Journal of Multidisciplinary Healthcare

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Differentiating Lung Adenocarcinoma from Tuberculous Nodules in HIV/AIDS Patients Using Preoperative CT-Based Intratumoral and Peritumoral Radiomics Combined with Clinical Features

Chang Song[1,2,*], Chun-Yan Zhao[1,2,*], Shu-Lin Song[3,*], Yan-Rong Lin[1], Chao-Yan Xu[1], Hang-Biao Qiang[1], Ren-Hao Liu[2], Qi Li[4], Qing-Dong Zhu[1]

[1]Department of Tuberculosis, The Fourth People's Hospital of Nanning, Nanning, People's Republic of China; [2]Clinical Medical School, Guangxi Medical University, Nanning, People's Republic of China; [3]Department of Radiology, The Fourth People's Hospital of Nanning, Nanning, People's Republic of China; [4]Department of Tuberculosis, Beijing Chest Hospital, Capital Medical University, Beijing, People's Republic of China

*These authors contributed equally to this work

Correspondence: Qing-Dong Zhu, Department of Tuberculosis, The Fourth People's Hospital of Nanning, No. 1, Lane 2, Changgang Road, Xingning District, Nanning, Guangxi Zhuang Autonomous Region, 530021, People's Republic of China, Email zhuqingdong2003@163.com; Qi Li, Department of Tuberculosis, Beijing Chest Hospital, Capital Medical University, No. 97, Machang, Tongzhou District, Beijing, 101100, People's Republic of China, Email liqilisa@126.com

Purpose: This study aimed to develop and validate a preoperative CT-based radiomics nomogram model incorporating intratumoral and peritumoral features to accurately differentiate lung adenocarcinoma (LUAD) from pulmonary tuberculosis (PTB) nodules in HIV/AIDS patients.

Patients and Methods: This retrospective study analyzed clinical and CT imaging data from 187 hIV/AIDS patients (84 with LUAD and 103 with PTB) treated at the Fourth People's Hospital of Nanning. Patients were randomly divided into training and validation cohorts in a 7:3 ratio. Radiomics features were extracted from both the intratumoral region and a 2 mm peritumoral region, then combined with clinical factors (eg, fever, C-reactive protein levels, and cardiac disease) to develop multiple predictive models, including clinical model, intra model, peri 2mm model, fusion model, and combined model (which integrates clinical and fusion models). Diagnostic performance was evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and other metrics.

Results: The combined model achieved the highest AUC in both the training (0.978) and validation cohorts (0.969) cohorts, significantly outperforming the other models while mitigating the overfitting observed in the clinical model. Hosmer-Lemeshow (HL) tests, Integrated Discrimination Improvement (IDI), Net Reclassification Index (NRI), and decision curve analysis (DCA) confirmed its superior performance.

Conclusion: The CT-based radiomics nomogram model, intratumoral and peritumoral radiomics features, enables accurate differentiation between LUAD and PTB in HIV/AIDS patients, providing a non-invasive tool for preoperative diagnosis.

Keywords: HIV/AIDS, lung adenocarcinoma, tuberculosis pulmonary nodules, radiomics, differential diagnosis

Introduction

Lung cancer is one of the most common non-AIDS-defining cancers (NADCs) in individuals with human immunodeficiency virus. HIV-related factors influence lung cancer risk and may adversely affect treatment and outcomes.[1,2] In addition to being an independent risk factor for lung cancer, HIV is also associated with a younger age of cancer onset.[3] Despite the widespread use of antiretroviral therapy (ART), HIV-associated pulmonary diseases remain a leading cause of morbidity and mortality in HIV/AIDS patients.[4] Studies, including international research, indicate that lung adenocarcinoma (LUAD) is the most

Received: 20 March 2025
Accepted: 4 May 2025
Published: 12 May 2025

prevalent subtype of lung cancer in this population.[1,3,5] LUAD accounts for approximately 50% of lung cancer diagnoses and is the most aggressive and rapidly fatal form of non-small cell lung cancer (NSCLC).[6–8] Concurrently, HIV infection increases the risk of active tuberculosis (TB) by 26-fold and contributes to atypical clinical presentations, complicating diagnosis, treatment, and prognosis.[9] LUAD and pulmonary tuberculosis (PTB) nodules are two common pulmonary conditions in HIV/AIDS patients. Due to immunosuppression, these patients often present with overlapping clinical and radiological features, posing significant challenges for differential diagnosis.

Imaging examination serves as an indispensable tool in clinical diagnosis. Among various imaging modalities, computed tomography (CT), as an advanced imaging technique, plays a pivotal role in disease diagnosis and treatment planning due to its high resolution and rapid imaging capabilities. Among patients undergoing lung cancer CT screening, approximately half present with pulmonary nodules, yet over 95% of these nodules are benign.[10] Distinguishing LUAD from PTB nodules on chest imaging remains highly challenging, as nodular or mass-like PTB is often misdiagnosed as peripheral lung cancer, despite fundamental differences in treatment and prognosis.[11,12]

Recent advances in radiomics have enabled computer-aided analysis of medical imaging data, allowing for the extraction of quantitative imaging features for improved disease diagnosis and prognosis.[13–15] By analyzing multimodal imaging features of tumors and their microenvironment, radiomics provides comprehensive biological insights and enhances diagnostic accuracy. In HIV/AIDS patients, integrating radiomic and clinical features may offer novel strategies for differentiating LUAD and TB. CT, the primary imaging modality for pulmonary diseases, captures detailed morphological, dimensional, and density characteristics of nodules. However, traditional CT diagnosis relies on subjective interpretation, lacking quantification and standardization. Radiomics overcomes this limitation by automating the extraction of quantitative features (eg, texture, shape, intensity), enabling the development of robust diagnostic models. Intratumoral and peritumoral radiomics extends this approach by analyzing both the tumor core and its surrounding microenvironment, offering richer biological insights and significantly improving diagnostic specificity and accuracy. Previous studies have demonstrated its potential to enhance model performance.[16–19]

This study pioneers the development of a preoperative diagnostic model that integrates CT-based intratumoral and peritumoral radiomics with clinical features to differentiate LUAD from PTB nodules in HIV/AIDS patients. This model aims to provide clinicians with a reliable, non-invasive diagnostic tool to aid preoperative decision-making.

Materials and Methods

Study Population

This retrospective study included 187 hIV/AIDS patients (84 with LUAD and 103 with PTB) treated at the Fourth People's Hospital of Nanning between January 2016 to August 2024. Inclusion criteria were: (1) preoperative imaging showing solitary solid nodules or masses; (2) PTB cases were confirmed by pathology or sputum culture, and LUAD cases were confirmed by pathology; (3) nodule diameter <8 cm (due to self-limiting growth of TB nodules); (4) availability of detailed clinical records and preoperative CT scans within 4 weeks; (5) no history of lung surgery, radiotherapy, or chemotherapy. Exclusion criteria: (1) presence of precancerous lesions (eg, atypical adenomatous hyperplasia, adenocarcinoma in situ); (2) prior cancer treatment (including surgical therapy, chemotherapy, radiotherapy, targeted therapy, and ablation therapy); (3) coexisting malignancies. (4) patients who received diagnostic anti-tuberculosis treatment. Ethical approval was obtained from the hospital's ethics committee (Approval No. (2022)64), in compliance with the Declaration of Helsinki. Informed consent was waived due to the retrospective nature and anonymized data.

