Biomolecules and Biomedicine

ISSN: 2831-0896 (Print) | ISSN: 2831-090X (Online)

Journal Impact Factor® (2023): 3.1

CiteScore®(2023):7.4
www.biomolbiomed.com | www.blog.bjbms.org

The BiomolBiomed publishes an "Advanced Online" manuscript format as a free service to authors in order to expedite the dissemination of scientific findings to the research community as soon as possible after acceptance following peer review and corresponding modification (where appropriate). An "Advanced Online" manuscript is published online prior to copyediting, formatting for publication and author proofreading, but is nonetheless fully citable through its Digital Object Identifier (doi®). Nevertheless, this "Advanced Online" version is NOT the final version of the manuscript. When the final version of this paper is published within a definitive issue of the journal with copyediting, full pagination, etc., the new final version will be accessible through the same doi and this "Advanced Online" version of the paper will disappear.

RESEARCH ARTICLE

Yuan et al: CT-based DL and inflammation in NSCLC

Deep learning and inflammatory markers predict early response to immunotherapy in unresectable NSCLC: A multicenter study

Lei Yuan[1,2#], Qi Wang[1,2#], Fei Sun[3#], Hongcan Shi[1,2*]

[1]Department of Thoracic Surgery, Northern Jiangsu People's Hospital Affiliated to Yangzhou University, Yangzhou, China;

[2]Institute of Translational Medicine, Medical College, Yangzhou University, Yangzhou, China;

[3]Department of Thoracic Surgery, The Affiliated Taizhou People's Hospital of Nanjing Medical University, Taizhou, China.

[*]Correspondence to **Hongcan Shi**: shihongcan@yzu.edu.cn

[#]**Lei Yuan, Qi Wang** and **Fei Sun** contributed equally to this work.

DOI: https://doi.org/10.17305/bb.2025.12324

EARLY ACCESS

ABSTRACT

Immune checkpoint inhibitors (ICIs) demonstrate substantial interpatient variability in clinical efficacy for unresectable non-small cell lung cancer (NSCLC), underscoring the unmet need for noninvasive biomarkers to predict early therapeutic responses and improve survival outcomes. To address this, we developed a CT-based deep learning model integrated with the systemic immune-inflammatory-nutritional index (SIINI) for early prediction of ICI response. In a retrospective multicenter study of 265 patients treated with ICIs (incorporating chest CT and laboratory data), the cohort was divided into training (70%), internal validation (30%), and external validation sets. The combined model—leveraging DenseNet121-derived deep radiomic features alongside SIINI—achieved strong predictive performance, with AUCs of 0.865 (95% CI: 0.7709–0.9595) in the internal validation cohort and 0.823 (95% CI: 0.6627–0.9827) in the external validation cohort. Gradient-weighted class activation mapping (Grad-CAM) highlighted key CT regions contributing to model predictions, enhancing interpretability for clinical application. These findings highlight the potential of integrating deep learning with inflammatory biomarkers to support personalized ICI therapy in unresectable NSCLC. Future directions include incorporating multi-omics biomarkers, expanding multicenter validation, and increasing sample sizes to further improve predictive accuracy and facilitate clinical translation.

Keywords: Artificial intelligence; deep learning; non-small cell lung cancer; NSCLC; inflammatory parameter; immunotherapy.

EARLY ACCESS

INTRODUCTION

Lung cancer remains a leading type of cancer and the foremost cause of cancer fatalities worldwide (1, 2). Non-small cell lung cancer (NSCLC) constitutes the majority of lung cancer cases (80-90%) and is often diagnosed at advanced stages (65%), frequently presenting with local or distant metastases (3), which often precludes surgical intervention. Recent progress in immunotherapy, especially the application of immune

checkpoint inhibitors (ICIs), has shown significant promise in improving outcomes for patients with unresectable NSCLC (4). However, the variable response to immunotherapy highlights the necessity for further investigation into predictive biomarkers that can forecast immune response. Accumulating evidence has implicated various biomarkers in predicting responsiveness to ICIs in NSCLC, including tumor mutational burden (TMB) (5), programmed death ligand-1 (PD-L1) expression (6), tumor-infiltrating lymphocyte (TIL) density (7), and inflammatory cytokine profiles (8). However, current biomarker assessment protocols predominantly depend on invasive tissue biopsies which present dual clinical challenges: procedure-related morbidity risks and limited capacity to map intratumoral heterogeneity due to inherent sampling constraints (9, 10). This critical methodological gap necessitates the development of robust non-invasive biomarkers capable of predicting therapeutic outcomes in patients with unresectable NSCLC undergoing ICI regimens.

Emerging evidence underscores the intricate interplay between tumor pathogenesis and host inflammatory response, immune status, and nutritional profile (11-15). The Systemic Immune-Inflammation-Nutritional Index (SIINI) is an innovative multidimensional biomarker that combines pre-treatment inflammatory indicators(16), immunocompetence metrics, and nutritional determinants, theoretically provides a more comprehensive evaluation of pretherapeutic host status compared to conventional unidimensional biomarkers. Nevertheless, the prognostic utility of SIINI in predicting clinical outcomes for NSCLC patients receiving ICIs remains unexplored.

Beyond conventional laboratory diagnostics, computed tomography (CT)-based imaging biomarkers have become indispensable in the diagnostic workflow of lung cancer (17). The integration of artificial intelligence with medical imaging has catalyzed the emergence of radiomics-driven deep learning (RDL) in thoracic oncology, enabling the quantitative extraction of high-dimensional imaging features imperceptible to human visual assessment (18). This computational approach facilitates the development of non-invasive predictive signatures for diverse clinical applications, including tumor characterization (19), therapeutic strategy optimization (20), and treatment response monitoring. Notably, foundational studies have established the prognostic relevance of conventional radiomic features in both localized and advanced NSCLC. For resectable

EARLY ACCESS

disease, radiomic signatures demonstrate predictive capacity for neoadjuvant chemotherapy response (21), while in advanced stages, specific imaging biomarkers correlate with immunotherapy outcomes (22). These findings underscore the evolving role of quantitative imaging biomarkers in precision oncology paradigms. Nevertheless, at present, there is limited evidence to substantiate the effectiveness of integrating clinical data, especially systemic immune-inflammatory-nutritional indexes such as SIINI, into deep learning models to predict the response of patients with unresectable NSCLC to immune checkpoint inhibitors. Moreover, CT-based RDL can reveal the heterogeneity within the tumor and provide a potential research direction for multi-dimensional interpretation of the tumor microenvironment(23, 24).

In this study, we aimed to investigate the early predictive capability of a CT-based deep learning model combined with the inflammation parameter SIINI for predicting the response of unresectable NSCLC patients to ICIs by utilizing clinical data from 265 patients across two independent medical centers.

MATERIALS AND METHODS

Data collection

In this study, patients with unresectable NSCLC treated with single-agent ICI at Northern Jiangsu People's Hospital (Center A) and Taizhou People's Hospital (Center B) were selected as subjects. (Ethical Review No. 2021ky211; KY 2024-093-01). Patients were administered 200 mg of Pembrolizumab every three weeks, or 3 mg/kg of Nivolumab every two weeks, or 200 mg of Sintilimab every three weeks. The planned time span is from January 2021 to December 2024. All procedures comply with the guidelines and ethical principles outlined in the 1964 Declaration of Helsinki.

Inclusion criteria:

(1) Eastern Cooperative Oncology Group (ECOG) performance status of 0-3; (2) Presence of measurable lung lesions as per Response Evaluation Criteria in Solid Tumors (RECIST V1.1), as determined by standard chest computed tomography (CT) scans; (3) Diagnosis of NSCLC confirmed by biopsy or bronchofibroscopy and histopathological examination, with staging based on imaging and pathology

EARLY ACCESS

according to the TNM (8th edition) classification as stage IIIB to IV (25) ; (4) The comprehensive availability of laboratory and imaging data for evaluating disease progression includes standard blood work and biochemical analyses carried out before the commencement of ICI therapy, along with chest CT scans performed every 6 to 8 weeks thereafter; (5) Comprehensive follow-up information available.

Exclusion criteria:

(1) Inadequate image quality, such as presence of artifacts; (2) History of thoracic surgery; (3) Loss to follow-up after receiving immunotherapy; (4) Inability to obtain complete laboratory and imaging data for pathological evaluation.

Clinical data

We collected baseline data of patients and laboratory test results, including age, gender(female==0/male==1), smoking history(Current or former smokers==1/Never smokers==0), basic disease(with basic disease==1/without basic disease==0), body mass index (BMI), treatment lines, medication regimen (Pembrolizumab==1/ Nivolumab==2/ Sintilimab==3), EGFR mutation (Postive==1/Negative==0), TNM(IIIB==3/IV==4), ECOG, pathological type(adenocarcinoma==1/squamocellular carcinoma==0), modality, PD-L1 expression(No record==0,Tumor Proportion Score, TPS <1%==1 、 TPS≥1%-49%==2 、 TPS≥50%==3), etc. Blood cell counts encompassed white blood cell enumeration, neutrophil, lymphocyte, monocyte, eosinophil, and basophil quantifications along with their respective percentages. Additionally, hemoglobin concentration, red blood cell count, hematocrit level, platelet count, proportion of larger platelets, and plateletcrit were determined. Blood biochemistry analyses comprised measurements of total protein, albumin, and levels of LDH, ALT, AST, urea, and creatinine. Based on clinical retrospective data, this study found that neutrophil count, lymphocyte count, platelet count, hemoglobin level, serum albumin level and BMI before treatment were calculated by neutrophil count × platelet count × hemoglobin level / (lymphocyte count × BMI × serum albumin level) to form a new index-systematic immune-inflammation-nutritional index (SIINI).At the same time, PNI, PLR, ALI, SII, and NLR calculated from these indicators were also included in the subsequent clinical prediction model studies, and the calculation formulas and data were in the supplementary materials. The

EARLY ACCESS

full names of some laboratory test results can be found in the abbreviation list at the end of this article.

