation; and showed poorer nutritional indices. Additionally, silicosis and drug-resistant tuberculosis were several-fold more prevalent in the LD group, and CT features - including atelectasis, cavitation, and multilobar involvement - were disproportionately frequent among eventual LD cases. These imbalances provided a clear clinical rationale for subsequent multivariable modeling.

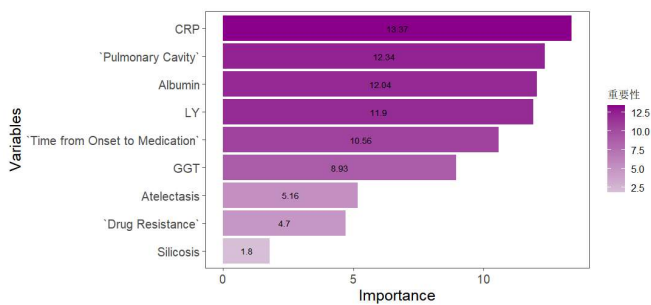
*Missing-data pattern and data quality.* Data completeness was excellent. Of 4,510 expected data cells, only 37 (0.8%) were missing, and no individual predictor exceeded 3.4% missingness (**Supplementary Table S1**). Given this minimal loss, we conducted primary complete-case analyses ( $n = 198$ ) and used multiply imputed datasets for sensitivity analyses to assess robustness.

*Derivation of the prediction model.* Variables were selected via a two-stage process. First, the least absolute shrinkage and selection operator (LASSO) regression with ten-fold cross-validation reduced the original 22 candidates to 16 variables with non-zero coefficients (**Figure 1**). These were then entered into a stepwise multivariable logistic regression guided by the Akaike information criterion (AIC), which ultimately retained nine independent predictors: silicosis, drug resistance, delay from symptom onset to treatment, lymphocyte count, C-reactive protein, aspartate aminotransferase,  $\gamma$ -glutamyl transferase, albumin, and the radiologic descriptor “atelectasis/cavity” (**Table 2**). A nomogram translating the final coefficients into bedside risk points is shown in **Figure 2**. This concise model balances interpretability and performance, providing the basis for internal validation.

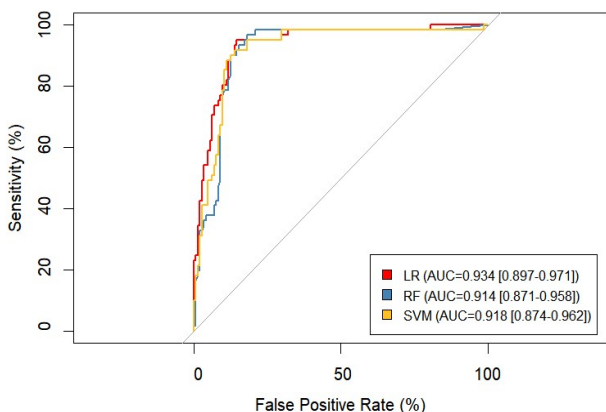
*Model performance and internal validation.* The nine-factor logistic model showed strong discrimination



**Figure 2. Nomogram for bedside prediction of lung destruction at completion of standard anti-tuberculosis therapy.** Nomogram to estimate LD risk at treatment completion from nine predictors (silicosis, drug-resistant TB, baseline atelectasis/cavity, symptom-to-treatment delay, lymphocyte count, CRP, AST, GGT, albumin). Sum points across axes to obtain the total and read the corresponding predicted probability.



**Figure 3. Variable importance plot from the random-forest comparator model.** Random-forest variable importance (mean decrease in Gini). Higher values indicate greater contribution to classification; CRP, cavity, and albumin rank highest.