Clinical Data Collection

Clinical data were extracted from electronic medical records and divided into training and validation cohorts at a 7:3 ratio. Continuous variables (eg, age, BMI, CD3, CD4, CD8, CD4/CD8 ratio, white blood cell count (WBC), red blood cell count (RBC), hemoglobin, platelet count, neutrophil count, lymphocyte count, monocyte count, eosinophil count, basophil count, and C-reactive protein (CRP) were expressed as mean ± standard deviation (SD). Categorical variables included gender, comorbidities (syphilis, hepatitis B, hypertension, cardiac disease, diabetes), residence, smoking/alcohol history, and fever.

Imaging Acquisition

CT scans were performed using a GE LightSpeed 64-slice spiral CT scanner. Patients were positioned supine and instructed to remain motionless during scanning. Non-contrast images were acquired with a 5-mm slice thickness and reconstructed for high-resolution analysis. All CT scans were performed from the lung apex to the base to ensure full lung coverage. The reconstruction parameters were as follows: tube voltage 120 kV, tube effective power 30 mAs, detector collimation 128×0.625 mm, and matrix size 512×512. Without contrast injection, the slice thickness was set to 5 mm, and the CTDIvol value was 2.03 mGy to ensure clear lung imaging. To ensure the reliability of the research data, we established strict quality control standards for the CT images included in the analysis: In terms of technical parameters, images with a slice thickness exceeding 5 mm were excluded to ensure sufficient spatial resolution, and images with a signal-to-noise ratio (SNR) lower than 20 were removed using objective measurement methods. Regarding image quality, images with obvious metal artifacts, linear artifacts, or beam-hardening artifacts that affect the assessment of lesions were excluded. For lesion display requirements, solid nodules must have a CT value difference of≥100 hU compared to the surrounding lung tissue, while ground-glass nodules must have clear boundary features. In terms of reconstruction methods, the standard algorithm was strictly used to avoid the loss of details caused by excessive smoothing. Additionally, all images must fully cover the target organ range, and images with a CTDIvol<1.0 mGy, which are considered underdosed or overexposed, were excluded. These measures collectively ensured the accuracy and reliability of the research data.

Image Processing and Region of Interest (ROI) Segmentation

CT images in DICOM-format were exported to ITK-SNAP software for standardization. Voxel spacing was normalized to 1×1×1 mm, and window settings were optimized (width: 400; level: 40). ROIs were manually annotated by two radiologists and reviewed by a senior radiologist with over >20 years of experience to resolve discrepancies. Peritumoral regions were defined by expanding the tumor mask outward in 2-mm increments (Figure 1). All ROIs were then merged into 3D volumes of interest (VOIs).

Radiomics Feature Extraction and Selection

Handcrafted radiological features were categorized into three groups: geometric (measuring tumor shape and size), intensity (assessing voxel brightness), and texture. A comprehensive radiomics analysis was performed on both the entire volume of interest (VOI) and specific tumor subregions. The K-Nearest Neighbors (KNN) method was used to resolved unclustered regions, ensuring consistent labeling of habitat zones. Feature extraction was conducted with Pyradiomics (version 3.0.1), in compliance with Imaging Biomarker Standardization Initiative (IBSI) guidelines. To enhance predictive accuracy, features from different subregions were integrated using pre-fusion techniques. To ensure feature reliability, we first performed an Intraclass Correlation Coefficient (ICC) analysis, retaining only features with ICC > 0.8 as evaluated by two physicians, ensuring high inter-observer consistency. For correlation analysis, Pearson's correlation analysis was used to identify and remove highly correlated features (threshold: 0.9). The Minimum Redundancy Maximum Relevance (mRMR) algorithm was used to balance relevance and redundancy, optimizing the feature set to 32 features. Least Absolute Shrinkage and Selection Operator (LASSO) regression was employed for final feature

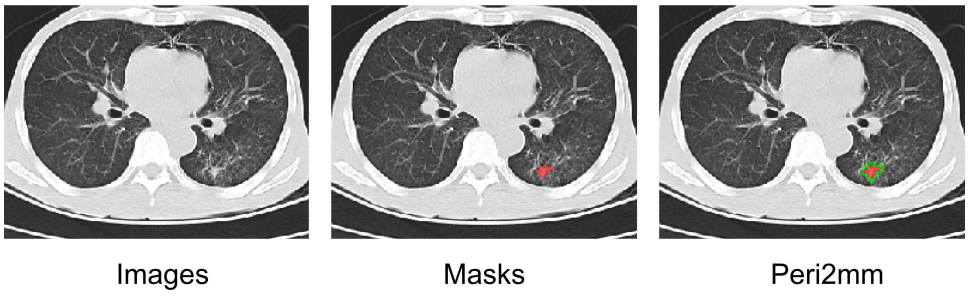

Figure 1 Presents the Peritumoral Regions Generated.

https://doi.org/10.2147/JMDH.S524527

selection, penalizing regression coefficients to eliminate non-essential features. The optimal regularization parameter (λ) was determined via 10-fold cross-validation, ensuring retention of the most predictive features. This rigorous approach multi-step process established a robust and predictive radiomics signature.

Model Development

This study characterizes intratumoral and peritumoral features to improve diagnostic and prognostic accuracy by identifying biomarkers linked to tumor behavior and treatment response. Advanced image processing and machine learning models were used to capture subtle morphological and functional variations. Peritumoral analysis provided further insights into the tumor microenvironment. Features were integrated within a multiple-instance learning framework, improving predictive specificity (Figure 2). For model development, machine learning models were trained to predict radiomics signatures (intra, peri, and fused) using LASSO-selected features. Hyperparameters were optimized via grid search and 5-fold cross-validation. Performance evaluation was conducted using 1000 bootstrap iterations of 5-fold cross-validation, with median AUC values compared across Intra, Peri, and Fusion models. Intra Radiomics Signature: Logistic Regression (LR) was used for linear modeling, while LR, Support Vector Machine (SVM), and Random Forest were employed for complex structures, integrating selected features into a detailed risk model that captures data nuances. Peri Radiomics Signature: The peritumoral model mirrored the intra-tumoral approach, utilizing identical feature selection and algorithms focused on peritumoral features. Fusion Signature: Feature fusion combined selected intra- and peri-tumoral features, maintaining consistency in feature selection and modeling. Clinical Signature: Clinical features were selected through univariate and multivariate analyses ($p < 0.05$) to develop a clinical model. Combined model: Significant clinical predictors ($p < 0.05$) were integrated with the strongest Peritumoral and Fusion Signatures via stepwise multivariate analyses, refining the final model to maximize predictive power.

Data Processing and Analysis

We assessed the normality of clinical features using the Shapiro–Wilk test. Continuous variables were evaluated for significance using the *t*-test or the Mann–Whitney *U*-test, depending on their distribution. Categorical variables were analyzed with Chi-square (χ^2) tests. The p-values between different cohorts were greater than 0.05, indicating no significant differences and confirming an unbiased division between groups. All data analyses were conducted on the Python 3.7.12. Statistical analyses were performed using Statsmodels version 0.13.2. Radiomics feature extraction was carried out using PyRadiomics version 3.0.1. Machine learning algorithms, including the Support Vector Machine (SVM), were implemented with Scikit-learn version 1.0.2.

