



OPEN Multicenter development of a deep learning radiomics and dosiomics nomogram to predict radiation pneumonia risk in non-small cell lung cancer

Xun Wang¹, Aiping Zhang², Huipeng Yang³, Guqing Zhang¹, Junli Ma⁴, Shucheng Ye⁴ & Shuang Ge⁴✉

Radiation pneumonia (RP) is the most common side effect of chest radiotherapy, and can affect patients' quality of life. This study aimed to establish a combined model of radiomics, dosiomics, deep learning (DL) based on simulated location CT and dosimetry images combining with clinical parameters to improve the predictive ability of ≥ 2 grade RP (RP2) in patients with non-small cell lung cancer (NSCLC). This study retrospectively collected 245 patients with NSCLC who received radiotherapy from three hospitals. 162 patients from Hospital I were randomly divided into training cohort and internal validation cohort according to 7:3. 83 patients from two other hospitals served as an external validation cohort. Multivariate analysis was used to screen independent clinical predictors and establish clinical model (CM). The radiomic and dosiomics (RD) features and DL features were extracted from simulated location CT and dosimetry images based on the region of interest (ROI) of total lung-PTV (TL-PTV). The features screened by the *t*-test and least absolute shrinkage and selection operator (LASSO) were used to construct the RD and DL model, and RD-score and DL-score were calculated. RD-score, DL-score and independent clinical features were combined to establish deep learning radiomics and dosiomics nomogram (DLRDN). The model performance was evaluated by area under the curve (AUC). Three clinical factors, including V20, V30, and mean lung dose (MLD), were used to establish the CM. 7 RD features including 4 radiomics features and 3 dosiomics features were selected to establish RD model. 10 DL features were selected to establish DL model. Among the different models, DLRDN showed the best predictions, with the AUCs of 0.891 (0.826–0.957), 0.825 (0.693–0.957), and 0.801 (0.698–0.904) in the training cohort, internal validation cohort and external validation cohort, respectively. DCA showed that DLRDN had a higher overall net benefit than other models. The calibration curve showed that the predicted value of DLRDN was in good agreement with the actual value. Overall, radiomics, dosiomics, and DL features based on simulated location CT and dosimetry images have the potential to help predict RP2. The combination of multi-dimensional data produced the optimal predictive model, which could provide guidance for clinicians.

Keywords Radiation pneumonia, Radiomic, Dosiomic, Deep learning, Non-small cell lung cancer

Abbreviations

RP	Radiation pneumonia
NSCLC	Non-small cell lung cancer
RD	Radiomic and dosiomics
DL	Deep learning
DLRDN	Deep learning radiomics and dosiomics nomogram

¹Department of Medical Imaging, Affiliated Hospital of Jining Medical University, Guhuai Road, Jining 272000, Shandong, China. ²Department of Radiation Oncology, Tumor Hospital of Jining, Jianshe North Road, Jining 272123, Shandong, China. ³Department of Radiation Oncology, Jining First People's Hospital, Jiankang Road, Jining 272029, Shandong, China. ⁴Department of Radiation Oncology, Affiliated Hospital of Jining Medical University, Guhuai Road, Jining 272000, Shandong, China. ✉email: geshuang616@163.com

ROI	Region of interest
LASSO	Least absolute shrinkage and selection operator
AUC	Area under the curve
MLD	Mean lung dose
CM	Clinical model
DCA	Decision curve analysis
GTV	Gross tumor volume
CTV	Clinical tumor volume
PTV	Planning target volume
DVH	Dose-volume histograms

Non-small cell lung cancer (NSCLC) is one of the most common causes of cancer-related death worldwide accounting for about 75–80% of all primary lung cancer patients^{1,2}. At present, the treatment of NSCLC mainly includes surgical resection, radiotherapy (RT), chemotherapy, targeted therapy, immunotherapy, microwave ablation and so on. RT plays a crucial role in the treatment of NSCLC, especially for locally advanced and unresectable cases^{3–5}. However, some patients may experience additional radiation-related adverse reactions such as radiotherapy induced lung injury, including radiation pneumonitis (RP) and pulmonary fibrosis (PF). RP is an acute inflammation of lung tissue caused by radiation exposure to normal tissues and is the most common side effect of thoracic radiotherapy, often occurring within 6 months post-RT, which may cause respiratory insufficiency, affect patients' quality of life, and even lead to death^{6–8}.

The pathogenesis of radiation-induced lung toxicity is described as multiple interacting cellular activities such as hypoxia, fibrogenesis, inflammation, and angiogenesis⁹. Patients with stage III/IV NSCLC have a 30–40% risk of developing RP⁸. Numerous studies have shown that some clinical risk factors, such as pulmonary function, smoking history, tumor location, lung interstitial disease, pulmonary emphysema, immunotherapy etc. are closely related to the occurrence of RP^{8,10–12}. Recent developments in quantitative analysis of medical images using artificial intelligence (AI) tools, such as machine learning, have created new frontiers in oncologic imaging. Radiomics, a new AI tool based on textural analysis extracting quantitative data from medical images, has been used to predict treatment response in different types of cancer, including patients with lung cancer treated with chemotherapy and/or RT^{13,14}. Several studies have demonstrated the potential of radiomics to predict RP^{15–17}. In addition, the occurrence of RP is directly related to radiation dose. Some studies have shown that RP is related to RT dose derived from dose-volume histograms (DVH), such as volume of the lung receiving 5 Gy (V5), volume of the lung receiving 10 Gy (V10), volume of the lung receiving 20 Gy (V20), volume of the lung receiving 30 Gy (V30), or mean lung dose (MLD)^{8,15}. However, we can only summarize the two-dimensional dose distribution in the target from the DVH parameters, and cannot obtain the spatial dose distribution from it¹⁸. Dosiomics derived from radiomics, can describe the spatial heterogeneity of dose distribution compared to DVH. Zhang¹⁵ and Huang¹⁸ extracted dosiomic features from 3D dose distribution for RP prediction, which played an important role in RP prediction. Besides, as a branch of artificial intelligence, deep learning (DL) technology can automatically learn representative mass and high-dimensional data information from raw medical image data, including decoding the radiomics representation of tumors, which shows great potential in the application of tumor medicine^{19,20}. Radiomics and DL features represent different modes of image analysis that are not redundant, but complementary²¹.

