Volume 20, No 4 | International Journal of Radiation Research, October 2022

The significance of peritumoral 5mm regions features for radiomics model in distinguishing the lung adenocarcinomas and granulomas

Y. Geng[1,2], L. Sun[1,2*], M. Sun[1,2], Z. Zhang[1,2], J. Liu[1,2]

[1]Department of Radiology, The fourth Affiliated Hospital of China Medical University, Shengyang, China
[2]Department of Cardiothoracic Surgery, The fourth Affiliated Hospital of China Medical University, Shengyang, China

► ***Original article***

Corresponding author:
Lingling Sun, Ph.D.,
E-mail: *1403952319@qq.com*

Received: October 2021
Final revised: February 2022
Accepted: April 2022

Int. J. Radiat. Res., October 2022; 20(4): 737-745

DOI: 10.52547/ijrr.20.4.2

Keywords: *lung adenocarcinomas, granulomas, radiomics, nomogram, machine learning.*

ABSTRACT

Background: To investigate whether features of 5-mm peritumoral regions could significantly improve the predictive efficacy of a radiomics model based on solid pulmonary tumors at distinguishing lung adenocarcinoma(LAC) from granuloma(GR). ***Materials and Methods***: We retrospectively evaluated 167 lung tumors pathologically proven to be LAC (96) or GR (71) and divided them into training group (116) and testing (51) group. We delineated each tumor with three different measures using the tumor and its 5-mm peritumoral region. Then, we extracted 465 features from each volume of interest(VOI) and chose the optimal features to build the diagnostic models. We built four different models using different methods. Finally, we compared the performance of the four models in the test set. ***Results***: The area under the curve (AUC) of each model in the test group was 0.765 (95% confidence interval(CI): 0.620–0.909), 0.797 (95%CI: 0.670–0.924), and 0.784 (95%CI: 0.647–0.920), respectively. Results of the DeLong test showed that the differences between model 2, model 3, and model 1 were not significant. Results of net reclassification improvement(NRI) showed that model 2 and model 3 had better differential diagnostic efficacy than model 1, with accuracies(ACCs) of 0.784, 0.745, and 0.686, respectively, but the differences were not significant (P>0.05). Moreover, the nomogram had good diagnostic and predictive abilities, with an AUC of 0.848 (95%CI: 0.736–0.961) and an ACC of 0.804. ***Conclusions***: Features of 5-mm peritumoral regions improved the predictive ability of the radiomics model based on the solid pulmonary tumor, but the difference was not significant.

INTRODUCTION

Lung cancer is one of the most common tumors and the leading cause of cancer-related death worldwide. Lung tumors are primarily divided into small cell lung cancer and non-small cell lung cancer, among which LAC is the most common pathological type [1-2]. Currently, the primary method of detecting LAC is chest computed tomography (CT) [3]. However, the imaging findings of CT (such as spiculation and lobulation) lack specificity, making it difficult to distinguish LAC from lung GR [4]. The "gold standard" for diagnosing LAC is pathological biopsy, but this is an invasive examination and is not suitable for all patients [5-6]. This problem not only causes a serious waste of medical resources worldwide but also increases the psychological burden of patients [7-8]. Therefore, clinical practice urgently needs a non-invasive and highly accurate method of distinguishing GR and LAC.

Radiomics, was first proposed by Philippe Lambin in 2012 [9], and has been subsequently confirmed by many studies to play an important role in the diagnosis, treatment, and prognosis of lung cancer. The differential diagnosis ability of radiomics is superior to that of traditional clinical and medical imaging methods and it provides new ideas and directions for solving several clinical problems [10]. Many studies have confirmed that radiomics features have advantages for distinguishing LAC and GR [11-12], and some of these features are linked to pathological results [13]. Recently, several studies have also found that the lung interstitium surrounding the tumor has good diagnostic features, especially regarding the aggressiveness of early LAC [14] and the likelihood of lymph node metastasis [15]. Beig *et al.* [16] also showed that the 5-mm peritumoral region has the strongest diagnostic features compared with other regions around the tumor for distinguishing LAC and GR.

Aim and the novelty, we aimed to build a better radiomics diagnostic model that incorporates the 5-mm peritumoral region to distinguish LAC and GR. To the best of our knowledge, few studies have investigated whether features of the 5-mm [16-17] peritumoral region can significantly improve the ability of the radiomics model using CT scans of the

[DOI: 10.52547/ijrr.20.4.2]

solid pulmonary tumor to distinguish LAC and GR. This study aims to explore this question and it is also the innovation of our research.

MATERIALS AND METHODS

Patients

We retrospectively analyzed the clinical data of 167 patients with LAC or lung GR confirmed by pathology between June 2017 and December 2020. The inclusion criteria were as follows: (1) cases with clear pathological diagnostic results (LAC or GR) by surgical or pathological biopsy; (2) a CT scan was performed within 1 week before treatment or biopsy; (3) solid or part-solid tumor with clear borders; (4) good image quality without metal or motion artifacts; and (5) complete clinical and imaging information of the patient was available. The exclusion criteria were as follows: (1) disqualifying image quality, including lack of uniform layer thickness (not 1 mm) or different algorithms were used to reconstruct the image; (2) ground-glass nodules and multiple tumors in one section of the lung (confused with other tumors) [11-12]; (3) large calcification or cavities and large fat components in the tumor; (4) tumors surrounded by excessive inflammation or atelectasis, or tumors stuck to the chest wall making it difficult to delineate the surrounding lung regions; and (5) incomplete clinical data. The inclusion and exclusion of patients are shown in figure 1.

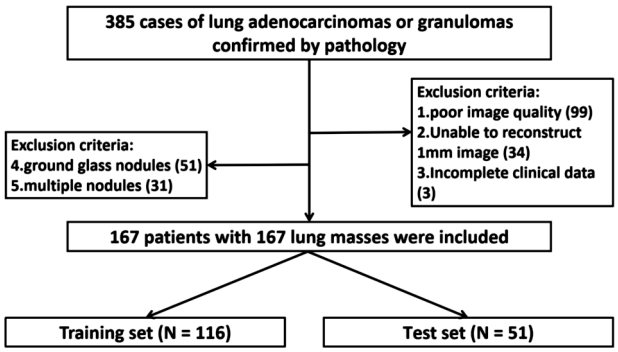

Figure 1. Patients' selection. Schematic diagram of case selection process with exclusion criteria.

