



OPEN ACCESS

EDITED BY

Paolo Borghetti,
Civil Hospital of Brescia, Italy

REVIEWED BY

Antonio Pontoriero,
University of Messina, Italy
Wang Fang,
University of Chinese Academy of Sciences,
China

*CORRESPONDENCE

Haitao Ma
✉ mht7403@163.com

RECEIVED 28 July 2025

ACCEPTED 21 October 2025

PUBLISHED 07 November 2025

CITATION

Geng N, Li Z, Deng L, Li Y and Ma H (2025)
Predicting immunotherapy response in stage
III–IV non-small cell lung cancer using
integrated radiomics and clinical features.
Front. Oncol. 15:1669469.
doi: 10.3389/fonc.2025.1669469

COPYRIGHT

© 2025 Geng, Li, Deng, Li and Ma. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](#)
(CC BY). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Predicting immunotherapy response in stage III–IV non-small cell lung cancer using integrated radiomics and clinical features

Na Geng¹, Zhijun Li², Langlang Deng¹, Yuxiang Li³
and Haitao Ma^{1*}

¹Department of Thoracic Surgery, The Fourth Affiliated Hospital of Soochow University, Suzhou, China, ²Department of Respiratory Medical, The Second Affiliated Hospital of Bengbu Medical University, Bengbu, China, ³Department of Rheumatology and Immunology, The Third Affiliated Hospital of Soochow University, Changzhou, China

Objective: To develop a combined predictive model based on CT radiomics and clinical features and evaluate its diagnostic value for predicting the efficacy prognosis of immunotherapy in stage III–IV non-small cell lung cancer (NSCLC).

Methods: A retrospective analysis was conducted on 106 patients with stage IIIa–IVB NSCLC who underwent immunotherapy at the Second Affiliated Hospital of Soochow University between December 2018 and December 2023. Patients were divided into two groups based on whether their progression-free survival (PFS) exceeded 12 months. The cohort was randomly split into a training set (75 patients) and a validation set (31 patients) in a 7:3 ratio. Clinical and imaging data were collected, and independent predictive factors were identified through univariate and multivariate logistic regression analysis to construct a clinical feature model. Radiomic features were extracted from contrast-enhanced chest CT images, and LASSO algorithm along with Pearson correlation coefficients were applied to select optimal features and calculate a radiomics score. A combined predictive model integrating clinical independent predictors and radiomic features was developed and visualized as a nomogram. Model performance was assessed by subject work characteristics (ROC) curves and area under the curve (AUC). Clinical utility was assessed via decision curve analysis (DCA), and calibration curves were used to evaluate the nomogram's predictive accuracy.

Results: Tumor location was an independent predictor of immunotherapy efficacy and formed the clinical model. Twelve contrast-enhanced CT radiomic features comprised the radiomics model. The combined model (clinical + radiomic) demonstrated superior diagnostic performance: training set AUCs (clinical: 0.705, radiomics: 0.835, combined: 0.896); validation set AUCs (clinical: 0.691, radiomics: 0.833, combined: 0.863). The combined model's AUC was significantly higher than either submodel alone in both sets. DCA confirmed its highest net clinical benefit, and calibration curves indicated good accuracy.

Conclusion: This study developed a predictive model based on clinical and radiomic features for assessing immunotherapy efficacy in NSCLC. The model demonstrated excellent performance, suggesting its potential as a clinical decision-support tool for prognosis prediction and treatment planning in NSCLC immunotherapy.

KEYWORDS

radiomics, immunotherapy response, non-small cell lung cancer, nomogram, decision curve analysis

1 Introduction

The advent of immune checkpoint inhibitors (ICIs) has revolutionized the treatment paradigm for advanced non-small cell lung cancer (NSCLC), with approximately 20-30% of patients achieving durable responses (1). However, the current gold-standard biomarker PD-L1 expression exhibits limited predictive accuracy (objective response rate concordance: 45-60%) (2, 3), while emerging biomarkers like tumor mutational burden (TMB) face challenges in clinical standardization (4–6). This predictive uncertainty leads to non-negligible rates of premature treatment discontinuation (34.7% in real-world studies (7)) and unnecessary immune-related adverse events (irAEs), highlighting the urgent need for more robust stratification tools. Radiomics, as a non-invasive approach decoding tumor heterogeneity through high-throughput imaging feature analysis, has demonstrated unique advantages in ICI response prediction. Recent studies revealed that CT-based radiomic signatures could reflect tumor microenvironment characteristics (e.g., CD8+ T-cell infiltration (8)) and predict progression-free survival (PFS) with AUCs of 0.71-0.79 (9, 10). By extracting quantitative radiomics features from patient CT images and computing the radiomics signature, combined with selecting predictive clinical indicators from patient medical records, we developed and validated a Combined Forecasting Model (CFM) to predict treatment response to immunotherapy in non-small cell lung cancer (NSCLC) patients.