Previous studies have shown that radiomics and DL can be used to predict RP, but most of them are single-model predictions, and few studies have conducted conjoint analyses. In this study, we aimed to develop and validate a combined model of deep learning radiomics and dosiomics nomogram (DLRDN) based on simulated location CT and dosimetry images to predict ≥ 2 grade RP (RP2) in NSCLC.

Materials and methods

Patient characteristics

This was a retrospective multi-center study approved and need to obtain informed consent was waived by the Ethics Committee of the Affiliated Hospital of Jining Medical University (Jining Medical University Ethics Committee for Human Research), China. This study was conducted according to the Declaration of Helsinki.

In this study, 162 patients with NSCLC who received intensity modulated radiation therapy (IMRT) or volumetric modulated arc therapy (VMAT) in Hospital I from June 2016 to December 2022 were retrospectively collected and randomly divided into a training cohort ($n = 113$, 74 without RP2 and 39 with RP2) and an internal validation cohort ($n = 49$, 32 without RP2 and 17 with RP2) at a ratio of 7:3. 83 patients (59 without RP2 and 24 with RP2) from two other hospitals were collected as an external validation cohort from January 2019 to December 2022. All patients underwent IMRT or VMAT and were administered radiotherapy at 1.8–3 Gy per day, 5 days a week, with 20–30 fractions and with a total dose of 40–60 Gy. The inclusion criteria of this study were as follows: (i) Definite pathological diagnosis of NSCLC. (ii) Patients who had completed radiotherapy with or without chemotherapy and had comprehensive clinical data. (iii) Follow up for at least 6 months after RT or for endpoint events. The endpoint was defined as the occurrence of RP2 within 6 months after the end of RT. The exclusion criteria of this study were as follows: (i) Surgically removed lung cancer patients. (ii) Patients received repeated chest RT or stereotactic body radiation therapy (SBRT). (iii) Patients with other malignant tumors or mediastinal lung cancer.

The severity of RP is graded according to the classification criteria for acute radiation pneumonia developed by the American Radiation Therapy Oncology Group (RTOG). A grade ≥ 2 was classified as symptomatic RP, which necessitated the use of steroids or the restriction of instrumental activities of daily living⁸. So, in this study, the risk of RP was divided into two levels, < 2 grade (without RP2) for the low-risk group, ≥ 2 grade (RP2) for the high-risk group. The grade of RP was determined by a radiologist and an oncologist separately, and by a

third physician if there was disagreement. The overall workflow of the RP2 predictive model development and validation was shown in Fig. 1.

Establishment of CM

Clinical data included age, sex, smoking status, emphysema, pathological type, T stage, peripheral blood markers, immunotherapy status, and DVH parameters. The peripheral blood markers mainly included lymphocytes, neutrophils, monocytes, platelets, neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR) and lymphocyte-to-monocyte ratio (LMR). The T stage was based on guidelines from the American Joint Committee on Cancer “Cancer Staging Manual, Eighth Edition”. The DVH parameters included V5, V10, V20, V30 and MLD. Multivariate analysis was used to screen independent clinical predictors and establish clinical model (CM).

CT acquisition and image segmentation

All patients were scanned and underwent tumour location by large aperture CT analog positioner (PHILIPS Brilliance CT Big Bore, and Canon Aquilion Prime TSX-303 A). The scanning voltage was 120 kV, the tube current was 250mAs, the helical sweep pitch was 1, and slice thickness was 5 mm. Experienced radiation oncologists delineated the gross tumor volume (GTV) of lung cancer and malignant lymph nodes in the Eclipse, Monaco, or Raystation treatment planning system (TPS). The clinical tumor volume (CTV), including potential areas of suspected microscopic tumor invasion and possible microscopic diffusion paths, was derived by GTV isotropic expansion of 5 mm, and the planning target volume (PTV) was generated by isotropic expansion of 5 mm on the basis of CTV. The target prescription dose covers at least 95% of PTV. Since RP primarily occurs in normal lung tissue rather than in the high-dose PTV region, selecting an appropriate region of interest (ROI) is critical for accurate prediction. Prior studies by Jiang et al.²² and Meng et al.²³ showed that models based on the total lung minus PTV (TL-PTV) region yielded better predictive performance for symptomatic RP compared to other ROIs such as TL-GTV, PTV, or GTV. Based on this evidence and the biological relevance of TL-PTV to RP development, it was chosen as the ROI in this study. The planning CT, ROI and 3D dose grids of planned radiotherapy were exported from TPS in digital imaging and communications in medicine (DICOM) format. Prior to feature extraction, CT and dosimetry images were isotropic spatially resampled (1 mm) to ensure comparability. Three-dimensional (3D) tumor segmentation was performed using 3D Slicer (software version 5.1.0, <https://www.slicer.org>). A sensitivity analysis was performed to evaluate the impact of inter-institutional variations in radiotherapy equipment and dose fractionation on the results.