Image acquisition

All patients underwent chest CT scans prior to the initiation of ICI therapy. Resultant imagery underwent moderate-detail reconstruction, yielding slice widths of 3–5 mm.Tumor segmentation was subsequently performed to delineate the primary NSCLC lesions. Using 3D-Slicer v4.11, two cancer specialists jointly delineated regions of interest (ROIs) and derived radiomic features from each detected nodule. The target lesion was defined as any tumor mass measuring ≥ 5 millimeters in diameter, which was distinctly marked at the baseline and consistently observed in follow-up CT scans. Response assessment based on the follow-up CT scans adhering to RECIST 1.1 criteria categorized patients into responders (label=1) exhibiting complete remission (CR), partial remission (PR), or stable disease (SD), while those with progressive disease (PD) were classified as non-responders (label=0). All CT scan interpretations were conducted by two independent oncologists to ensure objectivity.

To enhance reliability, two researchers (QW and FS) independently outlined the ROIs. Each researcher repeated this process for the same tumor at different time points. Intra-group consistency of the extracted radiomic features was evaluated using the intra-class correlation coefficient (ICC), thereby ensuring robustness in the data collected. After calculating the ICC within and between groups, the characteristics of ICC > 0.8 at both time points were selected. Any differences are resolved through discussions between the two researchers.

Methods

We have developed the workflow shown in Figure 1 to carry out this research. This study included 265 participants, comprising 207 patients from Center A; Center B: 58 cases, 145 cases as training set, 62 cases as validation set, and 58 cases in Center B as test set. (Supplementary Figure 1)

EARLY ACCESS

Data preprocessing

In our medical image analysis, voxel spacing was standardized across all volumes of interest to a uniform resolution of $1mm \times 1mm \times 1mm$ employing a fixed resolution resampling method. Concurrently, CT Hounsfield Units (HU) were constrained within the range of -400 to 600. This standardization process was critical for facilitating precise image comparisons, markedly enhancing the accuracy and dependability of our analytical outcomes.

Radiomics procedure

Feature extraction

In this study, we've neatly divided radiomic features into three main buckets: (I) Shape and Size Descriptors, (II) Intensity-Based Measures, and (III) Texture Features. Shape and size descriptors are all about capturing the 3D form of the tumors. Intensity-based measures look at the spread of voxel intensities using basic statistical tools. On the flip side, texture features dig into the patterns and how voxel intensities are arranged in space, using more complex statistical methods like second-order and higher-order analyses.

In analyzing the texture, we utilized well-established methods including the gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), and neighborhood gray-tone difference matrix (NGTDM). Each specified subregion underwent feature extraction using the PyRadiomics tool (version 3.0.1), following the protocols established by the Imaging Biomarker Standardization Initiative (IBSI) meticulously.

Feature selection

In the feature selection process, we took a layered approach. We kicked things off by standardizing the features using Z-scores, followed by running t-tests to gauge their significance. Any feature with a p-value under 0.05 was given the green light to move on to the next round. Subsequently, features exhibiting high reproducibility were assessed using Pearson's correlation coefficient. In cases where pairs of features demonstrated a

EARLY ACCESS

correlation greater than 0.9, a strategic recursive elimination process was applied to retain a single representative feature from each highly correlated pair, thereby minimizing redundancy. Thereafter, refinement of the radiomic signature was accomplished using Least Absolute Shrinkage and Selection Operator (LASSO) regression, which effectively reduced the influence of non-contributory features. The ideal regularization parameter (λ) was established using 10-fold cross-validation.

Radiomics signature

Following feature refinement via LASSO regression, risk evaluation was conducted utilizing both linear models (such as Logistic Regression) and tree-based models (including Random Forest and LightGBM). Model hyperparameter optimization was performed through 10-fold cross-validation within the training dataset, employing the GridSearch algorithm to fine-tune parameters. The parameters demonstrating the highest median efficacy were selected for the final model training.

Deep learning procedure

Data preparation

Crop ROI: Within our methodology, for each patient, we identified the slice exhibiting the largest ROI as the representative image. For streamline analysis and minimize interference, the region of interest (ROI) was limited to its minimal bounding rectangle, extended by 10 pixels. This expansion acknowledges the importance of peritumoral regions, as indicated by recent studies (26).

Data augmentation: The intensity distribution across RGB channels for the input images was standardized using Z-score normalization. During training stage, real-time data augmentation boosted model resilience using random crops and horizontal/vertical flips. For test images, processing was limited to normalization to ensure consistency.

Model training

Transfer Learning: Previous studies have shown that DenseNet121 (27, 28) , with its unique dense connection mechanism, is significantly superior to traditional CNN models (such as ResNet and VGG) in terms of feature reuse, parameter efficiency, training stability and task adaptability. Its advantages are particularly prominent in scenarios that

EARLY ACCESS

require high-precision and efficient feature extraction, such as medical imaging and target detection. This study leveraged advanced architectures DenseNet121 to surpass traditional CNN-based model performance. We performed comparative analyses of these networks to determine the most effective model for our specific research needs.

Hyperparameters: Our strategy incorporated transfer learning to accommodate diverse patient populations and variability. Models began with ImageNet-derived parameters for enhanced accommodation. Our methodology hinged on meticulous learning rate calibration, employing a cosine decay strategy to maximize generalization across diverse data:

$$\eta_t = \eta_{min}^i + \frac{1}{2}\left(\eta_{max}^i - \eta_{min}^i\right)\left(1 + \cos\left(\frac{T_{cur}}{T_i}\pi\right)\right)$$

Here, $\eta_{min}^i = 0$ represents the minimum learning rate, $\eta_{max}^i = 0.01$ sets the maximum learning rate, and $T_i = 30$ denotes the number of epochs for each training cycle. Additional critical hyperparameters included the use of Stochastic Gradient Descent (SGD) as the optimizer and softmax cross-entropy as the loss function.

Deep learning signature

In our model, the probabilities outputted by the DenseNet121 are defined as the deep learning signature, representing the model's predictive capabilities.

Clinical use

Clinical Signature: We employed the same model used for the Radiomics Signature to model our clinical task. We then selected the model that performed best on the test set for subsequent comparisons of the signatures. This approach ensured that the most effective predictive model was utilized for clinical evaluation.

Combined Model: To enhance its clinical utility, we carried out univariable and stepwise multivariable analyses on all clinical features to identify significant predictors. These selected clinical features were integrated with outputs from our deep learning model to develop a Logistic Regression (LR) linear model, resulting in the formation of the

EARLY ACCESS

Combined Signature. We employed a nomogram for effective visualization of this signature.

Metrics: We gauged how well our models could distinguish between true and false positives by using Receiver Operating Characteristic (ROC) curves. To see if our models were well-calibrated, we plotted calibration curves and then ran Hosmer-Lemeshow tests to really put them through their paces. On top of that, we performed Decision Curve Analysis (DCA) to figure out if our predictive models would actually be helpful in a clinical setting.

Statistical analysis

We randomly split the dataset, earmarking 70% for training and setting aside the remaining 30% for internal validation. To really put our model through its paces and see how well it generalized, we also tapped into data from an outside center, using it as an external validation set. Table 1 shows baseline characteristic of this study.

We ran our analyses, using Python 3.7.12 and the statsmodels package, version 0.13.2. When it came to building our machine learning models, we leaned on scikit-learn, specifically version 1.0.2. For the deep learning side of things, we harnessed the power of an NVIDIA 4090 GPU, along with the MONAI (version 0.8.1) and PyTorch (version 1.8.1) frameworks.

Ethical statement

This study was conducted according to the ' Helsinki Declaration '. Besides, this study was carefully reviewed by the Ethical Review Committee of Northern Jiangsu People 's Hospital and Taizhou People 's Hospital (Ethical Review No.2021ky211; No.KY 2024-093-01), unanimously agreed that the patient 's hospitalization data and images used in this retrospective study were exempted from the informed consent application in the ethics committees and approved by the committees.

EARLY ACCESS

RESULTS

Clinical features analysis

Univariable and Multivariable Analysis: In our research, we performed an extensive univariate analysis of all clinical features, calculating the Odds Ratio (OR) and associated p-values for each. Features with a p-value less than 0.05 were selected for inclusion in the nomogram construction (Figure 2). Additionally, we constructed a clinical model based on these clinical features, including EGFR, TNM, SIINI, gender. To some extent, these indicators present the body nutritional inflammation status and tumor heterogeneity of unresectable NSCLC patients using ICIs, which can be included in the prediction model after multivariate variable analysis (Supplementary Table 1 and Table 2).

