Original Article

Radiogenomics models for predicting prognosis in locally advanced non-small cell lung cancer patients undergoing definitive chemoradiotherapy

Xiaoyu Song[1,2#], Li Li[2,3,4#], Qingxi Yu[2], Ning Liu[3,4], Shouhui Zhu[2], Shuanghu Yuan[2,3,4]

[1]School of Clinical Medicine, Shandong Second Medical University, Weifang, China; [2]Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China; [3]Department of Radiation Oncology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China; [4]Department of Radiation Oncology, Anhui Provincial Cancer Hospital, Hefei, China

Contributions: (I) Conception and design: X Song, L Li; (II) Administrative support: S Yuan; (III) Provision of study materials or patients: Q Yu, N Liu, S Zhu; (IV) Collection and assembly of data: X Song, L Li; (V) Data analysis and interpretation: X Song; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Shuanghu Yuan, MD, PhD. Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, No. 440 Jiyan Road, Jinan 250117, China; Department of Radiation Oncology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230001, China; Department of Radiation Oncology, Anhui Provincial Cancer Hospital, Hefei 230031, China. Email: yuanshuanghu@sina.com.

Background: Definitive chemoradiotherapy (dCRT) is the cornerstone for locally advanced non-small cell lung cancer (LA-NSCLC). The study aimed to construct a multi-omics model integrating baseline clinical data, computed tomography (CT) images and genetic information to predict the prognosis of dCRT in LA-NSCLC patients.

Methods: The study retrospectively enrolled 105 stage III LA-NSCLC patients who had undergone dCRT. The pre-treatment CT images were collected, and the primary tumor was delineated as a region of interest (ROI) on the image using 3D-Slicer, and the radiomics features were extracted. The least absolute shrinkage and selection operator (LASSO) was employed for dimensionality reduction and selection of features. Genomic information was obtained from the baseline tumor tissue samples. We then constructed a multi-omics model by combining baseline clinical data, radiomics and genomics features. The predictive performance of the model was evaluated by the area under the curve (AUC) of the receiver operating characteristic (ROC) and the concordance index (C-index).

Results: The median follow-up time was 30.1 months, and the median progression-free survival (PFS) was 10.60 months. Four features were applied to construct the radiomics model. Multivariable analysis demonstrated the Rad-score, *KEAP1* and *MET* mutations were independent prognostic factors for PFS. The C-index of radiomics model, genomics model and radiogenomics model all performed well in the training group (0.590 *vs.* 0.606 *vs.* 0.663) and the validation group (0.599 *vs.* 0.594 *vs.* 0.650).

Conclusions: The radiomics model, genomics model and radiogenomics model can all predict the prognosis of dCRT for LA-NSCLC, and the radiogenomics model is superior to the single type model.

Keywords: Non-small cell lung cancer (NSCLC); radiomics; genomics; radiogenomics; prognosis

Submitted Mar 04, 2024. Accepted for publication Jul 17, 2024. Published online Aug 28, 2024.
doi: 10.21037/tlcr-24-145
View this article at: https://dx.doi.org/10.21037/tlcr-24-145

Introduction

Lung cancer ranks as the second most frequent cancer worldwide, non-small cell lung cancer (NSCLC) accounts for 81% of all cases. Approximately 30% of NSCLC patients are diagnosed with stages IIIA to IIIC. With the increasing use of immunotherapy, immune-consolidation after radiotherapy is the current standard treatment for unresectable locally advanced non-small cell lung cancer (LA-NSCLC) and the 5-year rate is 47.5% for overall survival (OS) and 33.1% for progression-free survival (PFS). In this treatment approach, definitive chemoradiotherapy (dCRT) is the cornerstone of comprehensive treatment. Studies indicate that 19.0% patients using concurrent chemoradiotherapy (CCRT) can be cured (1). Therefore, effective tools are needed to help screen patients with different prognoses after dCRT, avoid unnecessary additional consolidation therapy, and improve patient survival.

Traditional imaging techniques such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET) play a vital role in clinical NSCLC staging and predicting prognosis (2-6). Traditional imaging is intuitive and easy to evaluate, but there are also disadvantages of subjective differences between observers and less use of image information. Compared with traditional imaging techniques, radiomics can extract hundreds of quantitative features from images and obtain additional information about tumor phenotypes, genes and proteins, which more fully reflects the most essential features underlying medical images (7). In NSCLC, radiomics have been used to distinguish benign and malignant lung lesions (8-10) and forecasted tumor histological (11,12), treatment prognosis (13-19) or molecular properties (20-24). With the development of technology, more studies have begun to explore the relationship between genes and prognosis (25,26). For example, a retrospective study of osimertinib analyzed two phase III clinical studies (AURA3, FLAURA) and found that the clearance status of epidermal growth factor receptor (*EGFR*) mutation in plasma at different time points was correlated with the clinical outcome of NSCLC patients (27).

However, the majority of approaches are restricted to single-data models. By amalgamating multimodal data such as genetic, imaging and clinical information, it is helpful to understand the development and occurrence of diseases (28). The emergence of radiogenomics, which integrates radiomics and genetic information, can provide a deeper understanding of the biological nature of tumors (29). At present, the existing studies focus on the potential to predict postoperative recurrence of NSCLC (30,31). There are no studies on radiogenomic features to predict survival after dCRT for LA-NSCLC.

This study aimed to construct a multi-omics model to predict PFS by integrating baseline clinical data, CT images and genetic information from patients with LA-NSCLC, which could help identify high risk patients and formulate more accurate and personalized treatment plans for clinical practice. We present this article in accordance with the TRIPOD reporting checklist (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-145/rc).

Methods

Patients

This retrospective study included 105 patients with stage III NSCLC who had undergone dCRT in our hospital from October 2014 to March 2019. Inclusion criteria were as follows: (I) pathological diagnosed NSCLC; (II) stage IIIA-C NSCLC diagnosed according to the tumor, lymph node and metastasis (TNM) staging system; (III) dCRT; (IV) available pre-treatment CT images; (V) complete clinical records. Exclusion criteria included: (I) concurrent other primary malignancies; (II) inaccurate lesion segmentation, e.g., lesions combined with peripheral atelectasis; (III) loss to follow-up.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and endorsed by the ethical review committee of Shandong Cancer Hospital and Institute (No. SDTHEC202004042). The requirement for informed consent was waived owing to the retrospective nature of the study.