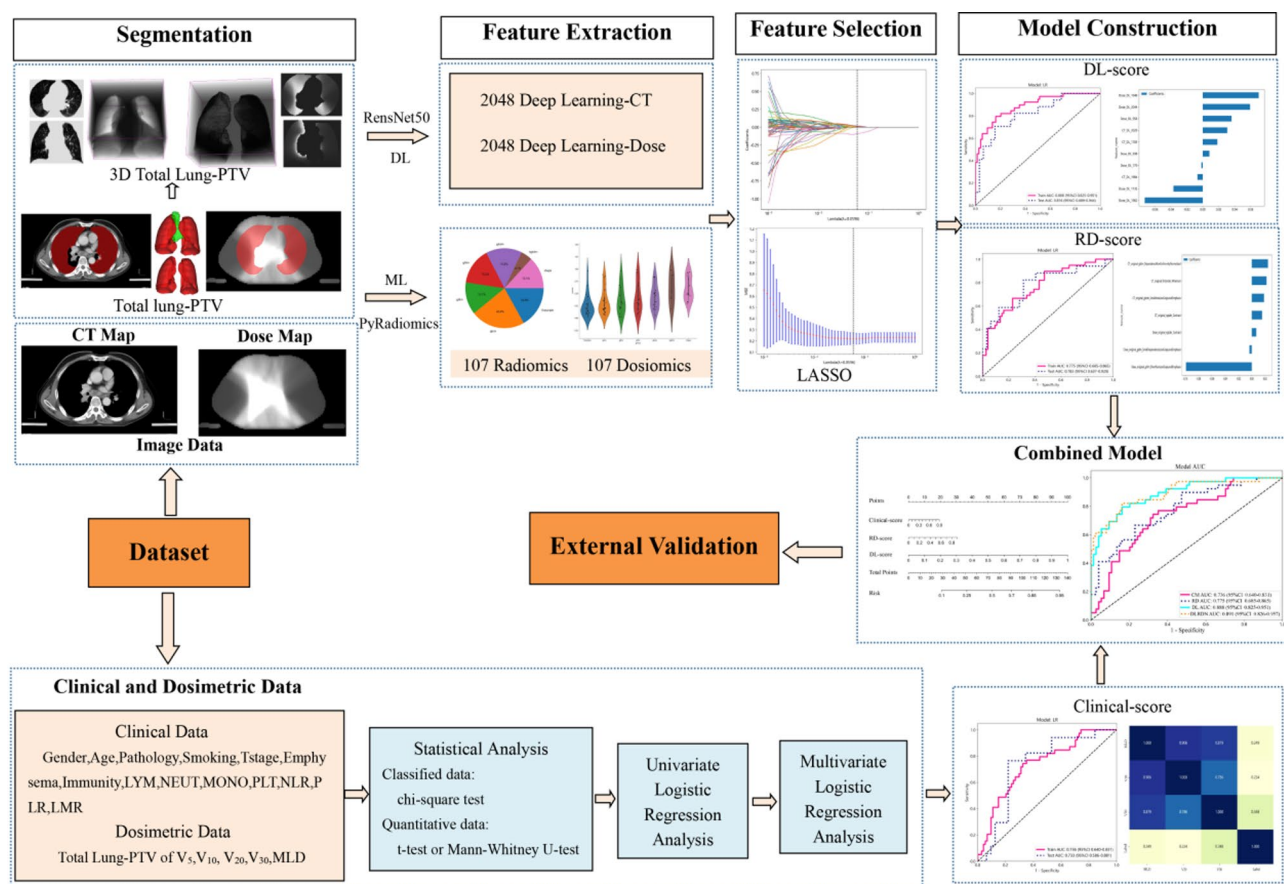


Fig. 1. The overall workflow of the RP2 predictive model development and validation.

Feature extraction and selection of RD and DL

A total of 214 radiomics and dosiomics(RD) features were extracted from each ROI (TL-PTV) using PyRadiomics, including first-order, texture and shape features. Additionally, 4096 DL features were extracted using a 3D ResNet50 architecture. The input images were cropped and resized to 96*96*96 voxels. To address limited data and improve model generalizability, data augmentation techniques—such as random clipping, rotation, flipping, and transfer learning—were applied during training. All features were standardized using Z-score normalization. Feature selection was performed using least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation to enhance model stability. RD-score and DL-score were calculated for each patient based on the multivariate logistic regression product of the selected features.

Development of the RD model, DL model and DLRDN

The selected RD features were input into supervised classifiers, including logistic regression (LR), support vector machine (SVM), and random forest (RF) to construct RD models, and the optimal model was selected based on the best-performing machine learning algorithm. Similarly, the DL model was developed using the selected DL features. A combined DLRDN was developed by combining the independent clinical features, RD-score and DL-score using multivariate logistic regression analysis.

Performance assessment of different models

The performances of each prediction model (CM, RD, DL, and DLRDN) were tested using the area under the curve (AUC) of the receiver operating characteristic (ROC). The optimum cut-off value was obtained by the Youden index, and the sensitivity, specificity, and accuracy were calculated in the training cohort, internal validation cohort and external validation cohort. The decision curve analysis (DCA) was used to determine the clinical benefit of each model. Calibration curves were drawn to evaluate the calibration of the models in the training cohort, internal validation cohort and external validation cohort by Hosmer-Lemeshow goodness-of-fit tests.

Statistics

To compare the differences in clinical characteristics of patients in different groups or cohorts, independent t test or Mann-Whitney U test was used for continuous variables, and chi-square test was used for categorical variables. SPSS software (version 20.0, IBM) was used for univariate and multivariate analysis (including chi-square test, *t*-test or Mann-Whitney U test). The R statistical software was used for LASSO regression analysis, ROC, and DCA analysis. $P < 0.05$ was considered statistically significant.

Results

Clinical characteristics of the patients

Clinical characteristics of patients in different cohorts are shown in Table 1. RP2 patients accounted for 34.5% (39/113) in the training cohort, 34.7% (17/49) in the internal validation cohort, and 28.9% (24/83) in the external validation cohort. Sensitivity analysis revealed no significant impact of inter-institutional variations in radiotherapy equipment or dose fractionation schedules on treatment outcomes ($p > 0.05$). No significant difference was detected in the age, gender, smoking, emphysema, pathology, neutrophils, monocytes, NLR, LMR, V5 and V10 between the RP2 and without RP2 groups in all three cohorts ($p > 0.05$). Multivariate analysis identified V20, V30, and MLD as independent predictors of RP2 ($p < 0.05$), which were subsequently used to establish the CM.

Feature selection and model construction of RD, DL, and DLRDN

Finally, 7 RD features including 4 radiomics features(1 first-order feature, 3 texture features) and 3 dosiomics features(3 texture features) were finally screened out which were closely associated with RP2 to build RD model (Suppl. Figure 1). We further evaluated the performance of three machine learning algorithms—LR, SVM, and RF—in predicting RP. All three models demonstrated favorable predictive performance, with LR slightly outperforming SVM and RF (Suppl. Figure 2). We ultimately selected LR as the primary modeling approach due to its stable performance, simplicity, and interpretability. 10 DL features were selected, among which 3 were from the CT images and 7 were from the dosimetry images to build DL model. The RD-score and DL-score were calculated based on the coefficients weighted by LR. Multivariate logistic regression analysis was performed on the independent clinical features, RD-score, and DL-score to construct the DLRDN (Fig. 2).

Performance comparison of different models

The 10-fold cross-validation results of the RD and DL models were presented in Suppl. Figure 3. The cross-validation showed that the RD model achieved a mean AUC of 0.775 ± 0.077 , while the DL model achieved a mean AUC of 0.888 ± 0.053 . The AUC, sensitivity, specificity, and accuracy of CM, RD, DL, and DLRDN in the training cohort, internal validation cohort and external cohort were shown in Table 2. DLRDN showed a good performance for RP2 prediction in the training cohort with an AUC of 0.891(95% CI 0.826–0.957), which was further confirmed in the internal and external validation cohorts with an AUC of 0.825 (95% CI 0.693–0.957) and 0.801 (95%CI, 0.698–0.904), respectively (Fig. 3). In both the training cohort and external validation cohort, DLRDN had significantly higher AUC than the CM ($p < 0.05$). DCA showed that DLRDN had a higher overall net benefit than the other models across most ranges of reasonable threshold probabilities, suggesting that DLRDN could be used to predict the occurrence of RP2 in NSCLC patients (Fig. 4A). The calibration curves of the DLRDN demonstrated that model-predicted RP2 was well-calibrated with the actual observation in all cohorts ($p > 0.05$) (Fig. 4B).