After applying the above inclusion and exclusion criteria, we included 167 cases. Among them, 96 cases were LAC and 71 were GR (92 male and 75 female patients; mean age: 59.16 ± 10.60 years). Then, all cases were randomly divided into a training group and test group at a ratio of 7:3. Finally, we collected the basic clinical information and serum tumor markers of patients (before treatment) through the hospital's records system, including age, location, family history, smoking history, Ferrtin, Alpha-fetoprotein(AFP), Carcinoembryonic antigen(CEA), Carbohydrate antigen125(CA125), Carbohydrate antigen153(CA153),Carbohydrate antigen(CA19-9), Neuron Specific Enolase(NSE), and Cytokeratin 19 fragment antigen(CYFRA21-1). The thresholds of the different tumor markers were 291.0 ng/mL (Ferrtin), 25 ng/mL (AFP), 5.0 ng/mL (CEA), 30 U/mL (CA125), 32.4 U/mL (CA153), 37.0 U/mL (CA19-9), 16.3 ng/mL (NSE), and 3.3 ng/mL (CYFRA21-1). If the values were higher than the threshold, it was recorded as 1, otherwise, it was 0. The acquisition of the imaging data and related clinical data of the cases was approved by the institutional ethics committee of our hospital (ethics approval number: 2020-122).

Acquiring CT images

We used a Siemens Somatom Sensa-Tion16 CT (Siemens Healthcare, Germany) or a GE LightSpeed VCT64 row spiral CT (GE Healthcare, USA) to collect CT images with the following flat sweep parameters: tube voltage, 120 kV; tube current, 100 or 150 mA; pitch, 1.15:1 or 1.375:1; layer thickness, 5 mm; interval, 5 mm; rotation time, 0.5 or 0.8 s/turn; and enhanced scan parameters: tube voltage, 120 kV; tube current, 250 mA; pitch, 0.984:1; the whole lung was a 1.00-mm layer, and the rotation time was 0.8 s/turn. The patient usually lay on their back with arms raised above the head. The scan range was from the tip of the lung to the bottom of the lung, and the scan was completed in a single held breath.

Tumor segmentation and extracting radiomics features

We imported all CT images into the open-source 3D-Slicer software (www.3D-Slicer.com, version: 4.10.2), and then two of our hospital's imaging department doctors (with 7 and 14 years' experience in chest imaging diagnosis) who were blinded to the pathological results independently analyzed and semi-automatically delineated three different VOIs for each tumor on the transverse section in the CT plain-phase. VOI1: only tumors were delineated; VOI2: tumors and their 5-mm peritumoral regions were delineated (using the "margin function" in 3D-slicer); VOI3: the 5-mm peritumoral regions alone (using the "hollow function" in 3D-slicer). In the process of delineating tumors, we tried to avoid non-tumor components such as calcifications and cavities. When there was disagreement between the two doctors, the chief physician of our department made the final decision [18]. To facilitate selecting features and evaluating their contribution to the identification of LAC or GR, we standardized (min-max normalization) the clinical data. The process of delineating the VOIs is shown in figure 2.

A total of 465 features were extracted from every VOI using Pyradiomics (Python 3.7.1, version: 3.0.1) and normalized (Max-min normalization) the data we have obtained.

Feature selection and building the radiomics models

The workflow of this study is shown in figure 3,

[DOI: 10.52547/ijrr.20.4.2]

we compared the intraclass and interclass correlation coefficients (ICCs) between different features in the training set to eliminate unstable features. ICCs greater than 0.75 indicated good consistency and stability. We also used the Mann–Whitney U test (P<0.05) to eliminate redundant features that did not contribute to the differential diagnosis.

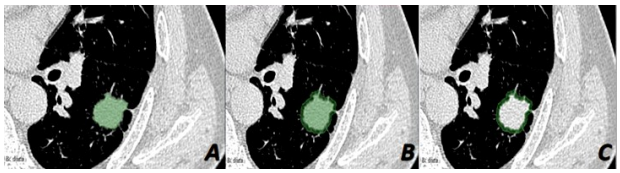

Figure 2. Three VOIs. Images show different measures of three VOIs. **A:** Only the tumor was delineated (VOI1). **B:** The tumor and surrounding 5-mm peritumoral region was delineated (VOI2). **C:** Only the 5-mm peritumoral region (VOI3).

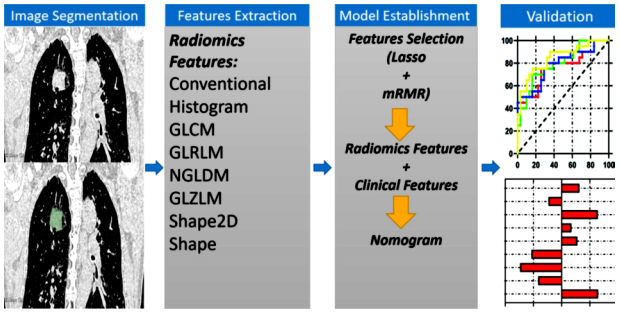

Figure 3. Workflow. Flow chart of our study, including lesion segmentation, feature extraction and selection, model construction, and validation of results.

Subsequently, the max-relevance and min-redundancy (mRMR) were used to select the groups of features that had the greatest correlation to the result but the least correlation between the features. Many studies [19-21] have confirmed that this method can remove redundant features and improve the final model performance. From the mRMR analysis, we only selected top-ranked features.