2 Materials and methods

2.1 General information

This study is a retrospective, multicenter investigation that collected clinical and imaging data from patients with advanced NSCLC who received ICIs treatment at two domestic centers: the Second Affiliated Hospital of Soochow University (Center 1) and the First Affiliated Hospital of Soochow University (Center 2) between December 2018 and December 2023. Baseline clinical data and computed tomography (CT) imaging were collected for all participants. This study was conducted in accordance with the

ethical principles of the Declaration of Helsinki (2013 revision) and received approval from the Ethics Committee of the First Affiliated Hospital of Soochow University and the Second Affiliated Hospital of Soochow University. Inclusion Criteria: ①Pathologically confirmed stage IIIa-IVb NSCLC according to the 9th edition of the TNM system (10) ②Treatment with at least two cycles of either anti-PD-1/PD-L1 monotherapy or combination therapy of immunotherapy with platinum-based doublet chemotherapy. Patients who received concurrent or sequential radiotherapy, targeted therapy, or any other anticancer regimens were excluded. ③Availability of contrast-enhanced CT scans obtained within 2 months prior to immunotherapy initiation ④Complete clinical documentation; Exclusion Criteria: ①Primary tumor undetectable or unsegmentable on chest CT imaging, or poor image quality ②Receipt of any additional anticancer therapy beyond specified immunotherapy or chemotherapy regimens ③Presence of active autoimmune diseases ④Loss to follow-up or incomplete postoperative surveillance records.

2.2 Indicator collection

Clinical data collected included patient demographics (age, sex, smoking history), tumor characteristics (non-small cell lung cancer histologic subtype, PD-L1 expression level, tumor stage, location), serum fibrinogen levels, and enhanced CT imaging findings. The administered anti-PD-1/PD-L1 monoclonal antibodies included Camrelizumab, Pembrolizumab, Sintilimab, Tislelizumab, and Toripalimab (see Table 1 for details). Treatment response was assessed per RECIST v1.1, classifying outcomes as progressive disease (PD), stable disease (SD), partial response (PR), or complete response (CR) using electronic medical records. The primary endpoint was progression-free survival (PFS), defined as the time from immunotherapy initiation until disease progression, death from any cause, or final follow-up. Patients were stratified into two groups based on whether their PFS exceeded 12 months. A PFS threshold of 12 months was chosen to define the durable clinical benefit (DCB) group, as it is a commonly used surrogate endpoint in immuno-oncology studies to reflect sustained treatment efficacy and is clinically meaningful for prognosis assessment.

TABLE 1 Baseline characteristics of patients in the training and external validation cohorts.

Characteristic	Training cohort (n=75) ¹	External validation cohort (n=34)	¹ P-value
Age (years)	70.0 (8.8)	72 (9.9)	0.315
Sex			0.684
Female	11 (15%)	4 (12%)	
Male	64 (85%)	30 (88%)	
PD-L1 Expression			0.017
<1%	9 (12%)	9 (26%)	
1-49%	36 (48%)	20 (59%)	
≥50%	30 (40%)	5 (15%)	
Tumor Location			0.826
Central	37 (49%)	16 (47%)	
Peripheral	38 (51%)	18 (53%)	
Cancer Type			0.576
Squamous	44 (59%)	18 (53%)	
Adenocarcinoma	31 (41%)	16 (47%)	
T Stage			0.844
3A	6 (8.0%)	3 (9%)	
3B	17 (23%)	9 (26%)	
3C	8 (11%)	3 (9%)	
4A	25 (33%)	8 (24%)	
4B	19 (25%)	11 (32%)	
Immunotherapy Drug			0.221
Camrelizumab	10 (13%)	9 (26%)	
Pembrolizumab	19 (25%)	8 (24%)	
Sintilimab	27 (36%)	12 (35%)	
Tislelizumab	18 (24%)	4 (12%)	
Toripalimab	1 (1.3%)	1 (3%)	
Treatment Regimen			0.773
Imm.	15 (20%)	6 (18%)	
Imm. + Chemo.	60 (80%)	28 (82%)	
Smoking Status			0.251
Non-smoker	33 (44%)	11 (32%)	
Smoker	42 (56%)	23 (68%)	
Fibrinogen	4.23 (1.18)	3.63 (0.98)	0.728

¹Data presented as n (%) or mean (SD).

2.3 Analysis and extraction of radiomics features

Enhanced CT images in DICOM format were retrieved from the institutional PACS system and anonymized. The contrast-enhanced CT images were reviewed, and tumor lesions were

manually segmented slice-by-slice on lung window settings (width: 1500 HU; level: -600 HU) using 3D Slicer software (v4.11) to define volumetric regions of interest (VOIs). The use of contrast-enhanced images facilitated more accurate delineation of tumor boundaries, particularly in distinguishing the tumor from adjacent vessels and atelectatic lung tissue. A total of 1,036

radiomics features were extracted from 3D VOIs using the PyRadiomics package (v3.0.1) in Python, encompassing four feature classes: 1. First-order statistics (e.g., intensity histogram metrics) 2. Shape-based features (3D morphological descriptors) 3. Texture features: Gray-Level Co-occurrence Matrix (GLCM): Correlation, Energy Gray-Level Run- Length Matrix (GLRLM): Long Run Emphasis (LRE) Gray-Level Size Zone Matrix (GLSZM): Size Zone Non-Uniformity 4. Higher-order features (filtered image approximations). To evaluate the reproducibility of radiomic feature extraction, both inter-observer and intra-observer reliability were assessed. For inter-observer reliability, volumetric segmentations were performed independently by two radiologists on a randomly selected subset of 30 patients. For intra-observer reliability, the primary radiologist repeated the segmentations on the same subset after a four-week interval to minimize memory bias. The intraclass correlation coefficient (ICC) was calculated for both assessments. Only features exhibiting excellent reproducibility ($ICC > 0.75$) in both evaluations were retained for subsequent analysis. Features with $ICC > 0.75$ (11, 12) were retained, indicating excellent agreement according to established criteria.