Clinical characteristics	Training cohort (n = 113)			Internal cohort (n = 49)			External cohort (n = 83)		
	Without RP2	With RP2	p	Without RP2	With RP2	p	Without RP2	With RP2	p
Age media (year)	59.99 ± 8.40	63.10 ± 6.04	0.07	61.12 ± 8.86	61.18 ± 8.60	0.98	64.61 ± 9.80	65.92 ± 7.49	0.56
Gender									
Male (%)	60 (81.08)	30 (76.92)	0.78	27 (84.38)	12 (70.59)	0.44	44 (74.58)	19 (79.17)	0.87
Female (%)	14 (18.92)	9 (23.08)		5 (15.62)	5 (29.41)		15 (25.42)	5 (20.83)	
Smoking									
Yes (%)	54 (72.97)	25 (64.10)	0.45	26 (81.25)	12 (70.59)	0.62	37 (62.71)	17 (70.83)	0.65
No (%)	20 (27.03)	14 (35.90)		6 (18.75)	5 (29.41)		22 (37.29)	7 (29.17)	
Emphysema									
Yes (%)	21 (28.38)	10 (25.64)	0.93	9 (28.12)	5 (29.41)	1.00	20 (33.90)	12 (50.00)	0.26
No (%)	53 (71.62)	29 (74.36)		23 (71.88)	12 (70.59)		39 (66.10)	12 (50.00)	
Pathology									
Adenocarcinoma	32 (43.24)	12 (30.77)	0.28	8 (25.00)	9 (52.94)	0.10	29 (49.15)	8 (33.33)	0.28
SCC	42 (56.76)	27 (69.23)		24 (75.00)	8 (47.06)		30 (50.85)	16 (66.67)	
T stage									
I	7 (9.46)	4 (10.26)	0.31	1 (3.12)	4 (23.53)	0.02	11 (18.64)	2 (8.34)	0.62
II	22 (29.73)	8 (20.51)		6 (18.75)	7 (41.18)		18 (30.51)	8 (33.33)	
III	18 (24.32)	6 (15.38)		7 (21.88)	1 (5.88)		10 (16.95)	6 (25.00)	
IV	27 (36.49)	21 (53.85)		18 (56.25)	5 (29.41)		20 (33.90)	8 (33.33)	
Lymphocytes	1.77 ± 0.56	1.65 ± 0.48	0.31	1.60 ± 0.60	1.54 ± 0.66	0.75	1.36 ± 0.54	1.80 ± 0.55	0.001
Neutrophils	3.88 ± 2.28	3.23 ± 1.25	0.23	3.67 ± 1.38	3.29 ± 1.62	0.21	3.63 ± 1.92	4.31 ± 2.74	0.56
Monocytes	0.46 ± 0.22	0.47 ± 0.17	0.89	0.52 ± 0.20	0.47 ± 0.18	0.44	0.52 ± 0.24	0.64 ± 0.37	0.29
Platelets	229.96 ± 62.78	229.67 ± 82.86	0.51	227.84 ± 81.26	215.00 ± 43.09	0.78	239.80 ± 107.43	236.29 ± 79.95	0.90
NLR	2.35 ± 1.60	2.12 ± 1.00	0.91	2.54 ± 1.37	2.62 ± 1.77	0.52	3.03 ± 2.06	2.58 ± 1.58	0.33
PLR	144.78 ± 78.52	150.44 ± 63.37	0.50	157.56 ± 65.53	177.73 ± 112.9	0.92	202.01 ± 132.63	143.08 ± 61.06	0.04
LMR	5.62 ± 5.97	3.87 ± 1.59	0.26	3.52 ± 1.94	3.74 ± 1.99	0.81	3.63 ± 3.76	4.30 ± 4.74	0.22
V5Gy	50.17 ± 12.07	51.17 ± 8.94	0.96	52.90 ± 9.67	55.01 ± 7.32	0.44	40.82 ± 14.16	39.81 ± 11.17	0.75
V10Gy	34.90 ± 10.03	37.39 ± 6.89	0.47	36.69 ± 8.41	39.52 ± 5.42	0.22	29.26 ± 9.74	30.40 ± 8.88	0.62
V20Gy	22.39 ± 6.61	25.30 ± 4.84	0.05	24.12 ± 4.84	27.41 ± 3.45	0.02	20.12 ± 6.50	21.85 ± 6.16	0.28
V30Gy	14.64 ± 4.93	17.74 ± 3.89	0.001	15.93 ± 3.07	19.98 ± 3.44	0.001	14.69 ± 5.10	16.43 ± 4.56	0.16
MLD	1233.88 ± 301.29	1366.37 ± 234.91	0.04	1291.64 ± 218.20	1482.74 ± 180.57	0.003	1126.24 ± 345.77	1194.07 ± 296.77	0.40

Table 1. Baseline clinical characteristics of patients in the training cohort, internal and external validation cohorts. Comparisons between patients with and without grade ≥ 2 radiation pneumonitis (RP2) were performed using appropriate statistical tests. Values were presented as mean $\pm$ standard deviation or n (%), and *p*-values indicate the significance of differences between groups.

Discussion

In our study, we investigated a novel comprehensive model (DLRDN) combining radiomics, dosiomics, DL features and independent clinical factors to predict RP risk stratification in NSCLC patients, and validated its ability to predict RP2 in the multicentre externally validation cohort. The model demonstrated robust predictive performance in both internal and external validation cohort (AUC: 0.825, 0.801, respectively). The predictive efficacy of DLRDN was better than that of clinical model, which demonstrated the added value of radiomics/dosiomics and DL features to traditional clinical and dosimetric parameters in prediction of RP2 in NSCLC.

RP is an important adverse event in NSCLC patients receiving chest radiotherapy. It is mainly managed with corticosteroids, and treatment failure can be fatal, thus offsetting the survival benefits of radiotherapy and affecting patients' quality of life^{8,24,25}. Patients with RP ≥ 2 represent a highly heterogeneous group. Early identification of high-risk individuals with symptomatic RP before treatment is crucial for effective prevention.