The LightGBM model exhibited the highest AUC of 0.820 in the test set (Table 2). This performance highlights its capability to differentiate between the classes, marking its importance in evaluating binary classification models in medical diagnostics (Supplementary Figure 2).

Rad signature

In this study, we compiled a comprehensive dataset of 1,834 handcrafted radiomic features, organized into three primary categories: shape, first-order, and texture. This compilation consists of 360 first-order metrics, 14 shape descriptors, and a broad array of texture characteristics. These features were extracted using a specialized program created with Pyradiomics, detailed at http://pyradiomics.readthedocs.io. These distribution of handcrafted features among the different categories are illustrated in Figure 3.

Radiomics Feature Selection: We employed the Least Absolute Shrinkage and Selection Operator cross-validation, LassoCV) methodology, integrating it with a rigorous 10-fold cross-validation framework, to select salient radiomic features. The intricate details of this feature selection process are vividly depicted in Figures 1, offering a comprehensive visual representation of our approach. Figure 4 showcases the coefficients obtained through Least Absolute Shrinkage and Selection Operator (LASSO) regression using 10-fold cross-validation, a technique we employed in both our Radiomics Signature and

EARLY ACCESS

INTRA Signature analyses.Left and right sub-figure portions display Lasso regularization paths, Mean Squared Error (MSE) values, and pertinent radiomic feature weighting.

Metrics

Looking at the AUC scores in Table 3, LightGBM comes out on top on the validation set with an AUC of 0.624. While it's not exactly blowing the competition out of the water, it still edges out Logistic Regression (LR) and RandomForest, which clocked in at 0.551 and 0.622 respectively on the validation set. This outcome suggests that the LightGBM model, a non-linear model, has a superior capability to fit and generalize the complex relationships in the dataset compared to the linear models like LR (Figure 5).

The higher AUC value in LightGBM underscores its enhanced ability to discriminate between the positive and negative classes under more varied and complex scenarios. This supports the assertion that non-linear models, due to their ability to model intricate interactions and non-linear dependencies, are often better suited for tasks where the relationships between features are not straightforward, thereby providing a more robust fit to the data.

Deep learning radiomics signature

Results

The performance of the DenseNet121 model, as indicated in the provided data (Table 4), shows promising results in terms of its ability to discriminate between classes, particularly highlighted by its AUC scores across different cohorts (Supplementary Figure 3).

- Training cohort: The DenseNet121 model attained an AUC of 0.846, with a 95% confidence interval (CI) between 0.7793 and 0.9126. This high AUC indicates strong discriminative capability during the training phase.
- Validation cohort: In the validation set, the model achieved an AUC of 0.751, with a CI from 0.6300 to 0.8722, still reflecting a relatively high predictive accuracy.

EARLY ACCESS

- Test cohort: On the test dataset, the AUC measured 0.691, with a confidence interval ranging from 0.4747 to 0.9071. Although lower than the training and validation phases, this score still suggests a moderate ability to distinguish between the positive and negative outcomes.

In all phases, especially notable are the high specificity and Positive Predictive Value (PPV) scores, reaching 1.000 in both the validation and test cohorts. This result illustrates that when the model forecasts a positive class, it is notably accurate, with no false positives documented. However, the sensitivity scores are comparatively lower, suggesting that while the model is excellent at confirming cases when present, it misses a significant number of positive cases (low true positive rate).

When contrasting these results with radiomics signature, the DenseNet121 deep learning approach potentially offers an improvement due to its capability to automatically learn and generalize from intricate image features across multiple levels of abstraction. This capability often translates into a more nuanced understanding and exploitation of the underlying patterns in medical images compared to more conventional radiomic approaches, which rely on pre-defined features. Thus, DenseNet121's performance, particularly in terms of its high specificity and PPV in the test cohort, underscores its potential for more accurate and reliable clinical applications, although there might be room for improvement in its sensitivity to ensure fewer positive cases are missed.

Grad-CAM visualization

To probe the deep learning models' recognition capabilities across different samples, we employed the Gradient-weighted Class Activation Mapping (Grad-CAM) technique for visualization. In the implementation of Grad-CAM, we focus on the analysis of the last convolutional layer feature map of DenseNet121 and use it to generate a heat map that reflects the metabolically active region at the edge of the tumor. Due to the dense connection mechanism of DenseNet, deep features can retain fine-grained semantic information through cross-layer aggregation. Experiments show that the convolutional layer at the end of the last dense block of DenseNet121 contributes the most to the final classification decision. After the global average pooling of the high-dimensional feature map output by this layer, the channel gradient weight directly reflects the degree of

EARLY ACCESS

attention of the model to the tumor area. Through Grad-CAM, we further localized the image areas associated with lung malignant tumors (such as irregularly enhanced areas on the edges, peripheral edema zones), while the activation areas of benign tumors were concentrated in the internal uniform texture areas. This analysis verifies the interpretability of the model decision, and the relevant heat map comparison will be presented in Figure 6 of the results section.

Clinical use

Analyzing the AUC scores across different models and cohorts (Table 5), the Combined model consistently demonstrates an improvement over single signature model (Figure 7). This trend is evident in training, validation, and test cohorts, underscoring the efficacy of integrating multiple types of data or analytical approaches.

Calibration Curve: The Hosmer-Lemeshow (HL) test plays a crucial role in evaluating the calibration of a predictive model by comparing the predicted probabilities with the actual outcomes. Higher HL p-values indicate better calibration, reflecting closer alignment between the model's predictions and observed outcomes. In our study, the Combined model exhibited outstanding calibration, as evidenced by HL test statistics of 0.964 in the training cohort, 0.633 in the validation cohort, and 0.140 in the test cohort. These results highlight the model's high effectiveness in accurately mirroring observed data (Supplementary Figure 4).

Delong Test: The DeLong test is a method for comparing whether there is a significant difference in the AUC of two or more models. In other words, it helps us to judge whether a model is significantly better than another model. If the *p*- value is less than 0.05, one of the models is significantly better than the other. While the Delong test confirmed no significant AUC difference between the Deep Learning model and Clinical model in the external test cohort (p= 0.298; AUC 95% CI overlap: Deep Learning model [0.475–0.907] vs. Clinical model [0.693–0.947]), the Combined model demonstrated superior overall performance in both the training and validation cohorts, as shown in Figure 8.

DCA: Figure 9 presents the Decision Curve Analysis (DCA) for the training and testing

EARLY ACCESS

sets. These curves demonstrate that our fusion model offers significant advantages in terms of its predictive probabilities.

Nomogram： Figure 10 Nomograms suggest that EGFR, TNM, SIINI, gender, and DeepLearning are incorporated into the combine model, and the corresponding scores can predict the unresectable NSCLC response to ICIs treatment in the early stage. Therefore, EGFR mutation negative, TNM stage III, low SIINI score, and high deep learning index indicated a greater tendency to have an early response to ICIs.

DISCUSSION

The persistent global burden of lung cancer, characterized by high incidence and mortality rates, has driven multidisciplinary efforts to identify clinically actionable biomarkers for predicting ICI response in unresectable NSCLC. Current biomarker discovery paradigms span traditional histopathological evaluation to molecular profiling (29-31). Prior studies have identified numerous hematological parameters as prospective prognostic markers, including PD-L1, TMB, the neutrophil-to-lymphocyte ratio (NLR), derived NLR (dNLR), the platelet-to-lymphocyte ratio (PLR), the prognostic nutritional index (PNI), the systemic immune-inflammation index (SII), the advanced lung cancer inflammation index (ALI), alongside hemoglobin concentrations, among others (15, 32-34). These parameters reflect distinct aspects of the tumor-host interface, yet their clinical application remains constrained by high costs, inherent biological variability, and limited capacity to capture the complex multidimensional nature of antitumor immunity.

Throughout the immunotherapy for NSCLC, lymphocytes play an instrumental role in tumor defense via inducing apoptosis and inhibiting tumor cell proliferation and migration (33). The reduction in lymphocytes may reflect a decrease in $CD4^+$ T lymphocytes, leading to a weakened lymphocyte-mediated immune response to malignant tumors (35). Furthermore, study by Lee et al.(36) suggests a possible link between serum hemoglobin levels and outcomes in lung cancer patients. Additionally, some studies indicate that NSCLC patients with lower baseline platelet (PLT) and NLR levels tend to have better prognoses (34).These findings suggest that the tumor inflammatory microenvironment may be closely related to anti-tumor immune responses, which can significantly impact the prognosis of NSCLC patients (37-39). Therefore, our

EARLY ACCESS

study identifies a novel indicator (SIINI), which considers various aspects of the body, offering a more comprehensive evaluation of immune, inflammatory, and nutritional indicators. The SIINI integrates neutrophil count, lymphocyte count, platelet count, hemoglobin level, serum albumin level, and BMI, offering a comprehensive evaluation of the nutritional, inflammatory, and immune status in patients with NSCLC. SIINI can be used not only to predict patient prognosis but also to assess treatment efficacy, and potentially offers greater clinical significance compared to established indicators such as NLR, PLR, PNI, SII, and ALI.