Treatment and assessments

All patients in this study underwent dCRT. During the follow-up period, the baseline CT images were compared to describe lesions and to assess the efficacy of treatment. The evaluation was conducted following the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 (32). PFS was defined as the duration from the initiation of treatment to the occurrence of disease progression or death.

DNA extraction

All tumor samples, obtained from the original biopsy

conducted before any treatment, were in the form of formalin-fixed paraffin-embedded (FFPE) specimens with a thickness of 10 μm, and all were confirmed by the pathologist to contain ≥10% tumor cells. Genomic DNA was extracted and isolated from the paraffin-embedded samples by DNA extraction kit, and the relevant information of *KEAP1*, *FGFR* family, *MET*, *PTEN* and *NOTCH2* genes, and *KEAP1-NRF2* pathway (*KEAP1*, *NFE2L2* or *CUL3*) status of each patient was obtained.

CT image acquisition

The contrast-enhanced CT images were exported and stored in DICOM format through the picture archiving and communication system (PACS) imaging system for image segmentation and outlining. The region of interest (ROI) was manually delineated by two experienced radiologists using 3D-Slicer software (Figure S1). All the contouring was performed on the mediastinal window, and the outlined area was the tumor edge and internal features, without peritumoral extension. The tumor areas of blood vessels and trachea were removed during delineation to reduce the impact on the radiomics features of the lesions. Both radiologists had extensive experience in the diagnosis of chest imaging and were unaware of the pathology and clinical data.

Radiomics feature extraction

A total of 851 features were extracted from each patient's manually segmented tumor using Pyradiomics, an open-source software package in 3D-Slicer. There were 851 radiomics features including 14 shape features, 18 first-order statistical features, 75 texture features and 744 higher-order features. The texture features include Gray-Level Co-Occurrence Matrix (GLCM), Gray-Level Dependence Matrix (GLDM), Gray-Level Run-Length Matrix (GLRLM), Gray-Level Size Zone Matrix (GLSZM) and Neighborhood Gray-Tone Difference Matrix (NGTDM). In addition, higher-order features include eight features processed by wavelets, namely, wavelet-Low-Low-Low (wavelet-LLL), wavelet-Low-Low-High (wavelet-LLH), wavelet-Low-High-Low (wavelet-LHL), wavelet-Low-High-High (wavelet-LHH), wavelet-High-Low-Low (wavelet-HLL), wavelet-High-Low-High (wavelet-HLH), wavelet-High-High-Low (wavelet-HHL) and wavelet-High-High-High (wavelet-HHH). The repeatability of these radiomics features was confirmed by repeated ROI delineation.

Screening and analysis of radiomics features

Radiomics feature selection and dimensionality reduction were performed using Python. Patient characteristics were standardized through Z-score normalization to enhance data comparability. To mitigate bias or overfitting resulting caused by too many features, initial features were filtered and dimensionality reduced using interclass correlation coefficient (ICC) and least absolute shrinkage and selection operator (LASSO) regression. ICC was calculated using radiomics features extracted from 20 patients by two radiologists. Two radiologists were unaware of the pathology and clinical data. Features with ICC ≥0.80 were chosen for subsequent analysis. Subsequently, all subjects were classified as validation and training groups at a ratio of 7:3, and the training group was used for feature screening. The best parameter lambda (λ) was determined by a 10-fold cross-validated LASSO model based on minimizing the mean square error (MSE). Subsequently, the most informative features and their respective weighting coefficients were identified for predicting PFS. Based on linear model, the characteristics of each coefficient and the characteristic value multiplication, generate the Rad-score.

Construction and evaluation of survival prediction models

According to the Rad-score, the receiver operating characteristic (ROC) curve for predicting PFS was plotted, and the diagnostic cut-off point with the largest Youden index was used as the best cut-off value of the Rad-score. The survival prediction models were constructed using Cox regression analysis based on the selected clinical factors, radiomics features and genetic characteristics. The area under the curve (AUC) and C-index were used to evaluate the prediction efficiency and accuracy of the model in the training group and the validation group.

Statistical analysis

Statistical analyses were performed by SPSS 25.0 software. The clinical characteristics of patients in the training and validation cohorts were assessed through the Chi-squared test. LASSO regression was performed using Python 3.4 software for feature screening and dimensionality reduction. Univariable and multivariable Cox proportional-hazard models and Kaplan-Meier curves were used for survival analysis. R4.2.1 software was used to calculate the C-index of the model and to plot the ROC curve to evaluate

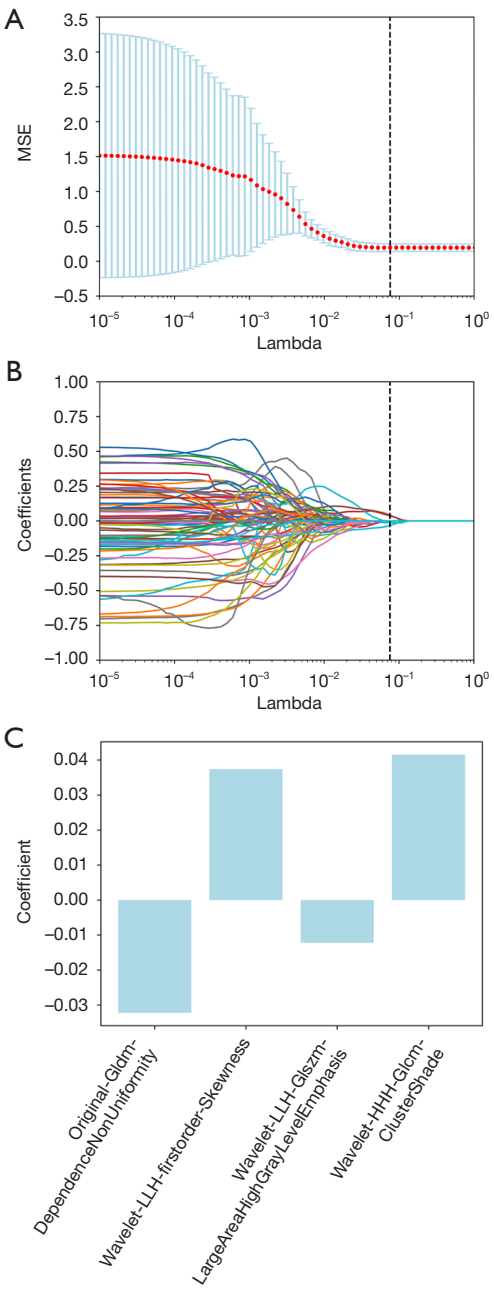

Figure 1 Feature dimension reduction and selection using the LASSO. (A) The cross-validation curve with mean square error on the vertical axis and lambda (λ) on the horizontal axis. The dotted vertical lines represent the optimal values based on both minimum criteria and the 1-SE criterion. The four features with the smallest binomial deviance were selected. (B) Coefficient curves for radiomic features, with the vertical axis illustrating the coefficients of these features and λ on the horizontal axis. (C) Four features selected are presented, with their coefficients indicating their contributions to the model. LASSO, least absolute shrinkage and selection operator; MSE, mean square error.