Results

Clinical Baseline Characteristics

Supplementary Table 1 presents the statistical descriptions and comparative analyses of various biochemical indicators and demographic features from the overall sample (ALL), the training set (Train), and the test set (Test). For numerical variables, the distributions in the training and test sets are similar, with P-values greater than 0.05, indicating no significant differences between the two groups. For categorical variables (such as gender, syphilis, hepatitis B, hypertension, heart disease, diabetes, smoking history, alcohol consumption history, and fever history), although the proportions vary across groups, all P-values exceed 0.05, indicating consistency between the training and the test. This confirms the rationality of data partitioning, ensuring reliable model training and validation.

Univariate analysis was performed on all clinical features, calculating odds ratios (OR) and corresponding P-values for each feature. Notably, the Fever, CRP, and heart disease had P-values below 0.05, indicating statistical significance. Therefore, these features were selected as clinical comparison factors in subsequent analyses (Supplementary Table 2). To assess predictive performance, we constructed clinical indicator models using three machine learning models: LR, SVM, and RandomForest. We conducted classification analyses on the dataset, and model performance was evaluated on both the training and test sets (Supplementary Table 3). Results showed that performance metrics on the test set were generally lower

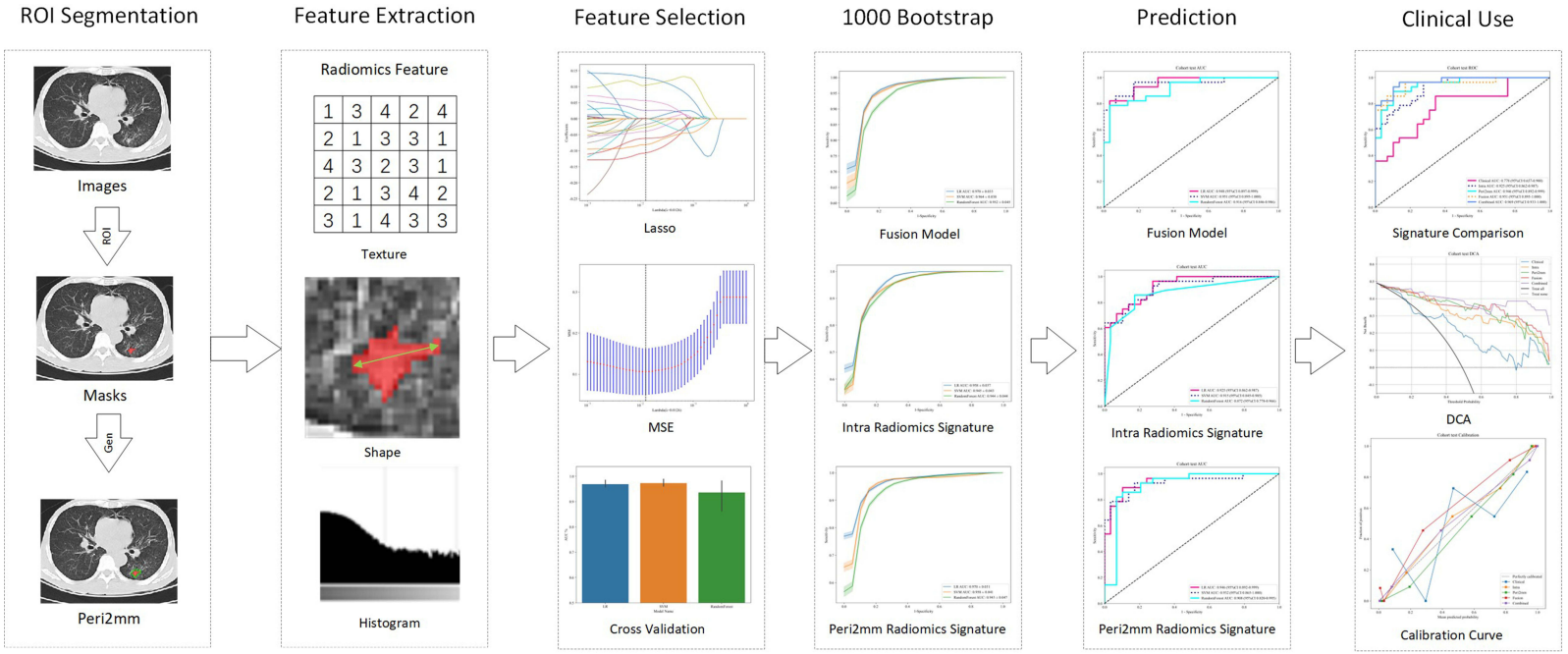

Figure 2 Flowchart of Research Study.

than those on the training set, suggesting a degree of overfitting in the clinical models. Supplementary Figure 1 illustrates the diagnostic 2×2 tables for the three machine learning models in both the training set and the test set.

Radiomics Feature Extraction and Selection

In our study, we extracted 360 first-order features, 14 shape features, and multiple texture features as handcrafted radiomics features. The final heterogeneity score was derived from both intratumoral and peritumoral subregions, comprising a total of 3668 features, with 1834 radiomics features and 1834 peritumoral features. Feature extraction was performed using a custom-developed tool in Pyradiomics. A visual representation of the feature distribution across different categories is provided (Supplementary Figure 2), offering an overview of their proportional composition within the dataset.

Intratumoral, Peritumoral, and Fusion Signatures

Figure 3A–C demonstrate the intratumoral feature screening process, including the Mean Standard Error (MSE) from 10-fold cross-validation, dimensionality reduction using LASSO regression, and the final selection of 11 optimal radiomics features. Figure 3D–F illustrate the peritumoral feature selection process, following the same 10-fold cross-validation, and LASSO regression-based screening, which resulted in the selection of 20 radiomics features. Figure 3G–I present the fusion process, which integrates intratumoral and peritumoral features, ultimately selecting 11 optimal radiomics features. Table 1 summarizes the performance metrics of the three machine learning models (intratumoral, peritumoral,

