



Original Research Article

CT-based different regions of interest radiomics analysis for acute radiation pneumonitis in patients with locally advanced NSCLC after chemoradiotherapy

Liqiao Hou^{a,1}, Kuifei Chen^{a,1}, Chao Zhou^{b,1}, Xingni Tang^a, Changhui Yu^a, Haijian Jia^b, Qianyi Xu^{c,*}, Suna Zhou^{a,*}, Haihua Yang^{a,*}

^a Key Laboratory of Radiation Oncology of Taizhou, Radiation Oncology Institute of Enze Medical Health Academy, Department of Radiation Oncology, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, NO.150 Ximen Street, Linhai, Taizhou City, 317000, Zhejiang Province, China

^b Department of Radiation Oncology, Enze Hospital Affiliated Hospital of Hangzhou Medical College, Zhejiang Province 317000, China

^c Department of Radiation Oncology, Thomas Jefferson Health System, 8081 Innovation Park Dr., Fairfax, VA, 22003, USA

ARTICLE INFO

Keywords:

Non-small cell cancer

Radiomics

Computed tomography simulated

Radiation pneumonitis

ABSTRACT

Purpose: To establish a radiomics model using radiomics features from different region of interests (ROI) based on dosimetry-related regions in enhanced computed tomography (CT) simulated images to predict radiation pneumonitis (RP) in patients with non-small cell lung cancer (NSCLC).

Methods: Our retrospective study was conducted based on a cohort of 236 NSCLC patients (59 of them with $RP \geq 2$) who were treated in 2 institutions and divided into the primary cohort ($n = 182,46$ of them with $RP \geq 2$) and external validation cohort ($n = 54,13$ of them with $RP \geq 2$). Radiomic features extracted from three ROIs were defined as the whole lung (WL), the dose volume histogram (DVH) of the lung V20 (V20 Lung) and the DVH of the V30 of lung minus the planning target volume (PTV) (V30 Lung-PTV). A total of 107 radiomics features were extracted from each ROIs. The *U* test, correlation coefficient and least absolute shrinkage and selection operator (LASSO) were performed for features selection. Six models based on different classification algorithms were developed to select the best radiomics model (R model). In addition, we built a dosimetry model then combined it with the best R model to create a mixed model (R+D model). The receiver operating characteristic (ROC) curve was delineated to assess the predictive efficacy of the models. Decision curve analysis could benefit from the model proposals through the assessment of clinical utility.

Results: Among the three ROIs, the best R model constructed from the LightGBM algorithm demonstrated the strongest discriminative ability in the ROI of V30 Lung-PTV. The corresponding area under the curve (AUC) value was 0.930 (95 % confidence interval (CI): 0.829–0.941). The D model, R model and R+D model achieved AUC values of 0.798 (95 % CI: 0.732–0.865), 0.930 (95 % CI: 0.829–0.941) and 0.940 (95 % CI: 0.906–0.974) in primary cohort, and in external validation cohort, the AUC values were 0.793 (95 % CI: 0.637–0.949), 0.887 (95 % CI: 0.810–0.993), 0.951 (95 % CI: 0.891–1.000). Decision curve demonstrate that R+D model could benefit for patients through the assessment of clinical utility.

Conclusion: The radiomics model was able to predict the acute RP more effectively in comparison with the traditional dosimetry model. Especially the radiomics model based on the V30 Lung-PTV region was able to achieve a higher accuracy when compared to the other regions.

Introduction

Lung cancer is the leading cause of cancer-related death in the world,

with an estimated 1.6 million deaths annually [1]. About 85 % of the lung cancer patients have a group of histological subtypes collectively known as non-small cell lung cancer (NSCLC) [2]. Chemoradiotherapy

* Corresponding authors.

E-mail addresses: xuqianyi@gmail.com (Q. Xu), annyzhou0913@163.com (S. Zhou), yhh93181@hotmail.com (H. Yang).

¹ These authors contribute equally to this work.

² Haihua Yang and Suna Zhou and Qianyi Xu contributed equally to this article.

<https://doi.org/10.1016/j.ctro.2024.100828>

Received 5 January 2024; Received in revised form 26 June 2024; Accepted 27 July 2024

Available online 31 July 2024

2405-6308/© 2024 The Author(s). Published by Elsevier B.V. on behalf of European Society for Radiotherapy and Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

has been a standard treatment approach playing an important role in the management of NSCLC [3]. However, during radiotherapy the healthy lungs were inevitably irradiated and radiation pneumonitis (RP) could be developed in consequence [4]. In thoracic chemoradiotherapy, the side effects of RP significantly reduced the effectiveness of treatment including a negative impact on quality of life, and in severe cases, could lead to death [5,6]. Therefore, the development of easy-to-use tools for the timely and effective detection of RP in patients with NSCLC after chemoradiotherapy could provide an optimal clinical decision making and the best personalized treatment option.

Radiomics, an emerging technology, has undergone rapid development in recent years. It is concerned with converting medical images into highly-specific high-throughput statistical features, known as radiomics feature, can provide additional information for therapeutic strategies [7,8]. Recent studies suggest that radiomics features is clinically useful in predicting RP [9–11]. Nevertheless, studies have shown that different regions of interest (ROIs) may have different roles in the prediction of RP. Zhen Zhang et al has been emphasised that the radiomics features based on whole lung, in combination with clinical and dosimetric risk factors, can be effective in predicting RP. The areas under the curve (AUC) were 0.793, 0.774, 0.855 in the training, bootstrapping and external test sets [9]. In other study, it was reported that the radiomics features of PTV-GTV(exclude gross tumor volume from planning tumor volume) and TL-PTV(exclude planning tumor volume from the total bilateral lung) outperformed those of the other ROIs. The accuracy were 76.7 % and 76.7 % [10]. Although the ability of different ROIs to predict RP has been demonstrated, however, since dosimetry has been shown to correlate with development of RP, we should focus not only on tumor-related regions but also on dosimetry-related regions. According to the quantitative analyses of normal tissue effects in the clinic(QUANTEC), on three-dimensional conformal radiation therapy (3D-CRT), the cutoff values of lung dose constraints to be the bilateral lung volume exceeding 20 Gy($V_{20} \leq 35\%$) and mean lung dose(MLD) ≤ 20 Gy [12]. However, as intensity modulated radiotherapy(IMRT) and volumetric modulated arc therapy(VMAT) become more and more mature, Meng et al proposed that the bilateral lung volume outside planning tumor volume(PTV) exceeding 30 Gy (V_{30} -PTV) is superior to predict RP [13].