Table 1 Radiomics features correlated with PFS and their respective coefficients chosen via LASSO regression

Radiomics features	Coefficients
Original-gldm-DependenceNonUniformity	–0.03224029
Wavelet-LLH-firstorder-Skewness	0.03739323
Wavelet-LLH-glszm-LargeAreaHighGrayLevelEmphasis	–0.01219957
Wavelet-HHH-glcm-ClusterShade	0.04151755

PFS, progression-free survival; LASSO, least absolute shrinkage and selection operator.

the predictive ability of the model in a comprehensive manner. All tests in this study were two-sided, and P<0.05 was considered statistically different.

Results

Patient overview

The patients' baseline characteristics in the training and validation groups are summarized in Table S1. The features of the training and validation groups were not significantly different. The median follow-up period for the entire cohort was 30.1 months, during which 79 patients experienced disease progression, and the median PFS was 10.60 months [95% confidence interval (CI): 8.42–12.79].

Construction of Rad-score

A total of 365 image features with ICC ≥0.80 were screened and included as stable feature parameters in the subsequent analysis (Figure S2). In the training group, the LASSO method was used for feature dimension reduction, and the optimal λ value was 0.0756. Four radiomics features related to the PFS were screened out (*Figure 1*). The features and their respective coefficients are shown in *Table 1*. The Rad-score for these features was calculated as follows:

Rad-score = −0.03224029 × original-gldm-DependenceNonUniformity + 0.03739323 × wavelet-LLH-firstorder-Skewness + −0.01219957 × wavelet-LLH-glszm-LargeAreaHighGrayLevelEmphasis + 0.04151755 × wavelet-HHH-glcm-ClusterShade.

Construction of predictive models

For radiomics model, 0.745 was determined as the optimal

Table 2 Univariable and multivariable analysis associated with PFS in the training group

Variables	Univariable analysis		Multivariable analysis	
	HR (95% CI)	P	HR (95% CI)	P
Gender (male *vs.* female)	0.939 (0.401–2.198)	0.88		
Age (<65 *vs.* ≥65 years)	0.458 (0.259–0.808)	0.007	0.553 (0.289–1.059)	0.07
Smoking status (never *vs.* former/current)	1.429 (0.762–2.683)	0.27		
Tumor location				
Left lung	Ref	0.70		
Right lung	1.217 (0.706–2.098)	0.48		
Others	0.692 (0.092–5.193)	0.72		
Histology (ADC *vs.* SCC)	0.821 (0.457–1.474)	0.51		
Clinical stage				
IIIA	Ref	0.19		
IIIB	0.637 (0.367–1.104)	0.11		
IIIC	0.486 (0.146–1.612)	0.24		
Radiation dose (<60 *vs.* ≥60 Gy)	1.043 (0.550–1.980)	0.90		
Chemo radiotherapy (SCRT *vs.* CCRT)	0.856 (0.502–1.461)	0.57		
Radiation therapy (3D-CRT *vs.* IMRT)	0.421 (0.198–0.895)	0.03	0.607 (0.276–1.338)	0.22
KEAP1 (wild *vs.* mutation)	3.123 (1.513–6.445)	0.002	2.843 (1.298–6.225)	0.009
KEAP1-NRF2 pathway (wild *vs.* deletion)	1.957 (0.954–4.012)	0.07		
FGFR family (wild *vs.* mutation)	1.228 (0.677–2.225)	0.50		
MET (wild *vs.* mutation)	4.151 (1.706–10.099)	0.002	4.651 (1.637–13.216)	0.004
NOTCH2 (Wild *vs.* mutation)	1.651 (0.776–3.516)	0.20		
PTEN (wild *vs.* mutation)	2.197 (1.027–4.699)	0.042	2.088 (0.891–4.892)	0.09
Rad-score	2.469 (1.338–4.556)	0.004	2.847 (1.499–5.387)	0.001

PFS, progression-free survival; HR, hazard ratio; CI, confidence interval; ADC, adenocarcinoma; SCC, squamous cell carcinoma; SCRT, sequential chemoradiotherapy; CCRT, concurrent chemoradiotherapy; CRT, chemoradiotherapy; IMRT, intensity modulated radiation therapy.

cut-off value of the Rad-score according to the ROC curve. Therefore, patients with a Rad-score ≥0.745 were categorised as being at high risk of progression whereas patients with a Rad-score <0.745 were categorised as at low risk for disease progression.

The univariable analysis indicated associations between age (<65 *vs.* ≥65 years), radiotherapy method [three-dimensional conformal radiotherapy (3D-CRT) *vs.* intensity modulated radiotherapy (IMRT)], *KEAP1* mutation, *MET* mutation, *PTEN* mutation, and Rad-score with PFS in the training group. These five factors were integrated into the multivariable analysis, revealing *KEAP1* mutation, *MET* mutation, and the Rad-score as independent prognostic indicators for PFS (*Table 2*). For the genomics model, the patients were divided into two groups: low risk group (*MET* wild and *KEAP1* wild) and high risk group (*MET* mutation and/or *KEAP1* mutation).

The radiogenomics model was constructed by combining radiomics model and genomics model. And the patients were divided into three groups of high, medium and low risk of progression: high risk group (Rad-score ≥0.745, *MET* mutation and/or *KEAP1* mutation), medium risk group (Rad-score ≥0.745, *MET* wild and *KEAP1* wild or Rad-score <0.745, *MET* mutation and/or *KEAP1* mutation)

and low risk group (Rad-score <0.745, *MET* wild and *KEAP1* wild).

Relationships between prediction models and PFS

In the radiomics model, there was a significant difference in the prognosis of the two patient groups. Specifically, the median PFS was 8.93 months in the low risk group and 18.53 months in the high risk group (*Figure 2A*, P=0.003). In the validation group, the same outcomes were observed (*Figure 2B*, 7.90 *vs.* 23.77 months, P=0.03).