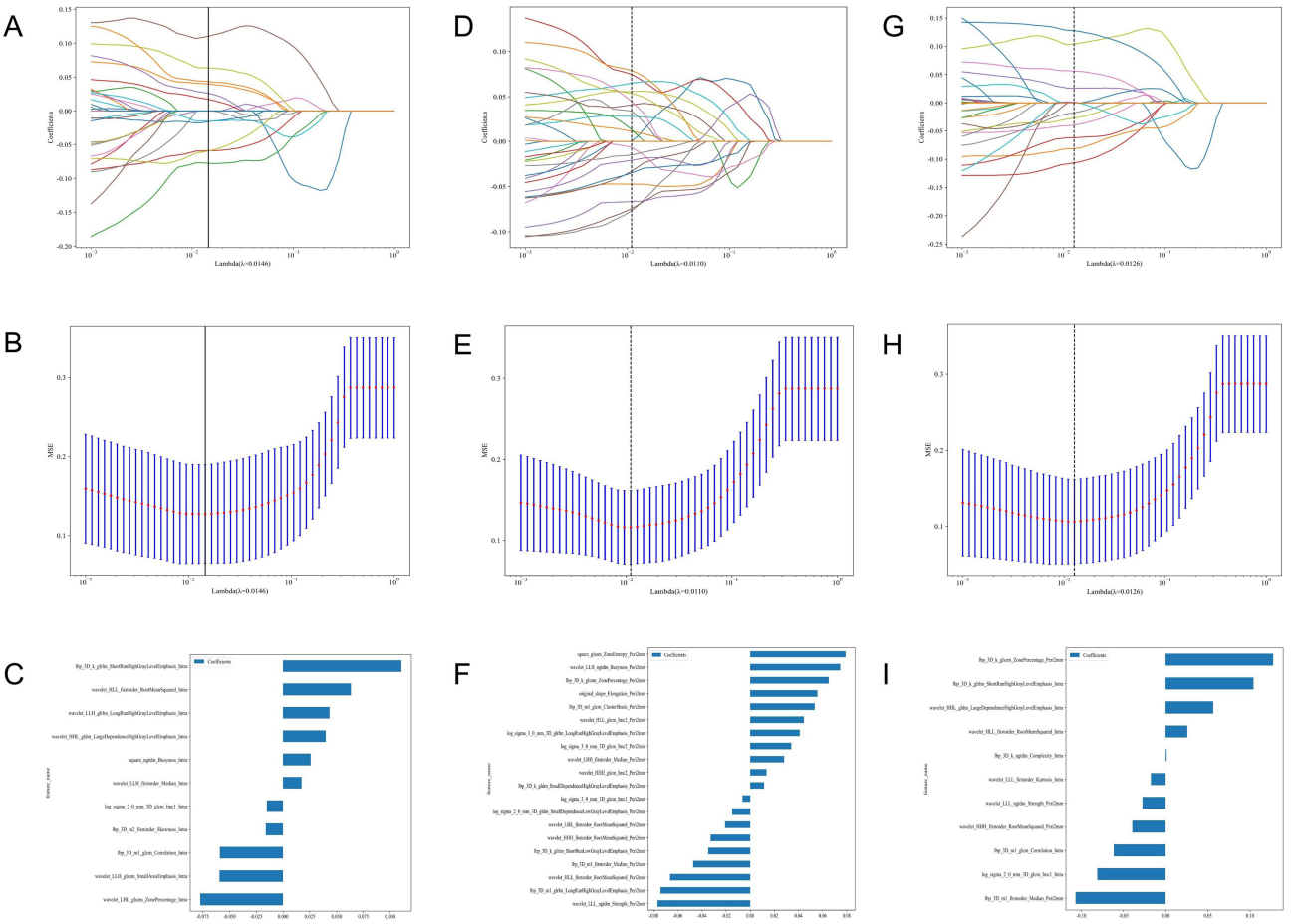

Figure 3 Coefficients of 10 Folds Cross Validation, MSE of 10 Folds Cross Validation and The Histogram of the Rad-score Based on the Selected Features (**A**–**C**) Intra Model; (**D**–**F**) Peri2mm Model; (**G**–**I**) Fusion Model.

Table 1 Model Performance of Different Machine Learning Algorithms in Each Cohorts

Model Name	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Accuracy
Intra Model							
LR (Train)	0.963	0.935 ~ 0.991	0.946	0.878	0.855	0.956	0.908
LR (Test)	0.925	0.862 ~ 0.987	0.786	0.793	0.786	0.793	0.789
SVM (Train)	0.954	0.919 ~ 0.990	0.946	0.878	0.855	0.956	0.908
SVM (Test)	0.915	0.845 ~ 0.985	0.750	0.862	0.840	0.781	0.807
RandomForest (Train)	0.946	0.902 ~ 0.990	0.875	0.932	0.907	0.908	0.908
RandomForest (Test)	0.872	0.778 ~ 0.966	0.679	0.897	0.864	0.743	0.789
Peri2mm Model							
LR (Train)	0.966	0.938 ~ 0.994	0.964	0.851	0.831	0.969	0.900
LR (Test)	0.946	0.892 ~ 0.999	0.929	0.793	0.812	0.920	0.860
SVM (Train)	0.987	0.973 ~ 1.000	0.982	0.932	0.917	0.986	0.954
SVM (Test)	0.932	0.865 ~ 1.000	0.821	0.862	0.852	0.833	0.842
RandomForest (Train)	0.956	0.925 ~ 0.988	0.911	0.851	0.823	0.926	0.877
RandomForest (Test)	0.908	0.820 ~ 0.995	0.857	0.862	0.857	0.862	0.860
Fusion Model							
LR (Train)	0.970	0.943 ~ 0.997	0.946	0.838	0.815	0.954	0.885
LR (Test)	0.948	0.897 ~ 0.999	0.821	0.828	0.821	0.828	0.825
SVM (Train)	0.975	0.948 ~ 1.000	0.911	0.932	0.911	0.932	0.923
SVM (Test)	0.951	0.895 ~ 1.000	0.786	0.931	0.917	0.818	0.860
RandomForest (Train)	0.974	0.953 ~ 0.994	0.893	0.905	0.877	0.918	0.900
RandomForest (Test)	0.916	0.846 ~ 0.986	0.786	0.966	0.957	0.824	0.877

and fusion models). Supplementary Figures 3–5 display the diagnostic contingency tables of three models in both the training and test sets.

Signature Comparison

In the training cohort, the Clinical model achieved an AUC of 0.862 (95% CI: 0.7994 ~ 0.9246), while the Intra model performed significantly better, with an AUC of 0.963 (95% CI: 0.9346 ~ 0.9910). The Peri 2-mm model further improved to an AUC of 0.966 (95% CI: 0.9375 ~ 0.9940), and the Fusion model demonstrated even greater predictive power, reaching an AUC of 0.975 (95% CI: 0.9478 ~ 1.0000). The combined model yielded the highest AUC of 0.978 (95% CI: 0.9544 ~ 1.0000). In the test cohort, the Clinical model showed a lower AUC of 0.778 (95% CI: 0.6568 ~ 0.8998), with the Intra model improving to 0.925 (95% CI: 0.8625 ~ 0.9873). The Peri 2-mm and Fusion models performed strongly, with AUCs of 0.946 (95% CI: 0.8923 ~ 0.9993) and 0.951 (95% CI: 0.8947 ~ 1.0000), respectively. The combined model achieved the highest AUC of 0.969 (95% CI: 0.9329 ~ 1.0000) (Table 2).

The Fusion model demonstrated superior performance compared to both Clinical and Radiomics models, confirming that a fusion of intra-tumoral and peri-tumoral features captures more information than either clinical data or radiomics alone. The combined model, which integrates Clinical, Intra, Peri, and Fusion features, further enhanced predictive accuracy, indicating that multi-modal fusion improves model performance.