The least absolute shrinkage and selection operator (Lasso) applied to eliminate redundant features. Lasso regression (L1 regularization) can reduce the coefficients of unimportant features to 0, which avoids overfitting and multicollinearity of the model. We used 5-fold cross-validation to select the most optimal features to build multiple linear regression models. Features from VOI1 were used to build model 1, features from VOI2 were used to build model 2, and then we combined features from VOI1 and VOI3 to build model 3. We used variance inflation factor (VIF) and heatmaps [22] to measure the multicollinearity of our models. In order to ensure the independence of the test set, all features selection and comparison are done in the training set. Finally, we evaluated their diagnostic capabilities in the independent test set.

Nomogram construction and comparing the performances of the models

The independent t-test (or Mann–Whitney U test) and Chi-square test were used to analyze clinical data, and model 3 incorporated the significant clinical information of the patients to develop the nomogram as model 4. We chose the calibration curve to measure the predictive effect of the nomogram [23]. After that, we calculated the AUC of the receiver operating curve (ROC) and the ACC of each group of models in the training group and test group to calculate their 95% CI.

Subsequently, we used the DeLong test [24] to compare whether there were statistically significant differences between the four different ROC curves in the test group. NRI was performed in the test set to evaluate whether model 2, model 3, or model 4 were better than model 1. Finally, we drew a decision curve (DCA) to observe the net benefit of the four models.

Statistical analysis

We used R studio (https:www.r-project.org, version: 4.0.5) for all statistical analyses and Prism (GraphPad Software, San Diego, CA, USA; version: 8.0.2) to draw all graphics. All continuous variables are expressed as mean ± standard deviation (SD). If the data followed a normal distribution, we used the independent t-test to compare statistical differences, otherwise, the Mann–Whitney U test was used for data with non-normal distributions. For categorical variables, we used the Chi-square test or Fisher's exact test. We used Pearson's correlation coefficient and Kendal's correlation coefficient to measure the correlation between different data. Heatmaps were used for visualizations. VIF was used to estimate the multicollinearity of our models, and the VIF of every feature in the model less than 10 was considered not to have strong multicollinearity. P<0.05 of two-sided tests was regarded as a statistically significant difference.

RESULTS

Basic clinical information of the patients

The pathological results of patients and the data regarding serum tumor markers are listed in table 1. Among the 167 cases there are 92 man and 75 women patients (mean age of 59.16 ± 10.60 years old). There were 71 cases (training group: 51 and test group: 20) in the benign group, while 96 cases (training group: 65 and test group: 31) were in the malignant group. More information is listed in table 1.

[DOI: 10.52547/ijrr.20.4.2]

Table 1. Baseline characteristics of patients in training group and test group.

Clinic characteristic	Training group (n = 116)			Test group (n = 51)		
	LAC* (n = 65)	GR* (n = 51)	*P* - value	LAC* (n =31)	GR* (n =20)	*P* - value
Age	52.84±9.49	63.94±8.39	0.000*	62.58±9.07	54.40±12.24	0.009*
Gender			0.837			0.493
Male	37 (56.92)	30 (58.82)		14 (45.16)	11 (55.00)	
Female	28 (43.08)	21 (41.18)		17 (54.84)	9 (45.00)	
Family History			0.07			0.633
Absent	56 (86.15)	49 (96.08)		23 (74.19)	16 (80.00)	
Present	9 (13.85)	2 (3.92)		8 (25.81)	4 (20.00)	
Smoking History			0.491			0.685
Absent	46 (90.20)	39 (76.47)		20 (64.52)	14 (70.00)	
Present	19 (29.23)	12 (23.53)		11 (35.48)	6 (30.00)	
Location			0.286			0.09
R upper*	26 (0.40)	22 (43.14)		11 (35.48)	11 (55.00)	
R middle*	5 (7.69)	1 (1.96)		5 (16.13)	1 (5.00)	
R lower*	15 (23.08)	8 (15.69)		1 (3.23)	4 (20.00)	
L upper*	14 (21.54)	11 (21.56)		10 (32.26)	3 (15.00)	
L lower*	5 (7.69)	9 (17.65)		4 (12.90)	1 (5.00)	
Ferrtin*(<291 ng/mL)			0.466			0.316
Normal	61 (93.85)	46 (90.20)		29 (93.55)	17 (85.00)	
Abnormal	4 (6.15)	5 (9.80)		2 (6.45)	3 (15.00)	
AFP* (<25ng/*mL*)			0.707			1.000
Normal	63 (96.92)	50 (98.04)		31 (100.00)	20 (100.00)	
Abnormal	2 (3.08)	1 (1.96)		0 (0.00)	0 (0.00)	
CEA* (<5ng/*mL*)			0.000*			0.005
Normal	44 (67.69)	49 (96.08)		21 (67.74)	20 (100.00)	
Abnormal	31 (32.31)	2 (3.92)		10 (32.26)	0 (0.00)	
CA125* (<30.0*U/mL*)			0.703			0.315
Normal	60 (92.31)	48 (94.12)		30 (96.77)	18 (90.00)	
Abnormal	5 (7.69)	3 (5.88)		1 (3.23)	2 (10.00)	
CA153* (<32.4*U/mL*)			0.374			0.417
Normal	64 (98.46)	51 (100)		30 (96.77)	20 (100.00)	
Abnormal	1 (1.54)	0 (0)		1 (3.23)	0 (0.00)	
CA19-9*(< 37.0*U/mL*)			0.976			0.094
Normal	60 (92.31)	47 (92.16)		27 (87.10)	20 (100.00)	
Abnormal	5 (7.69)	4 (7.84)		4 (12.90)	0 (0.00)	
NSE* (<16.3*ng/mL*)			0.424			0.427
Normal	29 (44.62)	19 (37.26)		12 (38.71)	10 (50.00)	
Abnormal	36 (55.38)	32 (62.74)		19 (61.29)	10 (50.00)	
CYFRA21-1*(<3.3*ng/mL*)			0.748			0.370
Normal	52 (0.80)	42 (82.35)		25 (80.65)	18 (90.00)	
Abnormal	13 (0.20)	9 (17.65)		6 (19.35)	2 (10.00)	

R upper: Right upper lobe; R middle: Right middle lobe; R lower: Right lower lobe; L upper: Left upper lobe; L lower: Lower lobe of left lung; LAC: Lung adenocarcinoma; GR: Granuloma; AFP: Alpha-fetoprotein; CEA: Carcinoembryonic antigen; CA125: Carbohydrate antigen125; CA153: Carbohydrate antigen153; CA19-9: Carbohydrate antigen19-9; NSE: Neuron specific enlase; CYFRA21-1: Cytokeratin 19 fragment antigen; p value < 0.05.