In the genomics model, the median PFS of patients with high risk in the training group was significantly shorter than the patients with low risk (*Figure 2C*, 4.83 *vs.* 13.17 months, P<0.001). In the validation group, similar results were also observed (*Figure 2D*, 6.73 *vs.* 11.47 months, P=0.01).

In the radiogenomics model, patients with high risk in the training group had significantly shorter median PFS than the medium risk group and the low risk group (*Figure 2E*, 4.60 *vs.* 10.13 months *vs.* NA, P<0.001). The same results were also observed in the validation group (*Figure 2F*, 7.32 *vs.* 8.27 *vs.* 23.77 months, P=0.01).

Comparison of model prediction performance

The ROC curve analysis depicting the performance of the radiomic models in predicting one-year, two-year, and three-year PFS probabilities is depicted in *Figure 3A,3B*. In the training group, the AUC for predicting one-year, two-year, and three-year PFS probabilities were 0.604 (95% CI: 0.492–0.717), 0.668 (95% CI: 0.530–0.806), and 0.871 (95% CI: 0.812–0.930), respectively. The AUC values were 0.671 (95% CI: 0.493–0.849), 0.720 (95% CI: 0.4335–1.005) and 0.637 (95% CI: 0.274–1.000) in the validation group, respectively. The C-index was 0.590 (95% CI: 0.520–0.660) and 0.599 (95% CI: 0.488–0.711) in the training and validation groups.

In the genomics model, the ROC curve in predicting one-year, two-year, and three-year PFS probabilities is depicted in *Figure 3C,3D*. In the training group, the AUC values for predicting one-year, two-year, and three-year PFS probabilities were 0.645 (95% CI: 0.566–0.724), 0.586 (95% CI: 0.500–0.672) and 0.533 (95% CI: 0.373–0.693), respectively. The AUC values were 0.684 (95% CI: 0.574–0.794), 0.641 (95% CI: 0.550–0.732) and 0.641 (95% CI: 0.550–0.732) in the validation group. The C-index was 0.606 (95% CI: 0.548–0.664) and 0.594 (95% CI: 0.510–0.678) in the training and validation groups.

Based on the radiogenomics model, the ROC curve in predicting one-, two-, and three-year PFS probabilities is depicted in *Figure 3E,3F*. In the training group, the AUC of the model for predicting one-year, two-year and three-year PFS were 0.705 (95% CI: 0.605–0.805), 0.715 (95% CI: 0.591–0.839) and 0.845 (95% CI: 0.706–0.984), respectively. The AUC values were 0.776 (95% CI: 0.630–0.921), 0.780 (95% CI: 0.554–1.007) and 0.717 (95% CI: 0.431–1.003) in the validation group. The C-index was 0.663 (95% CI: 0.600–0.726) and 0.650 (95% CI: 0.556–0.43), respectively.

These findings demonstrated that the radiomics model, genomics model, and radiogenomics model exhibit robust predictive efficacy in both the training and validation groups (*Table 3*).

Discussion

In this study, we analyzed the initial CT scans and baseline genetic characteristics of LA-NSCLC patients who received dCRT, and developed three models, including the radiomics model, genomics model and radiogenomics model, to predict the response of LA-NSCLC patients to dCRT. Multivariable analysis based on clinical data, gene mutations and radiomics features showed that MET, KEAP1 mutations and Rad-score were independent predictors of PFS in LA-NSCLC patients after chemoradiotherapy. The prediction of these models was subsequently assessed and confirmed in a validation cohort. The performance of radiogenomics model is better than that of radiomics model and genomics model, which is helpful in predicting the efficacy of chemoradiotherapy in LA-NSCLC patients.

With the development of medical imaging technology and genomics technology, radiogenomics has emerged. Radiogenomics is widely used in the diagnosis and treatment of tumors by correlating patients' imaging data with genomic data. Aonpong *et al.* (30) proposed a genotype-guided imaging genomics method that improved the accuracy of preoperative NSCLC recurrence prediction from 78.61% in the conventional method to 84.39%. Chen *et al.* (31) combined radiomics, genomics and clinical risk factors to construct a nomogram, and it was found that the C-index of the combined model for predicting OS was 0.850 and 0.736 in the training group and test group, respectively, which was superior to the single mode data. Therefore, the combination of radiomics and genomics may improve OS prediction in NSCLC patients. A single-institution retrospective analysis of 124 patients with NSCLC provides evidence that combining clinicopathological models with