Table 2 Metrics on Different Signature

Signature	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Accuracy
Train Cohort							
Clinical	0.862	0.7994 ~ 0.9246	0.804	0.770	0.726	0.838	0.785
Intra	0.963	0.9346 ~ 0.9910	0.946	0.878	0.855	0.956	0.908
Peri2mm	0.966	0.9375 ~ 0.9940	0.964	0.851	0.831	0.969	0.900
Fusion	0.975	0.9478 ~ 1.0000	0.911	0.932	0.911	0.932	0.923
Combined	0.978	0.9544 ~ 1.0000	0.946	0.946	0.930	0.959	0.946
Test Cohort							
Clinical	0.778	0.6568 ~ 0.8998	0.679	0.690	0.679	0.690	0.684
Intra	0.925	0.8625 ~ 0.9873	0.786	0.793	0.786	0.793	0.789
Peri2mm	0.946	0.8923 ~ 0.9993	0.929	0.793	0.812	0.920	0.860
Fusion	0.951	0.8947 ~ 1.0000	0.786	0.931	0.917	0.818	0.860
Combined	0.969	0.9329 ~ 1.0000	0.893	0.897	0.893	0.897	0.895

Model Calibration Performance and Superiority Analysis of the Combined Model

The HL test demonstrated excellent calibration performance, with HL values of 0.403 (training cohort) and 0.705 (test cohorts), confirming the accuracy and reliability of the model predictions (Figure 4A and B). To assess improvements in predictive performance, the DeLong test was conducted, showing that the combined model consistently outperformed most individual models (Figure 4C and D). Additionally, the Integrated Discrimination Improvement (IDI) and Net Reclassification Index (NRI) were both positive, indicating that the combined model improved classification accuracy and discriminative capability compared to other models (Figure 5).

Clinical Utility Evaluation

Figure 6A and B display the DCA curves for the training and test cohorts. The analysis demonstrated that the combined model provided a higher net benefit across different prediction probability ranges. To enhance clinical interpretability, we constructed a nomogram (Figure 6C), which visually represents the variables included in the combined model, enabling personalized probability estimation and highlighting the relative importance of each feature.

Discussion

In recent years, radiomics has revolutionized disease differentiation by enabling the high-throughput extraction of quantitative imaging features (eg, texture, morphology, and wavelet transform features) and the development of predictive models using machine learning algorithms. In distinguishing LUAD from PTB, studies have demonstrated the significant value of intratumoral heterogeneity features, such as contrast and correlation from the gray-level co-occurrence matrix.[10,11] However, existing research has three major limitations. Neglect of the peritumoral microenvironment: Immune cell infiltration, angiogenesis, and extracellular matrix remodeling in the peritumoral region differ significantly between lung cancer and granulomatous lesions, yet these critical features remain largely unexplored. Lack of specific data for HIV/AIDS populations: Current radiomics studies primarily focus on individuals with normal immune function, overlooking the altered immune microenvironment in HIV-infected patients, which may influence imaging feature expression. Insufficient exploration of multimodal fusion strategies: while clinical features with radiomics is expected to enhance model performance, current studies often rely on simple feature concatenation without in-depth investigation of feature interactions.

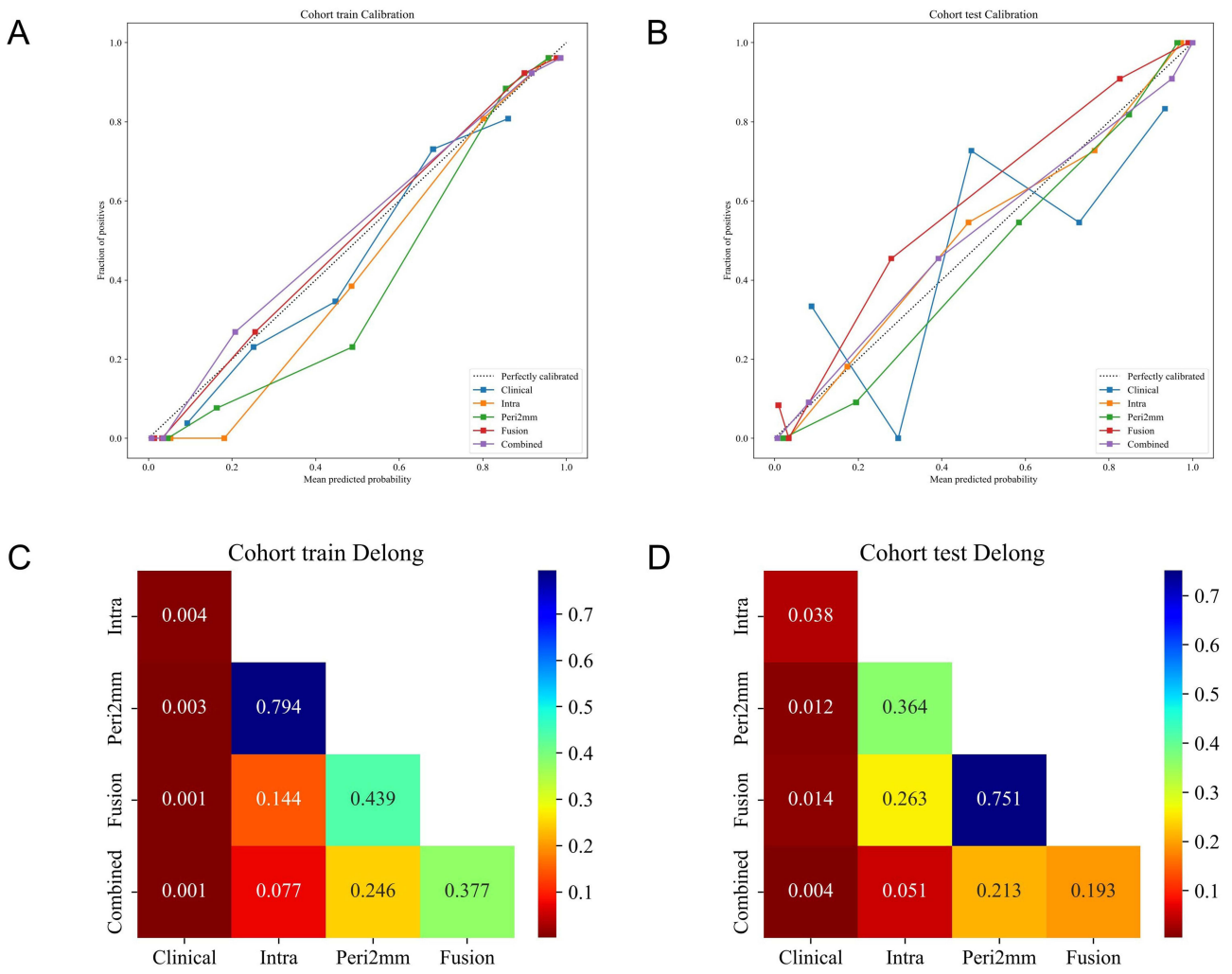

Figure 4 Model Evaluation (**A** and **B**) Different Signatures' Calibration Curve; (**C** and **D**) DeLong Test of Different.

To address these challenges, this study proposes an innovative "intratumoral and peritumoral radiomics + clinical features" three-dimensional integration. For the first time, we developed and validated a CT-based model capable of accurately differentiating LUAD and PTB nodules in HIV/AIDS patients.

Intratumoral radiomics features capture the tumor's internal structure and biological characteristics, while peritumoral radiomics features provide insight into the tumor's interaction with surrounding tissues, offering richer diagnostic information. Additionally, incorporating clinical features enhances the robustness and clinical practicality of the model, ensuring reliable diagnostic outcomes in the complex HIV/AIDS population. This integrated model serves as an intuitive decision-support tool for clinicians, enabling preoperative, non-invasive differentiation between LUAD and PTB. By facilitating more precise diagnoses, it supports the development of personalized treatment strategies, ultimately improving patient outcomes.