**Figure 4. Receiver-operating-characteristic (ROC) curves comparing predictive models for post-treatment lung destruction.** ROC curves comparing models for LD prediction: logistic regression (AUC 0.93), RF (0.91), and SVM (0.92). Shaded bands show 95% CIs

between patients with and without subsequent LD. The apparent area under the receiver-operating characteristic (ROC) curve (AUC) was 0.93 (95% CI 0.897-0.971), and 1,000-sample bootstrap correction yielded an optimism-adjusted AUC of 0.93, indicating negligible overfitting (**Figure 4**). Accuracy was 0.87, with sensitivity 0.80 and specificity 0.90; positive and negative predictive values and the F1 score reflected similarly balanced performance (**Table 3**). Compared with alternative machine-learning approaches evaluated on the same data, logistic regression outperformed both the support vector machine (AUC = 0.92) and random forest classifiers (AUC = 0.91) while offering clearer clinical interpretability. Calibration was satisfactory, with a Brier score of 0.13 and visually close agreement between predicted and observed risks across deciles (**Supplementary Figure S1**). Within the 10-40% threshold range, the nomogram yielded a higher net benefit than treat-all or treat-none (**Supplementary Figure S2**). For orientation, choosing thresholds of 10%, 20%, 30%, and 40% corresponds to accepting approximately 9, 4, 2.3, and 1.5 unnecessary escalations, respectively, for each correctly escalated patient, allowing sites to align triage with local capacity (**Supplementary Figure S2**).

*Sensitivity to missing-data handling.* To assess whether the small proportion of missing values biased inference, we refitted the model across ten multiply imputed datasets. Pooled odds ratios differed from complete-case estimates by no more than 6% for any predictor, and the pooled AUC (0.93) was virtually identical to that of the

**Table 3.** Predictive performance metrics of three modelling strategies.

Metric	Logistic regression	Random forest	Support-vector machine
Accuracy	0.87	0.86	0.88
Sensitivity	0.80	0.75	0.85
Specificity	0.90	0.90	0.89
Positive predictive value	0.78	0.77	0.76
Negative predictive value	0.92	0.909	0.93
F1-score	0.79	0.76	0.81
AUC (95% CI)‡	0.93 (0.897–0.971)	0.91 (0.871–0.958)	0.92 (0.874–0.962)

‡Area under the receiver-operating-characteristic curve.

**Table 4.** Variable importance derived from the random-forest model.

Rank	Predictor	Mean decrease in Gini
1	C-reactive protein	12.86
2	Pulmonary cavity	12.61
3	Albumin	11.80
4	Lymphocyte count	11.65
5	Symptom-to-treatment delay	10.89
6	γ-Glutamyl-transferase	9.73
7	Aspartate aminotransferase	8.92
8	Drug-resistant TB	7.51
9	Silicosis	6.48
10	Atelectasis	6.07

A higher mean-decrease-in-Gini scores indicates a greater contribution to classification accuracy.

complete-case model, with a bootstrap-based difference of only 0.002 (**Supplementary Table S2**). These findings indicate that the conclusions drawn from the complete-case analysis are robust to plausible assumptions about the missing-data mechanism.

*Reliability of the radiological outcome.* Finally, we assessed the objectivity of the LD outcome in a blinded reread of 25 randomly selected follow-up CT scans. Agreement between two experienced thoracic radiologists was excellent: Cohen's  $\kappa = 0.82$  (95% CI 0.63-1.00) for the binary presence of LD, with raw concordance of 92%. In the subset in which percentage parenchymal destruction was recorded, the intraclass correlation coefficient was 0.88 (95% CI 0.74-0.95). Bland-Altman analysis showed narrow limits of agreement (-5.0% to +4.0%), indicating that reader variability was unlikely to compromise the study's primary endpoint.

**Discussion.** This study developed a simple yet rigorous nine-factor nomogram for anticipating post-treatment LD in adults with pulmonary tuberculosis. Specifically, the model achieved an AUC = 0.934 with only a marginal optimism correction to 0.925 and yielded the highest net clinical benefit on decision-curve analysis. Notably, the underlying logistic regression framework outperformed more complex random forest and support vector machine classifiers while preserving full interpretability, offering clinicians a transparent,

evidence-based means of triaging follow-up imaging and early interventions.