2.4 Construction of clinical characterization model, CT imaging histology model

Using the “Pyradiomics” package in Python to extract radiomics features, we first conduct Z-score normalization on the extracted radiomics features. Then, we perform dimensionality reduction on the radiomics data using mRMR and LASSO. We use Logistic Regression (LR) as the classifier to build the radiomics model. Moreover, we adopt 10-fold cross-validation to enhance the stability of the model. Univariate logistic regression analysis was used to screen out the predictors. Subsequently, multivariate logistic regression analysis was used for the clinical indicators with statistical differences in univariate analysis to screen out the

independent predictors, and a clinical model was constructed based on this.

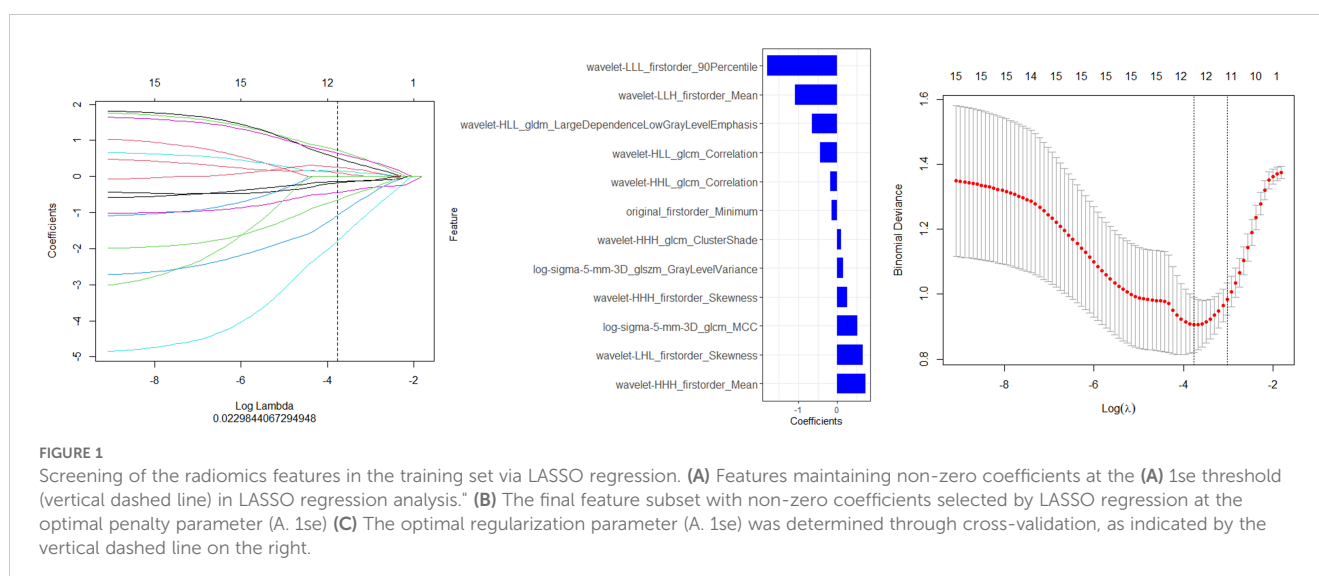
2.5 Construction and evaluation of the CFM

