


# Development and Validation of a Comprehensive Model for Predicting Distant Metastasis of Solid Lung Adenocarcinoma: 3D Radiomics, 2D Radiomics and Clinical Features

Zhiwei Peng\*, Ze Lin\*, Anjing He\*, Li Yi, Mengni Jin, Zhiyong Chen, Yahong Tao, Yuhua Yang, Can Cui, Ying Liu, Minjing Zuo 

Department of Radiology, The Second Affiliated Hospital of Nanchang University, Nanchang, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Minjing Zuo, Email [minjingzuo@163.com](mailto:minjingzuo@163.com)

**Objective:** To develop and validate models for predicting distant metastases in patients with solid lung adenocarcinomas using 3D radiomic features, 2D radiomic features, clinical features, and their combinations.

**Methods:** This retrospective study included 253 eligible patients with solid adenocarcinoma of the lung diagnosed at our hospital between August 2018 and August 2021. 3D and 2D regions of interest were segmented from computed tomography-enhanced thin-slice images of the venous phase, and 851 radiomic features were extracted in each region. The Least Absolute Shrinkage and Selection Operator (LASSO) algorithm was used to select radiomic features and calculate radiomic scores, and logistic regression was used to develop the model. Development of a 3D radiomics model (model 1), a 2D radiomics model (model 2), a combined 3D radiomics and 2D radiomics model (model 3), a clinical model (model 4), and a comprehensive model (model 5) for the prediction of distant metastases in patients with solid lung adenocarcinomas. Nomograms were drawn to illustrate model 5, and receiver operating characteristic (ROC) curve, calibration curve, and decision curve analysis (DCA) were used for model evaluation.

**Results:** The AUC (area under the curve) of model 1, model 2, model 3, model 4, and model 5 in the test set was 0.711, 0.769, 0.775, 0.829, and 0.892, respectively. The Delong test showed that AUC values were statistically different between model 5 and model 1 ( $p=0.001$ ), and there was no statistical difference in AUC between the other models. Based on a comprehensive review of DCA, ROC curve, and Akaike information criterion (AIC), Model 5 is demonstrated to have better clinical utility, goodness of fit, and parsimony.

**Conclusion:** A comprehensive model based on 3D radiomic features, 2D radiomic features, and clinical features has the potential to predict distant metastasis in patients with solid lung adenocarcinomas.

**Keywords:** computed tomography, radiomics, lung neoplasms, adenocarcinoma of lung, distant metastasis

## Introduction

According to the Global Cancer Statistics Report 2020, lung cancer is the second most common cancer after breast cancer, and its mortality rate ranks first among all cancers.<sup>1</sup> Adenocarcinoma of the lung is the most common pathological form of lung cancer and occurs in the lung tissue distal to the bronchi of the lung segment. Most imaging manifestations are peripheral lung carcinomas, and early symptoms are often not obvious.<sup>2</sup> Pure ground-glass and partially solid lung adenocarcinomas rarely develop lymph node and distant metastases.<sup>3,4</sup> Lung adenocarcinomas with a ground glass component have excellent survival rates compared with solid lesions.<sup>5,6</sup> Solid lung adenocarcinomas are more prone to distant metastases and are usually unresectable, requiring chemoradiotherapy, targeted therapy, or immunotherapy.<sup>7</sup> Therefore, accurately determining whether patients with solid lung adenocarcinomas have developed distant metastases prior to treatment is of great importance for clinical staging and treatment decisions.

Histopathologic examination is the gold standard for the diagnosis of distant metastases in lung cancer,<sup>8</sup> but it is invasive and cannot be used for screening. Many non-invasive imaging modalities, such as computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography/computed tomography (PET/CT), are widely used to evaluate distant metastases in lung cancer patients.<sup>9,10</sup> CT and MRI are the most commonly used clinical examination modalities for the diagnosis of distant metastases, but they often require systemic examination of the head, neck, chest and abdomen. The radiation exposure, high cost, and high false-positive rate of PET/CT limit its wide clinical application.<sup>11</sup> Therefore, there is an urgent need for a comprehensive, accurate, simple, and noninvasive method to assess distant metastasis in lung cancer patients.

In recent years, with the rapid development of artificial intelligence, radiomics has become established in the field of radiology. Radiomics refers to the extraction and analysis of quantitative image features, invisible to the naked eye, from medical images, combined with machine learning or deep learning algorithms to investigate their possible links to biological and clinical outcomes.<sup>12</sup> Radiomics has been studied in the aspects of pathological detection, diagnosis prediction, gene mutation, treatment strategy, follow-up, efficacy evaluation, and prognosis of lung cancer.<sup>13</sup> In addition, radiomics has obvious advantages in evaluating tumor heterogeneity and may be able to comprehensively and accurately predict distant metastasis of adenocarcinoma of the lung.

In some previous studies, Chen et al<sup>14</sup> limited themselves to predicting brain metastases in lung adenocarcinomas, and Tao et al<sup>15</sup> did not exclude subsolid lung carcinomas, which rarely have distant metastases. In addition, most previous studies extracted only three-dimensional radiomic features from CT images, and few studies extracted two-dimensional radiomic features. In the present study, we developed and validated a comprehensive model based on contrast-enhanced CT-based 3D and 2D radiomic features combined with clinical parameters to predict distant metastasis in patients with solid lung adenocarcinomas.