The incidence and severity of RP were directly related to the dose distribution in lung. Dosimetric factors such as MLD, V5, V10, V20, V30 have been widely used to predict RP^{8,11,15}. In our study, MLD, V20 and V30 were identified as independent predictors of RP2. However, their predictive performance was limited, with AUCs of 0.736, 0.733 and 0.631 in the training, internal validation and external validation cohort, respectively. These results suggested that traditional DVH parameters might not fully capture the complexity of RP risk. Liang et al.^{26,27} reported that increased local dose variation in the ipsilateral lung and larger low-dose regions in the total lung were significantly associated with higher RP incidence. Unlike spatial metrics, DVH parameters represent only the cumulative dose to specific lung volumes and fail to account for spatial heterogeneity, local dose gradients, or anatomical context. This highlights the need for more refined metrics that incorporate spatial dose distribution and organ-specific characteristics to improve RP risk prediction.

Radiomics is a high-throughput extraction of quantitative features from medical images, offering a non-invasive means to assess pulmonary heterogeneity before and during radiotherapy. Prior studies have

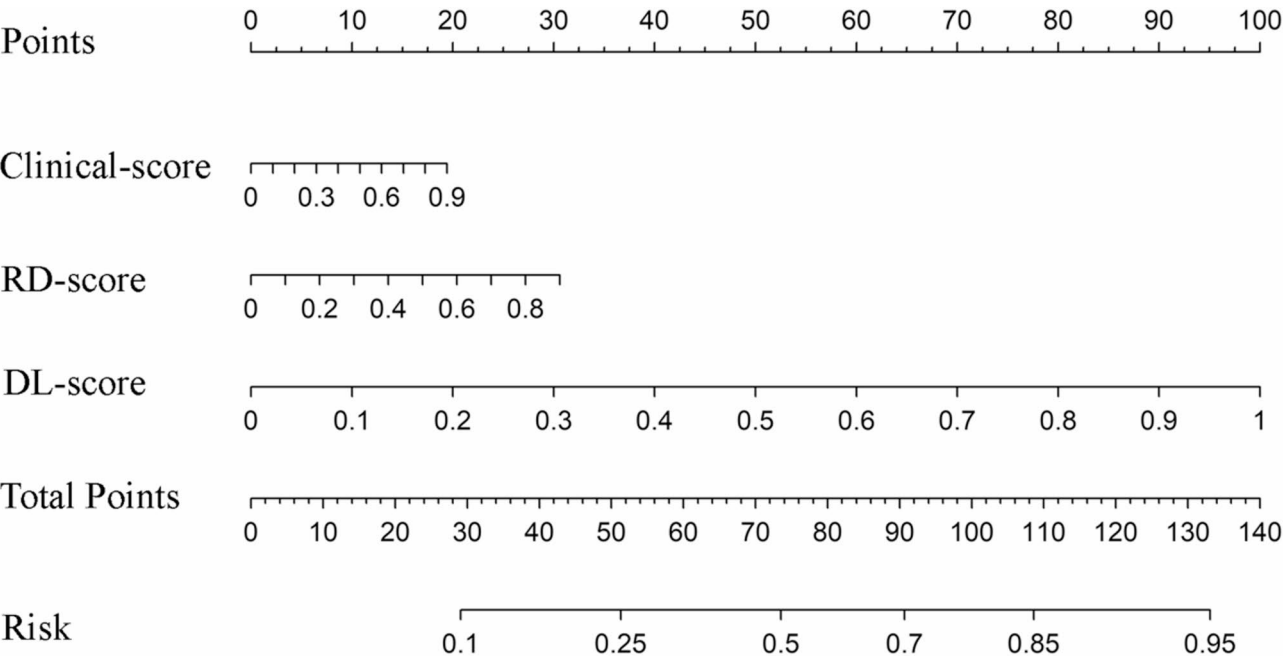


Fig. 2. The DLRDN, developed based on training cohort combining the RD-score, DL-score and clinical features.

Cohort	Model	AUC(95%CI)	Accuracy	Sensitivity	Specificity
Training cohort	CM	0.736 (0.640–0.831)	0.690	0.718	0.676
	RD	0.775 (0.685–0.865)	0.726	0.641	0.770
	DL	0.888 (0.825–0.951)	0.814	0.769	0.838
	DLRDN	0.891 (0.826–0.957)	0.823	0.795	0.838
Internal validation cohort	CM	0.733 (0.586–0.881)	0.755	0.706	0.781
	RD	0.783 (0.637–0.929)	0.714	0.765	0.687
	DL	0.816 (0.689–0.944)	0.776	0.647	0.844
	DLRDN	0.825 (0.693–0.957)	0.837	0.647	0.937
External validation cohort	CM	0.631 (0.502–0.759)	0.614	0.667	0.593
	RD	0.637 (0.503–0.771)	0.651	0.500	0.712
	DL	0.770 (0.658–0.883)	0.771	0.667	0.814
	DLRDN	0.801 (0.698–0.904)	0.735	0.792	0.712

Table 2. Performance of different models in training cohort, internal and external validation cohort. The diagnostic performance of clinical model (CM), radiomics–dosimetrics model (RD), deep learning model (DL), and deep learning radiomics–dosimetrics nomogram model (DLRDN) was evaluated by area under the curve (AUC), accuracy, sensitivity, and specificity. The results showed that DLRDN achieved the best overall performance across all cohorts.

demonstrated its potential in predicting symptomatic RP. Krafft et al.¹⁷ reported an AUC of 0.68 using whole-lung CT radiomics, outperforming clinical and dosimetric parameters (AUC: 0.51). Study by Nie et al.²⁸ also demonstrated that radiomics had good performance in predicting symptomatic RP, with an AUC of 0.740–0.802. Dosimetrics can provide quantitative measurements of intensity, shape, or heterogeneity within a defined volume of interest, capturing the inhomogeneity of dose distribution in a medical image^{18,29}. Liang et al.²⁶ showed that dosimetrics outperformed traditional dosimetric factors in RP prediction (AUCs: 0.709 and 0.782 vs. 0.665 and 0.676). Zhang et al.¹⁵ showed that the combined model composed of radiomics features, dosimetrics features and clinical parameters yielded the highest AUC (0.793 in the training set, and 0.855 in the testing set), which was higher than the single radiomics performance (AUC: 0.676 in the training set, and 0.671 in the testing set) and dosimetrics performance (AUC: 0.728 in the training set, and 0.684 in the testing set). In our study, the AUCs of RP2 prediction using radiomics and dosimetrics were 0.775, 0.783, and 0.637 in the training, internal validation, and external validation cohorts, respectively—all higher than those achieved using clinical features alone. These findings were consistent with previous studies. Additionally, we further evaluated the performance of SVM and RF in predicting RP. Although both algorithms showed good predictive performance, LR achieved