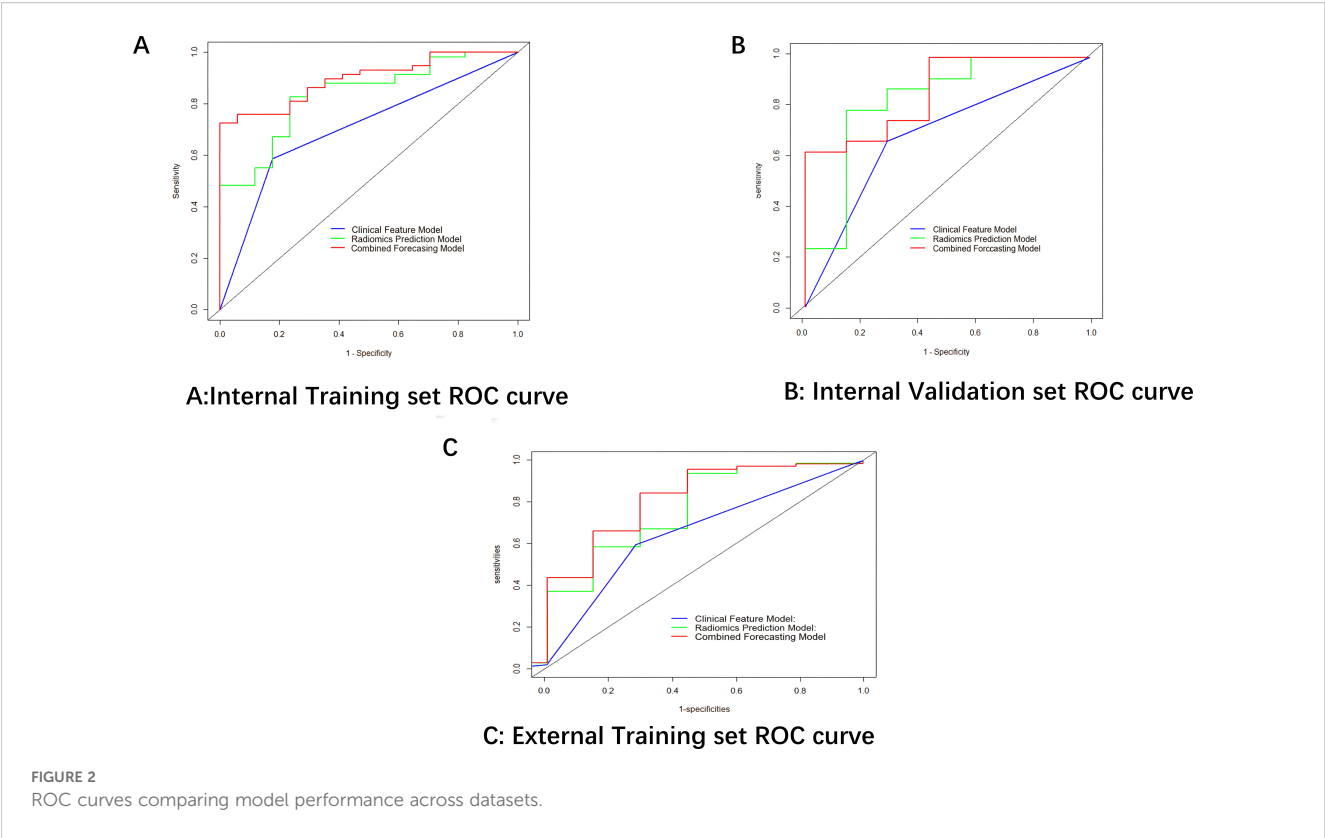
The independent clinical predictors and the radiomics score (Rad-score) were integrated into a final multivariate logistic regression analysis to construct the Combined Forecasting Model (CFM). A nomogram was generated to visualize this model. The performance of all models was evaluated using several metrics. Receiver Operating Characteristic (ROC) curves were used to assess discriminative ability, calibration curves were used to evaluate model reliability, and Decision Curve Analysis (DCA) was used to assess the net clinical benefit and clinical utility the independent predictors and radiomics scores were included in the multivariate logistic regression analysis, and a combined model was constructed based on this, and a nomogram was drawn. The ROC curve was used to evaluate CFM, the calibration curve was used to evaluate the reliability of the model, and the DCA was used to evaluate the clinical efficacy of the radiomics model and CFM.

3 Results

3.1 Baseline characteristics of patients in the training and external validation cohorts

No significant differences in baseline characteristics existed between the training and external test sets, as confirmed by independent samples t-tests/Mann-Whitney U tests for continuous variables and chi-square tests for categorical variables (Table 1). The balanced distribution of clinical features across both cohorts supports the validity of the external validation approach and ensures that performance differences reflect true model efficacy rather than cohort disparities.





3.2 Rad-scores and their predictive performance

Finally,12 radiomic features with non-zero coefficients were obtained from the 1036 radiomic features in the training set using the LASSO method. (Figure 1) The radiomics prediction models for the training and validation sets achieved AUC values of 0.835 and 0.833 under the ROC curve, respectively. The accuracy rates were 0.813 and 0.806,respectively, and the sensitivity rates were 0.828 and 0.792, respectively. (Figure 2) These results indicate that the radiomics model demonstrated robust performance in predicting

immunotherapy efficacy. The radiomics features of the above-mentioned radiomics model were linearly combined and weighted by their respective coefficient AUC = 0.896 (0.825-0.966) to calculate the rad-scores.

3.3 Construction of the clinical model

Collecting clinical data on all patients. Univariate and multifactorial logistic regression analyses showed that tumor location: peripheral or central was of predictive value (Table 2).

TABLE 2 Univariate and multivariate logistic regression analysis of predictive factors.

Variable	Univariate logistic regression analysis OR (95%CI)	p-value ²	Multivariate logistic regression analysis OR (95%CI)	p-value ²
Age (years)	2.242 (0.569-8.832)	0.143		
Sex	2.242 (0.569-8.832)	0.249		
PD-L1 Expression	1.897 (0.837-4.302)	0.125		
Tumor Location	0.151 (0.039-0.585)	0.006	0.127 (0.025-0.653)	0.013
Cancer Type	0.739 (0.249-2.194)	0.586		
Stage	1.596 (0.997-2.554)	0.052		
Treatment Regimen	0.462 (0.093-2.284)	0.343		
Smoking	1.594 (0.538-4.723)	0.4		
Fibrinogen	0.820 (0.524-1.284)	0.386	3.63 (0.98)	
rad_score	3.305 (1.662-6.569)	<0.001		<0.001

¹Data presented as n (%) or mean (SD).