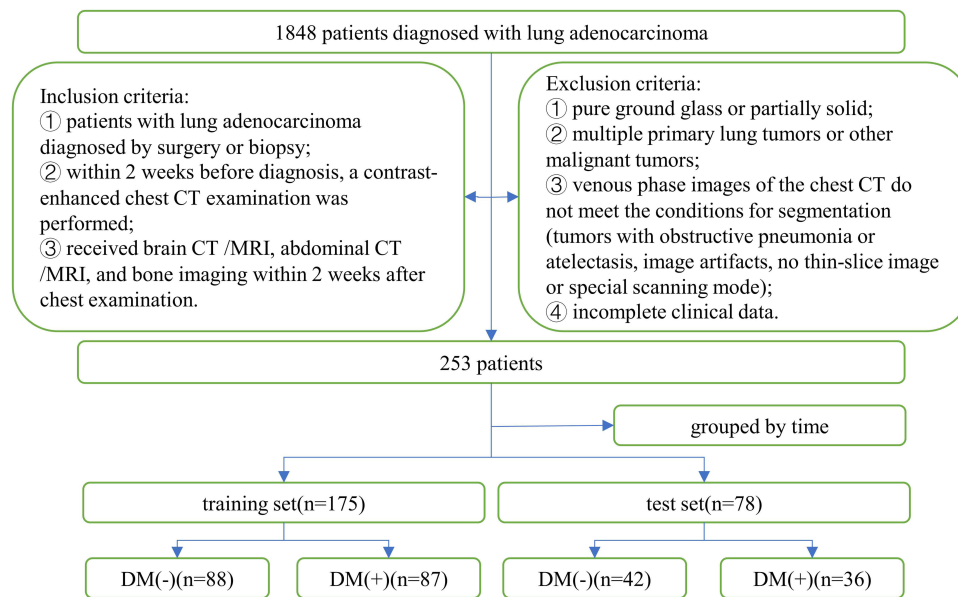
## Methods

### Patients and Clinical Parameters

This study retrospectively analyzed 1848 patients with lung adenocarcinoma who were pathologically diagnosed at our hospital between August 2018 and August 2021. Inclusion criteria: (1) patients with lung adenocarcinoma diagnosed by surgery or biopsy; (2) within 2 weeks before diagnosis, a contrast-enhanced chest CT examination was performed; (3) received brain CT/MRI, abdominal CT/MRI, and bone imaging within 2 weeks after chest examination. Exclusion criteria: (1) pure ground glass or partially solid lung adenocarcinoma; (2) multiple primary lung tumors or other malignant tumors; (3) venous phase images of the chest CT do not meet the conditions for segmentation (tumors with obstructive pneumonia or atelectasis, image artifacts, no thin-slice image or special scanning mode); (4) incomplete clinical data.

Finally, 253 patients with solid adenocarcinoma of the lung were included, and 123 patients (48.6%) developed distant metastases before treatment. A total of 175 patients before November 1, 2020, were classified as the training set, of which 87 patients had distant metastases and 88 patients did not. Seventy-eight patients after November 1, 2020, constituted the test set, of which 36 patients had distant metastases and 42 patients did not. [Figure 1](#) shows the inclusion criteria, exclusion criteria, and grouping of patients.

In addition, the clinical parameters of the patients were recorded: Age, sex, smoking status, carcinoembryonic antigen (CEA), carbohydrate antigen 199 (CA199), location and largest diameter of the primary tumor, short diameter of the largest lymph node, clinical TNM stage, and CT semantic features (lobulation, burr, pleural indentation, cavity). The CT semantic features were observed and analyzed by two radiologists (with 4 and 8 years of professional experience, respectively) without knowledge of the results of distant metastases. Lobulation sign: defined as at least three arcuate depressions at the tumor margin; burr sign: defined as a line extending from the tumor margin to the surrounding parenchyma without reaching the pleura; pleural depression sign: referring to the tumor pulling the adjacent pleura; cavity: at least 5mm air shadow inside the tumor. Clinical TNM staging is based on the 8th edition of the TNM staging criteria for lung cancer published by the International Association for the Study of Lung Cancer (IASLC).<sup>16</sup> Distant



**Figure 1** Patient recruitment process.  
**Abbreviation:** DM, distant metastases.

metastases are diagnosed based on histologic examination or apparent progression or shrinkage of lesions after antitumor treatment.

## CT Image Acquisition

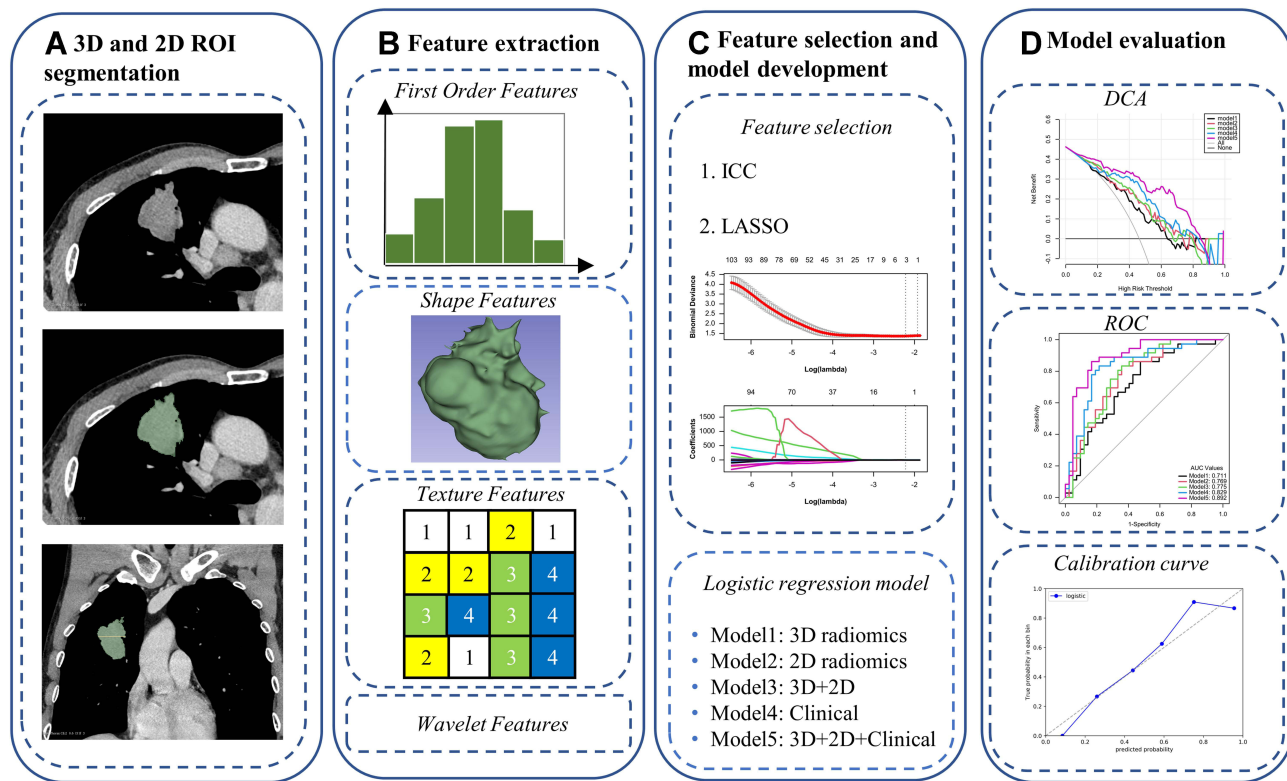
All patients underwent enhanced chest CT scan within 2 weeks before diagnosis. Scans were performed with the Siemens SOMATOM Definition Flash CT and GE REVOLUTION HD CT. The patient was examined in supine position. The scan was started after the patient stopped breathing at the end of inspiration, and the scan extended from the base of the lung to the level of the thoracic inlet. The slice thickness during image reconstruction was 1.0 or 1.25 mm. Other relevant scan parameters can be found in [Supplementary Table 1](#).