In our cohort, the prevalence of post-tuberculosis LD was 29.8%, higher than the long-term UK registry figure of 12% but between the 25% reported in a South African prospective cohort and the 32% seen among Chinese patients with multidrug-resistant pulmonary tuberculosis (MDR-TB).<sup>5,13-15</sup> We confirmed that silicosis, drug resistance, elevated C-reactive protein, hypoalbuminemia, and a prolonged symptom-to-treatment delay independently predicted LD, supporting associations documented for silica exposure,<sup>16,17</sup> MDR-TB,<sup>18</sup> systemic inflammation,<sup>19</sup> and malnutrition.<sup>18</sup> Whereas earlier tools relied on a single radiologic surrogate (e.g., initial cavity size or symptom duration),<sup>20-21</sup> our nine-factor nomogram increased the AUC to 0.934 and delivered the highest net clinical benefit across 10-40% risk thresholds, thereby providing substantially finer risk stratification while remaining transparent and feasible at the bedside.

Biologically, these predictors map onto well-described mechanisms of post-TB lung injury. Chronic inflammation-driven by CRP-induced matrix metalloproteinase activation and sustained oxidative stress-accelerates collagen degradation and cavity expansion; parallel elevations of AST and GGT in our high-risk patients support this inflammatory-fibrotic axis.<sup>22,23</sup> Conversely, robust host immunity and adequate nutrition promote repair: lymphopenia impairs collagen remodeling, and hypoalbuminemia delays cavity closure-findings consistent with Tanzanian and meta-

analytic data<sup>24-27</sup> and supported by the protective effects of higher lymphocyte counts and albumin in our model. Radiologically, baseline cavitation and subsegmental atelectasis herald later volume-loss fibrosis, as shown in Japanese and Turkish series.<sup>28-30</sup> Finally, the pronounced risk conferred by silicosis and MDR-TB illustrates both how inhaled silica amplifies fibrogenesis and how persistent bacillary burden in drug-resistant disease extends tissue destruction.<sup>31-33</sup> Together, these pathophysiological links lend mechanistic plausibility to the nomogram and underscore its potential to guide targeted surveillance and early intervention across diverse TB-burden settings.

Because the nomogram uses variables that are routinely collected at treatment initiation (symptom-to-treatment delay, lymphocyte count, CRP, AST, GGT, albumin, drug-resistant TB status, and a simple baseline imaging descriptor), it can be implemented as a paper tool or basic offline calculator without specialized infrastructure. In high-burden programs with constrained capacity, sites can select a risk threshold within the 10-40% range identified by decision-curve analysis to triage resources. For example, a higher threshold (e.g., nearer 30-40%) prioritizes follow-up CT/spirometry and early pulmonary-rehabilitation for only the highest-risk patients, whereas a lower threshold (e.g., nearer 10-20%) expands access when resources allow. A pragmatic workflow is to (i) compute risk at treatment initiation, (ii) allocate closer follow-up (earlier imaging/spirometry, rehabilitation, nutrition support) to those above the local threshold, and (iii) maintain standard follow-up otherwise. Where baseline CT is unavailable, programs may approximate the "atelectasis/cavity" indicator with chest-radiograph evidence of cavitation or volume-loss fibrosis; such adaptation should be accompanied by logistic recalibration (updating intercept and slope using local outcomes) once feasible, without re-estimating individual coefficients. These applications are intended to support risk-stratified surveillance while acknowledging that external validation remains necessary before formal clinical adoption.

The single-center, retrospective nature of our current study has certain limitations. Prospective, multicenter external validation is required to establish generalizability. We prioritized bootstrap internal

validation and parsimony, but with 61 events, residual bias is possible. Importantly, the endpoint was radiological since we did not collect spirometry or patient-reported outcomes. Thus, the nomogram estimates the risk of CT-defined lung destruction and should not be interpreted as a surrogate for functional limitation or health-related quality of life. Future work will prospectively validate the model in diverse settings, co-measure spirometry and a brief QoL instrument, and apply logistic recalibration when necessary for local transport without re-estimating individual coefficients.

In summary, we developed and internally validated a parsimonious, nine-factor nomogram that predicts post-treatment lung destruction with excellent discrimination and good calibration. By enabling early identification of high-risk individuals, the tool provides a practical pathway to risk-stratified follow-up imaging, timely referrals for pulmonary rehabilitation, and targeted preventive strategies, thereby improving long-term survivorship care for tuberculosis.