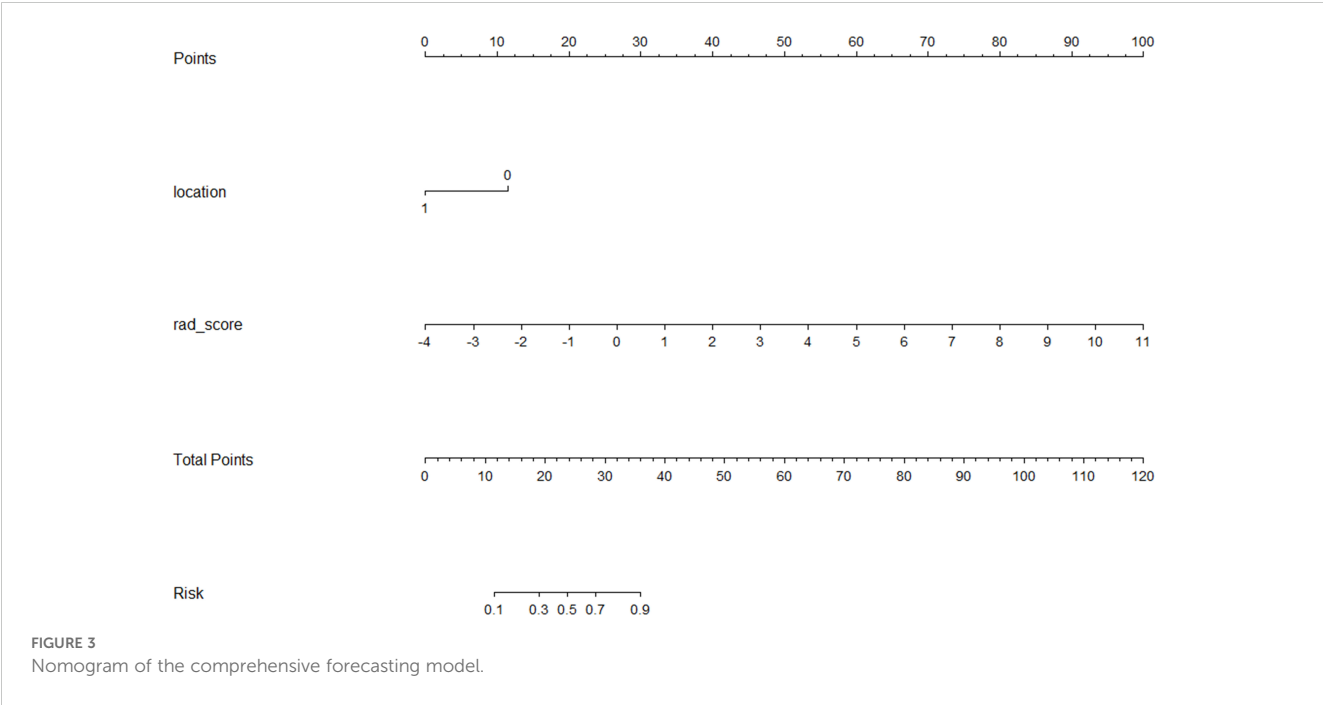
TABLE 3 Performance metrics of clinical, radiomics and combined models across datasets.

Model	Dataset	AUC (95% CI)	Accuracy	Sensitivity	Specificity
Clinical Model	Internal Training Set	0.705 (0.592-0.818)	0.64	0.586	0.824
	Internal Validation Set	0.691 (0.486-0.895)	0.677	0.667	0.714
	External Validation Set	0.653 (0.450-0.857)	0.618	0.593	0.714
	Internal Training Set	0.835 (0.733-0.937)	0.813	0.828	0.765
Radiomics Model	Internal Validation Set	0.833 (0.624-1.000)	0.806	0.792	0.857
	External Validation Set	0.831 (0.641-1.000)	0.912	1	0.571
	Internal Training Set	0.896 (0.825-0.966)	0.787	0.724	1
Combined Model	Internal Validation Set	0.863 (0.713-1.000)	0.71	0.625	1
	External Validation Set	0.884 (0.757-1.000)	0.735	0.667	1