## Segmentation and Feature Extraction

The region of interest (ROI) of the primary tumor was segmented by a radiologist using 3D Slicer software (version 4.11; <https://www.slicer.org/>), and the ROI included bronchi, blood vessels, necrosis and calcification within the lesion, etc, but not normal lung tissue. Because the included lesions were solid lung adenocarcinomas, delineation was performed at the mediastinal window (window width 350, window level 40) and a 3D ROI was obtained by setting the lower threshold of the mediastinal window (-135HU) to delineate each layer of the lesion layer by layer semi-automatically; in combination with coronal and sagittal images, the largest cross-section of the lesion was selected to outline a 2D ROI. Radiomic feature extraction using PyRadiomics in the 3D slicer extension extracted 851 features per ROI, including 14 shape features, 18 first-order features, 24 Gray Level Co-occurrence Matrix (GLCM) features, 14 Gray Level Dependence Matrix (GLDM) features, 16 Gray Level Run Length Matrix (GLRLM) features, 16 Gray Level Size Zone Matrix (GLSZM) features, 5 Neighboring Gray Tone Difference Matrix (NGTDM) features, and 744 wavelet features (all feature names are listed in [Supplementary Table 2](#)).

The settings for image preprocessing are as follows: Gray level discretization (the bin width is fixed at 25), and the image is resampled into voxels of size 1\*1\*1.

To assess the reproducibility of the radiomics signature, we randomly selected 30 patients who were segmented by another radiologist 1 month later. None of them had information about clinical and pathologic details. Subsequently, the features extracted from the two segmentations were subjected to inter-rater correlation coefficient (ICC) analysis, and the features with an ICC  $\geq 0.75$  were selected. [Figure 2](#) is a flowchart of this study.



**Figure 2** Flowchart of this study: 3D and 2D ROI segmentation (A), feature extraction (B), feature selection and model development (C), and model evaluation (D).

### Feature Selection, and Model Development and Validation

Radiomic features were normalized with the Z-score before feature screening. The 3D and 2D radiomic features were screened separately using the LASSO algorithm. Univariate analysis and multivariate logistic regression analysis were performed to search for meaningful clinical parameters. Training data were used for both feature screening and modeling. A 3D radiomics model (model 1), the 2D radiomics model (model 2), a combined 3D radiomics and 2D radiomics model (model 3), a clinical model (model 4), and a comprehensive model consisting of 3D radiomics, 2D radiomics, and clinical parameters (model 5) were developed.

To evaluate the predictive power of each model, the ROC curve was constructed and AUC, accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated. Delong’s test was used to compare AUC between models. DCA was used to evaluate the clinical utility of each model.<sup>17,18</sup> AIC is used to compare and rank multiple models and emphasizes comparison of the goodness of fit of the models, taking into account the principle of parsimony. In general, the smaller the AIC value, the better the model fit and the simpler the model.<sup>19</sup> The nomogram is used to represent model 5. The value assigned to each variable is scored on a scale from 0 to 100. The total score is obtained by adding the scores for each variable. The risk of distant metastases in patients with solid adenocarcinoma of the lung can be predicted by mapping the total score to the probability of the bottom. The calibration curve is used to evaluate the calibration degree of model 5.

### Statistical Analysis

Description of clinical statistical method: a chi-square test or Fisher’s exact test was used for the nominal variable and expressed by the number and proportion. Continuous variables conforming to normal distribution are expressed by two independent samples *t*-test and mean and standard deviation. For continuous variable with abnormal distribution, Mann–Whitney *U*-test was used and expressed by median and quartile. All statistical analysis for the present study were

performed using R (version 4.0.4; <https://www.r-project.org>) and Python (version 3.5.6; <https://www.python.org>). A two-sided p-value <0.05 indicated statistical significance.

## Results

### Clinical Characteristics

The patient baseline data sheet showed that there were essentially no significant differences in clinical characteristics between the training and test sets (see [Supplementary Table 3](#)). The results of the univariate analysis of the training set showed that there were statistically significant differences in burr sign, T stage, N stage, age, CEA, CA199, the largest diameter of the tumor, and the short diameter of the largest lymph node ([Table 1](#)). Immediately thereafter, multivariate logistic regression results showed that N stage, age, and CEA were independent predictors of distant metastases ([Table 2](#)).