**Ethics approval and consent to participate.** The protocol was approved by the Zhuji People's Hospital's Institutional Review Board, and written informed consent was obtained for all participants. All procedures conformed to the Declaration of Helsinki.

**Funding.** This study was supported by the Zhejiang Provincial Traditional Chinese Medicine Science and Technology Program (Traditional Chinese Medicine Health Service Research Plan), Grant No. 2023ZF180, and the Zhejiang Provincial Traditional Chinese Medicine Science and Technology Program (Scientific Research Fund Project), Grant No. 2022ZA177.

**Data availability statement.** Data sets generated during the current study are available from the corresponding author on reasonable request.

**Author Contribution Statement.** The authors confirm contribution to the paper as follows: study conception and design: M.X.; data collection: M.X., J.H., J.F.; analysis and interpretation of results: M.X., J.H., J.F.; draft manuscript preparation: M.X., J.H., J.F. All authors reviewed the results and approved the final version of the manuscript.

## References:

1. Yang, H. Y., Ruan, X. D., Li, W., Xiong, J., & Zheng, Y. (2024). Global, regional, and national burden of tuberculosis and attributable risk factors for 204 countries and territories, 1990-2021: a systematic analysis for the Global Burden of Diseases 2021 study. *BMC Public Health*, 24(1) <https://doi.org/10.1186/s12889-024-20664-w> PMID:39529028 PMCID:PMC11552311
2. World Health Organization. (2014). The burden of disease caused by TB. In *Global tuberculosis report 2014* (pp. 7-31). World Health Organization. [https://iris.who.int/bitstream/handle/10665/137094/9789241564809\\_eng.pdf](https://iris.who.int/bitstream/handle/10665/137094/9789241564809_eng.pdf)
3. Allwood, B. W., Byrne, A., Meghji, J., Rachow, A., van der Zalm, M. M., & Schoch, O. D. (2021). Post-Tuberculosis Lung Disease: Clinical Review of an Under-Recognised Global Challenge. *Respiration*, 100(8), 751-763. <https://doi.org/10.1159/000512531> PMID:33401266
4. Cupido, G., & Günther, G. (2024). Post tuberculosis lung disease and tuberculosis sequelae: A narrative review. *Indian Journal of Tuberculosis*, 71(1), 64-72. <https://doi.org/10.1016/j.ijtb.2023.04.001>