3.4 Performance and evaluation of the Combined Forecasting Model

A CFM was established by incorporating the Rad-score and the independent clinical factor (tumor location) into a logistic regression model, which was visualized as a nomogram (Figure 3). The performance of the CFM was superior to the individual models. As shown in Table 3, the CFM achieved an AUC of 0.902 (95% CI: 0.833-0.970) in the internal training cohort and 0.863 (95% CI: 0.813-1.000) in the internal validation cohort. The model's high performance was maintained in the independent external validation cohort (AUC = 0.863, 95% CI: 0.813-1.000), underscoring its robustness and generalizability. To quantitatively assess the model's performance and generalizability, DeLong tests were conducted to compare the AUCs among the three models within each cohort. In the training cohort, the combined model

significantly outperformed the clinical model ($p < 0.001$). However, no significant differences were found between the combined model and the radiomics model ($p = 0.095$), nor among any of the three models in both the internal validation (all $p > 0.30$) and external validation cohorts (all $p > 0.07$). These results indicate that the combined model achieves robust and stable predictive performance without overfitting, generalizing well to independent patient cohorts. Detailed results of the pairwise comparisons are provided in Supplementary Figure 1. Decision Curve Analysis (DCA) demonstrated that the CFM provided the highest net clinical benefit across a wide range of threshold probabilities compared to the clinical-only or radiomics-only models in all cohorts (Figure 4). Furthermore, the calibration curves indicated good agreement between the model's predicted probabilities and the actual observed outcomes in both the training and validation datasets, confirming the model's reliability (Figure 5).



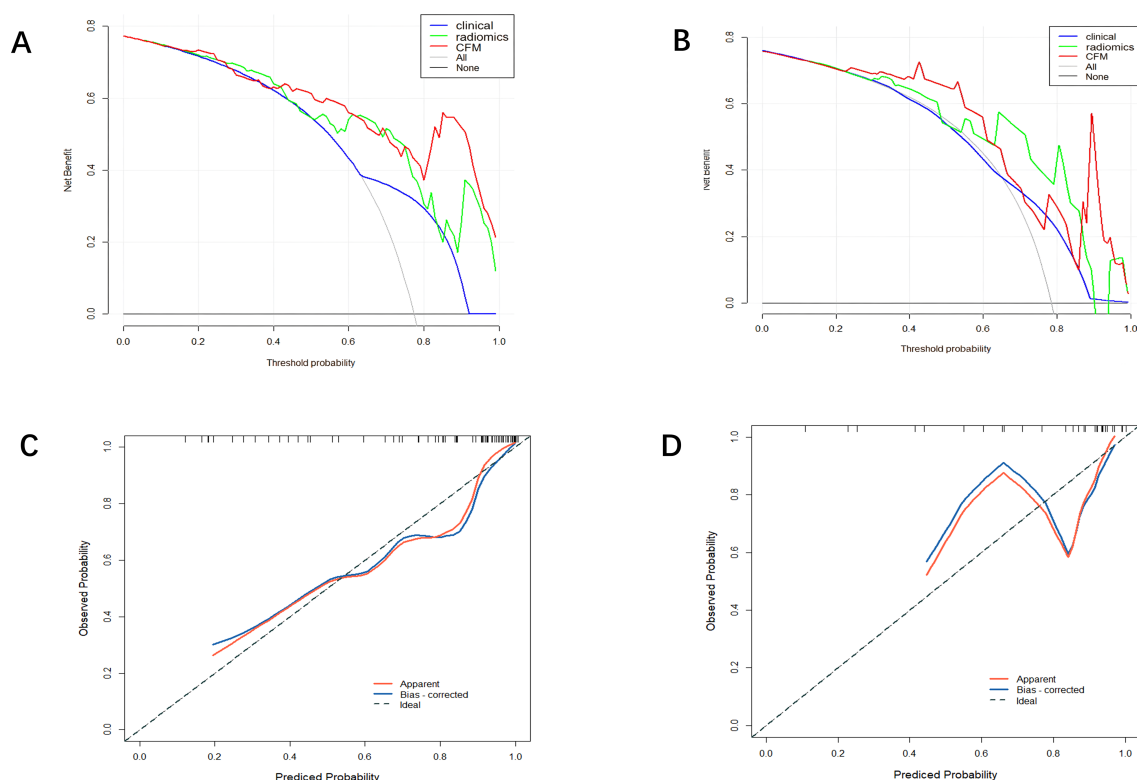


FIGURE 4

(A) Training set DCA curve (B) Validation set DCA curve (C) Calibration curve of the comprehensive forecasting model in the training set. (D) Calibration curve of the comprehensive forecasting model in the test set.

4 Discussion

This study integrated clinical characteristics and CT radiomics features of patients with stage III-IV non-small cell lung cancer (NSCLC) receiving immunotherapy to develop an integrated diagnostic model. This study first explored independent predictors of immunotherapy efficacy in the clinical setting and found that

central tumor location was adverse prognostic outcomes in NSCLC. Previous studies (13, 14) have shown that in terms of microenvironment, central-type tumors have 2.8-fold more TGF- β expression than peripheral-type, CD8+ T-cell density is decreased by 40% under multiple fluorescence, and central-type lung cancers have significantly worse (15) OS than peripheral-type lung cancers; there are also previous studies (16) that emphasized the peritumor texture-