**Table 1** Univariate Analysis of the Training Set and the Test Set

Variable	Training Set (n=175)			Test Set (n=78)		
	DM(-)	DM(+)	P-value	DM(-)	DM(+)	P-value
Sex			0.705			0.507
Male	43(48.86%)	45(51.72%)		23(54.76%)	17(47.22%)	
Female	45(51.14%)	42(48.28%)		19(45.24%)	19(52.78%)	
Smoking			0.375			0.213
No	67(76.14%)	71(81.61%)		30(71.43%)	30(83.33%)	
Yes	21(23.86%)	16(18.39%)		12(28.57%)	6(16.67%)	
Tumor location1			0.372			0.306
Right upper lobe	35(39.77%)	28(32.18%)		16(38.10%)	11(30.56%)	
Right middle lobe	7(7.95%)	7(8.05%)		3(7.14%)	6(16.67%)	
Right lower lobe	10(11.36%)	16(18.39%)		7(16.67%)	2(5.56%)	
Left upper lobe	19(21.59%)	25(28.74%)		9(21.43%)	12(33.33%)	
Left lower lobe	17(19.32%)	11(12.64%)		7(16.67%)	5(13.89%)	
Tumor location2			0.823			0.001
Peripheral	70(79.55%)	68(78.16%)		38(90.48%)	21(58.33%)	
Central	18(20.45%)	19(21.84%)		4(9.52%)	15(41.67%)	
Lobulation			0.398			0.001
No	42(47.73%)	36(41.38%)		28(66.67%)	11(30.56%)	
Yes	46(52.27%)	51(58.62%)		14(33.33%)	25(69.44%)	
Burr			0.014			0.421
No	66(75.00%)	50(57.47%)		28(66.67%)	27(75.00%)	
Yes	22(25.00%)	37(42.53%)		14(33.33%)	9(25.00%)	
Pleural_indentation*			0.99			0.722
No	3(3.41%)	2(2.30%)		3(7.14%)	1(2.78%)	
Yes	85(96.59%)	85(97.70%)		39(92.86%)	35(97.22%)	
Cavity			0.669			0.722
No	73(82.95%)	70(80.46%)		40(95.24%)	34(94.44%)	
Yes	15(17.05%)	17(19.54%)		2(4.76%)	2(5.56%)	
T stage			0.017			0.001
T1	30(34.09%)	20(22.99%)		19(45.24%)	4(11.11%)	
T2	43(48.86%)	35(40.23%)		19(45.24%)	20(55.56%)	
T3	6(6.82%)	20(22.99%)		4(9.52%)	7(19.44%)	
T4	9(10.23%)	12(13.79%)		0(0.00%)	5(13.89%)	
N stage			<0.001			<0.001
N0	60(68.18%)	23(26.44%)		33(78.57%)	13(36.11%)	
N1	11(12.50%)	10(11.49%)		4(9.52%)	2(5.56%)	
N2	13(14.77%)	39(44.83%)		3(7.14%)	13(36.11%)	
N3	4(4.55%)	15(17.24%)		2(4.76%)	8(22.22%)	

(Continued)

**Table 1** (Continued).

Variable	Training Set (n=175)			Test Set (n=78)		
	DM(-)	DM(+)	P-value	DM(-)	DM(+)	P-value
Age	62.50(54.45, 68.55)	66.00(59.00, 71.80)	0.028	61.43±8.29	63.64±10.61	0.315
CEA	3.10(1.59, 8.19)	15.23(4.26, 36.79)	<0.001	2.22(1.59, 4.47)	17.98(8.17, 100.00)	<0.001
CA199	13.34(9.41, 19.54)	20.19(11.95, 61.12)	0.001	13.16(9.64, 21.98)	24.52(10.76, 41.59)	0.018
T-max	33.00(25.00, 45.00)	40.00(28.00, 54.00)	0.028	24.00(18.00, 32.00)	42.00(35.35, 49.00)	<0.001
N-max	8.00(7.00, 10.55)	12.00(9.00, 15.00)	<0.001	8.00(7.00, 9.00)	10.50(8.00, 14.00)	<0.001

Note: \*Variable use corrected chi-square test.

Abbreviations: DM, distant metastasis; T-max, maximum diameter of primary tumor; N-max, short diameter of the largest lymph node.

**Table 2** Multivariate Logistic Regression Analysis of the Training Set

Variable	OR (95% CI)	P-value
Burr	1.833(0.825,4.072)	0.137
T stage	1.664(0.839,3.300)	0.145
N stage	1.896(1.127,3.189)	0.016
Age	1.042(1.001,1.084)	0.043
CEA	1.025(1.006,1.044)	0.008
CA199	1.003(0.999,1.007)	0.119
T-max	0.978(0.940,1.017)	0.271
N-max	1.083(0.958,1.225)	0.203

Abbreviations: OR, odds ratio; CI, confidence interval.

## Consistency Test

ICC was applied to measure the consistency between the two image segmentation and feature extraction processes. The ICC cut-off values recommended by Koo et al<sup>20</sup> are as follows: poor reliability:  $ICC < 0.5$ ; moderate reliability:  $0.5 \leq ICC < 0.75$ ; good reliability:  $0.75 \leq ICC < 0.9$ ; excellent reliability:  $0.9 \leq ICC$ . A total of 815 (95.8%) 3D radiomic features and 620 (72.9%) 2D radiomic features had good reliability ( $ICC \geq 0.75$ ), indicating that the extracted radiomic features were well reproducible. The results for ICC are described in more detail in [Supplemental Table 4](#).

## Selection of Radiomic Features and Establishment of Five Models

A total of 815 reliable 3D radiomic features were screened using LASSO, and the remaining 3 features were used to create Model 1. These three radiomic features were original\_gldm\_DependenceEntropy, original\_glszm\_ZoneEntropy, and wavelet-LLL\_glszm\_ZoneEntropy.

A total of 620 reliable 2D radiomic features were screened using LASSO ([Supplementary Figure 1](#)), and the remaining 6 features were used to build Model 2. These six radiomic features were original\_shape\_Elongation, original\_gldm\_DependenceEntropy, original\_glrIm\_RunEntropy, wavelet-LHL\_glcM\_Contrast, wavelet-LHL\_glrIm\_RunVariance, wavelet-LHH\_glcM\_Imc2.

We combined the above three radiomic features used to construct model 1 and the six radiomic features used to construct model 2 to construct model 3, and calculated the combined radiomics score (c-radscore) according to the weighting coefficient of the features. For more details, see [Supplemental Table 5](#) and [Supplemental Figure 2](#).

N stage, age, and CEA with statistical significance in univariate and multivariate analysis were selected to build the model 4.

Select N stage, age, CEA, and c-radscore to build model 5.