Results of the independent t-test (or Mann–Whitney U test) and Chi-square test showed that the age and CEA level were significantly different between the two groups (P=0.009 and P<0.001, respectively). The results of multiple logistic regression (clinical laboratory model;Clin-Lab) showed that the odds ratio(OR) of age and CEA were 1.155 (95%CI: 1.092–1.236, P<0.001) ,10.984 (95% CI: 2.409–83.781, P<0.001), respectively. The AUC of the Clin-Lab model is 0.806 (95%CI: 0.689–0.924) with an ACC of 0.706 in the test group.

Results of feature selection

In total, 32, 28, and 41 features of VOI1, VOI2, and VOI3, respectively, were eliminated due to their ICCs being less than 0.75. Subsequently, the Mann–Whitney U test results showed that 180, 211, and 207 features of VOI1, VOI2, and VOI3, respectively, were not significant in distinguishing LAC and GR. This left a total of 253, 226, 217 features of VOI1, VOI2, and VOI3, respectively. Then we selected the top 100 features of mRMR and performed Lasso regression (L1 Regularization) with 5-fold cross-validation in the training group (figure 4).

Ultimately, eight, nine, and five features of VOI1, VOI2, and VOI3, respectively, were screened out. Eight features from VOI1 were used to build model 1, nine features from VOI2 were used to build model 2, after that we combined the features from VOI1 and VOI3 to build model 3. We performed the Lasso operation due to some features from VOI1 and VOI3 (pre) having strong correlations (figure A.1). Therefore, we ultimately chose nine features (seven from VOI1 and two from VOI3) to build model 3. The

[DOI: 10.52547/ijrr.20.4.2]

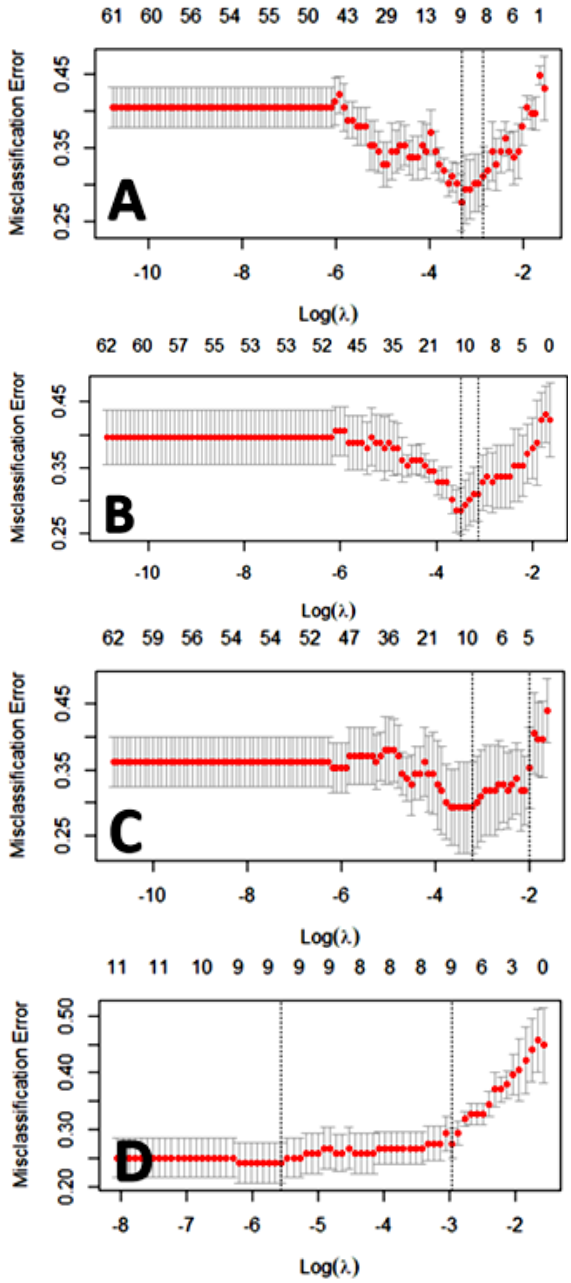

Figure 4. Lasso. Lasso regression results with 5-fold cross validation. The X-axis shows the log (Penalty coefficient λ), and the Y-axis shows the misclassification rate. A total of eight of VOI1 **(A)**, nine of VOI2 **(B)**, five of VOI3 **(C)**, and nine of VOI1 and VOI3 **(D)** non-zero coefficient features were selected to build the prediction models.

VIF values of all models were less than 10. The coefficients and features of the models are listed in (table A.1) and were used to calculate the radiomics score.

Mode-1: 0.6324 + 1.2920 × wavelet. HHL _ glcm _ DifferenceVariance - 0.8290 × wavelet. HHL _ glrlm _ ShortRunHighGrayLevelEmphasis - 1.4727 × wavelet. HLL _ glcm _ Correlation - 1.0684 × wavelet. HLH _ glcm _ MCC + 0.55631 × wavelet. LLL _ glcm _ Correlation + 0.34572 × original _ glcm _ ClusterProminence + 1.28243 × wavelet. LHH _ glrlm _ GrayLevelNonUniformity-0.46432 × wavelet. LHL _ glrlm _ LongRun LowGrayLevelEmphasis.