In the clinical model, the distribution characteristics of all numerical variables were similar between the training and testing sets, with P-values greater than 0.05, indicating no significant differences between the two groups. The further confirms the randomness and balance of data partitioning. Univariate and multivariate analyses were performed on all clinical features, revealing that fever, CRP, and heart disease had P-values less than 0.05, indicating statistical significance. Fever is a common clinical symptom in PTB, often associated with infections and inflammatory diseases. In this study, fever was identified as a significant clinical feature, potentially reflecting the patient's inflammatory response or immune status, which is crucial for disease diagnosis and prognosis assessment.[20–22] CRP, an acute-phase

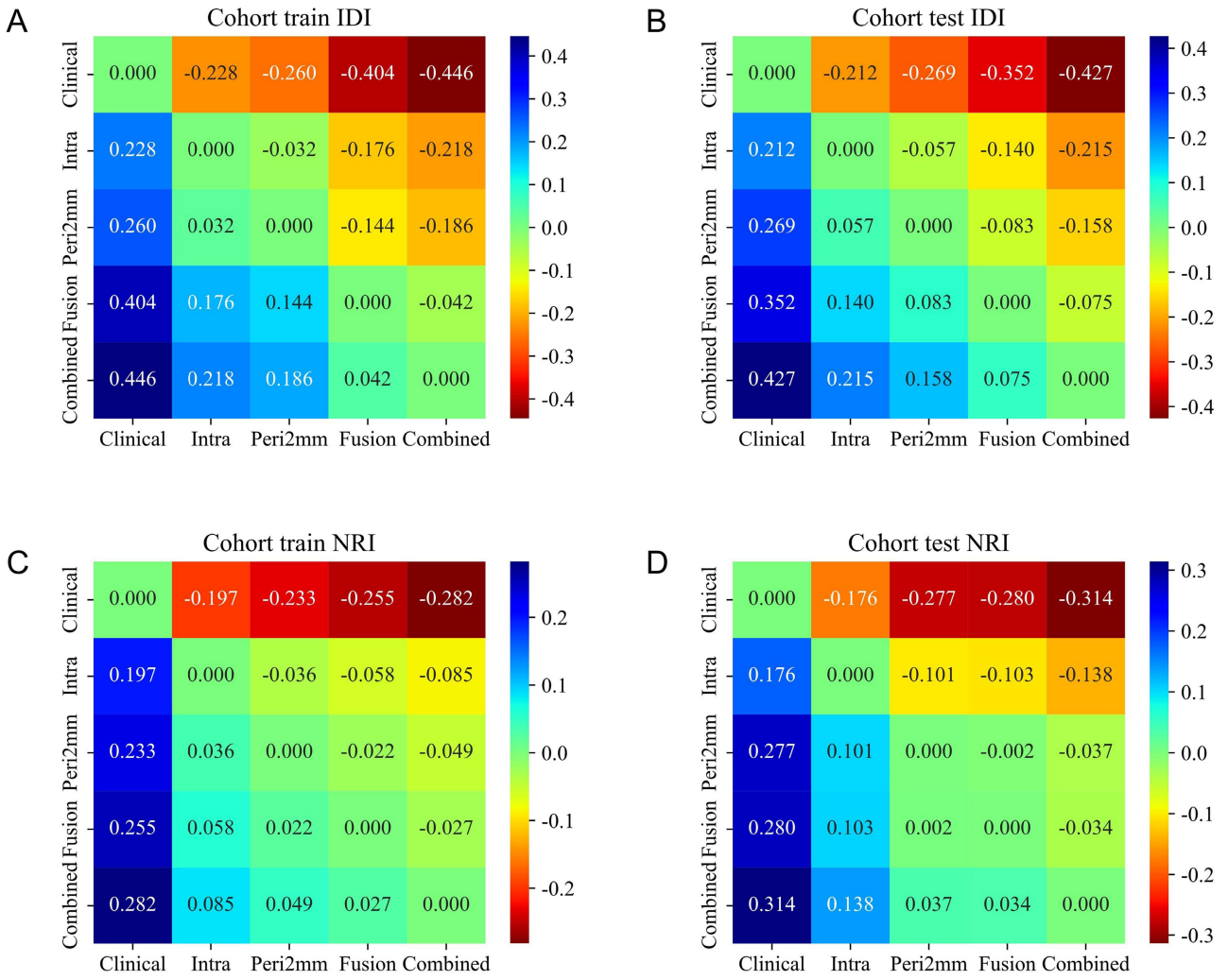

Figure 5 Performance improvement of models (**A** and **B**) Integrated Discrimination Improvement; (**C** and **D**) Net Reclassification Index.

reactant, reflects the intensity of inflammatory responses,[23] and previous studies have suggested its utility as a rapid screening tool for tuberculosis.[24–26] Our results unexpectedly identified cardiac disease as a predictive factor in differentiating pulmonary tuberculosis from lung cancer. Khoufi et al's findings of elevated latent tuberculosis rates in ischemic heart disease patients suggest that tuberculosis screening should be incorporated into cardiac risk assessments.[27] Intriguingly, He et al's Mendelian randomization analysis revealed an inverse relationship between coronary heart disease events and lung cancer risk.[28] Collectively, these studies indicate that cardiac disease markers may serve as valuable biomarkers, offering new perspectives for developing clinical diagnostic models for pulmonary tuberculosis and lung cancer. A clinical prediction model was developed using three machine learning methods, and its performance decline of the testing set a risk of overfitting, likely due to the limited dimensionality of clinical features, making it difficult for the model to capture complex nonlinear relationships. Inadequate handling of noise or multicollinearity in high-dimensional clinical data. These findings highlight the limitations of relying solely on static clinical features, suggesting that traditional indicators alone may not meet high-precision prediction requirements.

In comparing the Intra Model, Peri 2-mm Model, and Fusion Model, the Fusion Model exhibited higher AUC values in both the training and testing sets, suggesting the integration intra- and peritumoral radiomics provides more detailed information. The selected features demonstrated strong generalization ability, effectively avoiding overfitting. Consistent with Niu et al's study, which identified the 2-mm peritumoral region as the optimal differentiation area,[29] this study also adopted a 2-mm extension area for radiomics feature extraction, distinguishing it from previous research. Most previous