In this study, we first constructed a radiomics model (R model) for the prediction of acute RP based on radiomics features within three ROIs (whole lung, V_{20} Lung, and V_{30} Lung-PTV) in enhanced CT-simulation image. We employed machine learning algorithms to determine the best ROI and features to enhance prediction accuracy, and computed their respective proportions in the prediction model. Subsequently, we present a dosimetry model (D model) for predicting RP based on conventional dosimetry. Consequently, a dosimetry-radiomics model (R+D model) was constructed to further improve the accuracy of predicting RP. In the era of precision medicine, the use of these convenient and clinically applicable tools will allow more rational and timely adoption of proactively tailored follow-up strategies and interventions for NSCLC patients undergoing chemoradiotherapy with different RP risk levels.

Methods and materials

Patient recruiting criteria

This study was approved by the Institutional Review Board. From January 2014 to December 2021, A total of 182 NSCLC patients treated in our institution were retrospectively recruited as the primary cohort to develop models to predict RP. Meanwhile, a cohort of 54 NSCLC patients treated in another institution were enrolled for external validation of the propose models. The patient eligibility criteria were as the follows: 1) confirmed pathology of non-small cell lung cancer with stage III; 2) first time received thoracic chemoradiotherapy; 3) The radiation schedule was 64 to 66 Gy with 2.14 to 2.20 Gy per fraction to planning gross tumor volume (PGTV) and 53–54 Gy to planning treatment volume

(PTV) in 30 fractions. If the dose to organs at risk (OAR) was above the limits, a total of 54 Gy to PGTV and 45 Gy to PTV in 25 fractions could be given instead; 4) followed up for at least three months; The main exclusion criteria were: 1) other pathological types of lung cancer; 2) incomplete or non-standard chemoradiotherapy treatment; 3) loss of follow-up data; 4) patients treated with immunotherapy. The overall pipeline of the study was shown as Fig. 1.

Clinical data and dosimetry parameters

The clinical data were collected from the electronic medical record system including age, gender, smoke-state, pathology and chronic obstructive pulmonary disease (COPD). All the patients underwent a free-breathing scan with enhanced contrast for treatment planning. The dosimetry parameters were analyzed from plan DVH in treatment planning system (TPS). The lung V_{20} , the V_{30} of lung minus PTV and MLD were extracted to build dosimetry model. It's worth mentioning that Meng et al. demonstrated V_{30} lung-PTV was superior to predict RP than other parameters for lung cancer patients treated using IMRT technique [13]. That's why in this study, we also included V_{30} Lung – PTV in the dosimetry study.

Radiomic feature extraction

The process of radiomic features extraction was described as follows and more details could be found in Method 1 of the [supplementary material](#). As mentioned earlier, the ROIs were defined as the whole lung (WL), V_{20} Lung, V_{30} Lung-PTV, a example was given in Fig. 1 of [Supplementary material](#). A total of 107 radiomic features (RFs) were extracted for each ROI in the Pyradiomics package (version 3.0.1, Python version 3.7.6) [14], following the Image Biomarker Standardization Initiative guidelines [15]. The methods used for RFs extraction included *U* test, correlation coefficient and least absolute shrinkage and selection operator (LASSO). More details on the RFs extraction could be found in Method 2 of the [Supplementary material](#).

Models construction and validation

Once, the RFs were extracted, six machine leaning approaches were modeled in different ROIs to predict the occurrence of RR, including Support Vector Machines (SVM), Logistic Regression (LR), Light Gradient Boosting Machine (Light GBM), Naïve Bayes, Adaptive Boosting (AdaBoosting) and Multi-Layer Perception (MLP) neural network. The details of the algorithms could be found in Method 3 of the [supplementary material](#). For each ROI, six models were built with a supervised task based on the RP and non-RP label. The RP label was defined as graded two or higher RP and non-RP label was defined as no RP or grade one RP. The grading of RP was based on the National Cancer Institute's Common Terminology Criteria for Adverse Events (version 4.03). Models were built by each algorithm in the primary cohort of patient and validated by the external cohort of patients. The goal was to find the best combination of machine learning algorithm and ROI with the most powerful prediction capability. Once the combination was determined, it would be further combined with the dosimetry model to build a hybrid model (R+D model). The receiver operating characteristic (ROC) curve and decision curve were adopted to visually demonstrate the predictive ability of the models. The diagnostic indices of the WH, V_{20} Lung, V_{30} Lung-PTV radiomics models including area under the ROC curve (AUC, with the 95 % CI), specificity, sensitivity, accuracy, positive predictive value (PPV) and the negative predictive value (NPV) were also calculated. The radiomic models were evaluated in the scikit-learning package (version 0.18 with Python version 3.7.6).