Model-2: 0.4056 - 2.8414 × log.sigma.3.0.mm.3D _ glcm _ Autocorrelation + 0.57284×wavelet. LLH _ glrlm _ GrayLevelNonUniformity - 0.07003 × wavelet. LHL _ glrlm _ LongRunLowGrayLevelEmphasis + 0.01920 × wavelet. LHL _ glcm _ Idmn + 0.72341 × log.sigma.3.0.mm.3D _ glcm _Imc1 - 0.78388 × log.sigma.1.0.mm.3D _ glcm _ MCC + 0.58241 × wavelet. LHH _ glrlm _ GrayLevelNonUniformity + 0.74316 × wavelet. HHL _ glrlm _ ShortRunHighGrayLevelEmphasis - 0.02061 × log.sigma.3.0.mm.3D _ glcm _ ClusterProminence.

Model-3: 0.6365 + 1.02961 × wavelet. HHL _ glcm _ DifferenceVariance - 1.25279 × wavelet. HLL _ glcm _ Correlation - 0.94404 × wavelet. HLH _ glcm _ MCC - 0.58965 × wavelet. LLL _ glcm _ Correlation + 0.21258 × original _ glcm _ ClusterProminence + 0.96581 × wavelet. LHH _ glrlm _ GrayLevelNonUniformity - 0.46123 × wavelet. LHL _ glrlm _ LongRunLowGrayLevelEmphasis - 0.30427 × wavelet. HHL _ glrlm _ ShortRunHighGrayLevelEmphasis (Pre) + 0.01792 × wavelet. LLL _ glcm _ MaximumProbability (Pre).

The results of ROC and ACC of different models

AUCs of the three models in the training group is 0.884 (95%CI: 0.824–0.943), 0.852 (95%CI: 0.786–0.919), and 0.874 (95%CI: 0.812–0.936) with ACCs of 82.0%, 78.4%, and 79.3%, respectively.

The AUCs of the three models in the test group is 0.765 (95%CI: 0.620–0.909), 0.797 (95%CI: 0.670–0.924), and 0.784 (95%CI: 0.647–0.920) with ACCs of 68.6%, 78.4%, and 74.5%, respectively. More details of these results are shown in (table 2).

Table 2. Performance of different models in training group and test group.

Training group	AUC* (95%CI)	*P* - value*	ACC (%)	SEN (%)	SPE (%)
Model1*	0.884 (0.824-0.943)		82.0 (95/116)	83.1	80.4
Model2*	0.852 (0.786-0.919)	0.338	78.4 (91/116)	87.7	68.6
Model3*	0.874 (0.812-0.936)	0.156	79.3 (92/116)	84.5	78.4
Test group	AUC (95%CI)	*P* - value*	ACC (%)	SEN (%)	SPE (%)
Model1*	0.765 (0.620-0.909)		68.6 (35/51)	74.2	75.0
Model2*	0.797 (0.670-0.924)	0.652	78.4 (40/51)	83.9	70.1
Model3*	0.784 (0.647-0.920)	0.280	74.5 (38/51)	67.7	80.1

P-value: The result of Delong test, CI: Confidence interval, pvalue<0.05, AUC: Area Under Curve, ACC: accuracy, SEN: Sensitivity, SPE: Specificity

Radiomics nomogram construction and validation

On the basis of the radiomics score (rad.score) of model 3, we combined the Clin-lab model to build the

[DOI: 10.52547/ijrr.20.4.2]

logistic regression model. The ORs of rad.score, age, and CEA is 0.936 (95%CI: 0.488–1.384, P<0.001), 0.121 (95%CI: 0.046–0.197, P=0.0016), and 2.906 (95%CI: 0.789–5.023, P=0.0071) (table 3), respectively. Then we drew the nomogram to better distinguish LAC and GR (figure 5).

Table 3. Formulas and performance of the clin-lab model and nomogram in the group set.

Model	Formula	AUC*	95%CI*	ACC*	SEN*	SPE*
Clin-Lab	0.144age+2.396 CEA -8.550	0.806	0.689-0.924	0.706	0.613	0.950
Nomogram	0.936rad-score+0.121age+ 2.906CEA-7.575	0.848	0.736-0.961	0.804	0.839	0.750

AUC:area under the curve; CI:Confidence interval; ACC:accuracy; SEN:Sensitivity; SPE: Specificity; CEA: Carcinoembryonic antigen; Clin-Lab: Clinical laboratory model

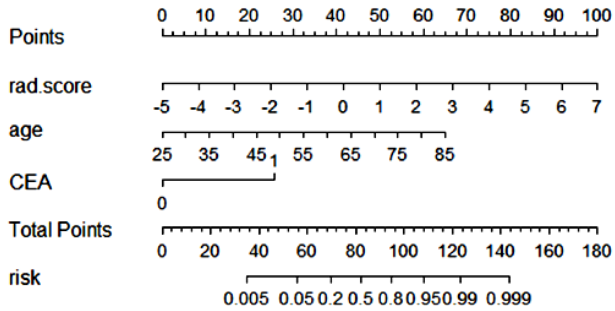

Figure 5. Nomogram. Nomogram to distinguish lung adenocarcinomas and granulomas based on model 4. A total of three indicators were included in the study: rad-score, age,CEA. First, the scores corresponding to score of model3 (rad-score), age, and CEA level were added to get the total score, which was then used to obtain the corresponding risk probability.

The AUC of the nomogram in the test group is 0.848 (95%CI: 0.736–0.961) with an ACC of 80.4%. Results of the DeLong test showed that the difference in ROC of model 3 and the nomogram in the test group is not statistically significant (P=0.143). Finally, calibration of the nomogram (figure 6) in the training group and test group showed that the nomogram could distinguish LAC and GR.

Comparisons of the different models

The AUC results showed that model 2, model 3, and the nomogram were better than model 1, but results of the Delong test indicated that the differences in the ROCs of model 2, model 3 and the nomogram were not significant in the test group (P=0.652, 0.280, and 0.114, respectively) (figure 7).

The NRI results (table 4) in the test group also showed that model 2 and model 3 were better than model 1 but the differences were not significant (P=0.148 and 0.083, respectively). However, the difference between the nomogram (model4) and model 1 in the test group was significant (P=0.029). This showed that the nomogram may have better discrimination ability than model 1. The decisive curve (figure 8) also indicated that once the risk threshold was greater than 0.3, the net benefit of the nomogram was greater than model 1, model 2, and model 3.