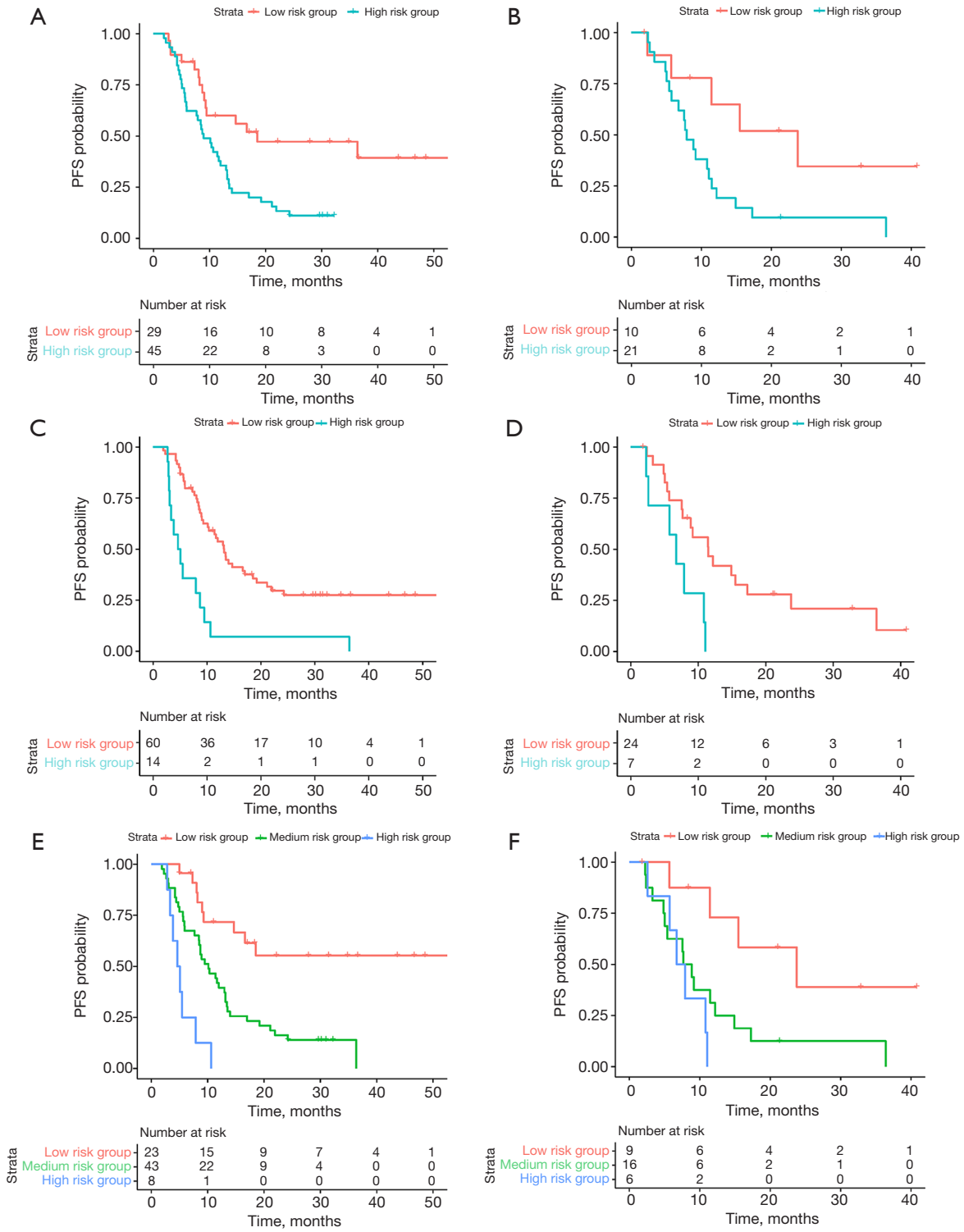

Figure 2 Comparison analysis of Kaplan-Meier survival curves of patients stratified based on radiomics, genomics and radiogenomics models. Applying the radiomics model, PFS curves for patients with the high risk group and low risk group in the training (A) and validation (B) groups. Applying the genomics model, PFS curve for patients with the high risk group and low risk group in the training (C) and validation (D) groups. Applying the integrated model, PFS curves for patients with the high risk group, medium risk group and low risk group in the training (E) and validation (F) groups. PFS, progression-free survival.

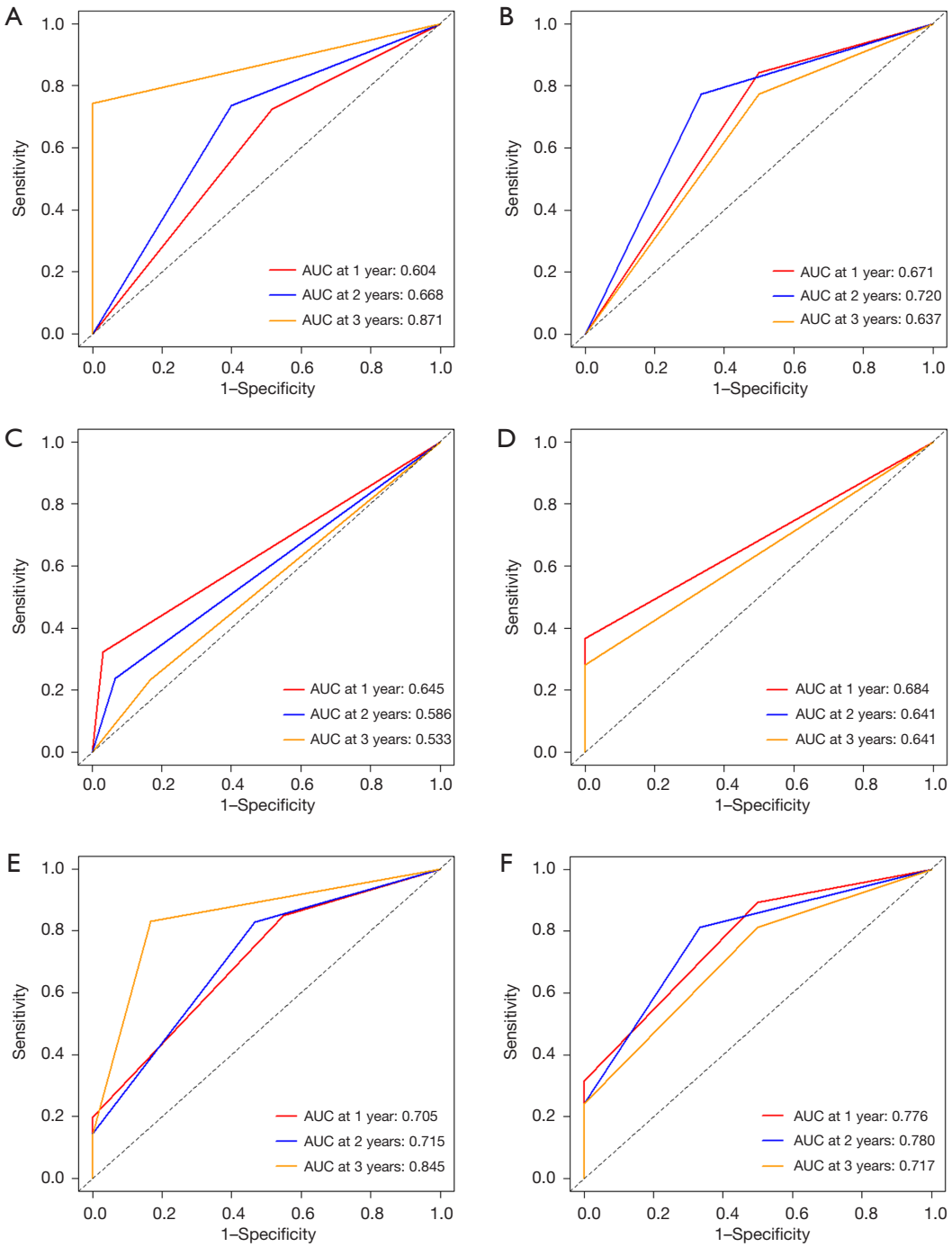

Figure 3 Evaluation of predictive value of radiomics models, genomics models and radiogenomics models for PFS after dCRT in LA-NSCLC patients. Receiver operating characteristic curves of the radiomics models (A,B), genomics models (C,D) and radiogenomics models (E,F) used to evaluate PFS in LA-NSCLC patients after dCRT in the training group and validation group. PFS, progression-free survival; dCRT, definitive chemoradiotherapy; LA-NSCLC, locally advanced non-small cell lung cancer; AUC, area under the curve.