Statistical analysis

The baseline patient data were analyzed in the statistical packages of

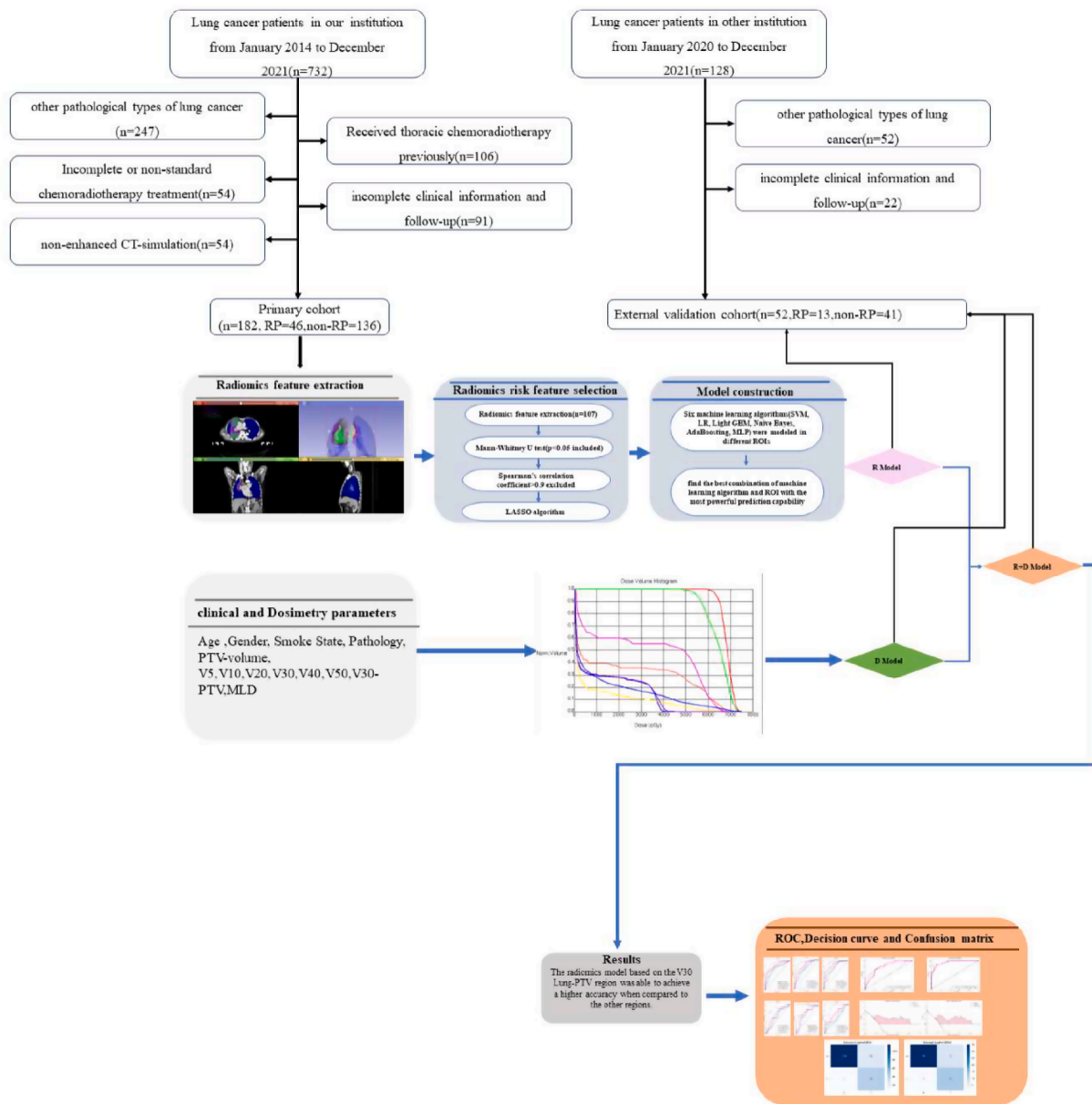


Fig. 1. The flowchart of this study.

SPSS (version 23.0) and Python (version 3.7.6). The continuous variables were expressed as mean ± standard deviation (std) and the Student's *t*-test or Mann-Whitney *U* test was used to compare their inter-group differences. The categorical variables were expressed as frequencies and percentages and the Chi-squared test was used to compare their difference. A two-tailed *p* test was used for statistical analysis and a *p* value of < 0.05 indicated statistical significance. The DeLong testing method was used to compare the AUCs in differences models.

Results

Baseline patient characteristics

During follow-up, 46 (25.3 %) patients in the primary cohort and 13 (24.1 %) patients in the validation cohort were diagnosed with RP. In the primary cohort, the gender, PTV volume, V30 Lung-PTV and MLD have the significant differences between the RP and non-RP patient groups (*p* values were 0.032,0.026,0.001). No significant differences were observed in the clinical and dosimetric parameters between the RP

and non-RP groups in the validation cohort. The detailed characteristics of the study patients were summarized in Table 1.

Feature extraction and selection

A total of 107 radiomics features were extracted for each ROI. The Mann-Whitney *U* test showed that 9 RFs in WL, 20 RFs in V20_Lung and 12 RFs in V30 Lung-PTV were significantly associated with RP and Non-RP groups. The Spearman's correlation coefficient method was used to further remove the linear correlation between the RFs and to reduce the dimension of the RFs. In consequence, the numbers RFs were reduced to 6 in WL, 11 in V20_Lung and 9 in V30 Lung-PTV. The LASSO algorithm was used to determine the final optimal numbers of RFs for each ROI (4 in WL, 6 in V20_Lung and 5 in V30 Lung-PTV). The features selection process by LASSO were showed in Fig. 2 (a, b, c). All the selected RFs demonstrated significant differences (*p* < 0.05) using the Mann-Whitney *U* test between the RP and Non-RP groups in the primary cohort (Supplement Table 1). The Spearman correlation coefficient map showed that all the selected RFs were not correlated to the other features (the

Table 1
Patients characteristics in primary cohort and external validation cohort.

Characteristics	Primary cohort (N=182)		P	Validation cohort (N=54)		P	P
	Non-RP2 (n = 136)	RP2 (N=46)		Non-RP2 (N=41)	RP2 (N=13)		
Age (Year)			0.462			0.555	0.163
Mean ± SD	63.5 ± 8.3	64.6 ± 8.2		66.2 ± 7.1	64.6 ± 8.6		
Gender			0.032			1.000	1.000
Female	13(9.6 %)	10(21.7 %)		5(12.2 %)	2(15.4 %)		
Male	123(90.4 %)	36(78.3 %)		36(87.8 %)	11(84.6 %)		
Smoke state			0.199			0.961	0.308
Non-smoke	34(25.0 %)	16(34.8 %)		15(36.6 %)	4(30.8 %)		
Smoke	102(75.0 %)	30(65.2 %)		26(63.4 %)	9(69.2 %)		
Pathology			0.502			0.468	0.390
Squamous	68(50.0 %)	23(50.0 %)		13(31.7 %)	3(23.1 %)		
Adenocarci-noma	41(30.1 %)	17(37.0 %)		26(63.4 %)	8(61.5 %)		
Other	27(19.9 %)	6(13.0 %)		2(4.9 %)	2(15.4 %)		
PTV_volume (cc)			0.026			0.448	0<.001
Mean ± SD	546.9 ± 285.5	451.9 ± 211.5		357.3 ± 195.0	403.0 ± 162.3		
V30-PTV (%)			0.001			0.429	0.117
Mean ± SD	11.6 ± 2.3	12.9 ± 2.4		11.8 ± 3.1	12.5 ± 2.2		
V5(%)			0.989			0.367	0.253
Mean ± SD	48.0 ± 10.5	48.1 ± 13.2		40.5 ± 12.4	44.0 ± 11.1		
V10(%)			0.558			0.384	0.175
Mean ± SD	34.8 ± 7.1	35.7 ± 9.2		30.4 ± 8.7	32.8 ± 7.2		
V20(%)			0.262			0.318	0.138
Mean ± SD	24.30 ± 3.5	22.30 ± 3.7		20.4 ± 4.9	21.9 ± 3.3		
V30(%)			0.347			0.634	0.136
Mean ± SD	116.3 ± 3.0	16.8 ± 2.9		14.9 ± 4.0	15.5 ± 2.5		
V40(%)			0.198			0.906	0.535
Mean ± SD	11.0 ± 3.6	11.8 ± 3.2		10.9 ± 3.6	11.0 ± 2.5		
V50(%)			0.163			0.831	0.104
Mean ± SD	5.5 ± 3.6	6.4 ± 3.5		6.5 ± 3.2	6.7 ± 2.3		
MLD(cGy)			0.002			0.434	0<.001
Mean ± SD	295.5 ± 524.2	621.8 ± 633.4		1152.8 ± 302.9	1224.0 ± 206.5		
Lung volume (cc)			0.820			0.865	0.567
Mean ± SD	3515.3 ± 957.9	3553.3 ± 1037.2		3505.7 ± 1068.5	3449.3 ± 920.4		
COPD			0.781			1.000	0.806
116(85.3 %)	40(87.0 %)		36(87.8 %)	11(87.0 %)			
20(14.7 %)	6(13.0 %)		5(12.2 %)	2(13.0 %)			