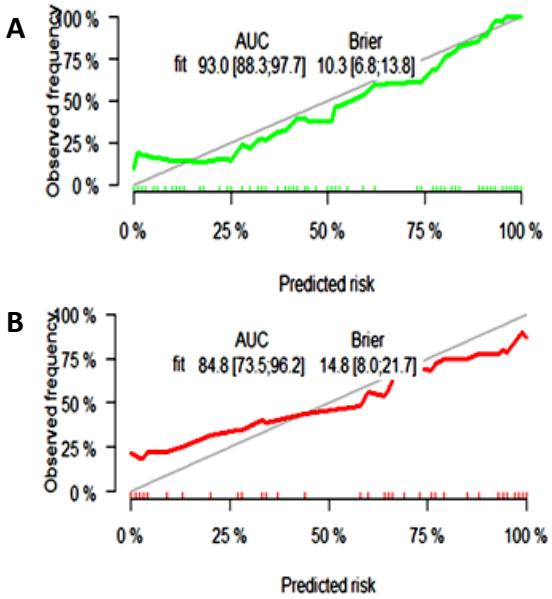

Figure 6. Calibration curves. Calibration curve of the nomogram in the training (green) and test (red) groups. The solid gray line represents the best ideal situation, while the green and red lines represent the predicted probabilities. The X-axis represents the score calculated by the nomogram, while the Y-axis represents the true value. The closer the two lines were to each other, the better the predictive ability. The results of the calibration curve show that nomogram has good predictive ability.

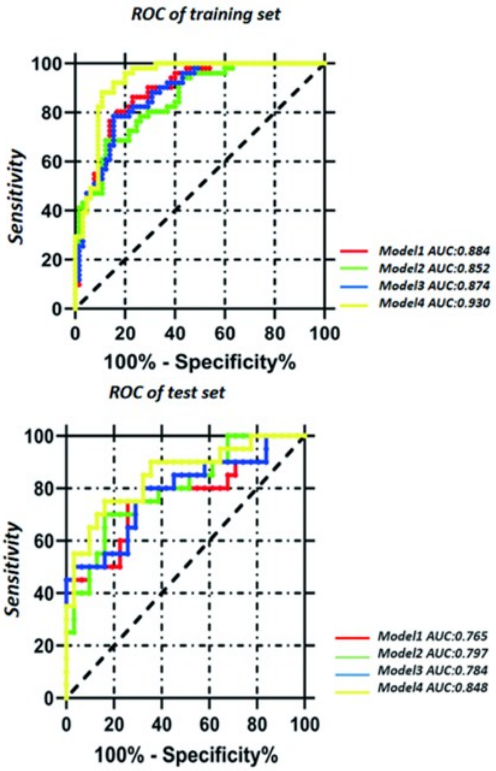

Figure 7. ROC curves. Receiver operating characteristic (ROC) curves of the training and test groups with the respective area under the curve (AUC) values for each group of models at distinguishing lung adenocarcinoma and granuloma.

[DOI: 10.52547/ijrr.20.4.2]

Table 4. Results of NRI in the test group.

	ACC	Improvement	AUC*	Improved AUC	*P* - value*
Model1	35/51	*Reference*	0.765		
Model2	40/51	5	0.797	0.032	0.143
Model3	38/51	3	0.784	0.019	0.067
Model4	41/51	6	0.848	0.083	0.028*

AUC: Area Under Curve; ACC: accuracy; NRI: Net Reclassification Improvement; p - value < 0.05

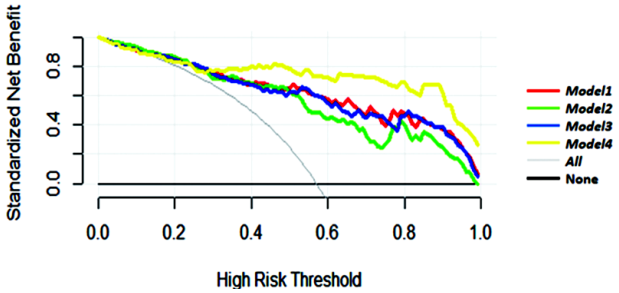

Figure 8. Decision curve analysis. Decision curve analysis (DCA) of the four prediction models in the training groups. The gray and black lines represent the hypotheses that all cases are benign and malignant, respectively. The result of the curve shows that the nomogram has better predictive ability than model 1(red), model 2(green), and model 3(blue) when the high-risk threshold (X-axis) is >0.3.

DISCUSSION

In this study, we compared the results of using three different methods of delineating VOIs for the differential diagnosis of LAC and GR and also developed and validated a nomogram that included serum tumor markers. This study showed that radiomics features of the 5mm peritumoral region could enhance the predictive ability of a radiomics model based on the tumor alone for distinguishing LAC and GR. Finally, we attempted to construct a better diagnostic model. In the test group, model 3 outperformed model1, which shows that features of the 5-mm peritumoral region are helpful for the radiomics model that was originally based on solitary pulmonary tumor features to distinguish LAC and GR. However, results of the DeLong test and NRI showed that this difference was not statistically significant (P>0.05), so the effect of the 5-mm peritumoral region is limited. As a control group, model 2 also outperformed model 1 with AUCs of 0.797 and 0.784 compared with 0.765 in the test group. Different delineation methods produced these different results.

Apart from this, we also found that the radiomics model had better differential diagnosis ability. In the test group, the AUC values of the models based on the radiomics features were 0.765, 0.797, and 0.784, respectively. Our results of AUC are similar to those of Chen *et al.* [25], who showed that the radiomics model achieved an AUC value of 0.798 in differentiating small granulomas and lung adenocarcinomas (less than 10 mm); But our results are smaller than that of Zhuo et al. [26], whose AUC value is close to 1(AUC=0.99). This may be because we are more stringent in the selection of features and strive for reproducibility. Furthermore, our results are similar to the study by Beig et al. [16] with the AUC value of 0.76 in the test set.