**Table 3** Performance Evaluation of Five Models in the Training Set and the Test Set

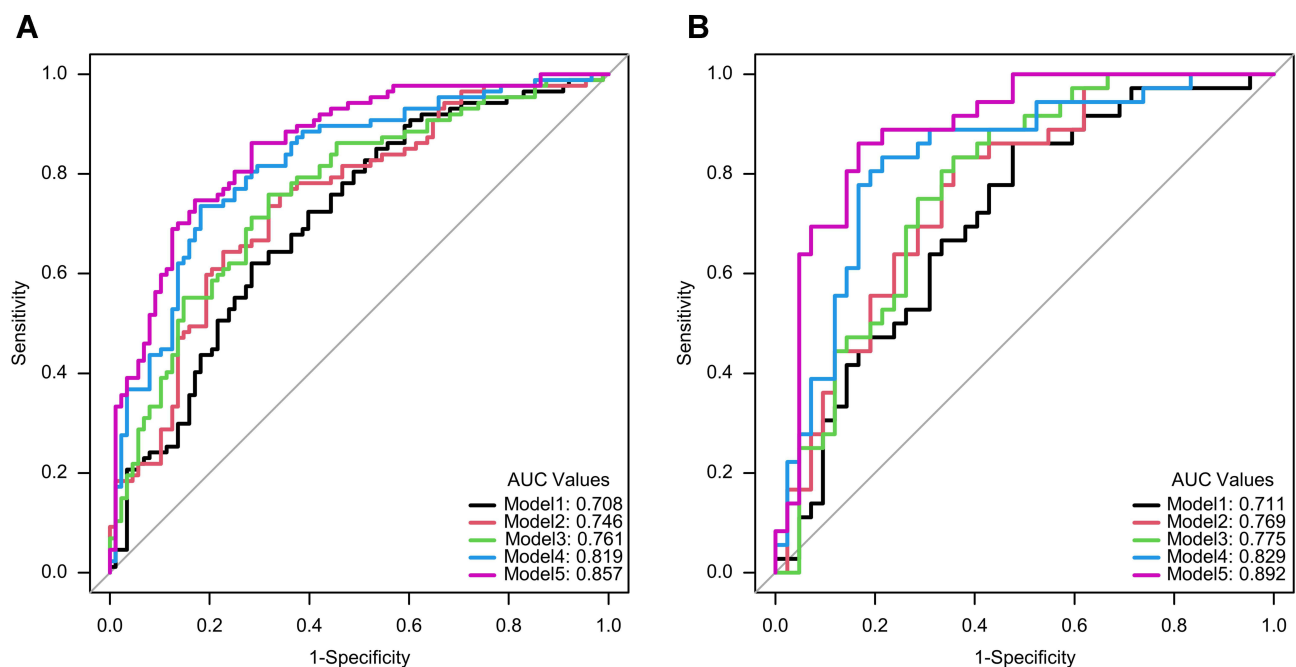
Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV
Training set						
Model1	0.708 (0.645, 0.773)	0.663	0.609	0.716	0.679	0.649
Model2	0.746 (0.682, 0.805)	0.703	0.747	0.659	0.684	0.725
Model3	0.761 (0.701, 0.819)	0.714	0.747	0.682	0.699	0.732
Model4	0.819 (0.763, 0.872)	0.771	0.724	0.818	0.797	0.75
Model5	0.857 (0.808, 0.901)	0.783	0.851	0.716	0.747	0.829
Test set						
Model1	0.711 (0.612, 0.804)	0.667	0.639	0.69	0.639	0.69
Model2	0.769 (0.681, 0.852)	0.692	0.694	0.69	0.658	0.725
Model3	0.775 (0.686, 0.86)	0.705	0.75	0.667	0.659	0.757
Model4	0.829 (0.752, 0.904)	0.769	0.694	0.833	0.781	0.761
Model5	0.892 (0.829, 0.949)	0.833	0.889	0.786	0.78	0.892

**Abbreviations:** AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

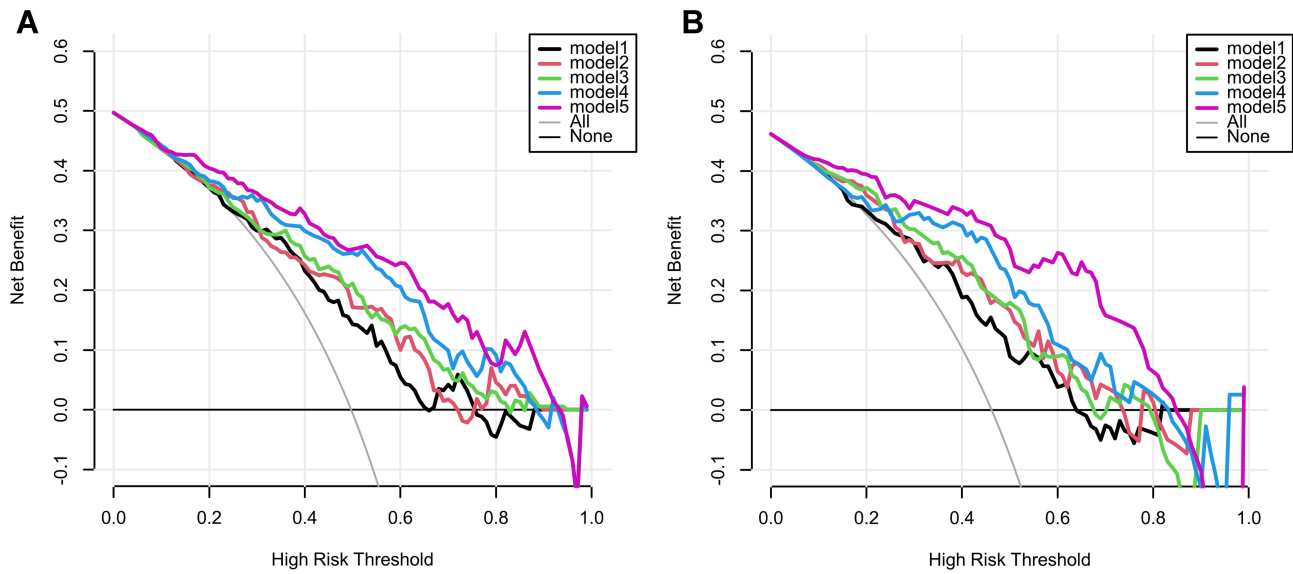
## Performance of the Five Models

In the training set and test set, the AUC values of model 1 for predicting distant metastasis were 0.708 (0.645, 0.773) and 0.711 (0.612, 0.804), respectively, and the AUC values of model 2 were 0.746 (0.682, 0.805) and 0.769 (0.681, 0.852), the AUC values of model 3 were 0.761 (0.701, 0.819) and 0.775 (0.686, 0.86), the AUC values of model 4 were 0.819 (0.763, 0.872) and 0.829 (0.752, 0.904), and the AUC values of model 5 were 0.857 (0.808, 0.901) and 0.892 (0.829, 0.949), respectively. The predictive performance of each model is shown in Table 3, and the ROC curve is shown in Figure 3.