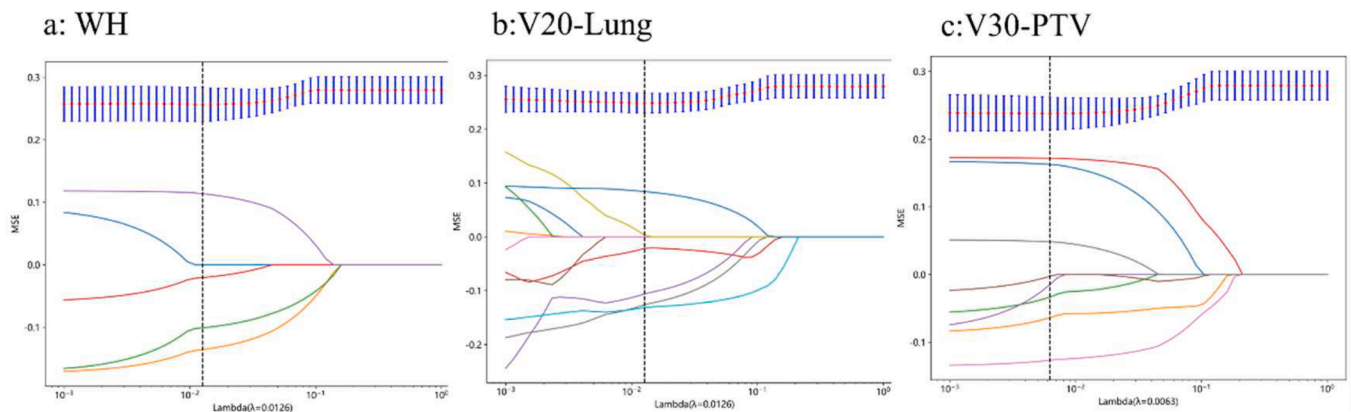


Fig. 2. Texture features were selected using the least absolute shrinkage and selection operator (LASSO) regression model. A 5-fold cross validation was used to select the best parameter (lambda) in the LASSO model.

correlation coefficient values was ± 0.9 (Fig. 2 in Supplementary material) [16].

Model performance

For each ROI, six robust supervised models were developed to predict RP and their performances were compared to determine the most optimal model. The prediction performance was quantified as AUC and summarized for all the models in Fig. 3. The other diagnostic indices of the models were provided in the Table 2 in the Supplementary material. We found the model based on the LightGBM algorithm outperformed the other models in all ROIs. In the primary cohort, the corresponding model yielded AUCs of 0.898 (95 %CI: 0.849–0.946), 0.914 (95 %CI:

0.866–0.962) and 0.930 (95 %CI:0.889–0.972) and in the external validation cohort, the model yielded AUCs of 0.846 (95 %CI: 0.711–0.982), 0.835 (95 %CI: 0.702–0.968), 0.887 (95 %CI: 0.799–0.976) in WL, V20_Lung and V30 Lung-PTV, respectively.

For dosimetry model, the dosimetry parameters were comprised of V20 of the Lung, MLD and V30 of the Lung-PTV. The AUCs of the dosimetry model were 0.798 (95 %CI: 0.732–0.865) in the primary cohort and 0.793 (95 %CI: 0.637–0.949) in the external validation cohort, both of which were lower than the optimal radiomics model (the ROCs was shown in Fig. 4(a)).

Since the LightGBM algorithm showed the best predictive performance, the same algorithm was associated with the dosimetry model to predict RP (R+D model). It turned out that the R+D model