As in previous study demonstrated [16], features from 5-mm regions outside the tumor have the best efficacy for distinguishing LAC and GR (compared with different distances). This may be due to the different cell morphology and arrangement of the surrounding lymphocytes and macrophages at the interface of the tumor or GR [16, 26]. These results are somewhat similar to those of some studies, which respectively confirmed that the peritumoral area is helpful for differentiating the pathological type [16], invasiveness [27] and lymph node metastasis [15] of lung cancer. However, we found that the improvement effect of the surrounding area of the tumor was not significant (P<0.05).Similarly, there are also studies [17] showing that the peritumoral region not helpful for the diagnosis of lung cancer. Different research methods produce different research results. In the future, further in-depth research is necessary.

We also included serum tumor markers when establishing the nomogram because they are relatively objective indicators and are commonly measured in clinical practice [28-30]. We found that CEA levels were generally elevated in patients with LAC. This is a clinically meaningful indicator which can also be used to predict the metastasis of lung cancer. In the study by Wang et al. [15], the level of CEA decreased in patients with lung adenocarcinoma with lymph node metastasis comparing the cases without lymph node metastasis.

Other serum tumor markers also have clinical significance, the previous studies [11] have shown that CA125 and CA153 levels are elevated in lung cancer patients compared to the benign control group. Ren *et al.* [23] have also found that compared with lung adenocarcinoma, the levels of NSE and CYFRA21-1 in lung squamous cell carcinoma are significantly higher, so they can be used to assist diagnosis different types of lung cancer. Consequently, as the study of lung cancer, we strongly recommend the addition of serum tumor markers because it is of great significance for the diagnosis and prognosis of lung cancer. In addition to this, the emergence and development of radiomics has also accelerated the individualization and precision of medical care [31]. I believe that it can give a good basis for clinicians' diagnosis and treatment in the future.

There were several limitations to this study. First, this was a single-center retrospectively study with a small number of patients. All of the cases were from our hospital, and almost all the cases were from the northeast region of China. Our study may have included selection bias, which may have led to inadequate extrapolation of the model. Second, the

[DOI: 10.52547/ijrr.20.4.2]

method we used to outline the regions of interest were based on the semi-automatic protocol outlined by our clinician [32-33]. Although we trained physicians before the trial, there was still some observer bias. Some features (32, 28, and 41) were deleted due to insufficient consistency. Our research methods still can be improved. Finally, owing to limitations in the study methodology and the questionable consistency of clinical classifications, we excluded many early and advanced lung cancers because they manifested as pure ground-glass nodules or had a large amount of inflammation and fluid accumulation around them. This narrows the scope of our research. A more comprehensive study will be needed in the future to further expand the field of radiomics [34].

CONCLUSION

In summary, we concluded that the features of 5-mm peritumoral regions improved the predictive ability of the radiomics model based on the solid pulmonary tumor, but the difference was not significant.

ACKNOWLEDGMENTS

Thanks to China Medical University for technical support of this study.

***Declaration of Competing Interest*:** The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of this article.

***Ethical statement*:** This retrospective study was approved by the medical ethics committee of our hospital. The need for informed consent from all individual participants included in study was waived.

***Funding*:** This research did not receive any specific grant from funding agencies in the public, commercials, or not-for-profit sectors.

***Authors contribution*:** Yujun Geng: Methodology, Conceptualization, Data curation, Formal analysis, Writing (original draft), and Validation; Lingling Sun: Supervision and Resources; Muchuan Sun: Supervision, Methodology, Investigation, and Conceptualization; Zhaogang Zhang: Data curation and Formal analysis; Jiaxuan Liu: Data curation and Formal analysis.