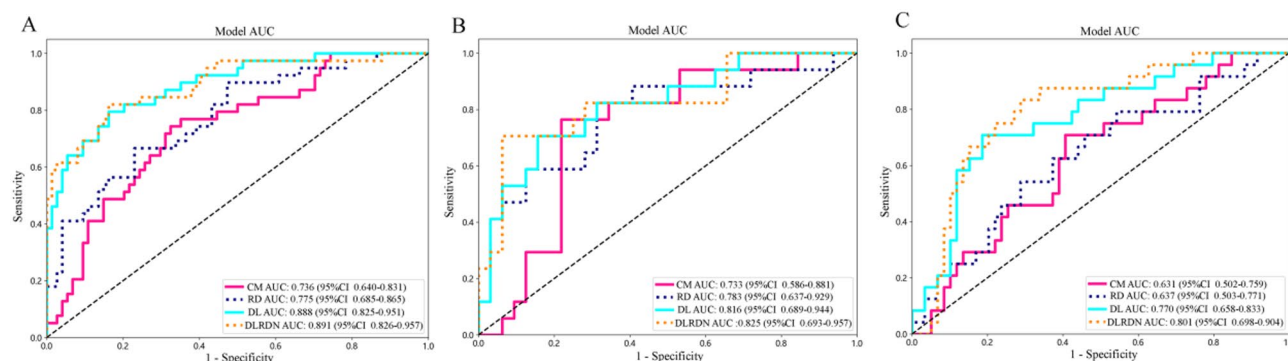


Fig. 3. Receiver operating characteristic (ROC) curves of the predictive models. ROC curves of DLRDN, DL, RD and CM for predicting RP2 in (A) training cohort, (B) internal validation cohort, and (C) external validation cohort.

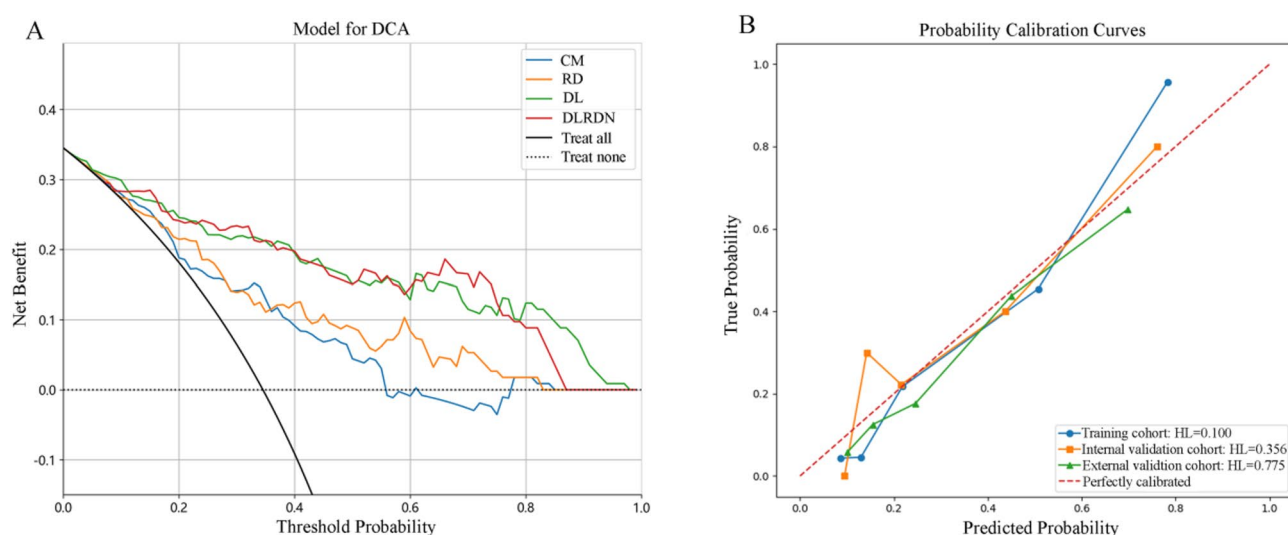


Fig. 4. (A) Decision curve analysis of the four models. (B) Calibration curves of DLRDN in all the three cohorts.

slightly superior results. Compared to the more complex and parameter-sensitive SVM and RF, LR offers greater interpretability, lower computational complexity, and enhanced clinical applicability. This study primarily aimed to assess the incremental value of integrating multi-dimensional data, and LR was employed as an illustrative modeling tool to highlight this benefit. The selected radiomics and dosiomics features were predominantly texture-based, in agreement with earlier finding³⁰. Among them, the Gray Level Dependence Matrix (GLDM)-derived feature Dependence Non-Uniformity Normalized (DNNUN) quantifies the heterogeneity of gray-level dependencies within the ROI³¹. In our study, higher DNNUN values were associated with an increased risk of RP, potentially reflecting greater structural inhomogeneity or underlying tissue vulnerability. These findings suggest that lung regions exhibiting uneven textural dependency might be more susceptible to radiation-induced damage, highlighting the potential of DNNUN as a predictive biomarker for RP. Conversely, Short Run Low Gray Level Emphasis (SRLGLE), derived from the Gray Level Run Length Matrix (GLRLM), showed a negative correlation with RP risk. Higher SRLGLE values, indicative of more uniform and aerated lung tissue, were linked to a lower likelihood of RP, whereas lower SRLGLE values—characteristic of denser and more heterogeneous textures—were more commonly observed in RP-prone regions. These results underscore the potential protective role of preserved lung structure against radiation injury.