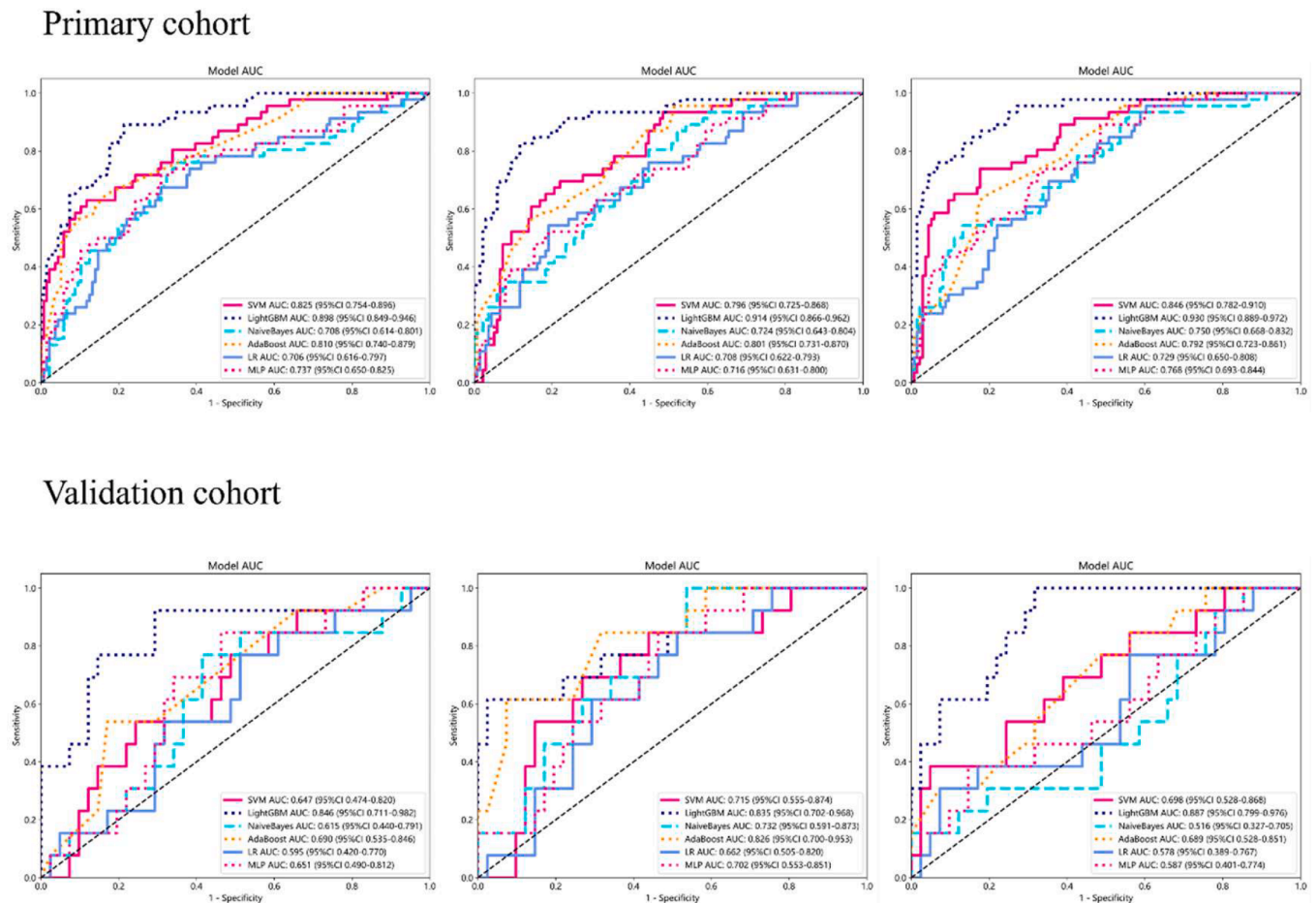


Fig. 3. ROC of the models built using different classification algorithms in the primary cohort and validation cohort.

Table 2
Performance of different models for predicting TP.

Model name	cohort	AUC (95 %CI)	ACC	SEN	SPE	PPV	NPV	Precision	Recall	F1
Dosimetry model	Primary	0.798 (0.732–0.865)	0.753	0.739	0.757	0.507	0.896	0.507	0.739	0.602
	validation	0.793 (0.637–0.949)	0.778	0.692	0.805	0.529	0.892	0.529	0.692	0.600
Radiomics model	Primary	0.930 (0.829–0.941)	0.868	0.717	0.919	0.750	0.906	0.750	0.717	0.733
	validation	0.887 (0.810–0.993)	0.852	0.923	0.829	0.632	0.971	0.632	0.923	0.750
D+R model	Primary	0.940 (0.906–0.974)	0.879	0.826	0.897	0.731	0.938	0.731	0.826	0.776
	validation	0.951 (0.891–1.000)	0.907	0.923	0.902	0.750	0.974	0.750	0.923	0.828

outperformed R model based on V30 Lung-PTV and dosimetry model for prediction of RP (Fig. 4 (b) and Table 2). The Delong test showed that the R+D model improved the model performance to predict RP ($p < 0.05$). The sensitivities of R+D model also outperformed the individual model. The recall of R+D model were also high enough to identify the RP or non-RP patients. We assigned the ranking of significance for each feature in the model and the results were shown in the Fig. 3 of supplementary material. Once the model has been built, the decision curve analysis was used to determine the range of model scores at which patients could benefit from the proposed model through the assessment of clinical utility. For the R+D model, when the threshold was set at 0.19–0.81, their clinical net benefits were higher than 0 in the external validation cohort. Fig. 5(a) showed the decision curves of the R+D model. The confusion matrix of the model also showed that R+D model

accurately justified whether patients had predicted RP or not (Fig. 5(b)). Moreover, we also provide the calibration curves of the predictive model, illustrating the R+D model’s potential use at primary and validation cohorts (Fig. 5(c)). To illustrate the feasibility of our best model, we provide two typical patients in the Fig. 4 of supplementary material.

Discussion

The classical definition of dose limits for peripheral OARs was more adapted to the tumor patients treated with radiotherapy alone. However, more and more combined treatment modes are practically applied in a clinic, meaning lung cancer patients may be bound to receive chemotherapy or immunotherapy or targeted therapy during, before, or following radiotherapy. These suggest that it is urgent to explore new

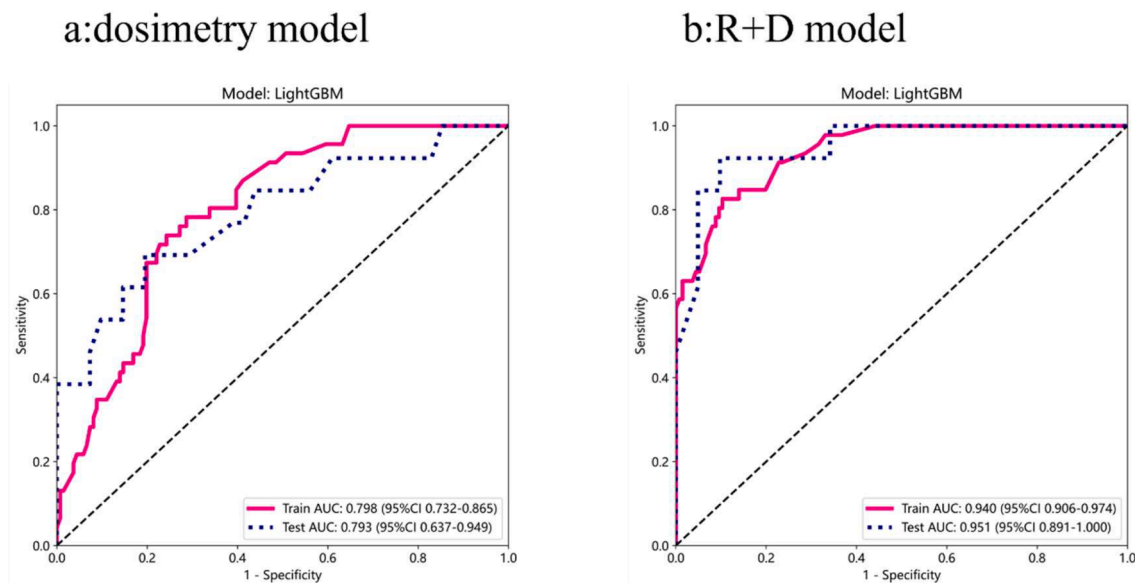


Fig. 4. (a) ROC of D model in primary cohort and validation cohort; (b) ROC of R+D model in primary cohort and validation cohort.