Furthermore, within this landscape, CT-based radiological biomarkers hold unique translational potential due to their intrinsic non-invasive nature and universal acquisition during standard diagnostic workflows(40). Unlike invasive tissue sampling techniques, which are limited by spatial sampling bias, advanced imaging modalities enable comprehensive three-dimensional tumor characterization, capturing both intralesional heterogeneity and peritumoral microenvironmental features with millimeter-level spatial resolution (41). A paramount advantage of deep learning in radiomics feature extraction lies in its adaptability and proficiency in discerning patterns from image data(42, 43). Presently, it epitomizes the pinnacle of image analysis and categorization, consistently surpassing antecedent image analysis methodologies (44, 45). Rakaee et al.(46) constructed a machine learning model based on tumor-infiltrating lymphocyte (TIL) scoring to forecast the response of NSCLC to immune checkpoint inhibitors. In a parallel vein, Vanguri et al. amalgamated radiological, histopathological, and genomic attributes to gauge the predictive potential of immune therapy responses in NSCLC. Through the application of machine learning, they consolidated multimodal attributes into a risk prediction paradigm. The investigation revealed that the multimodal framework attained an AUC of 0.80, surpassing any solitary variable. These discoveries provide a quantitative foundation for harnessing multimodal integrated attributes in conjunction with machine learning to augment the precision of anticipating immune therapy responses in NSCLC patients (47). In our research, the DenseNet121 model manages to categorize responses at each evaluation point, essentially augmenting the training dataset, even with a small sample size. Furthermore, the validation set proves to be an unexpectedly good predictor in distinguishing responders from non-responders. Moreover,

EARLY ACCESS

external validation demonstrated good generalizability (AUC = 0.823), confirming the universality of the core predictive factors.

This study primarily focused on patients with unresectable NSCLC at advanced TNM stages. Moreover, for patients with EGFR-positive NSCLC, targeted therapy remains the preferred treatment approach. Nevertheless, the emergence of resistance to targeted therapy is an inevitable challenge. In this context, ICIs have emerged as a promising therapeutic avenue for unresectable NSCLC patients. This study suggests that EGFR mutation status is associated with the response to ICIs, with EGFR-negative patients being more likely to exhibit an early response to ICI treatment, which is consistent with the findings of Jiang and colleagues (48, 49). The SIINI serves as both a clinical efficacy biomarker and prognostic indicator, offering distinct advantages in accessibility, safety, cost-efficiency, reproducibility, and adaptability for longitudinal monitoring. These strengths stem from its calculation using routine clinical parameters, including complete blood count, biochemical profiles, and BMI. However, some inflammatory components in SIINI are susceptible to a variety of confounding factors, which may lead to differences in model performance across different data sets and introduce bias into the research results. To mitigate such limitations, integrating SIINI with complementary inflammatory biomarkers and adopting longitudinal assessments could serve as effective strategies to enhance diagnostic accuracy and reduce measurement variability.

Although results are promising, this research faced several constraints. Initially, participant numbers were limited. Although the data of two medical centers in China are included, further validation through larger prospective studies is needed. Secondly, real-world data collection reveals that there are challenges in obtaining indicators such as TMB, circulating tumor DNA (ctDNA) (50), which may be due to high detection costs or inconsistent detection standards in local medical centers. Future investigations should synergistically integrate multi-omics biomarker panels, expand prospective multicenter validation frameworks, and achieve statistically powered cohort sizes ($n \geq 400$), collectively addressing current limitations in predictive robustness and clinical implementation scalability.

EARLY ACCESS

CONCLUSION

This study investigated the early predictive capability of a CT-based deep learning model combined with the inflammation parameter SIINI for predicting the response of unresectable NSCLC patients to ICIs. By aiding in the selection of suitable candidates for ICI treatment, this research aims to reduce unnecessary financial and time burdens on patients while providing a feasible approach for precision therapy.

ACKNOWLEDGMENTS

All authors acknowledge the Department of Thoracic Surgery, Northern Jiangsu People's Hospital Affiliated to Yangzhou University; Medical College, Yangzhou University; Department of Thoracic Surgery, The Affiliated Taizhou People's Hospital of Nanjing Medical University, and OnekeyAI platform, for their assistance with this work.

Conflict of interest: Authors declare no conflicts of interest.

Funding: This work was supported by the Postgraduate Research & Practice Innovation Program of Jiangsu Province grant KYCX23_3623 (L.Y.); The clinical project of Taizhou Clinical Medical School of Nanjing Medical University grant TZKY20240204 (F.S.) and National Natural Science Foundation of China (No. 82070020; 82470101)

Data availability: The datasets presented in this study are available in the Supplementary Material and Zenodo DOI: https://doi.org/10.5281/zenodo.15546756.

Submitted: 04 March 2025

Accepted: 03 June 2025

Published online: 10 June 2025

EARLY ACCESS

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49. https://doi.org/10.3322/caac.21660.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021;71(1):7-33. https://doi.org/10.3322/caac.21654.

3. Brozos-Vázquez EM, Díaz-Peña R, García-González J, León-Mateos L, Mondelo-Macía P, Peña-Chilet M, et al. Immunotherapy in nonsmall-cell lung cancer: current status and future prospects for liquid biopsy. Cancer Immunol Immun. 2021;70(5):1177-88. https://doi.org/10.1007/s00262-020-02752-z.
4. Howlader N, Forjaz G, Mooradian MJ, Meza R, Kong CY, Cronin KA, et al. The Effect of Advances in Lung-Cancer Treatment on Population Mortality. New Engl J Med. 2020;383(7):640-9. https://doi.org/10.1056/NEJMoa1916623.
5. Leslie M. High TMB Predicts Immunotherapy Benefit. Cancer Discov. 2018;8(6):668. https://doi.org/Doi 10.1158/2159-8290.Cd-Nb2018-048.
6. Kerr K, Tsao M, Nicholson A, Yatabe Y, Wistuba I, Hirsch F. Programmed Death-Ligand 1 Immunohistochemistry in Lung Cancer: In what state is this art? J Thorac Oncol. 2015;10(7):985-9. https://doi.org/10.1097/jto.0000000000000526.
7. Zhang Y, Cheng R, Ding T, Wu J. Discrepancies in PD-L1 expression, lymphocyte infiltration, and tumor mutational burden in non-small cell lung cancer and matched brain metastases. Transl Lung Cancer Res. 2024;13(12):3590-602. https://doi.org/10.21037/tlcr-24-735.
8. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman D, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017;127(8):2930-40. https://doi.org/10.1172/jci91190.
9. Ilie M, Long-Mira E, Bence C, Butori C, Lassalle S, Bouhlel L, et al. Comparative study of the PD-L1 status between surgically resected specimens and matched biopsies of NSCLC patients reveal major discordances: a potential issue for anti-PD-L1 therapeutic strategies. Ann Oncol. 2016;27(1):147-53. https://doi.org/10.1093/annonc/mdv489.
10. McLaughlin J, Han G, Schalper KA, Carvajal-Hausdorf D, Pelekanou V, Rehman J, et al. Quantitative Assessment of the Heterogeneity of PD-L1 Expression in Non-Small-Cell Lung Cancer. JAMA Oncol. 2016;2(1):46-54. https://doi.org/10.1001/jamaoncol.2015.3638.
11. Valero C, Lee M, Hoen D, Weiss K, Kelly DW, Adusumilli PS, et al. Pretreatment neutrophil-to-lymphocyte ratio and mutational burden as biomarkers of tumor response to immune checkpoint inhibitors. Nat Commun. 2021;12(1):729. https://doi.org/10.1038/s41467-021-20935-9.
12. Guo W, Cai SH, Zhang F, Shao F, Zhang GC, Zhou Y, et al. Systemic immune-inflammation index (SII) is useful to predict survival outcomes in patients with surgically resected non-small cell lung cancer. Thorac Cancer. 2019;10(4):761-8. https://doi.org/10.1111/1759-7714.12995.
13. Diem S, Schmid S, Krapf M, Flatz L, Born D, Jochum W, et al. Neutrophil-to-Lymphocyte ratio (NLR) and Platelet-to-Lymphocyte ratio (PLR) as prognostic markers in patients with non-small cell lung cancer (NSCLC) treated with nivolumab. Lung Cancer. 2017;111:176-81. https://doi.org/10.1016/j.lungcan.2017.07.024.
14. Zhai BQ, Chen J, Wu JC, Yang L, Guo XL, Shao JJ, et al. Predictive value of the hemoglobin, albumin, lymphocyte, and platelet (HALP) score and lymphocyte-to-monocyte ratio (LMR) in patients with non-small cell lung cancer after radical lung cancer surgery. Ann Transl Med. 2021;9(12):976. https://doi.org/10.21037/atm-21-2120.
15. Shao HC, Zhu J, Shi L, Yao J, Wang YX, Ma CG, et al. Value of computed tomography radiomics combined with inflammation indices in predicting the efficacy of