- PMid:38296392
5. Irfan, M. (2016). Post-tuberculosis pulmonary function and noninfectious pulmonary disorders. *The International Journal of Mycobacteriology*, 5. <https://doi.org/10.1016/j.ijmyco.2016.08.015> PMID:28043612
  6. Yarbrough, C., Miller, M. S., Zulu, M., Sharp, D., Andom, A. T., Ndayizigiye, M., Seung, K. J., & Sonenthal, P. D. (2024). Post-tuberculosis lung disease: Addressing the policy gap. *PLOS Global Public Health*, 4(9), e0003560. <https://doi.org/10.1371/journal.pgph.0003560> PMID:39236033 PMCID:PMC11376554
  7. Seo, W., Kim, H.-W., Kim, J. S., & Min, J. (2024). Long term management of people with post-tuberculosis lung disease. *The Korean Journal of Internal Medicine*, 39, 7-24. <https://doi.org/10.3904/kjim.2023.395> PMID:38225822 PMCID:PMC10790047
  8. Paez, R., Kammer, M. N., Kammer, M. N., Massion, P. P., & Massion, P. P. (2021). Risk stratification of indeterminate pulmonary nodules. *Current Opinion in Pulmonary Medicine*, 27(4), 240-248. <https://doi.org/10.1097/MCP.0000000000000780> PMID:33973553
  9. Hsu, W., & Sohn, J. H. (2022). Using radiomics for risk stratification: Where we need to go. *Radiology*, 302(2), 435-437. <https://doi.org/10.1148/radiol.2021212085> PMID:34726541 PMCID:PMC8805519
  10. Kim, R. Y. (2024). Radiomics and artificial intelligence for risk stratification of pulmonary nodules: Ready for primetime? *Cancer Biomarkers*, 1-11. <https://doi.org/10.3233/CBM-230360> PMID:38427470 PMCID:PMC11300708
  11. Oke, J. L., Pickup, L. C., Declerck, J., Callister, M. E., Baldwin, D., Gustafson, J., Peschl, H., Ather, S., Tsakok, M., Exell, A., & Gleeson, F. (2018). Development and validation of clinical prediction models to risk stratify patients presenting with small pulmonary nodules: a research protocol. *Diagnostic and prognostic research*, 2, 22. <https://doi.org/10.1186/s41512-018-0044-3> PMID:31093569 PMCID:PMC6460802
  12. Liang, W., Tao, J., Cheng, C., Sun, H., Ye, Z., Wu, S., Guo, Y., Zhang, Q., Chen, Q., Liu, D., Liu, L., Tian, H., Lin, T., Zhong, N., Fan, J., & He, J. (2024). A clinically effective model based on cell-free DNA methylation and low-dose CT for risk stratification of pulmonary nodules. *Cell Reports Medicine*, 101750. <https://doi.org/10.1016/j.xcrm.2024.101750> PMID:39341207 PMCID:PMC11513810
  13. De Valliere, S., & Barker, R. D. (2004). Residual lung damage after completion of treatment for multidrug-resistant tuberculosis. *International Journal of Tuberculosis and Lung Disease*, 8(6), 767-771.
  14. Cupido, G., & Günther, G. (2024). Post tuberculosis lung disease and tuberculosis sequelae: A narrative review. *The Indian journal of tuberculosis*, 71(1), 64-72. <https://doi.org/10.1016/j.ijtb.2023.04.001> PMID:38296392
  15. Ma, Y., Ye, X. P., Fu, X. Y., & Wu, G. H. (2022). Zhonghua jie he he hu xi za zhi = Zhonghua jiehe he huxi zazhi = Chinese journal of tuberculosis and respiratory diseases, 45(10), 955-959.
  16. Ehrlich, R., Akugizibwe, P., Siegfried, N., & Rees, D. (2021). The association between silica exposure, silicosis and tuberculosis: a systematic review and meta-analysis. *BMC Public Health*, 21(1), 953. <https://doi.org/10.1186/s12889-021-10711-1> PMID:34016067 PMCID:PMC8136154
  17. Cheepsattayakorn, A., & Cheepsattayakorn, R. (2018). Silicosis-Associated Tuberculosis: Management and Control. *American Journal of Public Health Research*, 6(2), 125-129.
  18. Rupani, M. P. (2006). Silicosis predicts drug resistance and retreatment among tuberculosis patients in India: a secondary data analysis from Khambhat, Gujarat (2006-2022). *BMC Pulmonary Medicine*, 24(1). <https://doi.org/10.1186/s12890-024-03338-6> PMID:39425124 PMCID:PMC11490045
  19. D'Sa, L., Pezzuto, F., Lunardi, F., Scalvenzi, F., Tinè, M., Comacchio, G. M., Giraudo, C., & Calabrese, F. (2023). A Complex Case of Pulmonary Silico-Tuberculosis and Review of Literature. *Diagnostics*, 13(17). <https://doi.org/10.3390/diagnostics13172728> PMID:37685266 PMCID:PMC10486577