risk factors of dose limits or dose area, and it is preferred to establish a predicted model including the more sensitive areas of the peripheral normal lung and the specific volume dose rather than the traditional focused volume dose. As per our previous report [13], V30 should be more relevant to RP than V20. So, we further conducted this study to find the more sensitive region for evaluating V30-related risk of RP. This study showed that the ROI of V30 Lung-PTV is recommended to be preferentially concerned in clinical applications. Moreover, the established radiomics model in our study was identified to predict acute RP more effectively than the traditional dosimetry model, especially the radiomics model based on the V30 Lung-PTV region. Therefore, we suggested promoting the clinical application of this radiomics model based on the V30 Lung-PTV region. However, for the high-risk patients who are defined by our predictive model, the future is needed to design a series of prospective studies to explore the optimum protective strategy to decrease RP incidence, such as further optimizing the plan, amifostine application, adaptive radiation therapy, a new mode of dose fractionation, pulmonary function monitoring, and so on.

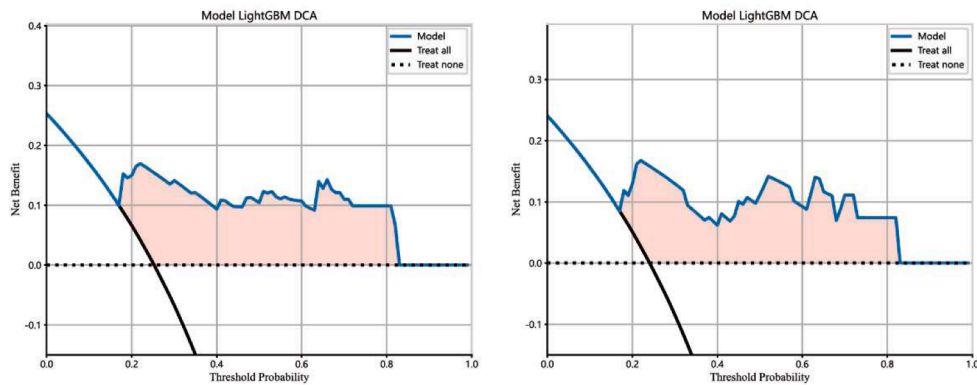
Non-invasive medical images are economical and frequently utilized in clinical practice. Radiomics is an advanced and automated image analysis technique that overcomes the limitations of manual interpretation by the human eye [17,18]. The application value of radiomics in diagnostic staging [19,20], treatment response [21–23] and the monitoring of prognosis [24,25], which has been already explored. The emergence and development of radiomics as a major breakthrough that challenges traditional measurements and assessment criteria with outstanding power has been supported by several studies. Further researches found that the extraction of medical image data to characterize normal tissue regions using radiomics approach could lead to clinically relevant improvements in the prediction of treatment-related toxicities such as acute RP. However, few studies have focused on this field. Our study was conducted using traditional dosimetry to identify normal tissue regions, which modelling done on various regions. The research addresses this gap and provides new insights into identifying acute RP, distinct from traditional dosimetry and clinical methods.

The primary innovation in this paper is the usage of conventional dosimetry to determine the ROIs for radiomics research. In the era of 3D-CRT, we focus on dosimetric indicators such as V20 and MLD of the whole lung [12,26–28], whereas in the IMRT era, we need to focus on another dosimetric indicator, namely V30-PTV [13]. Several previous studies have indirectly corroborated some of what we have found [10,11,29]. Wei Jiang et al reported that the prediction of symptomatic

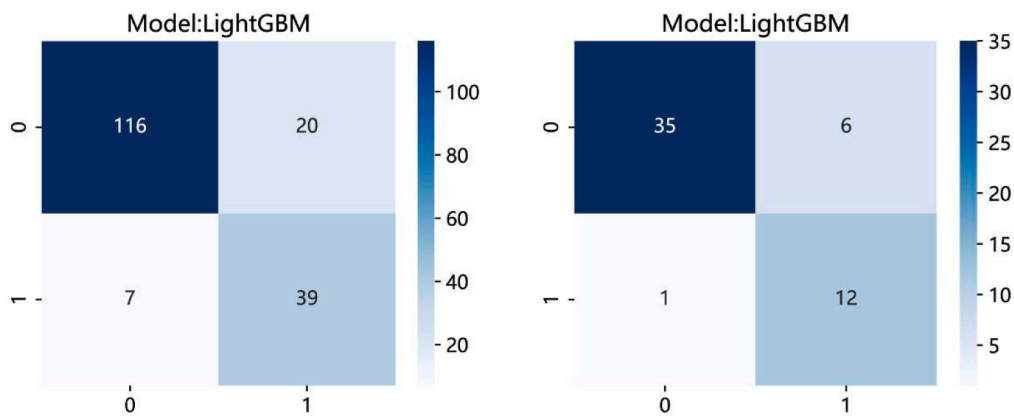
RP can be improved by using a machine learning model that uses dosimetric factors and radiomics features in different ROIs from planning computed tomography images [10]. Although the power of the radiomics approach to assess symptomatic RP was increased in this study, there still had significant limitations. One of the disadvantages is that the delineation of the ROIs in this study did not take into account delineation by dose distribution. In addition, this study only retrospectively collected 79 patients with lung cancer, the sample size is too small and there is no external cohort for validation of the model's robustness and generalizability. Nevertheless, our study was able to circumvent these limitations. As the development of acute RP is historically based on the volume of the dose received by the whole lung, we separated the various ROIs according to dosimetry parameters that are currently relevant in clinical for radiomics analysis. Our research demonstrated that radiomics models, divided into different ROIs based on dosimetry parameters, proved to be a more effective tool for discriminating acute RP compared to dosimetry models. In particular, the radiomics model based on the ROI of V30 Lung-PTV achieves the best predictive performance and provides a basis for the V30-PTV dosimetry parameter to better predict RP in our previous study. Our sample data, on the other hand, includes external validation data that provides evidence for the stability of the model. The DCA and calibration curves demonstrate the potential usefulness of the model presented in our study for clinical decision making. For example, if the personal threshold probability of a patients is 40 %, then the net benefit is 0.10 when using the R+D model to make the decision of whether acute RP occurs after treatment.