EARLY ACCESS

immunotherapy in patients with locally advanced and metastatic non-small cell lung cancer. J Thorac Dis. 2024;16(5):3213-27. https://doi.org/10.21037/jtd-24-526.
16. Xia YY, Xia CL, Wu LD, Li Z, Li H, Zhang JX. Systemic Immune Inflammation Index (SII), System Inflammation Response Index (SIRI) and Risk of All-Cause Mortality and Cardiovascular Mortality: A 20-Year Follow-Up Cohort Study of 42,875 US Adults. J Clin Med. 2023;12(3):1128. https://doi.org/10.3390/jcm12031128.
17. Zheng J, Xu S, Wang GY, Shi YM. Applications of CT-based radiomics for the prediction of immune checkpoint markers and immunotherapeutic outcomes in non-small cell lung cancer. Front Immunol. 2024;15:1434171. https://doi.org/10.3389/fimmu.2024.1434171.
18. Parekh V, Jacobs M. Deep learning and radiomics in precision medicine. Expert Rev Precis Med Drug Dev. 2019;4(2):59-72. https://doi.org/10.1080/23808993.2019.1585805.
19. Dong D, Fang MJ, Tang L, Shan XH, Gao JB, Giganti F, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. Ann Oncol. 2020;31(7):912-20. https://doi.org/10.1016/j.annonc.2020.04.003.
20. Zhang LW, Dong D, Zhang WJ, Hao XH, Fang MJ, Wang S, et al. A deep learning risk prediction model for overall survival in patients with gastric cancer: A multicenter study. Radiother Oncol. 2020;150:73-80. https://doi.org/10.1016/j.radonc.2020.06.010.
21. Coroller TP, Agrawal V, Narayan V, Hou Y, Grossmann P, Lee SW, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. Radiother Oncol. 2016;119(3):480-6. https://doi.org/10.1016/j.radonc.2016.04.004.
22. Coroller TP, Agrawal V, Huynh E, Narayan V, Lee SW, Mak RH, et al. Radiomic-Based Pathological Response Prediction from Primary Tumors and Lymph Nodes in NSCLC. J Thorac Oncol. 2017;12(3):467-76. https://doi.org/10.1016/j.jtho.2016.11.2226.
23. Yuan L, Shen ZM, Shan YB, Zhu JW, Wang Q, Lu Y, et al. Unveiling the landscape of pathomics in personalized immunotherapy for lung cancer: a bibliometric analysis. Front Oncol. 2024; 14:1432212. https://doi.org/10.3389/fonc.2024.1432212.
24. Yin XY, Lu Y, Cui YB, Zhou ZC, Wen JX, Huang ZQ, et al. CT-based radiomics-deep learning model predicts occult lymph node metastasis in early-stage lung adenocarcinoma patients: A multicenter study. Chin J Cancer Res. 2025;37(1):12-27. https://doi.org/10.21147/j.issn.1000-9604.2025.01.02.
25. Detterbeck FC, Franklin WA, Nicholson AG, Girard N, Arenberg DA, Travis WD, et al. The IASLC Lung Cancer Staging Project: Background Data and Proposed Criteria to Distinguish Separate Primary Lung Cancers from Metastatic Foci in Patients with Two Lung Tumors in the Forthcoming Eighth Edition of the TNM Classification for Lung Cancer. J Thorac Oncol. 2016;11(5):651-65. https://doi.org/10.1016/j.jtho.2016.01.025.
26. Shang YL, Zeng Y, Luo SW, Wang YS, Yao JQ, Li M, et al. Habitat Imaging With Tumoral and Peritumoral Radiomics for Prediction of Lung Adenocarcinoma Invasiveness on Preoperative Chest CT: A Multicenter Study. Am J Roentgenol. 2024;223(4):e2431675. https://doi.org/10.2214/AJR.24.31675.

EARLY ACCESS

27. Dehbozorgi P, Ryabchykov O, Bocklitz T. A comparative study of statistical, radiomics, and deep learning feature extraction techniques for medical image classification in optical and radiological modalities. Computers in biology and medicine. 2025;187:109768. https://doi.org/10.1016/j.compbiomed.2025.109768.
28. Kumar Y, Shrivastav S, Garg K, Modi N, Wiltos K, Wozniak M, et al. Automating cancer diagnosis using advanced deep learning techniques for multi-cancer image classification. Sci Rep. 2024;14(1):25006. https://doi.org/10.1038/s41598-024-75876-2.
29. Mu W, Tunali I, Gray J, Qi J, Schabath M, Gillies R. Radiomics of ^{18}F-FDG PET/CT images predicts clinical benefit of advanced NSCLC patients to checkpoint blockade immunotherapy. Eur J Nucl Med Mol Imaging. 2020;47(5):1168-82. https://doi.org/10.1007/s00259-019-04625-9.
30. He BX, Dong D, She YL, Zhou CC, Fang MJ, Zhu YB, et al. Predicting response to immunotherapy in advanced non-small-cell lung cancer using tumor mutational burden radiomic biomarker. J Immunother Cancer. 2020;8(2):e000550. https://doi.org/10.1136/jitc-2020-000550.
31. Xie YL, Chen HL, Tian M, Wang ZQ, Wang LY, Zhang J, et al. Integrating multi-omics and machine learning survival frameworks to build a prognostic model based on immune function and cell death patterns in a lung adenocarcinoma cohort. Front Immunol. 2024;15:1460547. https://doi.org/10.3389/fimmu.2024.1460547.
32. Lei W, Wang W, Qin S, Yao W. Predictive value of inflammation and nutritional index in immunotherapy for stage IV non-small cell lung cancer and model construction. Sci Rep. 2024;14(1):17511. https://doi.org/10.1038/s41598-024-66813-4.
33. Xu FH, Zhu H, Dong YJ, Li L, Liu N, Yuan SH. Combined inflammatory parameters and tertiary lymphoid structure predict prognosis in patients with resectable non-small cell lung cancer treated with neoadjuvant chemoimmunotherapy. Front Immunol. 2023;14:1244256. https://doi.org/10.3389/fimmu.2023.1244256.
34. Wang Z, Li XX, Chen JH, Hua Y, Zhong XQ, Tang CH, et al. The significance of inflammatory markers in prognosticating the effectiveness and safety of immunotherapy in conjunction with chemotherapy during the primary intervention of advanced non-small cell lung carcinoma. Lung Cancer. 2024;192:107817. https://doi.org/10.1016/j.lungcan.2024.107817.
35. Berthe J, Poudel P, Segerer FJ, Jennings EC, Ng F, Surace M, et al. Exploring the impact of tertiary lymphoid structures maturity in NSCLC: insights from TLS scoring. Front Immunol. 2024;15:1422206. https://doi.org/10.3389/fimmu.2024.1422206.
36. Lee S, Jeon H, Shim B. Prognostic Value of Ferritin-to-Hemoglobin Ratio in Patients with Advanced Non-Small-Cell Lung Cancer. J Cancer. 2019;10(7):1717-25. https://doi.org/10.7150/jca.26853.
37. Zhai WY, Zhang C, Duan FF, Xie JD, Dai SQ, Lin YB, et al. Dynamics of peripheral blood inflammatory index predict tumor pathological response and survival among patients with locally advanced non-small cell lung cancer who underwent neoadjuvant immunochemotherapy: a multi-cohort retrospective study. Front Immunol. 2024;15:1422717. https://doi.org/10.3389/fimmu.2024.1422717.
38. Lei WQ, Wang W, Qin SX, Yao WR. Predictive value of inflammation and nutritional index in immunotherapy for stage IV non-small cell lung cancer and model construction. Sci Rep. 2024;14(1):17511.https://doi.org/10.1038/s41598-024-66813-4.