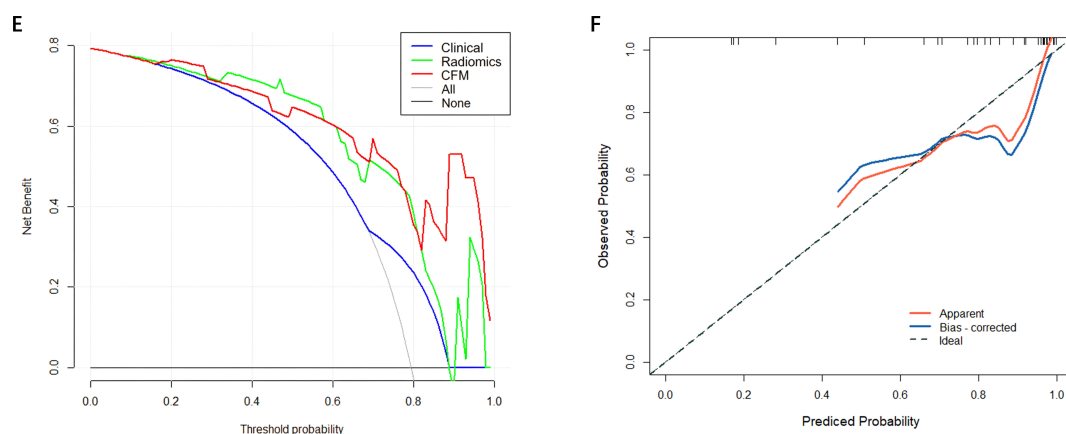


FIGURE 5

(E) External Validation DCA curve (F) Calibration curve of the comprehensive forecasting in the External Validation.

entropy imaging histological features of central-type lung cancers compared with the peripheral-type lung cancers. Based on the above independent influences on the efficacy of immunotherapy, the present study firstly modeled the clinical characteristics of prognosis. Utilizing data from the Second Affiliated Hospital of Soochow University, a cohort of 106 patients was included. The constructed integrated diagnostic model demonstrated (17) superior diagnostic efficacy, achieving areas under the AUC of 0.902 and 0.863 in the training and validation sets. The nomogram developed in this study serves as an intuitive tool with potential for integration into clinical decision support systems (CDSS). In future clinical practice, prior to initiating immunotherapy, clinicians could input the patient's tumor location and CT-based Rad-score to obtain an individualized probability of treatment response. This approach would help identify patients likely to derive suboptimal benefit, enabling consideration of alternative treatment strategies or more intensive monitoring, ultimately advancing toward precision immunotherapy for lung cancer. Furthermore, grayscale normalization techniques (18) were applied during preprocessing to mitigate image variations attributable to differences in imaging acquisition parameters. Comparative analysis using ROC curves evaluated the clinical feature model, CT radiomics model, and the integrated diagnostic model. The results revealed that the AUC of the integrated model was significantly higher than the other two models in both the training and validation sets, indicating its robust diagnostic performance and favorable generalizability (19). DCA demonstrated that the integrated diagnostic model yielded the highest net clinical benefit across both the training and validation cohorts compared to the clinical feature model and the CT radiomics model. Calibration curves indicated good agreement between predicted and observed outcomes for the integrated model, although variations in performance across different predicted probability ranges were noted among all models. This study has several limitations that should be considered. First, despite being a two-center study, its retrospective design may still introduce selection bias (20). To mitigate this, we applied strict, consistent inclusion and exclusion criteria across both participating centers. Second, the sample size, particularly for the external validation cohort ($n=34$), remains relatively small (21, 22), which may affect the stability of the estimates. Future large-sample studies needed to confirm our findings. We acknowledge that the sample size, particularly of the external validation cohort, is relatively modest. While larger-scale prospective studies are certainly needed to further solidify our findings, a key strength of our work is the very inclusion of an independent external validation set. The model's excellent and stable performance in this cohort (AUC: 0.863) provides strong preliminary evidence of its generalizability, addressing a common limitation in many radiomics studies. Third, heterogeneity in immunotherapy regimens exists. Although our univariate analysis showed that the specific treatment regimen was not a significant predictor, the heterogeneity in drugs and combinations (e.g., mono-immunotherapy vs. chemo-immunotherapy) remains a limitation. We subsequently performed a *post hoc* exploratory analysis by stratifying patients according to treatment regimen

(immunotherapy monotherapy versus combined immunotherapy-chemotherapy). The combined model maintained robust discriminatory capacity across all subgroups (all AUCs > 0.80); however, formal statistical comparisons were not conducted due to limited sample sizes within the subgroups. Future studies with larger, more homogeneous cohorts could further refine the model's applicability. Fourth, manual segmentation inherently carries a risk of inter-observer variability. We minimized this by having segmentations performed by an experienced radiologist following a standardized protocol and by rigorously assessing feature reproducibility (ICC > 0.75). Fifth, our study currently lacks delta-radiomics (23–25) from longitudinal imaging, which represents a promising future direction to capture dynamic changes in tumor biology during therapy and potentially further enhance predictive performance (26, 27).

Data availability statement

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

Ethics statement

The studies involving humans were approved by Ethics Review Committee of the Second Affiliated Hospital of Soochow University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

NG: Writing – original draft, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. ZL: Formal analysis, Investigation, Writing – review & editing. YL: Software, Validation, Writing – review & editing. LD: Data curation, Formal analysis, Writing – original draft. HM: Conceptualization, Formal analysis, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, and/or publication of this article.