The majority of the previous studies that investigate RP employed only one machine learning algorithm. Wei Jiang et al reported that a mixed model combining clinical, dosimetry parameters and radiomics features was created using a support vector machine algorithm to discriminate RP [10]. The results demonstrate that the mixed model has an AUC value of 0.94 (95CI:0.85–1), and its model discriminations were significantly higher than those of the individual models, which includes clinical model(AUC[95 %CI]:0.73[0.54–0.92]), dosimetry model(AUC [95 %CI]:0.53[0.31–0.75]), radiomics model(AUC[95 %CI]:0.82 [0.65–0.99]). Anthony et al showed that a logistic regression classifier for radiomics model, which combines lung texture features from CT with a normalised uptake value from 18F-fluorodeoxyglucose positron emission tomography, could be used to predict RP in patients with esophageal cancer after radiation therapy [30]. However, Differences in the algorithms may introduce some bias into the model. To ensure model accuracy and stability, our study employs six machine learning

a: DCA for R+D model in primary cohort and validation cohort



b: Confusion matrix for R+D model in primary cohort and validation cohort



c: Calibration curves for R+D model in primary cohort and validation cohort

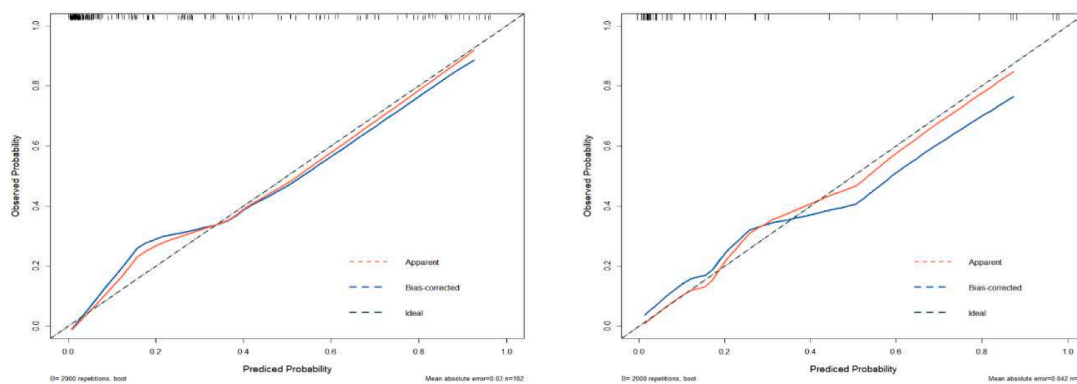


Fig. 5. (a):DCA for R+D model in primary cohort and validation cohort. The y-axis means the net benefit.The blue line represents the R+D model. The thick solid black line represents the assumption that all patients have acute RP after treatment and the black dotted line represents the assumption that no patients have acute RP after treatment. (b): Confusion matrix for R+D model in primary cohort and validation cohort. A confusion matrix is a table that is often used to describe the classification model on a set of data for which the true values are known. In our study, the classifier made a total 182 predictions in primary cohort, out of those 182 cases, the classifier predicted 123 patients may not have acute RP and 59 patients may have acute RP after treatment. In reality,136 patients in the cohort do not have acute RP and 46 patients have acute RP after treatment.(c):Calibration curves of the R+D model in primary cohort and validation cohort. Calibration curve is a visual tool to assess the agreement between predictions and acute outcome in different percentiles of the predicted values. The solid black line represents the reference line where a model would like, the blue line corrects for any bias in the model, and the red line represents the performance of the model.

algorithms to build the model. This approach reduces the model's instability caused by the singularity of algorithms and helps identify the optimal algorithm to build a reliable and stable model. Our results released that LightGBM models produce optimal discriminative performance compared to other algorithms for both the primary and external validation sets. The reason may be explained that LightGBM utilizes a Leaf-wise growth strategy, whereby it selects one leaf from all existing leaves that offers the largest splitting gain, splits it, and repeats this process [31]. This approach offers the benefit of reducing the error and providing better accuracy with the same number of splits. However, Leaf-wise might grow a deeper decision tree and lead to overfitting. To prevent overfitting while ensuring high efficiency, Light GBM sets a maximum depth limit atop of the Leaf-wise strategy.

This study has several limitations. First, the heterogeneity of the clinical data or dosimetry parameters may be biased as a real-world retrospective study of lung toxicity. Second, the occurrence of RP is objectively low. In our study, its incidence is reported as 25.3 % in primary cohort and 24.1 % in external validation cohort, respectively. This low incidence results in an uneven dataset when modelled. In order to address the issue of category imbalance, our study used the Synthetic Minority Over-sampling Technique (SMOTE) algorithm [32,33]. SMOTE operates by selecting examples that are close together in feature space, drawing a line between these examples in feature space, and then generating new examples at points along that line. Random under-sampling was initially employed to reduce the number of samples in the majority class. This was followed by the use of SMOTE to oversample the minority class, in order to equalize the class distribution. The effectiveness of this approach stems from the creation of new synthetic samples in the minority class that are closely aligned with existing minority samples in the feature space. Third, this study only examined dosimetry variables and CT-based radiomic characteristics; however, we did not observe the biological influence (such as genetic phenotypes and inflammatory cytokines) and clinical characteristics, although it has been suggested to be significant on RP occurrence in previous studies. Future research could incorporate additional data into the model to assess whether there is potential to improve the accuracy of the predictions. Finally, more work is required to enhance the comprehensibility of our radiomics model.

CRedit authorship contribution statement

Liqiao Hou: Conceptualization, Data curation, Formal analysis, Software, Writing – original draft, Writing – review & editing. **Kuifei Chen:** Data curation, Investigation, Software, Writing – original draft, Writing – review & editing. **Chao Zhou:** Data curation, Investigation, Writing – original draft, Writing – review & editing. **Xingni Tang:** Data curation, Investigation, Writing – review & editing. **Changhui Yu:** Software, Formal analysis, Writing – review & editing. **Haijian Jia:** Data curation, Writing – review & editing. **Qianyi Xu:** Conceptualization, Software, Writing – review & editing. **Suna Zhou:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Haihua Yang:** Conceptualization, Supervision, Funding acquisition, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Haihua Yang reports financial support was provided by the Chinese National Science Foundation Projects. the Basic Public Welfare Research Project of Zhejiang Province. Suna Zhou was provided by the Natural Science Foundation of Shaanxi Province (2023-JC-YB-645).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ctro.2024.100828>.