The DL algorithm, which takes the raw images as input and calculate the output signal using multilayer transformations, can be a powerful and promising tool for studying complex patterns in the field of radiation therapy³². Studies of Huang, Kong, and Zhang et al.^{18,33,34} showed that DL or DL-radiomics/dosiomics could improve the predictive performance of RP occurrence or risk stratification. Although radiomics and DL features represent different modalities for image analysis, they are not redundant but complementary. All of them can reflect the spatial heterogeneity and microenvironment within the tumor²¹. However, most of their studies were based on single-omics deep learning method and were mostly single-center studies. Our study is multicentric, and based on multi-omics (radiomics and dosiomics) deep learning to predict RP2. Whether in

the training, internal validation, or external validation cohort, the combined model (DLRDN) all showed good predictive performance, with AUCs above 0.8, which were higher than other predictive models. Our results also showed that DL and multi-omics features complement each other, and their integration could provide more useful information. In addition, the cross-validation results (RD: mean AUC = 0.775 ± 0.077 ; DL: mean AUC = 0.888 ± 0.053) demonstrated consistent performance across folds, suggesting that the model's predictive capacity was not driven by overfitting to a specific subset of the data.

However, there are still limitations in this study. First, although this study was a multicenter study, the sample size was still small. DL models with complex neural networks require large data sets to avoid overfitting. Only 245 patients were included in this study, although the 10-fold cross-validation results indicated stable performance of the RD and DL model, further expansion of the sample size is essential to enhance the model's generalizability and ensure more reliable predictions in broader clinical applications. Second, although multicenter data were included, all participating centers were located within the same geographic region, which may limit the generalizability of our findings to broader populations. Additionally, the external validation cohort consisted of only 83 patients from two independent centers, which, while providing a preliminary assessment of the model's generalizability, remains relatively small. To further validate the robustness and applicability of our model, larger-scale external validation in geographically diverse cohorts is warranted. Ongoing collaborations with institutions in other regions are underway to facilitate prospective validation studies. Third, although all PTV delineations in this study followed standardized guidelines and were reviewed by experienced radiation oncologists, inter-observer and inter-institutional variability remains a challenge. Future studies should explore automated or consensus-based contouring methods to improve the reproducibility and generalizability of ROI-based modeling. Finally, this retrospective study lacked prospective data, which might introduce bias. Some patients were lost to follow-up, potentially affecting the results. Immunotherapy and chemotherapy, which had been shown to impact RP^{15,35}, were not included due to treatment heterogeneity and inconsistent recording across centers. Future research will involve prospective studies with standardized protocols and a broader range of clinical variables to enhance model accuracy.

Conclusion

In summary, our study found that radiomics, dosiomics, and 3D DL-derived features based on simulated location CT and dosimetry images were effective in predicting risk stratification of RP. The combined model (DLRDN) established in our study also achieved good results in the external validation cohort. In future clinical practice, this model could serve as a decision-support tool to stratify patients based on their predicted risk of RP. For those identified as high-risk, clinicians may consider individualized radiotherapy dose adjustments, more frequent imaging surveillance, early preventive interventions (such as the administration of glucocorticoids), or even alternative treatment strategies to mitigate toxicity and adverse effects.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 25 December 2024; Accepted: 12 May 2025

Published online: 16 May 2025

References

- Jia, B. et al. A nomogram model to predict death rate among non-small cell lung cancer (NSCLC) patients with surgery in surveillance, epidemiology, and end results (SEER) database. *BMC Cancer*. **20**, 666. <https://doi.org/10.1186/s12885-020-07147-y> (2020).
- Siegel, R. L., Miller, K. D., Wagle, N. S. & Jemal, A. Cancer statistics, 2023. *Cancer J. Clin.* **73**, 17–48. <https://doi.org/10.3322/caac.21763> (2023).
- Sung, H. et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J. Clin.* **71**, 209–249. <https://doi.org/10.3322/caac.21660> (2021).
- Yang, W. C., Hsu, F. M. & Yang, P. C. Precision radiotherapy for non-small cell lung cancer. *J. Biomed. Sci.* **27** <https://doi.org/10.1186/s12929-020-00676-5> (2020).
- Vinod, S. K. & Hau, E. Radiotherapy treatment for lung cancer: current status and future directions. *Respirol. (Carlton Vic)*. **25** (Suppl 2), 61–71. <https://doi.org/10.1111/resp.13870> (2020).
- Arroyo-Hernández, M. et al. Radiation-induced lung injury: current evidence. *BMC Pulm. Med.* **21** <https://doi.org/10.1186/s12890-020-01376-4> (2021).
- Liu, X., Shao, C. & Fu, J. Promising biomarkers of Radiation-Induced lung injury: A review. *Biomedicine* **9** <https://doi.org/10.3390/biomedicine9091181> (2021).
- Yang, S. et al. Risk analysis of grade ≥ 2 radiation pneumonitis based on radiotherapy timeline in stage III/IV non-small cell lung cancer treated with volumetric modulated Arc therapy: a retrospective study. *BMC Pulm. Med.* **22**, 402. <https://doi.org/10.1186/s12890-022-02211-8> (2022).
- Yu, H. et al. Machine learning to build and validate a model for radiation pneumonitis prediction in patients with Non-Small cell lung Cancer. *Clin. cancer Research: Official J. Am. Association Cancer Res.* **25**, 4343–4350. <https://doi.org/10.1158/1078-0432.Ccr-18-1084> (2019).
- Zhang, X. J. et al. Prediction of radiation pneumonitis in lung cancer patients: a systematic review. *J. Cancer Res. Clin. Oncol.* **138**, 2103–2116. <https://doi.org/10.1007/s00432-012-1284-1> (2012).
- Palma, D. A. et al. Predicting radiation pneumonitis after chemoradiation therapy for lung cancer: an international individual patient data meta-analysis. *Int. J. Radiat. Oncol. Biol. Phys.* **85**, 444–450. <https://doi.org/10.1016/j.ijrobp.2012.04.043> (2013).
- Glick, D. 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*. **19**, e219–e226. <https://doi.org/10.1016/j.clcc.2017.06.021> (2018).
- Chen, M., Copley, S. J., Viola, P., Lu, H. & Aboagye, E. O. Radiomics and artificial intelligence for precision medicine in lung cancer treatment. *Sem. Cancer Biol.* **93**, 97–113. <https://doi.org/10.1016/j.semcancer.2023.05.004> (2023).