EARLY ACCESS

39. Li N, Yu ZW, Zhang X, Liu TM, Sun YX, Wang RT, et al. Elevated mean platelet volume predicts poor prognosis in colorectal cancer. Sci Rep. 2017;7(1):10261. https://doi.org/10.1038/s41598-017-11053-y.
40. Wen Q, Yang Z, Dai HH, Feng A, Li Q. Radiomics Study for Predicting the Expression of PD-L1 and Tumor Mutation Burden in Non-Small Cell Lung Cancer Based on CT Images and Clinicopathological Features. Front Oncol. 2021;11:620246. https://doi.org/10.3389/fonc.2021.620246.
41. Wang JH, Wahid KA, van Dijk L, Farahani K, Thompson RF, Fuller CD. Radiomic biomarkers of tumor immune biology and immunotherapy response. Clin Transl Rad Onco. 2021;28:97-115. https://doi.org/10.1016/j.ctro.2021.03.006.
42. Qiu QT, Xing LG, Wang Y, Feng AL, Wen Q. Development and Validation of a Radiomics Nomogram Using Computed Tomography for Differentiating Immune Checkpoint Inhibitor-Related Pneumonitis From Radiation Pneumonitis for Patients With Non-Small Cell Lung Cancer. Front Immunol. 2022;13:870842. https://doi.org/10.3389/fimmu.2022.870842.
43. Cao PD, Jia X, Wang X, Fan LY, Chen Z, Zhao YY, et al. Deep learning radiomics for the prediction of epidermal growth factor receptor mutation status based on MRI in brain metastasis from lung adenocarcinoma patients. BMC cancer, 25(1):443. https://doi.org/10.1186/s12885-025-13823-8.
44. Wang H, Chen YZ, Li WH, Han Y, Li Q, Ye ZX. Pretreatment Thoracic CT Radiomic Features to Predict Brain Metastases in Patients With ALK-Rearranged Non-Small Cell Lung Cancer. Front Genet. 2022;13:772090. https://doi.org/10.3389/fgene.2022.772090.
45. Robertson S, Azizpour H, Smith K, Hartman J. Digital image analysis in breast pathology-from image processing techniques to artificial intelligence. Transl Res. 2018;194:19-35. https://doi.org/10.1016/j.trsl.2017.10.010.
46. Rakaee M, Adib E, Ricciuti B, Sholl LM, Shi WW, Alessi J, et al. Association of Machine Learning-Based Assessment of Tumor-Infiltrating Lymphocytes on Standard Histologic Images With Outcomes of Immunotherapy in Patients With NSCLC. JAMA Oncol. 2023;9(1):51-60. https://doi.org/10.1001/jamaoncol.2022.4933.
47. Vanguri RS, Luo J, Aukerman AT, Egger J, Fong CJ, Horvat N, et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. Nat Cancer. 2022;3(10):1151-1164. https://doi.org/10.1038/s43018-022-00416-8.
48. Ling J, Zhiyi L, Na L, Jinfang J, Cengceng L, Shenghang D, et al. [Correlation Study on Expression of PD-1 and PD-L1 in Non-small Cell Lung Cancer and Epidermal Growth Factor Receptor Mutations]. Zhongguo Fei Ai Za Zhi. 2021;24(9):623-631. https://doi.org/10.3779/j.issn.1009-3419.2021.102.31.
49. Shi SL, Wang YY, Wu JJ, Zha BY, Li PH, Liu YK, et al. Predictive value of PD-L1 and TMB for short-term efficacy prognosis in non-small cell lung cancer and construction of prediction models. Front Oncol. 2024;14:1342262. https://doi.org/10.3389/fonc.2024.1342262.
50. Ryssel H, Ahlborn L, Dejanovic D, Keller S, Pøhl M, Østrup O, et al. Circulating Tumor DNA and [F]FDG-PET for Early Response Assessment in Patients with Advanced NSCLC. Diagnostics (Basel). 2025;15(3):247. https://doi.org/10.3390/diagnostics15030247.

EARLY ACCESS

TABLES AND FIGURES WITH LEGENDS

Table 1. Baseline characteristics

Feature_name	train-label=0	train-label=1	*P*-value	val-label=0	val-label=1	*P*-value
age	65.97±10.21	67.61±8.34	0.499	66.87±10.22	65.71±10.02	0.702
BMI	22.76±3.75	22.47±3.94	0.618	24.21±4.55	22.67±4.58	0.253
modality	3.17±1.92	3.11±1.60	0.992	3.07±1.98	3.40±1.74	0.605
NLR	4.83±2.93	3.41±1.01	0.002	3.99±1.61	3.27±0.95	0.298
PLR	202.05±107.29	191.07±94.95	0.784	189.75±75.05	197.81±127.05	0.67
ALI	290.00±236.72	403.32±1086.61	0.717	284.93±141.93	317.74±250.06	0.905
SII	1109.62±924.00	922.70±852.36	0.255	994.24±591.15	794.51±648.57	0.116
PNI	45.73±5.97	46.97±5.65	0.272	47.91±5.24	48.77±5.65	0.604
SIINI	179.80±69.18	120.42±59.06	<0.001	169.75±49.88	123.02±70.43	0.003
gender			0.025			0.835
female	10(28.57)	11(10.78)		3(20.00)	6(13.33)	
male	25(71.43)	91(89.22)		12(80.00)	39(86.67)	
smoking			0.826			0.343
0	13(37.14)	42(41.18)		7(46.67)	13(28.89)	
1	22(62.86)	60(58.82)		8(53.33)	32(71.11)	
basic_disease			0.317			0.536
0	16(45.71)	35(34.31)		7(46.67)	15(33.33)	
1	19(54.29)	67(65.69)		8(53.33)	30(66.67)	
ECOG			0.095			0.013
0	null	3(2.94)		null	2(4.44)	
1	13(37.14)	58(56.86)		4(26.67)	25(55.56)	
2	15(42.86)	31(30.39)		6(40.00)	16(35.56)	
3	7(20.00)	10(9.80)		5(33.33)	2(4.44)	
EGFR			<0.001			0.002
0	6(17.14)	86(84.31)		7(46.67)	40(88.89)	
1	29(82.86)	16(15.69)		8(53.33)	5(11.11)	
PD_L1			0.247			0.339
0	32(91.43)	78(76.47)		12(80.00)	38(84.44)	
1	2(5.71)	11(10.78)		1(6.67)	5(11.11)	
2	1(2.86)	8(7.84)		1(6.67)	2(4.44)	
3	null	5(4.90)		1(6.67)	null	

EARLY ACCESS

Feature_name	train-label=0	train-label=1	*P*-value	val-label=0	val-label=1	*P*-value
medication_regimen			0.291			0.318
1	13(37.14)	29(28.43)		7(46.67)	13(28.89)	
2	19(54.29)	69(67.65)		8(53.33)	29(64.44)	
3	3(8.57)	4(3.92)		null	3(6.67)	
treatment_lines			0.389			0.342
1	20(57.14)	69(67.65)		8(53.33)	28(62.22)	
2	12(34.29)	23(22.55)		6(40.00)	10(22.22)	
3	3(8.57)	10(9.80)		1(6.67)	7(15.56)	
type			0.578			0.732
1	33(94.29)	100(98.04)		15(100.00)	42(93.33)	
2	2(5.71)	2(1.96)		null	3(6.67)	
pathological_type			0.399			1.0
0	16(45.71)	57(55.88)		7(46.67)	20(44.44)	
1	19(54.29)	45(44.12)		8(53.33)	25(55.56)	
TNM			0.004			0.408
3	3(8.57)	37(36.27)		6(40.00)	11(24.44)	
4	32(91.43)	65(63.73)		9(60.00)	34(75.56)	

Notes: gender(female==0/male==1), smoking history(Current or former smokers==1/Never smokers==0), basic disease(with basic disease==1/without basic disease==0), medication regimen (Pembrolizumab==1/ Nivolumab==2/ Sintilimab==3), EGFR mutation (Postive==1/Negative==0), TNM(IIIB==3/IV==4), pathological type(adenocarcinoma==1/squamocellular carcinoma==0), PD-L1 expression(No record==0, TPS <1%==1、TPS≥1%-49%==2、TPS≥50%==3).

Table 2. Metrics of clinical model

Model_name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
LR	0.905	0.929	0.871 - 0.986	0.931	0.829	0.941	0.806	train
LR	0.767	0.741	0.597 - 0.884	0.800	0.667	0.878	0.526	val
LR	0.776	0.868	0.758	0.750	1.000	1.00	0.31	test

EARLY ACCESS

Model_name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
			- 0.978			0	2	
SVM	0.883	0.965	0.934 - 0.995	0.863	0.943	0.978	0.702	train
SVM	0.700	0.794	0.672 - 0.916	0.644	0.867	0.935	0.448	val
SVM	0.796	0.855	0.750 - 0.959	0.773	1.000	1.000	0.333	test
LightGBM	0.891	0.938	0.892 - 0.984	0.912	0.829	0.939	0.763	train
LightGBM	0.650	0.830	0.726 - 0.933	0.533	1.000	1.000	0.417	val
LightGBM	0.571	0.820	0.693 - 0.947	0.523	1.000	1.000	0.192	test

Table 3. Rad signature results

Model_name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
LR	0.679	0.801	0.722 - 0.880	0.618	0.857	0.926	0.435	train
LR	0.717	0.551	0.360 - 0.742	0.822	0.400	0.804	0.429	val
LR	0.755	0.673	0.457 - 0.888	0.773	0.600	0.944	0.231	test
SVM	0.839	0.912	0.847 - 0.977	0.804	0.943	0.976	0.623	train
SVM	0.700	0.517	0.319 - 0.715	0.822	0.333	0.787	0.385	val

EARLY ACCESS

Model_name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
SVM	0.490	0.718	0.523 - 0.913	0.432	1.000	1.000	0.167	test
LightGBM	0.591	0.817	0.748 - 0.886	0.461	0.971	0.979	0.382	train
LightGBM	0.333	0.624	0.469 - 0.779	0.133	0.933	0.857	0.264	val
LightGBM	0.469	0.652	0.411 - 0.894	0.432	0.800	0.950	0.138	test
Random Forest	0.803	0.861	0.788 - 0.935	0.794	0.829	0.931	0.580	train
Random Forest	0.683	0.622	0.445 - 0.800	0.733	0.533	0.825	0.400	val
Random Forest	0.755	0.632	0.332 - 0.931	0.773	0.600	0.944	0.231	test

Table 4. Metric results for deep learning radiomics signature

Model_name	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
Densenet121	0.730	0.846	0.7793 - 0.9126	0.667	0.914	0.958	0.485	train
Densenet121	0.633	0.751	0.6300 - 0.8722	0.511	1.000	1.000	0.405	val
Densenet121	0.490	0.691	0.4747 - 0.9071	0.432	1.000	1.000	0.167	test