References

- [1] Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature* 2018;553(7689):446–54.
- [2] Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. *Nat Rev Clin Oncol* 2023;20(9):624–39.
- [3] Conibear J, AstraZeneca UKL. Rationale for concurrent chemoradiotherapy for patients with stage III non-small-cell lung cancer. *Br J Cancer* 2020;123(Suppl 1): 10–7.
- [4] Bledsoe TJ, Nath SK, Decker RH. Radiation Pneumonitis. *Clin Chest Med* 2017;38(2):201–8.
- [5] Palma DA, Senan S, Tsujino K, et al. Predicting radiation pneumonitis after chemoradiation therapy for lung cancer: an international individual patient data meta-analysis. *Int J Radiat Oncol Biol Phys* 2013;85(2):444–50.
- [6] Tonison JJ, Fischer SG, Viehriq M, et al. Radiation Pneumonitis after Intensity-Modulated Radiotherapy for Esophageal Cancer: Institutional Data and a Systematic Review. *Sci Rep* 2019;9(1):2255.
- [7] Wang T, She Y, Yang Y, et al. Radiomics for Survival Risk Stratification of Clinical and Pathologic Stage IA Pure-Solid Non-Small Cell Lung Cancer. *Radiology* 2022; 302(2):425–34.
- [8] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14(12):749–62.
- [9] Huang Y, Feng A, Lin Y, et al. Radiation pneumonitis prediction after stereotactic body radiation therapy based on 3D dose distribution: dosiomics and/or deep learning-based radiomics features. *Radiat Oncol* 2022;17(1):188.
- [10] Jiang W, Song Y, Sun Z, Qiu J, Shi L. Dosimetric Factors and Radiomics Features Within Different Regions of Interest in Planning CT Images for Improving the Prediction of Radiation Pneumonitis. *Int J Radiat Oncol Biol Phys* 2021;110(4): 1161–70.
- [11] Zhang Z, Wang Z, Yan M, et al. Radiomics and Dosiomics Signature From Whole Lung Predicts Radiation Pneumonitis: A Model Development Study With Prospective External Validation and Decision-curve Analysis. *Int J Radiat Oncol Biol Phys* 2023;115(3):746–58.
- [12] Marks LB, Bentzen SM, Deasy JO, et al. Radiation dose-volume effects in the lung. *Int J Radiat Oncol Biol Phys* 2010;76(3 Suppl):S70–6.
- [13] Meng Y, Luo W, Wang W, et al. Intermediate Dose-Volume Parameters, Not Low-Dose Bath, Is Superior to Predict Radiation Pneumonitis for Lung Cancer Treated With Intensity-Modulated Radiotherapy. *Front Oncol* 2020;10:584756.
- [14] van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* 2017;77(21):e104–7.
- [15] 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(2):328–38.
- [16] Forrest IS, Petrazzini BO, Duffy A, et al. Machine learning-based marker for coronary artery disease: derivation and validation in two longitudinal cohorts. *Lancet* 2023;401(10372):215–25.
- [17] Lafata KJ, Wang Y, Konkel B, Yin FF, Bashir MR. Radiomics: a primer on high-throughput image phenotyping. *Abdom Radiol (NY)* 2022;47(9):2986–3002.
- [18] Mayerhoefer ME, Materka A, Langs G, et al. Introduction to Radiomics. *J Nucl Med* 2020;61(4):488–95.
- [19] Wu L, Yang X, Cao W, et al. Multiple Level CT Radiomics Features Preoperatively Predict Lymph Node Metastasis in Esophageal Cancer: A Multicentre Retrospective Study. *Front Oncol* 2019;9:1548.
- [20] Wu L, Wang C, Tan X, et al. Radiomics approach for preoperative identification of stages I-II and III-IV of esophageal cancer. *Chin J Cancer Res* 2018;30(4):396–405.
- [21] Ligerio M, Garcia-Ruiz A, Viaplana C, et al. A CT-based Radiomics Signature Is Associated with Response to Immune Checkpoint Inhibitors in Advanced Solid Tumors. *Radiology* 2021;299(1):109–19.
- [22] Zhou C, Hou L, Tang X, et al. CT-based radiomics nomogram may predict who can benefit from adaptive radiotherapy in patients with local advanced-NSCLC patients. *Radiother Oncol* 2023;183:109637.
- [23] Yang B, Zhou L, Zhong J, et al. Combination of computed tomography imaging-based radiomics and clinicopathological characteristics for predicting the clinical benefits of immune checkpoint inhibitors in lung cancer. *Respir Res* 2021;22(1): 189.
- [24] Xie C, Yang P, Zhang X, et al. Sub-region based radiomics analysis for survival prediction in oesophageal tumours treated by definitive concurrent chemoradiotherapy. *EBioMedicine* 2019;44:289–97.
- [25] Foley KG, Shi Z, Whybra P, et al. External validation of a prognostic model incorporating quantitative PET image features in oesophageal cancer. *Radiother Oncol* 2019;133:205–12.
- [26] Simone 2nd CB. Thoracic Radiation Normal Tissue Injury. *Semin Radiat Oncol* 2017;27(4):370–7.
- [27] Glick D, Lyen S, Kandel S, et al. Impact of Pretreatment Interstitial Lung Disease on Radiation Pneumonitis and Survival in Patients Treated With Lung Stereotactic Body Radiation Therapy (SBRT). *Clin Lung Cancer* 2018;19(2):e219–26.
- [28] Lee YH, Kim YS, Lee SN, et al. Interstitial Lung Change in Pre-radiation Therapy Computed Tomography Is a Risk Factor for Severe Radiation Pneumonitis. *Cancer Res Treat* 2015;47(4):676–86.
- [29] Wang L, Gao Z, Li C, et al. Computed Tomography-Based Delta-Radiomics Analysis for Discriminating Radiation Pneumonitis in Patients With Esophageal Cancer After Radiation Therapy. *Int J Radiat Oncol Biol Phys* 2021;111(2):443–55.

- [30] Anthony GJ, Cunliffe A, Castillo R, et al. Incorporation of pre-therapy (18) F-FDG uptake data with CT texture features into a radiomics model for radiation pneumonitis diagnosis. *Med Phys* 2017;44(7):3686–94.
- [31] Ke G, Meng Qi, Finley T, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Adv Neural Inf Proces Syst* 2017;30.
- [32] Chawla NV, Bowyer KW, HI O. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res* 2002;16:321–57.
- [33] Cheng W-C, Mai T-H, Lin H-T. From SMOTE to Mixup for Deep Imbalanced Classification. *arXiv Preprint arXiv* 2023;2308(15457).