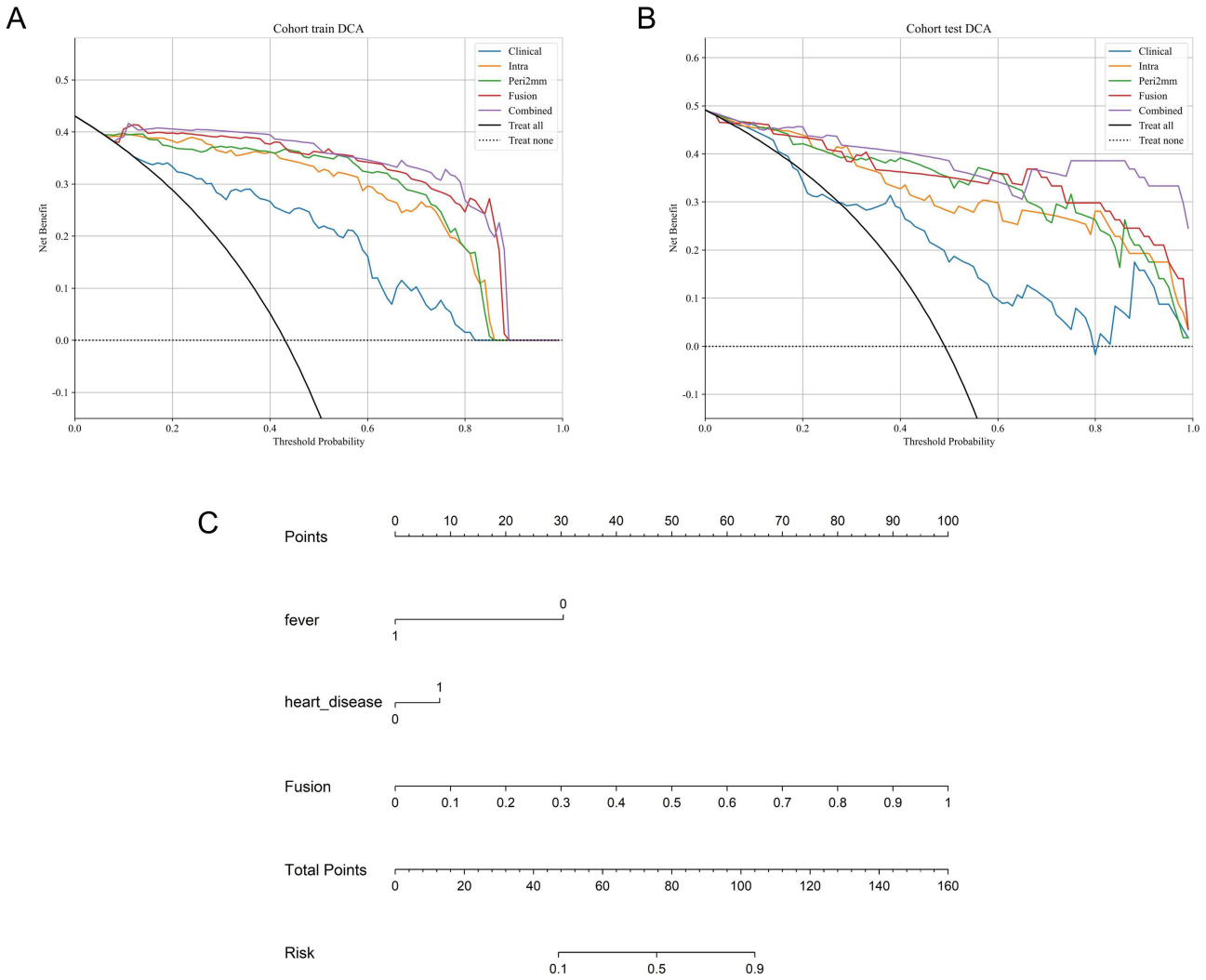

Figure 6 Performance improvement of models (**A**) DCA Curve of the Train Cohort; (**B**) DCA curve of the Test Cohort; (**C**) Nomogram for Clinical Use.

studies on LUAD-PTB differentiation constructed radiomics models based on a single intra-tumoral region, capturing internal tumor characteristics while overlooking peritumoral microenvironmental information. For example, Dong et al reported AUC values of 0.8344 and 0.751 in the training and testing datasets, respectively, for a model combining intra-tumoral radiomics and clinical features.[12] By contrast, the Fusion Model in this study not only considers intratumoral heterogeneity but also integrates 2-mm peritumoral features, providing a more comprehensive tumor assessment. Peritumoral features may reflect tumor-host interactions, including inflammatory responses, immune responses, and angiogenesis, which are crucial for disease diagnosis and prognosis prediction. The enhanced performance of the hybrid model indicates the complementary nature of t intratumoral and peritumoral features, collectively enhancing the model's predictive accuracy. This data fusion strategy offers a more complete tumor characterization, effectively captures the complexity of disease pathology and improving the model's robustness and reliability.

Furthermore, integrating clinical indicators with the Fusion Model further increased the AUC values to 0.978 in the training set and 0.969 in the testing set, demonstrating the substantial potential of multimodal data fusion in disease prediction. Clinical indicators provide insights into patients' health status and risk factors, while radiomics features reveal tumor characteristics at the microscopic level. The combined model not only improves predictive accuracy but also enhances the clinical interpretability and practicality. By leveraging the model's predictions alongside individual patient presentations, clinicians can develop more personalized treatment strategies, thereby improving treatment outcomes and patient quality of life.

Although the model developed in this study demonstrated good predictive performance in internal validation, an important limitation must be acknowledged: the study lacks external validation. As the data were derived from a single-center cohort with a limited sample size, the generalizability of the model to other populations, different medical centers, or diverse ethnic groups requires further verification. Future studies should incorporate multicenter collaborations and include more diverse population samples for external validation to better assess the model's generalizability and clinical translational value. Additionally, multimodal data integration requires advanced data processing and modeling techniques to ensure the complementarity and synergy between different feature sets. Future research could explore deep learning approaches, such as Convolutional Neural Networks and Recurrent Neural Networks, to further enhance the model's predictive capabilities and interpretability.

Conclusion

This study developed and validated a CT-based radiomics nomogram model capable of accurately differentiating LUAD and PTB nodules in HIV/AIDS patients. The findings systematically demonstrate the superiority of multimodal fusion strategies in disease prediction. By integrating clinical features with multi-regional radiomics data, the combined model not only significantly improves predictive accuracy but also exhibits excellent calibration and clinical applicability. These results provide methodological support for precision medicine and establish a theoretical foundation for the future development of intelligent decision-support systems.

Abbreviations

AI, Artificial Intelligence; ART, Antiretroviral Therapy; AUC, Area Under the Curve; BMI, Body Mass Index; CNN, Convolutional Neural Network; CRP, C-Reactive Protein; CT, Computed Tomography; DCA, Decision Curve Analysis; DICOM, Digital Imaging and Communications in Medicine; DL, Deep Learning; HL, Hosmer-Lemeshow; HRCT, High-Resolution Computed Tomography; HIV/AIDS, Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome; ICC, Intraclass Correlation Coefficient; IDI, Integrated Discrimination Improvement; IBSI, Imaging Biomarker Standardization Initiative; KNN, K-Nearest Neighbors; LASSO, Least Absolute Shrinkage and Selection Operator; LR, Logistic Regression; LUAD, Lung Adenocarcinoma; mRMR, Minimum Redundancy Maximum Relevance; MRI, Magnetic Resonance Imaging; NADCs, Non-AIDS-Defining Cancers; NRI, Net Reclassification Index; NSCLC, Non-Small Cell Lung Cancer; OR, Odds Ratio; PTB, Pulmonary Tuberculosis; RBC, Red Blood Cell; ROC, Receiver Operating Characteristic; ROI, Region Of Interest; SD, Standard Deviation; SVM, Support Vector Machine; TB, Tuberculosis; VOI, Volume of Interest; WBC, White Blood Cell.

Data Sharing Statement

All data generated or analyzed during this study are included in this article. Further enquiries can be directed to the corresponding author.