Table 5. Metrics on different signature

Signature	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort	
Clinical	0.891	0.938	0.8924 - 0.9840	0.912	0.829	0.939	0.763	train	
Radiomics	0.591	0.817	0.7478 - 0.8861	0.461	0.971	0.979	0.382	train	
DeepLearning	0.730	0.846	0.7793 -	0.667	0.914	0.958	0.485	train	

EARLY ACCESS

Signature	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort	
			0.9126						
Combined	0.942	0.966	0.9322 - 0.9995	0.971	0.857	0.952	0.909	train	
Clinical	0.650	0.830	0.7258 - 0.9334	0.533	1.000	1.000	0.417	val	
Radiomics	0.333	0.624	0.4687 - 0.7787	0.133	0.933	0.857	0.264	val	
DeepLearning	0.633	0.751	0.6300 - 0.8722	0.511	1.000	1.000	0.405	val	
Combined	0.800	0.865	0.7709 - 0.9595	0.778	0.867	0.946	0.565	val	
Clinical	0.571	0.820	0.6934 - 0.9475	0.523	1.000	1.000	0.192	test	
Radiomics	0.469	0.652	0.4109 - 0.8937	0.432	0.800	0.950	0.138	test	
DeepLearning	0.490	0.691	0.4747 - 0.9071	0.432	1.000	1.000	0.167	test	
Combined	0.612	0.823	0.6627 - 0.9827	0.568	1.000	1.000	0.208	test	

EARLY ACCESS

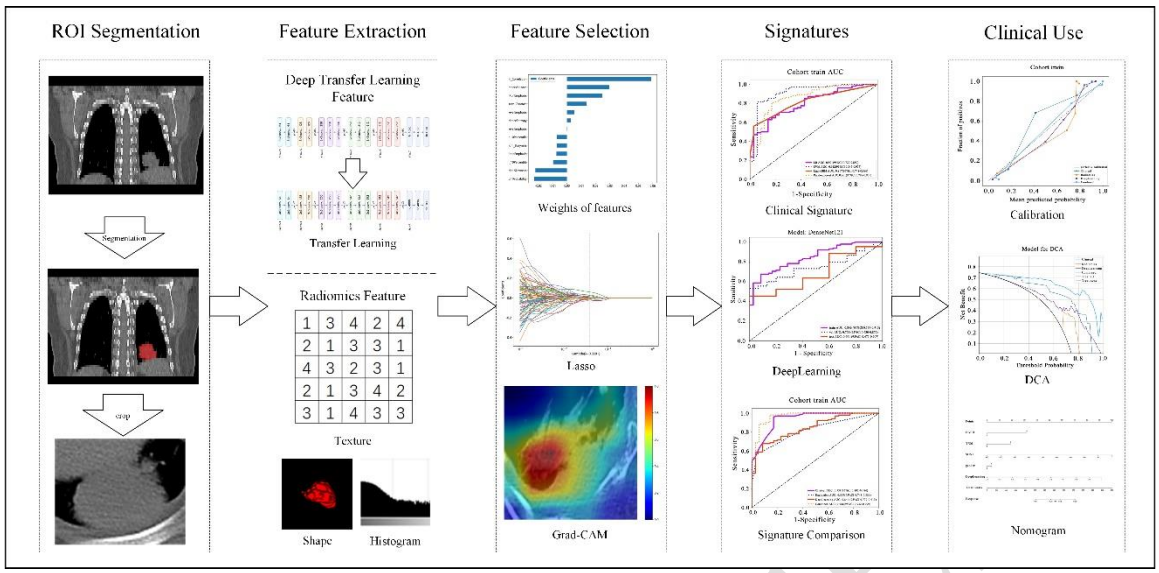


Figure 1. Workflow of this study

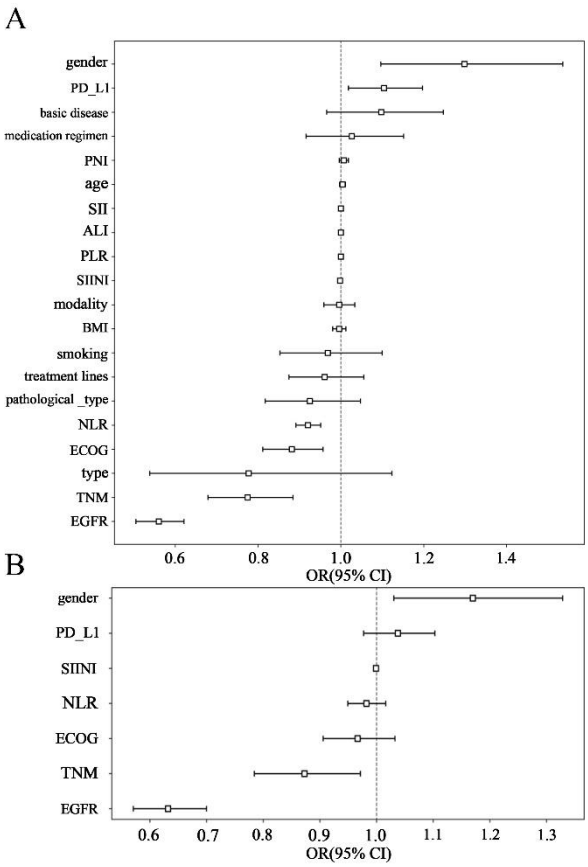


Figure 2. OR of clinical features: **(A)** OR of clinical features in univariable analysis; **(B)** OR of clinical features in multivariable analysis.

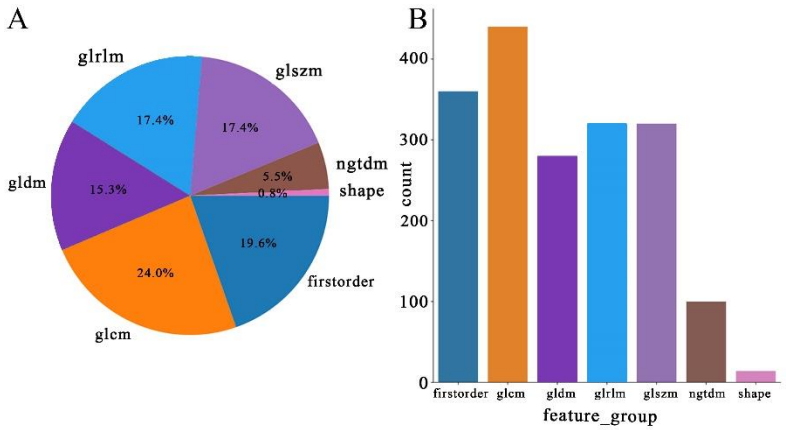
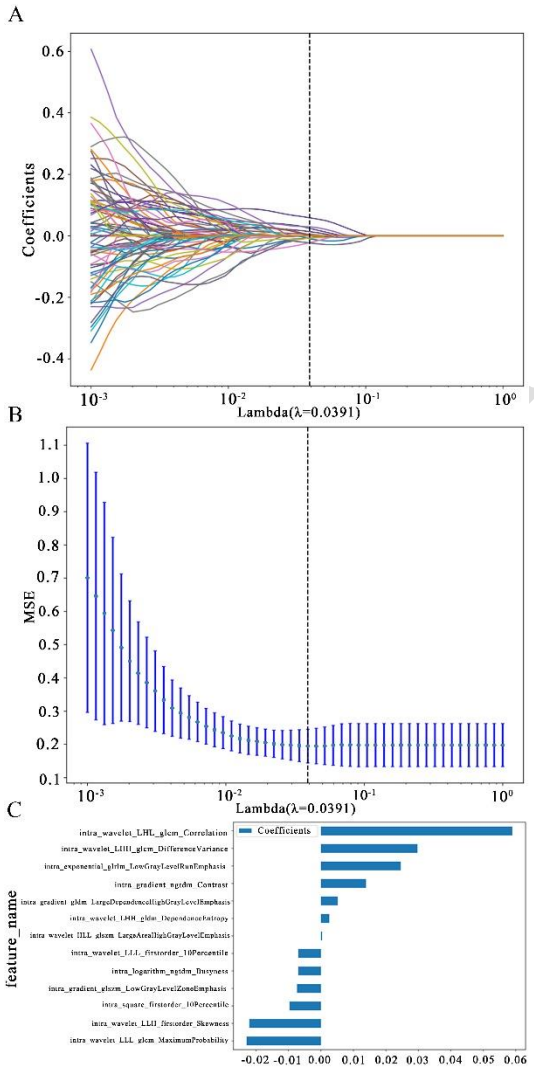


Figure 3. The distribution of these handcrafted features: **(A)** Number of handcrafted features; **(B)** Ratio of handcrafted features.



ACCESS

Figure 4. showcases the coefficients obtained through Least Absolute Shrinkage and Selection Operator (LASSO) regression using 10-fold cross-validation: **(A)** Lasso regularization paths; **(B)** Mean Squared Error (MSE); **(C)** Weights of the selected radiomic features.