Table 3 Prediction performance of radiomics model, genomics model and radiogenomics model in training group and validation group

Model	AUC (95% CI)			C-index (95% CI)
	One-year PFS	Two-year PFS	Three-year PFS	
Training group				
Radiomics model	0.604 (0.492–0.717)	0.668 (0.530–0.806)	0.871 (0.812–0.930)	0.590 (0.520–0.660)
Genomic model	0.645 (0.566–0.724)	0.586 (0.500–0.672)	0.533 (0.373–0.693)	0.606 (0.548–0.664)
Radiogenomics model	0.705 (0.605–0.805)	0.715 (0.591–0.839)	0.845 (0.706–0.984)	0.663 (0.600–0.726)
Validation group				
Radiomics model	0.671 (0.493–0.849)	0.720 (0.435–1.005)	0.637 (0.274–1.000)	0.599 (0.488–0.711)
Genomic model	0.684 (0.574–0.794)	0.641 (0.550–0.732)	0.641 (0.550–0.732)	0.594 (0.510–0.678)
Radiogenomics model	0.776 (0.630–0.921)	0.780 (0.554–1.007)	0.717 (0.431–1.003)	0.650 (0.556–0.743)

AUC, area under the curve; CI, confidence interval; PFS, progression-free survival.

radiological and genomic characteristics may improve the accuracy of predicting prognosis compared with traditional clinicopathological data (33). The above radiogenomics studies on lung cancer mainly focus on predicting the recurrence and survival of postoperative NSCLC, while our study is the first exploration of a radiogenomics model for predicting the prognosis of LA-NSCLC patients who underwent dCRT. The radiogenomics model in our study demonstrated significantly better predictive performance than the radiomics model (C-index, 0.663 *vs.* 0.590) and genomics model (C-index, 0.663 *vs.* 0.606) in the training group. It is demonstrated that the radiogenomics model has potential strengths in predicting the progression risk of NSCLC patients undergoing dCRT.

Within the realm of lung cancer, the radiomics is developing rapidly, attributed to the widespread availability of chest CT scans and the integration of artificial intelligence (AI) (34). Radiomics refers to extracting a significant number of image features from the ROI of radiological images for analysis and accurate quantitative evaluation of lesions. At present, a number of studies have confirmed the worth of radiomics to predict the prognosis of LA-NSCLC patients treated with radiotherapy. Zhang *et al.* (17) integrated the imaging features of the combined primary tumor and lymph nodes and demonstrated that 1441 were significantly better than conventional imaging features (C-index: 0.77–0.79 *vs.* 0.53–0.73). Chen *et al.* (35) divided LA-NSCLC patients undergoing definitive CCRT into the high risk group and low risk group. The low risk group achieved a significantly higher 3-year OS rate (68.4% *vs.* 3.3%, P<0.001), and may benefit more from dCRT and further adjuvant immunotherapy. Chen *et al.* (36) found that a nomogram combining radiomics features and clinical prognostic factors for predicting locoregional failure (LRF) in LA-NSCLC treated with dCRT showed good performance in testing and validation, with C-indexes of 0.796 and 0.756. With research development, the seminal PACIFIC trial, which investigated durvalumab revealed a substantial extension of PFS among those who received durvalumab following CRT. Therefore, a subsequent study has found that tumor radiomics on pre-treatment CT images is a predictor of PFS and OS in patients with unresectable stage III NSCLC treated with CRT sequential durvalumab and CRT alone (37). In view of the above findings, the baseline CT images of LA-NSCLC patients before dCRT treatment were also included in our study for radiomics analysis. We found that the radiomic model can predict the PFS of LA-NSCLC patients receiving dCRT, with the C-index of 0.590. The predictive ability of the model was further verified in the validation group, where the C-index was 0.599. Therefore, our research supports existing evidence indicating the potential value of radiomics to predict the progression of LA-NSCLC receiving dCRT.

One study, unlike the above study, could not confirm that radiomics features extracted by cone beam CT contributed to improved prognostic information (38). This may be due to poor performance of baseline radiomics features or clinical parameters, affecting within-dataset and between-dataset heterogeneity. Besides, early identification of important driver gene mutations is essential to the formation of therapeutic strategies for NSCLC patients. Therefore, the introduction of genomic models can help to better

predict survival. Previous genomics models have confirmed established correlations between genetic mutations and the prognosis of NSCLC patients. Jiao *et al.* (39) found that *TP53* mutations serve as a negative prognostic factor in 1,441 metastatic NSCLC patients, and when combined with *EGFR* mutations, they can more accurately predict the prognosis of advanced NSCLC patients. A German multicenter retrospective study observed that truncating *TP53* mutations and *KEAP1* mutations were related to shorter disease-free survival (DFS) in the 1,518 stage I–IIIA patients undergoing surgical treatment (40). In 4,779 stage IIIB–IV patients, *TP53* mutations were not predictive of prognosis, while *KEAP1* mutations were associated with prognosis. Additionally, tumors with *KEAP1* mutations and co-occurring *TP53* missense mutations demonstrated longer OS compared to *KEAP1* mutant tumors with wildtype or truncating *TP53* mutations. In summary, *TP53* and *KEAP1* mutations serve as prognostic factors in NSCLC. Similarly, researchers previously investigated the clinical application value of large panel in precision radiotherapy using real-world clinical samples and found that the *Keap1-NRF2* pathway, *KEAP1*, *FGFR* family, *MET*, *PTEN*, *NOTCH2* mutations can independently predict the risk of recurrence after dCRT, which is an important guideline for clinical screening of patients sensitive to radiotherapy and chemotherapy (41). In this study, we included the results of the previous study and on this basis selected *KEAP1* and *MET* mutations by Cox analysis to construct the genomics model. The genomic model demonstrated good performance in predicting the PFS of LA-NSCLC, with the C-index of 0.606. Furthermore, this conclusion was also confirmed in the validation group with a C-index of 0.594. Therefore, the addition of genomics could further improve the reliability and accuracy of the radiomics model.