Conflict of interest

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

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

References

- Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med*. (2016) 375(19):1823–33. doi: 10.1056/NEJMoa1606774
- Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. (2015) 372:2018–28. doi: 10.1056/NEJMoa1501824
- Cottrell TR, Thompson ED, Forde PM, Stein JE, Duffield AS, Anagnostou V, et al. Biomarkers of response to immunotherapy in early stage non-small cell lung cancer. *J Immunother Cancer*. (2021) 9:e002657. doi: 10.1136/jitc-2021-002657
- Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N Engl J Med*. (2018) 378:2093–104. doi: 10.1056/NEJMoa1801946
- Mhandu P, McGonigle N. Lung cancer. *Surg (Oxford)*. (2023) 41:142–7. doi: 10.1016/j.mpsur.2023.02.008
- Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor JF, et al. Oncogene-specific differences in tumor mutational burden, PD-L1 expression, and outcomes from immunotherapy in non-small cell lung cancer. *Oncogene*. (2018) 37:1589–600. doi: 10.1038/s41388-017-0060-8
- Brahmer JR, Lacchetti C, Schneider BJ, Atkins MB, Brassil KJ, Caterino JM, et al. Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: American Society of Clinical Oncology clinical practice guideline. *J Clin Oncol*. (2018) 36:1714–68. doi: 10.1200/JCO.2017.77.6385
- Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol*. (2018) 19:1180–91. doi: 10.1016/S1470-2045(18)30413-3
- Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*. (2018) 24:1550–8. doi: 10.1038/s41591-018-0136-1
- Zeng D, Qin Y, Li Z, Li Y, Wang Y, Zhou Q, et al. Identification of a tumour-immunity cycle for individualized prediction of immunotherapy response in patients with advanced cancer. *Nat Commun*. (2023) 14:5099. doi: 10.1038/s41467-023-40879-6
- Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull*. (1979) 86:420–8. doi: 10.1037/0033-2909.86.2.420
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. (2016) 15:155–63. doi: 10.1016/j.jcm.2016.02.012
- Halpenny DF, Chen WC, Deng L, Zhao Y, Feng M, Wang Q, et al. Tumor location-dependent immune microenvironment in NSCLC. *J Thorac Oncol*. (2020) 15:e80–2. doi: 10.1016/j.jtho.2020.02.010
- Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. *Spatial heterogeneity of CD8+ T cells in NSCLC revealed by multiplex IF*. *Eur J Cancer*. (2020) 139:76–85. doi: 10.1016/j.ejca.2020.07.036
- Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C, et al. Central vs peripheral NSCLC survival disparity. *Lung Cancer*. (2021) 160:25–33. doi: 10.1016/j.lungcancer.2021.07.015
- Liu Y, Zhang Y, Cheng H, Liu S, Cui H, Zhang L, et al. CT texture entropy as a surrogate for tumor immunosuppression. *Cancers*. (2022) 14:712. doi: 10.3390/cancers14030712
- Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. (2006) 26:565–74. doi: 10.1177/0272989X06295361
- Zwanenburg A, Vallieres M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative. *Radiology*. (2020) 295:328–38. doi: 10.1148/radiol.2020191145
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing AUCs of multiple diagnostic tests. *Biometrics*. (1988) 44:837–45. doi: 10.2307/2531595
- Kocak B, Kus EA, Kilickesmez O. How to read and review papers on machine learning and artificial intelligence in radiology: a survival guide to key methodological concepts. *Eur Radiol*. (2021) 31:1819–30. doi: 10.1007/s00330-020-07324-4
- Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. (2015) 162:55–63. doi: 10.7326/M14-0697
- van Smeden M, Moons KGM, de Groot JAH, Collins GS, Altman DG, Eijkemans MJC, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. *J Clin Epidemiol*. (2022) 144:203–21. doi: 10.1016/j.jclinepi.2021.02.003
- Fave X, Mackin D, Yang J, Zhang J, Fried D, Balter P, et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Sci Rep*. (2017) 7:588. doi: 10.1038/s41598-017-00665-z
- Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. Radiomics to assess tumour-infiltrating CD8 T cells and response to PD-1/PD-L1 blockade in non-small cell lung cancer. *Nat Commun*. (2021) 12:3189. doi: 10.1038/s41467-021-23446-9
- Trebeschi S, Drago SG, Birkbak NJ, Kurilova I, Călin AM, Pizzi AD, et al. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol*. (2021) 32:80–8. doi: 10.1016/j.annonc.2020.10.008
- Choe J, Lee SM, Do KH, Lee G, Lee JG, Lee SM, et al. Deep learning-based radiomics classifier for immunotherapy response prediction in advanced non-small cell lung cancer. *Radiology*. (2023) 307:e222268. doi: 10.1148/radiol.222268
- Xu Y, Hosny A, Zeleznik R, Parmar C, Coroller T, Franco I, et al. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin Cancer Res*. (2021) 27:3296–306. doi: 10.1158/1078-0432.CCR-20-3455

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2025.1669469/full#supplementary-material>