Ethics Approval and Informed Consent

The Ethics Committee of Human Research Ethics Committee of the Fourth People's Hospital of Nanning (Approval No: [2022]64). The requirement for patient consent was waived, granting ethical clearance for this study. The study removed all the relevant personal information of the individuals except age and gender. This research based on human samples have been performed in accordance with the principles stated in the Declaration of Helsinki.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by the Guangxi Zhuang Autonomous Region Health Commission Self-Financed Scientific Research Project (Z-A20231211); Nanning Scientific Research and Technology Development Project (20233069).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Sigel K, Pitts R, Crothers K. Lung Malignancies in HIV infection. *Semin Resp Crit Care Med*. 2016;37(2):267–276. doi:10.1055/s-0036-1578803
2. Sigel K, Makinson A, Thaler J. Lung cancer in persons with HIV. *Curr Opin HIV AIDS*. 2017;12(1):31–38. doi:10.1097/COH.0000000000000326
3. Bhikoo R, Koegelenberg CFN. Lung cancer in people living with HIV: a different kettle of fish? *Afr J Thorac Crit Care Med*. 2022;28(2). doi:10.7196/AJTCCM.2022.v28i2.245
4. Cribbs SK, Crothers K, Morris A. Pathogenesis of HIV-related lung disease: immunity, infection, and inflammation. *Physiol Rev*. 2020;100(2):603–632. doi:10.1152/physrev.00039.2018
5. Bearz A, Vaccher E, Martellotta F, et al. Lung cancer in HIV positive patients: the GICAT experience. *Eur Rev Med Pharmacol Sci*. 2014;18(4):500–508.
6. Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. *Ann Glob Health*. 2019;85(1). doi:10.5334/aogh.2419
7. Bade BC, Dela Cruz CS. Lung cancer 2020: epidemiology, etiology, and prevention. *Clinics Chest Med*. 2020;41(1):1–24. doi:10.1016/j.ccm.2019.10.001
8. Thai AA, Solomon BJ, Sequist LV, Gainor JF, Heist RS. Lung cancer. *Lancet*. 2021;398(10299):535–554. doi:10.1016/S0140-6736(21)00312-3
9. Yang Q, Han J, Shen J, Peng X, Zhou L, Yin X. Diagnosis and treatment of tuberculosis in adults with HIV. *Medicine*. 2022;101(35):e30405. doi:10.1097/MD.0000000000030405
10. Li Y, Lyu B, Wang R, et al. Machine learning-based radiomics to distinguish pulmonary nodules between lung adenocarcinoma and tuberculosis. *Thoracic Cancer*. 2024;15(6):466–476. doi:10.1111/1759-7714.15216
11. Yang L, Jiang Z, Tong J, Li N, Dong Q, Wang K. Development and validation of a preoperative CT-based radiomics nomogram to differentiate tuberculosis granulomas from lung adenocarcinomas: an external validation study. *BMC Cancer*. 2024;24(1):670. doi:10.1186/s12885-024-12422-3
12. Dong Q, Wen Q, Li N, et al. Radiomics combined with clinical features in distinguishing non-calcifying tuberculosis granuloma and lung adenocarcinoma in small pulmonary nodules. *PeerJ*. 2022;10:e14127. doi:10.7717/peerj.14127
13. Avanzo M, Stancanello J, Pirrone G, Sartor G. Radiomics and deep learning in lung cancer. *Strahlentherapie und Onkologie*. 2020;196(10):879–887. doi:10.1007/s00066-020-01625-9
14. Chen M, Copley SJ, Viola P, Lu H, Aboagye EO. Radiomics and artificial intelligence for precision medicine in lung cancer treatment. *Semi Cancer Biol*. 2023;93:97–113. doi:10.1016/j.semcancer.2023.05.004
15. Guiot J, Vaidyanathan A, Deprez L, et al. A review in radiomics: making personalized medicine a reality via routine imaging. *Med Res Rev*. 2022;42(1):426–440. doi:10.1002/med.21846
16. Liao T, Yang Y, Lin X, Ouyang R, Deng Y, Ma J. High-risk breast lesions: a combined intratumoral and peritumoral radiomics nomogram model to predict pathologic upgrade and reduce unnecessary surgical excision. *Front Oncol*. 2024;14:1479565. doi:10.3389/fonc.2024.1479565
17. Schmitz F, Voigtländer H, Strauss D, et al. Differentiating low- and high-proliferative soft tissue sarcomas using conventional imaging features and radiomics on MRI. *BMC Cancer*. 2024;24(1):1589. doi:10.1186/s12885-024-13339-7
18. Wang B, Guo H, Zhang M, et al. Prediction of soft tissue sarcoma grading using intratumoral habitats and a peritumoral radiomics nomogram: a multi-center preliminary study. *Front Oncol*. 2024;14:1433196. doi:10.3389/fonc.2024.1433196
19. Zhao W, Hou M, Wang J, Song D, Niu Y. Interpretable machine learning model for predicting clinically significant prostate cancer: integrating intratumoral and peritumoral radiomics with clinical and metabolic features. *BMC Med Imaging*. 2024;24(1):353. doi:10.1186/s12880-024-01548-2
20. Jeong YJ, Park JS, Kim HW, et al. Deaths from tuberculosis: differences between tuberculosis-related and non-tuberculosis-related deaths. *Front Public Health*. 2023;11:1207284. doi:10.3389/fpubh.2023.1207284
21. Nguyen TK, Nguyen YH, Nguyen HT, Khong QM, Tran NK. Etiologies of fever of unknown origin in HIV/AIDS patients, Hanoi, Vietnam. *BMC Infect Dis*. 2022;22(1):61. doi:10.1186/s12879-022-07049-3
22. Saini A, Yadav G, Gothwal M, Singh P, Kathuria P, Elhence P. Tuberculosis and ovarian malignancy: sometimes mimics, sometimes coexists. *J Obstet Gynaecol Res*. 2020;46(6):945–949. doi:10.1111/jog.14223
23. Rizo-Téllez SA, Sekheri M, Filep JG. C-reactive protein: a target for therapy to reduce inflammation. *Front Immunol*. 2023;14:1237729. doi:10.3389/fimmu.2023.1237729
24. Fusani L, Tersigni C, Chiappini E, Venturini E, Galli L. Old biomarkers in tuberculosis management: are they still useful? A systematic review. *Exp Rev Anti-Infect Ther*. 2021;19(10):1191–1203. doi:10.1080/14787210.2021.1898945
25. Zimmer AJ, Lainati F, Aguilera Vasquez N, et al. Biomarkers that correlate with active pulmonary tuberculosis treatment response: a systematic review and meta-analysis. *J Clin Microbiol*. 2022;60(2):e0185921. doi:10.1128/jcm.01859-21
26. Kagujje M, Mwanza W, Somwe P, Chilukutu L, Creswell J, Muyoyeta M. Sensitivity and specificity of CRP and symptom screening as tuberculosis screening tools among HIV-positive and negative outpatients at a primary healthcare facility in Lusaka, Zambia: a prospective cross-sectional study. *BMJ Open*. 2023;13(4):e061907. doi:10.1136/bmjopen-2022-061907
27. Khoufi EAA. Association between latent tuberculosis and ischemic heart disease: a hospital-based cross-sectional study from Saudi Arabia. *Pan Afr Med J*. 2021;38:362. doi:10.11604/pamj.2021.38.362.28110
28. He D, Lu H, Ou X, et al. Exposure to major coronary heart disease events reduces lung cancer risk: a Mendelian randomization study based on a European population. *BMC Cancer*. 2025;25(1):152. doi:10.1186/s12885-025-13485-6
29. Niu S, Yu T, Cao Y, Dong Y, Luo Y, Jiang X. Digital breast tomosynthesis-based peritumoral radiomics approaches in the differentiation of benign and malignant breast lesions. *Diagn Interv Radiol*. 2022;28(3):217–225. doi:10.5152/dir.2022.20664