Despite the advantages of radiogenomics in prognostic prediction of lung cancer, there are still many obstacles limiting its use in clinical practice, with reproducibility standing as the chief challenge. Reproducibility can be affected by a multitude of factors, including variations in scanning devices (42,43), tumor segmentation processes (44,45), and acquisition protocols (43,46). These variations may subsequently impact the radiomic prediction models (47). In addition, the study was carried out with a relatively small patient sample, which constrained our capacity for performing extensive analyses. Further multi-center and large-sample prospective studies may be an important direction for imaging genomics research.

Conclusions

In summary, this study investigated the feasibility of radiomics, genomics, and radiogenomics models to predict PFS in LA-NSCLC patients undergoing dCRT. The CT-based radiomics model, the genomic model and the radiogenomics model can all predict the prognosis of dCRT in LA-NSCLC, and the radiogenomics model is superior to the single data type model. This indicates that the radiogenomics model can be used as a non-invasive and accurate pre-treatment assessment tool for LA-NSCLC patients, and that better identification of NSCLC patients at high risk of progression can assist clinicians to optimize individualized treatment regimens.

Acknowledgments

Funding: This study was supported in part by the National Natural Science Foundation of China (No. NSFC82073345), Natural Science Innovation and Development Joint Foundation of Shandong Province (No. ZR202209010002), Jinan Clinical Medicine Science and Technology Innovation Plan (No. 202019060) and the Taishan Scholars Program to S.Y.; and the Natural Science Youth Foundation of Shandong Province (No. ZR2023QH155) to L.L..

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-145/rc

Data Sharing Statement: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-145/dss

Peer Review File: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-145/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-145/coif). L.L. reports that this study was supported by the Natural Science Youth Foundation of Shandong Province (No. ZR2023QH155). S.Y. reports that this study was supported by the National Natural Science Foundation of China (No. NSFC82073345), Natural Science Innovation and Development Joint Foundation of Shandong Province

(No. ZR202209010002), Jinan Clinical Medicine Science and Technology Innovation Plan (No. 202019060) and the Taishan Scholars Program. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Our investigation was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was endorsed by the ethical review committee of Shandong Cancer Hospital and Institute (No. SDTHEC202004042), which waived the requirement for informed consent owing to the retrospective nature of the study.

References

1. Spigel DR, Faivre-Finn C, Gray JE, et al. Five-Year Survival Outcomes From the PACIFIC Trial: Durvalumab After Chemoradiotherapy in Stage III Non-Small-Cell Lung Cancer. J Clin Oncol 2022;40:1301-11.
2. Duan X, Zhao Q, Li F. Association of CT features with TNM stage and pathology of patients with rectal cancer and their significance in evaluation of efficacy and prognosis. J BUON 2020;25:1430-5.
3. Jagoda P, Fleckenstein J, Sonnhoff M, et al. Diffusion-weighted MRI improves response assessment after definitive radiotherapy in patients with NSCLC. Cancer Imaging 2021;21:15.
4. Groheux D, Hindié E, Delord M, et al. Prognostic impact of (18)FDG-PET-CT findings in clinical stage III and IIB breast cancer. J Natl Cancer Inst 2012;104:1879-87.
5. Dunlap NE, Yang W, McIntosh A, et al. Computed tomography-based anatomic assessment overestimates local tumor recurrence in patients with mass-like consolidation after stereotactic body radiotherapy for early-stage non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2012;84:1071-7.
6. Cochet A, Dygai-Cochet I, Riedinger JM, et al. ^{18}F-FDG PET/CT provides powerful prognostic stratification in the primary staging of large breast cancer when compared with conventional explorations. Eur J Nucl Med Mol Imaging 2014;41:428-37.
7. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology 2020;295:328-38.
8. Beig N, Khorrami M, Alilou M, et al. Perinodular and Intranodular Radiomic Features on Lung CT Images Distinguish Adenocarcinomas from Granulomas. Radiology 2019;290:783-92.
9. Yang X, He J, Wang J, et al. CT-based radiomics signature for differentiating solitary granulomatous nodules from solid lung adenocarcinoma. Lung Cancer 2018;125:109-14.
10. Jing R, Wang J, Li J, et al. A wavelet features derived radiomics nomogram for prediction of malignant and benign early-stage lung nodules. Sci Rep 2021;11:22330.
11. Wu G, Woodruff HC, Shen J, et al. Diagnosis of Invasive Lung Adenocarcinoma Based on Chest CT Radiomic Features of Part-Solid Pulmonary Nodules: A Multicenter Study. Radiology 2020;297:451-8.
12. Zhu X, Dong D, Chen Z, et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. Eur Radiol 2018;28:2772-8.
13. Hindocha S, Charlton TG, Linton-Reid K, et al. Gross tumour volume radiomics for prognostication of recurrence & death following radical radiotherapy for NSCLC. NPJ Precis Oncol 2022;6:77. Erratum in: NPJ Precis Oncol 2022;6:87.
14. Li S, Yang N, Li B, et al. A pilot study using kernelled support tensor machine for distant failure prediction in lung SBRT. Med Image Anal 2018;50:106-16.
15. Yang H, Wang L, Shao G, et al. A combined predictive model based on radiomics features and clinical factors for disease progression in early-stage non-small cell lung cancer treated with stereotactic ablative radiotherapy. Front Oncol 2022;12:967360.
16. Jiao Z, Li H, Xiao Y, et al. Integration of Deep Learning Radiomics and Counts of Circulating Tumor Cells Improves Prediction of Outcomes of Early Stage NSCLC Patients Treated With Stereotactic Body Radiation Therapy. Int J Radiat Oncol Biol Phys 2022;112:1045-54.
17. Zhang N, Liang R, Gensheimer MF, et al. Early response evaluation using primary tumor and nodal imaging features