The Delong test showed that the AUC values of model 5 and model 1 were statistically different in the test set ( $p=0.001 < 0.05/10$ , corrected by Bonferroni method<sup>21</sup>), and there was no statistical difference in AUC between the other models (see Supplementary Table 6 for details). DCA was used to evaluate the clinical utility of five predictive models by calculating the net gains at different probability thresholds. DCA (Figure 4) shows that model 5 predicts a higher net



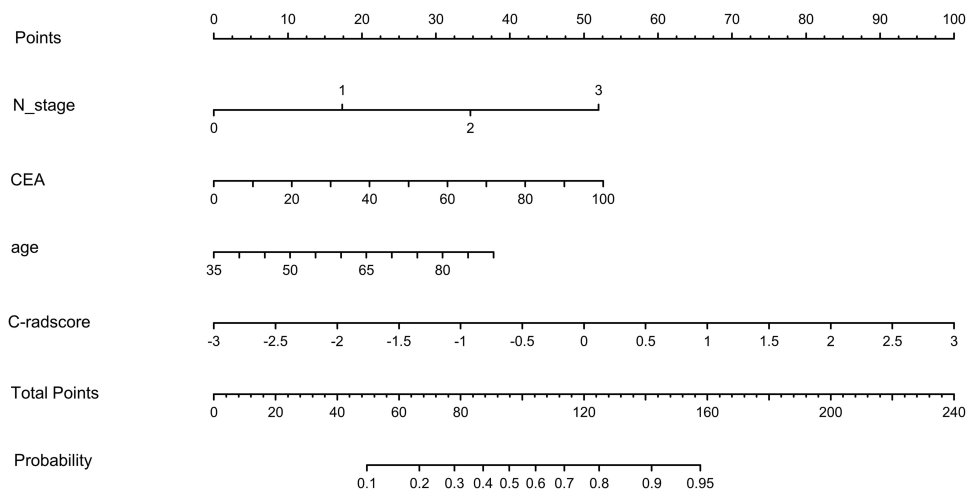
**Figure 3** Receiver operating characteristic curves of five models in training set (A) and test set (B).



**Figure 4** Decision curve analysis for five models in training set (A) and test set (B). The y-axis represents the net benefit and the x-axis represents the threshold probability. The black, red, green, blue, and purple lines represent the net gains for model 1, model 2, model 3, model 4, and model 5, respectively. The gray line represents the assumption that all patients had distant metastases. The black line represents the assumption that all patients had no distant metastases.

benefit for distant metastases than the other four models and all or no treatment strategy in both training and test sets when the threshold probability is approximately between 10% and 90%. The AIC values of the five models were 228.521, 223.886, 224.871, 192.447, and 176.645, with the AIC value of model 5 being the smallest. Based on a comprehensive consideration of DCA, ROC curve and AIC, it is proved that model 5 has better clinical utility, goodness of fit and parsimony.

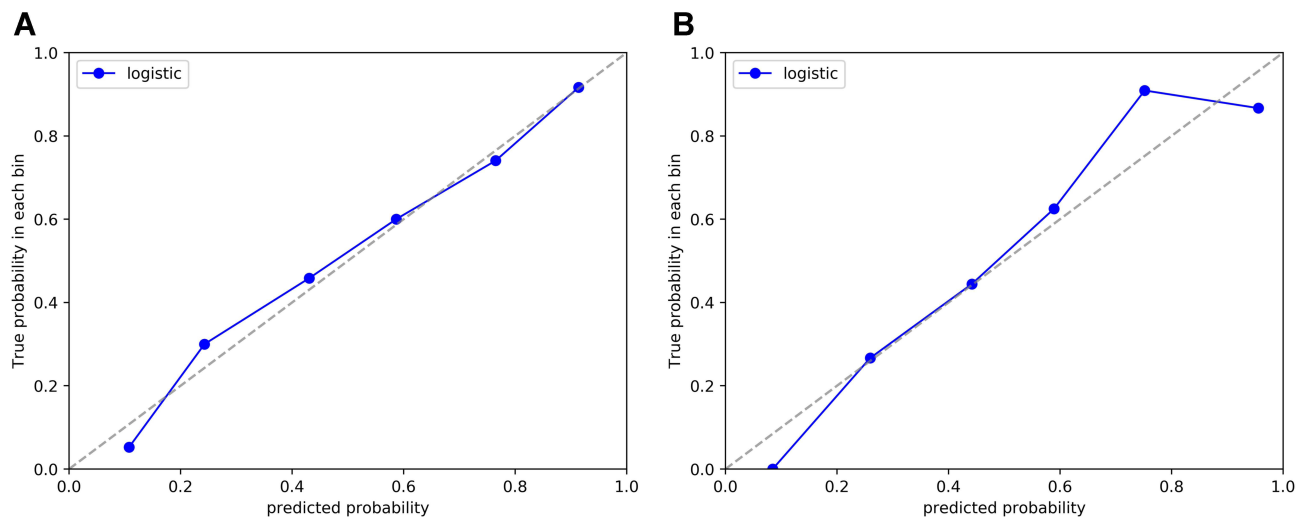
Therefore, we created a nomogram of model 5 to facilitate clinical application (Figure 5). In addition, the calibration curve for model 5 (Figure 6) showed good agreement between its prediction of distant metastases and the actual situation. The Hosmer–Lemeshow test showed that model 5 was a good fit and was not statistically different from the ideal perfect model ( $p=0.642 > 0.05$ ).