REFERENCES

1. Bosetti C, Traini E, Alam T, Allen CA, Carreras G, Compton K, *et al.* (2020) National burden of cancer in Italy, 1990-2017: a systematic analysis for the global burden of disease study 2017. *Sci Rep,* ***10*** *(1): 22099*.
2. Yang X, Man J, Chen H, Zhang T, Yin X, He Q, *et al.* (2021) Temporal trends of the lung cancer mortality attributable to smoking from 1990 to 2017: A global, regional and national analysis. *Lung Cancer,* ***152****: 49-57*.
3. Oudkerk M, Liu S, Heuvelmans MA, Walter JE, Field JK (2021) Lung cancer LDCT screening and mortality reduction - evidence, pitfalls and future perspectives. *Nat Rev Clin Oncol,* ***18****(3): 135-51*.
4. Chen CH, Chang CK, Tu CY, Liao WC, Wu BR, Chou KT, *et al.* (2018) Radiomic features analysis in computed tomography images of lung nodule classification. *PLoS One,* ***13****(2): e0192002*.
5. Thorsteinsson H, Alexandersson A, Oskarsdottir GN, Skuladottir R, Isaksson HJ, Jonsson S, *et al.* (2012) Resection rate and outcome of pulmonary resections for non-small-cell lung cancer: a nationwide study from Iceland. *J Thorac Oncol,* ***7****(7): 1164-9*.
6. Qvick A, Stenmark B, Carlsson J, Isaksson J, Karlsson C, Helenius G (2021) Liquid biopsy as an option for predictive testing and prognosis in patients with lung cancer. *Molecular Medicine,* ***27****(1): 68*.
7. Kim H, Park CM, Goo JM, Wildberger JE, Kauczor HU (2015) Quantitative Computed Tomography Imaging Biomarkers in the Diagnosis and Management of Lung Cancer. *Invest Radiol,* ***50****(9): 571-83*.
8. Bach PB, Mirkin JN, Oliver TK, Azzoli CG, Berry DA, Brawley OW, *et al.* (2012) Benefits and harms of CT screening for lung cancer: a systematic review. *JAMA,* ***307****(22): 2418-29*.
9. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, *et al.* (2012) Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer,* ***48****(4): 441-6*.
10. Gillies R, Kinahan P, Hricak H (2016) Radiomics: Images Are More than Pictures, They Are Data. *Radiology,* ***278****(2): 563-77*.
11. Yang X, He J, Wang J, Li W, Liu C, Gao D, *et al.* (2018) CT-based radiomics signature for differentiating solitary granulomatous nodules from solid lung adenocarcinoma. (Amsterdam, Netherlands) *Lung cancer,* ***125****: 109-14*.
12. Feng B, Chen X, Chen Y, Liu K, Li K, Liu X, *et al.* (2020) Radiomics nomogram for preoperative differentiation of lung tuberculoma from adenocarcinoma in solitary pulmonary solid nodule. *Eur J Radiol,* ***128****: 109022*.
13. Alvarez-Jimenez C, Sandino AA, Prasanna P, Gupta A, Viswanath SE, Romero E (2020) Identifying Cross-Scale Associations between Radiomic and Pathomic Signatures of Non-Small Cell Lung Cancer Subtypes: Preliminary Results. *Cancers (Basel),* ***12****(12)*.
14. Wu G, Woodruff HC, Shen J, Refaee T, Sanduleanu S, Ibrahim A, *et al.* (2020) Diagnosis of Invasive Lung Adenocarcinoma Based on Chest CT Radiomic Features of Part-Solid Pulmonary Nodules: A multicenter study. *Radiology,* ***297****(2): E282*.
15. Wang X, Zhao X, Li Q, Xia W, Peng Z, Zhang R, *et al.* (2019) Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? *European radiology,* ***29****(11): 6049-58*.
16. Beig N, Khorrami M, Alilou M, Prasanna P, Braman N, Orooji M, *et al.* (2019) Perinodular and Intranodular Radiomic Features on Lung CT Images Distinguish Adenocarcinomas from Granulomas. *Radiology,* ***290****(3): 783-92*.
17. Wu S, Zhang N, Wu Z, Ren J, Lining E (2022) Can Peritumoral Radiomics Improve the Prediction of Malignancy of Solid Pulmonary Nodule Smaller Than 2 cm. *Acad Radiol,* ***29****(2): S47-S52*.
18. Wang R, Liu H, Liang P, Zhao H, Li L, Gao J (2021) Radiomics analysis of CT imaging for differentiating gastric neuroendocrine carcinomas from gastric adenocarcinomas. *Eur J Radiol,* ***138****: 109662*.
19. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts H (2015) Machine Learning methods for Quantitative Radiomic Biomarkers. *Sci Rep,* ***5****: 13087*.
20. Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJ (2015) Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. *Front Oncol,* ***5****: 272*.
21. Peng H, Long F, Ding C (2005) Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell,* ***27****(8): 1226-38*.
22. Lu X, Li M, Zhang H, Hua S, Meng F, Yang H, *et al.* (2020) A novel radiomic nomogram for predicting epidermal growth factor receptor mutation in peripheral lung adenocarcinoma. *Phys Med Biol,* ***65*** *(5): 055012*.
23. Ren C, Zhang J, Qi M, Zhang J, Zhang Y, Song S, *et al.* (2021) Machine learning based on clinico-biological features integrated (18)F-FDG PET/CT radiomics for distinguishing squamous cell carcinoma from adenocarcinoma of lung. *Eur J Nucl Med Mol Imaging,* ***48*** *(5):1538-49*.
24. Wu J, Sun X, Wang J, Cui Y, Kato F, Shirato H, *et al.* (2017) Identifying relations between imaging phenotypes and molecular subtypes of breast cancer: Model discovery and external validation. Journal of magnetic resonance imaging: *JMRI,* ***46****(4): 1017-27*.
25. Xiangmeng C, Bao F, Yehang C, Kunfeng L, Kunwei L, Xiaobei D, *et al.* (2020) A CT-based radiomics nomogram for prediction of lung adenocarcinomas and granulomatous lesions in patient with solitary sub-centimeter solid nodules. %J Cancer imaging: *The official publication of the International Cancer Imaging Society,* ***20****(1)*.

[DOI: 10.52547/ijrr.20.4.2]

26. Zhuo Y, Zhan Y, Zhang, Shan F, Shen J, Wang D, Yu M (2021) Clinical and CT radiomics nomogram for preoperative differentiation of pulmonary adenocarcinoma from tuberculoma in solitary solid nodule. *Frontiers in Oncology,* ***2021****: 11*.
27. Wu L, Gao C, Xiang P, Zheng S, Pang P, Xu M (2020) CT-Imaging based analysis of invasive lung adenocarcinoma presenting as ground glass nodules using Peri- and Intra-nodular radiomic features. *Frontiers in Oncology,* ***2020****: 10*.
28. Salem A, Asselin MC, Reymen B, Jackson A, Lambin P, West CML, *et al.* (2018) Targeting hypoxia to improve Non-Small cell lung cancer outcome. *J Natl Cancer Inst,* ***110****(1)*.
29. Cheng C, Yang Y, Yang W, Wang D, Yao C (2021) The diagnostic value of CEA for lung cancer-related malignant pleural effusion in China: a meta-analysis. *Expert Rev Respir Med,* ***2021****: 1-10*.
30. Xie H, Kong YX, Zhang Q (2019) Value of Serum Tumor Marker Isocitrate Dehydrogenase 1 in the Diagnosis of Lung Cancer. *Lung Cancer,* ***41****(6): 813-7*.
31. Rocco G, Morabito A, Leone A, Muto P, Fiore F, Budillon A (2016) Management of non-small cell lung cancer in the era of personalized medicine. *Int J Biochem Cell Biol,* ***78****: 173-9*.
32. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, *et al.* (2012) Radiomics: the process and the challenges. *Magn Reson Imaging,* ***30****(9): 1234-48*.
33. Zwanenburg A (2019) Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *European Journal of Nuclear Medicine and Molecular Imaging,* ***46****(13): 2638-55*.
34. Lee G, Bak S, Lee H, Choi J, Park H, Lee S, *et al.* (2019) Measurement Variability in Treatment Response Determination for Non-Small Cell Lung Cancer: Improvements Using Radiomics. *Journal of Thoracic Imaging ,* ***34****(2): 103-15*.

[DOI: 10.52547/ijrr.20.4.2]