14. Tunali, I., Gillies, R. J. & Schabath, M. B. Application of radiomics and artificial intelligence for lung Cancer precision medicine. *Cold Spring Harbor Perspect. Med.* **11** <https://doi.org/10.1101/cshperspect.a039537> (2021).
15. Zhang, Z. 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.* **115**, 746–758. <https://doi.org/10.1016/j.ijrobp.2022.08.047> (2023).
16. Wang, L. 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.* **111**, 443–455. <https://doi.org/10.1016/j.ijrobp.2021.04.047> (2021).
17. Krafft, S. P. et al. The utility of quantitative CT radiomics features for improved prediction of radiation pneumonitis. *Med. Phys.* **45**, 5317–5324. <https://doi.org/10.1002/mp.13150> (2018).
18. Huang, Y. et al. Radiation pneumonitis prediction after stereotactic body radiation therapy based on 3D dose distribution: dosiomics and/or deep learning-based radiomics features. *Radiation Oncol. (London England)*. **17**, 188. <https://doi.org/10.1186/s13014-022-02154-8> (2022).
19. Esteve, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118. <https://doi.org/10.1038/nature21056> (2017).
20. Li, X. et al. Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study. *Lancet Oncol.* **20**, 193–201. [https://doi.org/10.1016/s1470-2045\(18\)30762-9](https://doi.org/10.1016/s1470-2045(18)30762-9) (2019).
21. Nie, P. et al. A CT-based deep learning radiomics nomogram outperforms the existing prognostic models for outcome prediction in clear cell renal cell carcinoma: a multicenter study. *Eur. Radiol.* **33**, 8858–8868. <https://doi.org/10.1007/s00330-023-09869-6> (2023).
22. 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.* **110**, 1161–1170. <https://doi.org/10.1016/j.ijrobp.2021.01.049> (2021).
23. Meng, Y. et al. Excluding PTV from lung volume May better predict radiation pneumonitis for intensity modulated radiation therapy in lung cancer patients. *Radiation Oncol. (London England)*. **14**, 7. <https://doi.org/10.1186/s13014-018-1204-x> (2019).
24. Xiao, L. et al. Comparison of predictive powers of functional and anatomic dosimetric parameters for radiation-induced lung toxicity in locally advanced non-small cell lung cancer. *Radiotherapy Oncology: J. Eur. Soc. Therapeutic Radiol. Oncol.* **129**, 242–248. <https://doi.org/10.1016/j.radonc.2018.09.005> (2018).
25. Zha, Y. et al. A dynamic nomogram predicting symptomatic pneumonia in patients with lung cancer receiving thoracic radiation. *BMC Pulm. Med.* **24** <https://doi.org/10.1186/s12890-024-02899-w> (2024).
26. Liang, B. et al. Extracting 3D Spatial features from dose distribution to predict incidence of radiation pneumonitis. *Front. Oncol.* **9** <https://doi.org/10.3389/fonc.2019.00269> (2019).
27. Liang, B. et al. Prediction of radiation pneumonitis with dose distribution: A convolutional neural network (CNN) based model. *Front. Oncol.* **9** <https://doi.org/10.3389/fonc.2019.01500> (2019).
28. Nie, T. et al. Integration of dosimetric parameters, clinical factors, and radiomics to predict symptomatic radiation pneumonitis in lung cancer patients undergoing combined immunotherapy and radiotherapy. *Radiotherapy Oncology: J. Eur. Soc. Therapeutic Radiol. Oncol.* **190**, 110047. <https://doi.org/10.1016/j.radonc.2023.110047> (2024).
29. Mylona, E. et al. Voxel-Based analysis for identification of urethrovessical subregions predicting urinary toxicity after prostate Cancer radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **104**, 343–354. <https://doi.org/10.1016/j.ijrobp.2019.01.088> (2019).
30. Kawahara, D. et al. Prediction of radiation pneumonitis after definitive radiotherapy for locally advanced non-small cell lung cancer using multi-region radiomics analysis. *Sci. Rep.* **11**, 16232. <https://doi.org/10.1038/s41598-021-95643-x> (2021).
31. Ressa, G. et al. AI differentiates radionecrosis from true progression in brain metastasis upon stereotactic radiosurgery: analysis of 124 histologically assessed lesions. *Neuro Oncol.* <https://doi.org/10.1093/neuonc/noaf090> (2025).
32. Bin, L. et al. A deep learning-based dual-omics prediction model for radiation pneumonitis. *Med. Phys.* **48**, 6247–6256. <https://doi.org/10.1002/mp.15079> (2021).
33. Kong, Y. et al. Enhancing the prediction of symptomatic radiation pneumonitis for locally advanced non-small-cell lung cancer by combining 3D deep learning-derived imaging features with dose-volume metrics: a two-center study. *Strahlentherapie Und Onkologie: Organ. Der Deutschen Rontgengesellschaft ... Et Alj.* <https://doi.org/10.1007/s00066-024-02221-x> (2024).
34. Zhang, Z. et al. Computed tomography and radiation dose images-based deep-learning model for predicting radiation pneumonitis in lung cancer patients after radiation therapy. *Radiother. Oncol. J. Eur. Soc. Therapeutic Radiol. Oncol.* **182**, 109581. <https://doi.org/10.1016/j.radonc.2023.109581> (2023).
35. Chen, F., Niu, J., Wang, M., Zhu, H. & Guo, Z. Re-evaluating the risk factors for radiation pneumonitis in the era of immunotherapy. *J. Translational Med.* **21**, 368. <https://doi.org/10.1186/s12967-023-04212-5> (2023).

Acknowledgements

Some of our experiments were carried out on the Onekey AI platform. Thank Onekey AI and its developers' help in this scientific research work.

Author contributions

X.W., Conceptualization, Formal analysis, Methodology, Writing – original draft. A.-P. Z., H.-P. Y., Editing, Data curation, Investigation. G.-Q. Z., Resources, Methodology, Investigation. J.-L. M., Investigation, Data curation. S.-C. Y., Validation, Resources. S. G., Project administration, Supervision, Funding acquisition, Resources. All authors reviewed the manuscript.

Funding

This work was supported by the Jining City of Science and Technology Bureau (Key research and development project: 2023YXNS052), Shandong Province Medical Health Science and Technology Development Plan Project (202409030221), and, China International Medical Foundation (Cancer Precision Radiotherapy Spark Program: 2019-N-11-22).

Declarations

Competing interests

The authors declare no competing interests.

Ethical recognition and informed consent

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee (2023-12-C013) of the Affiliated Hospital of Jining Medical University, Tumor Hospital of Jining, and Jining First People's Hospital. Informed consent was waived due to the retrospective nature of this study.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-02045-4>.

Correspondence and requests for materials should be addressed to S.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025