**Figure 5** Nomogram for model 5.

**Abbreviation:** c-radscore, radiomic score combining 3D and 2D radiomic features.





**Figure 6** Calibration curves of model 5 in training set (A) and test set (B).

## Discussion

In this study, we developed five models for predicting distant metastasis in patients with solid lung adenocarcinoma based on 3D radiomics, 2D radiomics of contrast-enhanced CT, and clinical features and evaluated them with the test set. Based on a comprehensive consideration of DCA, the ROC curve and AIC, it is proved that model 5 has better clinical utility, goodness of fit and parsimony than the other four models, and a nomogram is constructed for the convenience of clinicians. In some previous studies, Chen et al<sup>14</sup> retrospectively included 89 patients with T1 lung adenocarcinoma to predict the occurrence of brain metastases. The AUC of the combined model of radiomics and clinical parameters was 0.871, and the sample size of this study was small with only 10-fold cross-validation. Tao et al<sup>15</sup> extracted radiomic features from preoperative CT images of 356 non-small cell lung cancer patients to predict distant metastases. The combined model of radiomic and clinical parameters had an AUC of only 0.77 in the test set. The study did not exclude subsolid lung cancers that rarely develop distant metastases.

Our study follows the recommendations of previous studies to avoid the fragility of radiomics model development<sup>22</sup>: (1) The code package for radiomic feature extraction is the Python-based code package PyRadiomics, which has been used by many studies and validated by research.<sup>23,24</sup> (2) We combined 3D radiomic features, 2D radiomic features, and clinical parameters to create a more comprehensive model. (3) This study improved the repeatability of radiomic features by setting threshold for semi-automatic delineation. The studies by Cohen et al<sup>25</sup> and Scholten et al<sup>26</sup> also confirm the accuracy of segmentation using the threshold method. (4) Multiple segmentations are performed to further check the repeatability of the radiomic features and remove the radiomic features with poor repeatability. (5) Pre-process the images from CT to reduce the influence of different devices and different slice thicknesses. (6) Minimize features in model development to reduce the risk of overfitting. (7) Temporal validation: according to Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD), temporal grouping is better than random grouping for evaluating model performance because it allows non-random variation between the two data sets.<sup>27,28</sup> (8) This study was assessed using the Radiomics Quality Score (RQS), with a score of 19/36, higher than most studies.<sup>29,30</sup>

In consistency tests, we found considerable reliability in the extraction of radiomic features. Overall, good inter-observer reliability ( $ICC \geq 0.75$ ) was achieved for 95.8% of 3D radiomic features and 72.9% of 2D radiomic features. This might reflect to some extent the stability and generalization of the radiomics model. According to the ICC results, the 3D radiomic features were more reproducible than the 2D radiomic features. However, according to the results of the DeLong test, the AUC values of model 1 and model 2 are not statistically different, which could be because the features with good reliability in the 2D radiomic features also have high predictive performance. Yang et al<sup>31</sup> also demonstrated the predictive performance of 2D texture features in predicting angiolymphatic vessel invasion in lung adenocarcinoma with AUCs of 0.856 and 0.701 for 2D and 3D features in the test set, respectively. Wang et al<sup>32</sup> combined the radiomic

features of tumor 2D, tumor 3D, and peritumoral regions to predict recurrence-free survival in patients with stage IA solid non-small cell lung cancer. The 3-year and 5-year AUCs were 0.76 and 0.75, respectively, but our study combined 3D and 2D radiomic features and showed good predictive performance in predicting distant metastasis in patients with lung adenocarcinoma (model 3 had an AUC of 0.775 on the test set).

The histopathologic grade and number of lymph node metastases in patients with solid lung adenocarcinoma may be important predictors of distant metastasis. However, this information needs to be confirmed by surgery or biopsy and cannot be used as parameters to predict distant metastasis before treatment. Therefore, we collected as many pre-treatment clinical parameters of patients as possible: age, sex, smoking status, CEA, CA199, location and largest diameter of the primary tumor, short diameter of the largest lymph node, clinical T and N stage, and CT semantic features. The results of univariate analysis of the training set showed that there were statistically significant differences in burr sign, T stage, N stage, age, CEA, CA199, largest diameter of tumor, and short diameter of largest lymph node, which was relatively similar to the results of Tao et al.<sup>15</sup> In multivariate analysis, N stage, age, and CEA were found to be independent predictors of distant metastasis in patients with solid lung adenocarcinomas. The constructed model 4 also showed good predictive performance with an AUC of 0.829 on the test set.

This study has some limitations. First, it is a retrospective study and selection bias may occur. Second, the data are from a single institution and the sample size was not very large, although we chose to use temporal grouping for validation. Third, we segmented only CT venous phase images that were more suitable for tumor boundary segmentation because of the workload. Fourth, the predictive performance of the model in different subgroups was not evaluated. In the future, prospective multicenter studies in more populations are needed to further validate our results.

## Conclusion

This study demonstrates that a comprehensive model based on 3D radiomics, 2D radiomics, and clinical features has the potential to predict distant metastasis in patients with solid lung adenocarcinomas.

## Abbreviations

CT, computed tomography; MRI, magnetic resonance imaging; 3D, three-dimensional; 2D, two-dimensional; LASSO, The Least Absolute Shrinkage and Selection Operator; ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval; DCA, decision curve analysis; AIC, Akaike information criterion; ROI, region of interest; ICC, inter-rater correlation coefficient; c-radscore, combined radiomics score; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen199.

## Ethics Approval and Consent to Participate

This study was conducted with approval from the Second Affiliated Hospital of Nanchang University Medical Research Ethics Committee (The Examination and Approval No. Review 2018 No.031 and date of approval is June 10, 2018). This study was conducted in accordance with the declaration of Helsinki. As this study was a retrospective cohort study, the data were anonymous and informed consent was not required.

## Acknowledgments

Thanks to all the authors for their contributions to the study.

## Funding

This study was funded by the Key Project of Science and Technology Program of Jiangxi Provincial Department of Education (GJJ200106) and the Applied Research Cultivation Program of Jiangxi Provincial Department of Science and Technology (20212BAG70048).