  20. Pandey, S. K., Singh, K. U., Dingankar, R. S., Jadhav, K., Gupta, K., & Yadav, R. K. (2023). Prediction of tuberculosis disease progression with AI analysis of clinical data. 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), 1-6. <https://doi.org/10.1109/ICAIIHI57871.2023.10489091>
  21. Lv, X., Li, Y., Cai, B., He, W., Wang, R., Chen, M., Pan, J., & Hou, D. (2023). Utility of Machine Learning and Radiomics Based on Cavity for Predicting the Therapeutic Response of MDR-TB. *Infection and drug resistance*, 16, 6893-6904. <https://doi.org/10.2147/IDR.S435984> PMID:37920476 PMCID:PMC10619461
  22. Lavrova, A. I., Esmedljaeva, D. S., Belik, V., & Postnikov, E. B. (2019). Matrix Metalloproteinases as Markers of Acute Inflammation Process in the Pulmonary Tuberculosis. *Data*, 4(4), <https://doi.org/10.3390/data4040137>
  23. Amaral, E. P., Vinhaes, C. L., Oliveira-de-Souza, D., Nogueira, B. M. F., Akrami, K., Akrami, K., & Andrade, B. B. (2021). The Interplay Between Systemic Inflammation, Oxidative Stress, and Tissue Remodeling in Tuberculosis. *Antioxidants & Redox Signaling*, 34(6), 471-485. <https://doi.org/10.1089/ars.2020.8124> PMID:32559410 PMCID:PMC8020551
  24. Squeglia, F., Ruggiero, A., & Berisio, R. (2018). Collagen degradation in tuberculosis pathogenesis: the biochemical consequences of hosting an undesired guest. *Biochemical Journal*, 475(19), 3123-3140. <https://doi.org/10.1042/BCJ20180482> PMID:30315001
  25. Kumar, N. P., Moideen, K., Nancy, A., Viswanathan, V., Thiruvengadam, K., Sivakumar, S., Hissar, S., Kornfeld, H., & Babu, S. (2021). Acute Phase Proteins Are Baseline Predictors of Tuberculosis Treatment Failure. *Frontiers in immunology*, 12, 731878. <https://doi.org/10.3389/fimmu.2021.731878> PMID:34867953 PMCID:PMC8634481
  26. Moores, R., Brilha, S., Brilha, S., Schutgens, F., Elkington, P. T., Elkington, P. T., & Friedland, J. S. (2017). Epigenetic Regulation of Matrix Metalloproteinase-1 and -3 Expression in Mycobacterium tuberculosis Infection. *Frontiers in Immunology*, 8, 602. <https://doi.org/10.3389/fimmu.2017.00602> PMID:28596772 PMCID:PMC5442172
  27. Källenius, G., Correia - Neves, M., & Sundling, C. (2024). Diagnostic markers reflecting dysregulation of the host response in the transition to TB disease. *International Journal of Infectious Diseases*, 106984. <https://doi.org/10.1016/j.ijid.2024.106984> PMID:38417614
  28. Im, J. G., Itoh, H., Lee, K. S., & Han, M. C. (1995). CT-pathology correlation of pulmonary tuberculosis. *Critical Reviews in Diagnostic Imaging*, 36(3), 227-285.
  29. Lee, J., Lee, K. S., Jung, K.-J., Han, J., Kwon, O., Kim, J. H., & Kim, T. (2000). Pulmonary tuberculosis: CT and pathologic correlation. *Journal of Computer Assisted Tomography*, 24(5), 691-698. <https://doi.org/10.1097/00004728-200009000-00005> PMID:11045687
  30. Deshpande, S., Joshi, A., & Shah, A. (2020). Aftermath of pulmonary tuberculosis: computed tomography assessment. *Polish Journal of Radiology*, 85(1), 137-147. <https://doi.org/10.5114/pjr.2020.93714> PMID:32322321 PMCID:PMC7172931
  31. Im, J. G., Itoh, H., & Han, M. C. (1995). CT of pulmonary tuberculosis. *Seminars in Ultrasound, CT and MRI*, 16(5), 420-434. [https://doi.org/10.1016/0887-2171\(95\)90029-2](https://doi.org/10.1016/0887-2171(95)90029-2) PMID:8527173
  32. Cheng, N., Wu, S., Luo, X., Xu, C., Lou, Q., Zhu, J., You, L., & Li, B. (2021). A Comparative Study of Chest Computed Tomography Findings: 1030 Cases of Drug-Sensitive Tuberculosis versus 516 Cases of Drug-Resistant Tuberculosis. *Infection and Drug Resistance*, 14, 1115-1128. <https://doi.org/10.2147/IDR.S300754> PMID:33776457 PMCID:PMC7987723
  33. Kim, S., Min, J., & Lee, J. (2014). Radiological Findings of Primary Multidrug-resistant Pulmonary Tuberculosis in HIV-seronegative Patients. *Hong Kong Journal of Radiology*, 17(1), 4-8. <https://doi.org/10.12809/hkjr1413213>