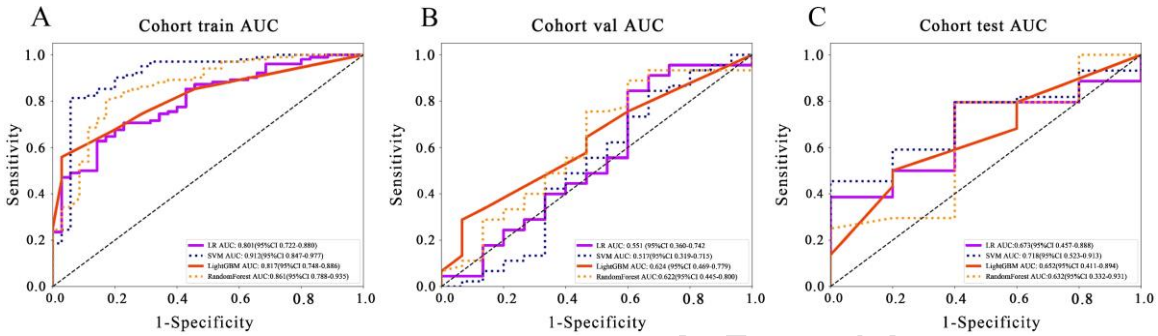


Figure 5. Performance of the models for rad signature on different datasets: (A) training set; (B) validation set; (C) test set.

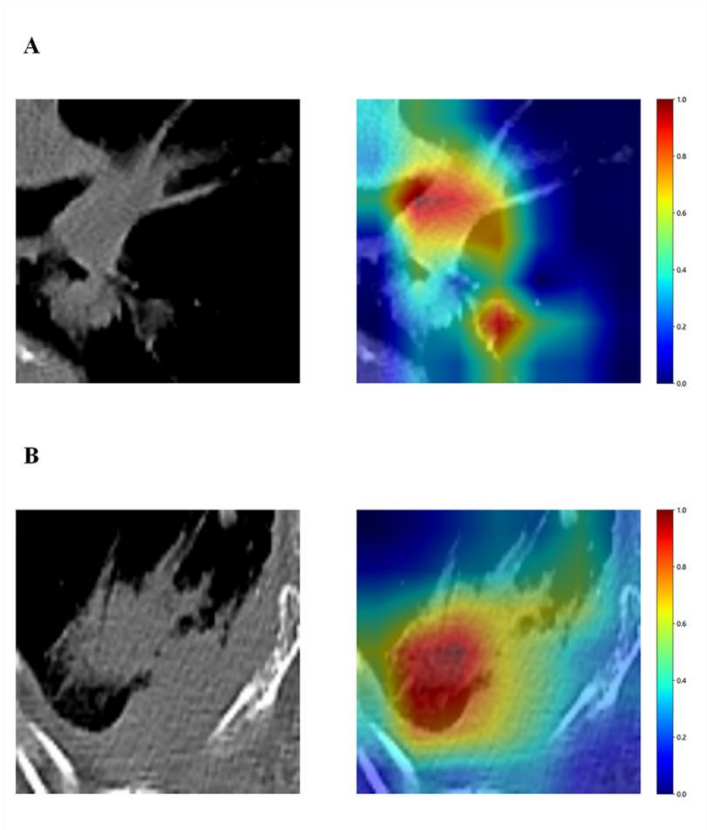


Figure 6. Present the Grad-CAM visualizations for two typical samples:(**A)** identified as "689166"; **(B)** identified as "866891". These visualizations are instrumental in demonstrating how the model focuses on different regions of the images for making its predictions.

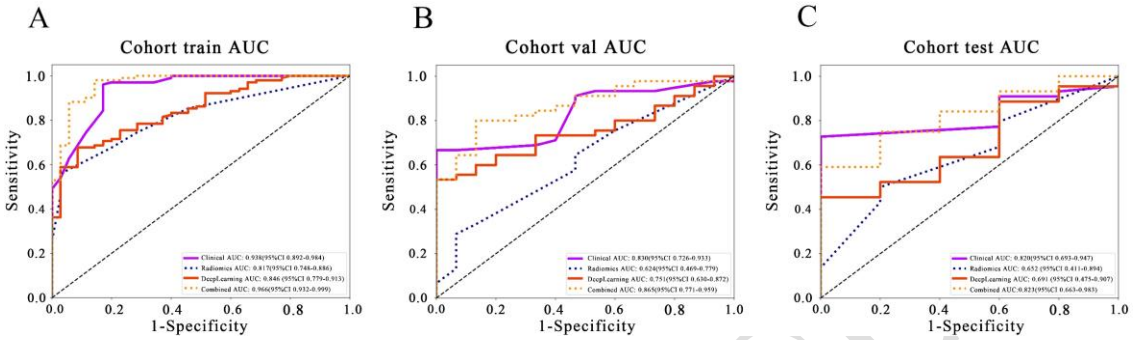


Figure 7. Illustration of the ROC curves for different model across various cohorts, offering a visual comparison of their diagnostic abilities: **(A)** training set; **(B)** validation set; **(C)** test set.

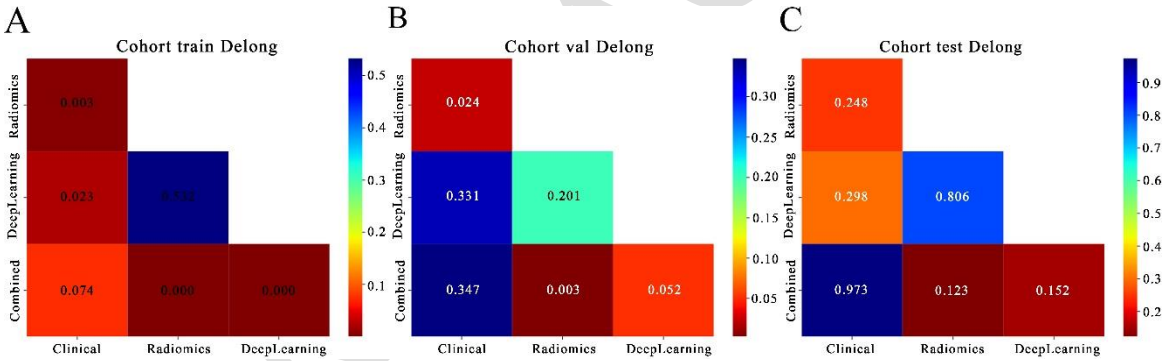


Figure 8. Presentation of the Delong test results of different model:**(A)** training set; **(B)** validation set; **(C)** test set.

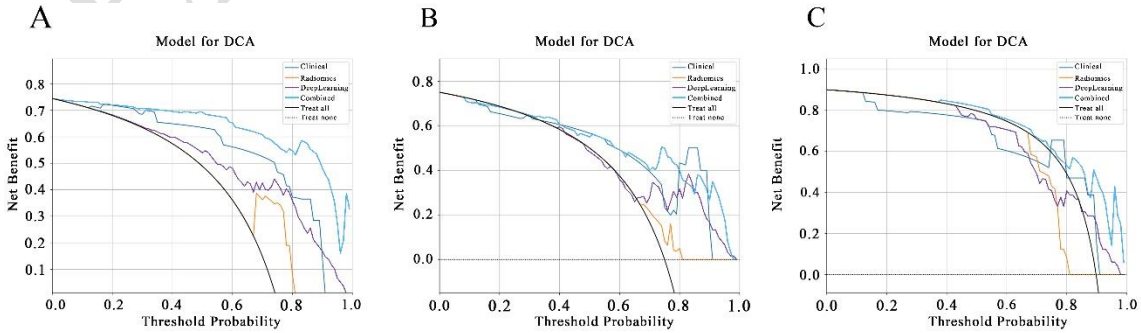


Figure 9. Decision curve of different model: **(A)** training set; **(B)** validation set; **(C)** test set.

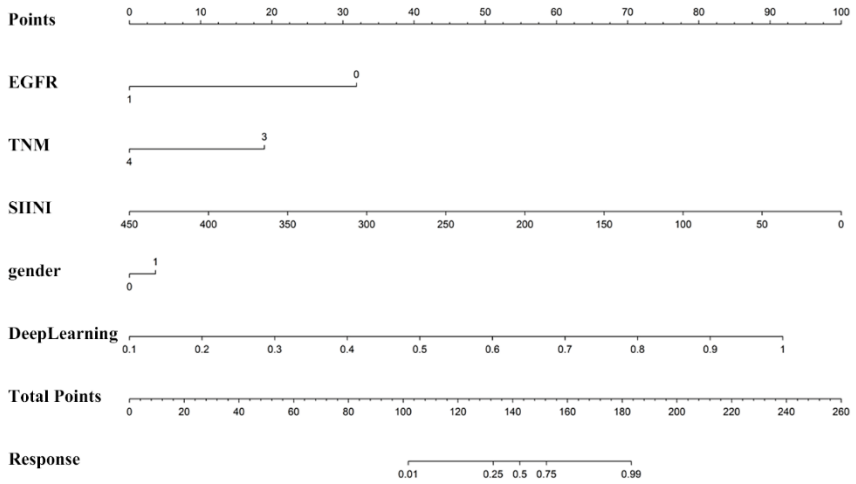


Figure 10. Nomogram.

SUPPLEMENTAL DATA

Supplemental data are available at the following link:

https://www.bjbms.org/ojs/index.php/bjbms/article/view/12324/3926

EARLY ACCESS