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–249. doi:10.3322/caac.21660
2. Succony L, Rassl DM, Barker AP, et al. Adenocarcinoma spectrum lesions of the lung: detection, pathology and treatment strategies. *Cancer Treat Rev.* 2021;99:102237. doi:10.1016/j.ctrv.2021.102237
3. Ye T, Deng L, Wang S, et al. Lung adenocarcinomas manifesting as radiological part-solid nodules define a special clinical subtype. *J Thorac Oncol.* 2019;14(4):617–627. doi:10.1016/j.jtho.2018.12.030
4. Cho H, Lee HY, Kim J, et al. Pure ground glass nodular adenocarcinomas: are preoperative positron emission tomography/computed tomography and brain magnetic resonance imaging useful or necessary? *J Thorac Cardiovasc Surg.* 2015;150(3):514–520. doi:10.1016/j.jtcvs.2015.06.024
5. Zhang Y, Fu F, Chen H. Management of ground-glass opacities in the lung cancer spectrum. *Ann Thorac Surg.* 2020;110(6):1796–1804. doi:10.1016/j.athoracsur.2020.04.094
6. Shigefuku S, Shimada Y, Hagiwara M, et al. Prognostic significance of ground-glass opacity components in 5-year survivors with resected lung adenocarcinoma. *Ann Surg Oncol.* 2021;28(1):148–156. doi:10.1245/s10434-020-09125-x
7. Ettinger DS, Wood DE, Aisner DL, et al. NCCN guidelines insights: non-small cell lung cancer, version 2.2021. *J Natl Compr Canc Netw.* 2021;19(3):254–266. doi:10.6004/jncn.2021.0013
8. Sica GL, Gal AA. Lung cancer staging: pathology issues. *Semin Diagn Pathol.* 2012;29(3):116–126. doi:10.1053/j.semdp.2012.06.001
9. Kandathil A, Kay FU, Butt YM, et al. Role of FDG PET/CT in the eighth edition of TNM Staging of non-small cell lung cancer. *Radiographics.* 2018;38(7):2134–2149. doi:10.1148/rg.2018180060
10. Sánchez DCJ, Hernández JH, López MF, et al. SEPAR guidelines for lung cancer staging. *Arch Bronconeumol.* 2011;47(9):454–465. doi:10.1016/j.arbres.2011.06.013
11. Liam CK, Andarini S, Lee P, et al. Lung cancer staging now and in the future. *Respirology.* 2015;20(4):526–534. doi:10.1111/resp.12489
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology.* 2016;278(2):563–577. doi:10.1148/radiol.2015151169
13. Wu G, Jochems A, Refaee T, et al. Structural and functional radiomics for lung cancer. *Eur J Nucl Med Mol Imaging.* 2021;48(12):3961–3974. doi:10.1007/s00259-021-05242-1
14. Chen A, Lu L, Pu X, et al. CT-based radiomics model for predicting brain metastasis in category T1 lung adenocarcinoma. *Am J Roentgenol.* 2019;213(1):134–139. doi:10.2214/AJR.18.20591
15. Tao J, Lv R, Liang C, et al. Development and validation of a CT-based signature for the prediction of distant metastasis before treatment of non-small cell lung cancer. *Acad Radiol.* 2021. doi:10.1016/j.acra.2020.12.007
16. Travis WD, Asamura H, Bankier AA, et al. The IASLC lung cancer staging project: proposals for coding T categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the TNM classification of lung cancer. *J Thorac Oncol.* 2016;11(8):1204–1223. doi:10.1016/j.jtho.2016.03.025
17. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ.* 2016;352:i6. doi:10.1136/bmj.i6
18. Van Calster B, Wynants L, Verbeek J, et al. Reporting and interpreting decision curve analysis: a guide for investigators. *Eur Urol.* 2018;74(6):796–804. doi:10.1016/j.eururo.2018.08.038
19. Yates LA, Richards SA, Brook BW. Parsimonious model selection using information theory: a modified selection rule. *Ecology.* 2021;102(10):e3475. doi:10.1002/ecy.3475
20. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155–163. doi:10.1016/j.jcm.2016.02.012
21. Noble WS. How does multiple testing correction work? *Nat Biotechnol.* 2009;27(12):1135–1137. doi:10.1038/nbt1209-1135
22. Welch ML, McIntosh C, Haibe-Kains B, et al. Vulnerabilities of radiomic signature development: the need for safeguards. *Radiother Oncol.* 2019;130:2–9. doi:10.1016/j.radonc.2018.10.027
23. van Griethuysen J, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77(21):e104–e107. doi:10.1158/0008-5472.CAN-17-0339
24. Bettinelli A, Marturano F, Avanzo M, et al. A novel benchmarking approach to assess the agreement among radiomic tools. *Radiology.* 2022;303(3):533–541. doi:10.1148/radiol.211604
25. Cohen JG, Goo JM, Yoo RE, et al. Software performance in segmenting ground-glass and solid components of subsolid nodules in pulmonary adenocarcinomas. *Eur Radiol.* 2016;26(12):4465–4474. doi:10.1007/s00330-016-4317-3
26. Scholten ET, Jacobs C, van Ginneken B, et al. Detection and quantification of the solid component in pulmonary subsolid nodules by semiautomatic segmentation. *Eur Radiol.* 2015;25(2):488–496. doi:10.1007/s00330-014-3427-z
27. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Eur Urol.* 2015;67(6):1142–1151. doi:10.1016/j.eururo.2014.11.025
28. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162(1):W1–W73. doi:10.7326/M14-0698
29. Nardone V, Reginelli A, Grassi R, et al. Delta radiomics: a systematic review. *Radiol Med.* 2021;126(12):1571–1583. doi:10.1007/s11547-021-01436-7
30. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017;14(12):749–762. doi:10.1038/nrclinonc.2017.141
31. Yang G, Nie P, Zhao L, et al. 2D and 3D texture analysis to predict lymphovascular invasion in lung adenocarcinoma. *Eur J Radiol.* 2020;129:109111. doi:10.1016/j.ejrad.2020.109111
32. 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):425–434. doi:10.1148/radiol.2021210109

Cancer Management and Research

Dovepress

### Publish your work in this journal

Cancer Management and Research is an international, peer-reviewed open access journal focusing on cancer research and the optimal use of preventative and integrated treatment interventions to achieve improved outcomes, enhanced survival and quality of life for the cancer patient. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/cancer-management-and-research-journal>