to predict progression-free survival of locally advanced non-small cell lung cancer. Theranostics 2020;10:11707-18.
18. Liu Y, Qi H, Wang C, et al. Predicting Chemo-Radiotherapy Sensitivity With Concordant Survival Benefit in Non-Small Cell Lung Cancer via Computed Tomography Derived Radiomic Features. Front Oncol 2022;12:832343.
19. Meng X, Xu H, Liang Y, et al. Enhanced CT-based radiomics model to predict natural killer cell infiltration and clinical prognosis in non-small cell lung cancer. Front Immunol 2023;14:1334886.
20. Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. Eur Respir J 2019;53:1800986.
21. Baranyi M, Buday L, Hegedűs B. K-Ras prenylation as a potential anticancer target. Cancer Metastasis Rev 2020;39:1127-41.
22. Song L, Zhu Z, Mao L, et al. Clinical, Conventional CT and Radiomic Feature-Based Machine Learning Models for Predicting ALK Rearrangement Status in Lung Adenocarcinoma Patients. Front Oncol 2020;10:369.
23. Mahajan A, Kania V, Agarwal U, et al. Deep-Learning-Based Predictive Imaging Biomarker Model for EGFR Mutation Status in Non-Small Cell Lung Cancer from CT Imaging. Cancers (Basel) 2024;16:1130.
24. Kohan A, Hinzpeter R, Kulanthaivelu R, et al. Contrast Enhanced CT Radiogenomics in a Retrospective NSCLC Cohort: Models, Attempted Validation of a Published Model and the Relevance of the Clinical Context. Acad Radiol 2024;31:2953-61.
25. Zuo S, Wei M, Zhang H, et al. A robust six-gene prognostic signature for prediction of both disease-free and overall survival in non-small cell lung cancer. J Transl Med 2019;17:152.
26. Li J, Wang H, Li Z, et al. A 5-Gene Signature Is Closely Related to Tumor Immune Microenvironment and Predicts the Prognosis of Patients with Non-Small Cell Lung Cancer. Biomed Res Int 2020;2020:2147397.
27. Gray JE, Ahn MJ, Oxnard GR, et al. Early Clearance of Plasma Epidermal Growth Factor Receptor Mutations as a Predictor of Outcome on Osimertinib in Advanced Non-Small Cell Lung Cancer; Exploratory Analysis from AURA3 and FLAURA. Clin Cancer Res 2023;29:3340-51.
28. Boehm KM, Khosravi P, Vanguri R, et al. Harnessing multimodal data integration to advance precision oncology. Nat Rev Cancer 2022;22:114-26.
29. Kuo MD, Jamshidi N. Behind the numbers: Decoding molecular phenotypes with radiogenomics--guiding principles and technical considerations. Radiology 2014;270:320-5.
30. Aonpong P, Iwamoto Y, Han XH, et al. Improved Genotype-Guided Deep Radiomics Signatures for Recurrence Prediction of Non-Small Cell Lung Cancer. Annu Int Conf IEEE Eng Med Biol Soc 2021;2021:3561-4.
31. Chen W, Qiao X, Yin S, et al. Integrating Radiomics with Genomics for Non-Small Cell Lung Cancer Survival Analysis. J Oncol 2022;2022:5131170.
32. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228-47.
33. Kim KH, Kim J, Park H, et al. Parallel comparison and combining effect of radiomic and emerging genomic data for prognostic stratification of non-small cell lung carcinoma patients. Thorac Cancer 2020;11:2542-51.
34. Huynh E, Hosny A, Guthier C, et al. Artificial intelligence in radiation oncology. Nat Rev Clin Oncol 2020;17:771-81.
35. Chen NB, Xiong M, Zhou R, et al. CT radiomics-based long-term survival prediction for locally advanced non-small cell lung cancer patients treated with concurrent chemoradiotherapy using features from tumor and tumor organismal environment. Radiat Oncol 2022;17:184.
36. Chen X, Tong X, Qiu Q, et al. Radiomics Nomogram for Predicting Locoregional Failure in Locally Advanced Non-small Cell Lung Cancer Treated with Definitive Chemoradiotherapy. Acad Radiol 2022;29 Suppl 2:S53-61.
37. Jazieh K, Khorrami M, Saad A, et al. Novel imaging biomarkers predict outcomes in stage III unresectable non-small cell lung cancer treated with chemoradiation and durvalumab. J Immunother Cancer 2022;10:e003778.
38. van Timmeren JE, van Elmpt W, Leijenaar RTH, et al. Longitudinal radiomics of cone-beam CT images from non-small cell lung cancer patients: Evaluation of the added prognostic value for overall survival and locoregional recurrence. Radiother Oncol 2019;136:78-85.
39. Jiao XD, Qin BD, You P, et al. The prognostic value of TP53 and its correlation with EGFR mutation in advanced non-small cell lung cancer, an analysis based on cBioPortal data base. Lung Cancer 2018;123:70-5.
40. Saleh MM, Scheffler M, Merkelbach-Bruse S, et al. Comprehensive Analysis of TP53 and KEAP1 Mutations and Their Impact on Survival in Localized- and Advanced-Stage NSCLC. J Thorac Oncol 2022;17:76-88.
41. He K, Zhang S, Pang J, et al. Genomic Profiling Reveals Novel Predictive Biomarkers for Chemo-Radiotherapy Efficacy and Thoracic Toxicity in Non-Small-Cell Lung

Cancer. Front Oncol 2022;12:928605.
42. Mackin D, Fave X, Zhang L, et al. Measuring Computed Tomography Scanner Variability of Radiomics Features. Invest Radiol 2015;50:757-65.
43. Berenguer R, Pastor-Juan MDR, Canales-Vázquez J, et al. Radiomics of CT Features May Be Nonreproducible and Redundant: Influence of CT Acquisition Parameters. Radiology 2018;288:407-15.
44. van Velden FH, Kramer GM, Frings V, et al. Repeatability of Radiomic Features in Non-Small-Cell Lung Cancer [(18)F]FDG-PET/CT Studies: Impact of Reconstruction and Delineation. Mol Imaging Biol 2016;18:788-95.
45. Pfaehler E, Beukinga RJ, de Jong JR, et al. Repeatability of (18) F-FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method. Med Phys 2019;46:665-78.
46. Lecler A, Duron L, Balvay D, et al. Combining Multiple Magnetic Resonance Imaging Sequences Provides Independent Reproducible Radiomics Features. Sci Rep 2019;9:2068.
47. He L, Huang Y, Ma Z, et al. Effects of contrast-enhancement, reconstruction slice thickness and convolution kernel on the diagnostic performance of radiomics signature in solitary pulmonary nodule. Sci Rep 2016;6:34921.

Cite this article as: Song X, Li L, Yu Q, Liu N, Zhu S, Yuan S. Radiogenomics models for predicting prognosis in locally advanced non-small cell lung cancer patients undergoing definitive chemoradiotherapy. Transl Lung Cancer Res 2024;13(8):1828-1840. doi: 10.21037/tlcr-24-145