

# A Radiomics Nomogram for Preoperative Prediction of Clinical Occult Lymph Node Metastasis in cT1-2N0M0 Solid Lung Adenocarcinoma

Ran Zhang,<sup>1,2</sup> Ranran Zhang,<sup>3</sup>  
Ting Luan,<sup>4,5</sup> Biwei Liu,<sup>6</sup>  
Yimei Zhang,<sup>6</sup> Yaping Xu,<sup>1</sup>  
Xiaorong Sun,<sup>5</sup> Ligang Xing<sup>6</sup>

<sup>1</sup>Department of Radiation Oncology, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai, People's Republic of China; <sup>2</sup>Tongji University, Shanghai, People's Republic of China; <sup>3</sup>Department of Medical Imaging, Linyi Cancer Hospital, Linyi, Shandong, People's Republic of China; <sup>4</sup>Department of Graduate, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong, People's Republic of China; <sup>5</sup>Department of Nuclear Medicine, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong, People's Republic of China; <sup>6</sup>Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, Shandong, People's Republic of China

Correspondence: Xiaorong Sun  
Department of Nuclear Medicine,  
Shandong Cancer Hospital and Institute,  
Shandong First Medical University and  
Shandong Academy of Medical Sciences,  
Jiyuan Road 440, Jinan, Shandong, 250000,  
People's Republic of China  
Tel/Fax +86 53167626767  
Email 251400067@qq.com

Yaping Xu  
Department of Radiation Oncology,  
Shanghai Pulmonary Hospital, Tongji  
University School of Medicine, Zhengmin  
Road 507, Shanghai, 200443, People's  
Republic of China  
Tel/Fax +86 2165115006  
Email xuyaping1207@163.com

**Background:** Clinical occult lymph node metastasis (cOLNM) means that the lymph node is negatively diagnosed by preoperative computed tomography (CT), but has been proven to be positive by postoperative pathology. The aim of this study was to establish and validate a nomogram based on radiomics features for the preoperative prediction of cOLNM in early-stage solid lung adenocarcinoma patients.

**Methods:** A total of 244 patients with clinical T1-2N0M0 solid lung adenocarcinoma who underwent preoperative contrast-enhanced chest CT were divided into a primary group (n = 160) and an independent validation group from another hospital (n = 84). The records of 851 radiomics features of each primary tumor were extracted. LASSO analysis was used to reduce the data dimensionality and select features. Multivariable logistic regression was utilized to identify independent predictors of cOLNM and develop a predictive nomogram. The performance of the predictive model was assessed by its calibration and discrimination. Decision curve analysis (DCA) was performed to estimate the clinical usefulness of the nomogram.

**Results:** The predictive model consisted of a clinical factor (CT-reported tumor size) and a radiomics feature (Rad-score). The nomogram presented good discrimination, with a C-index of 0.782 (95% CI, 0.768–0.796) in the primary cohort and 0.813 (95% CI, 0.787–0.839) in the validation cohort, and good calibration. DCA showed that the radiomics nomogram was clinically useful.

**Conclusion:** This study develops and validates a nomogram that incorporates clinical and radiomics factors. It can be tailored for the individualized preoperative prediction of cOLNM in early-stage solid lung adenocarcinoma patients.

**Keywords:** solid lung adenocarcinoma, prediction, clinical occult lymph node, computed tomography, radiomics

## Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-specific death worldwide.<sup>1</sup> Among these patients, the proportion of early-stage NSCLC has risen in recent years thanks to the popularity of low-dose computed tomography (LDCT).<sup>2</sup> Patients with early-stage NSCLC are mainly treated with surgery, showing a 70% 5-year survival rate and a 55% to 75% recurrence rate. For medically inoperable patients, stereotactic body radiotherapy (SBRT) is preferred, as it yields more than 85% local control and 60% 3-year overall survival.<sup>3–6</sup> However, the rate of local/regional failure (LRF) following definitive operation for early-stage NSCLC is as high



as 26%.<sup>7</sup> Compared to surgery, SABR achieves worse locoregional tumor control because of more lymph nodal failures.<sup>8</sup> Therefore, it is essential to emphasize the poor prognosis caused by local/regional failure (LRF) in the treatment of early-stage NSCLC. The main reason for the high rate of LRF may be the limited ability of current evaluation methods to detect clinical occult lymph node metastasis (cOLNM), a key deficiency of node-staging techniques.

cOLNM means that the short axis of the hilar or mediastinal lymph nodes is less than 10 mm on axial preoperative computed tomography (CT) images,<sup>9</sup> but postoperative pathology results of lymph node demonstrate to be metastatic. The preoperative prediction of cOLNM comprises invasive and noninvasive methods. Noninvasive methods, such as CT and 18F fluorodeoxyglucose positron emission tomography combined with computed tomography (18FDG-PET/CT),<sup>10,11</sup> play an indispensable role in pretreatment evaluation. Despite improving the accuracy of staging, CT and PET/CT still have defects due to their relatively low resolution. Furthermore, the high false positive rate, possibly caused by lymphadenitis or tuberculosis, prevents PET from becoming the routine preoperative evaluation of lung cancer.<sup>12</sup> Thus, it is crucial to find a novel method to increase the discovery rate of cOLNM.

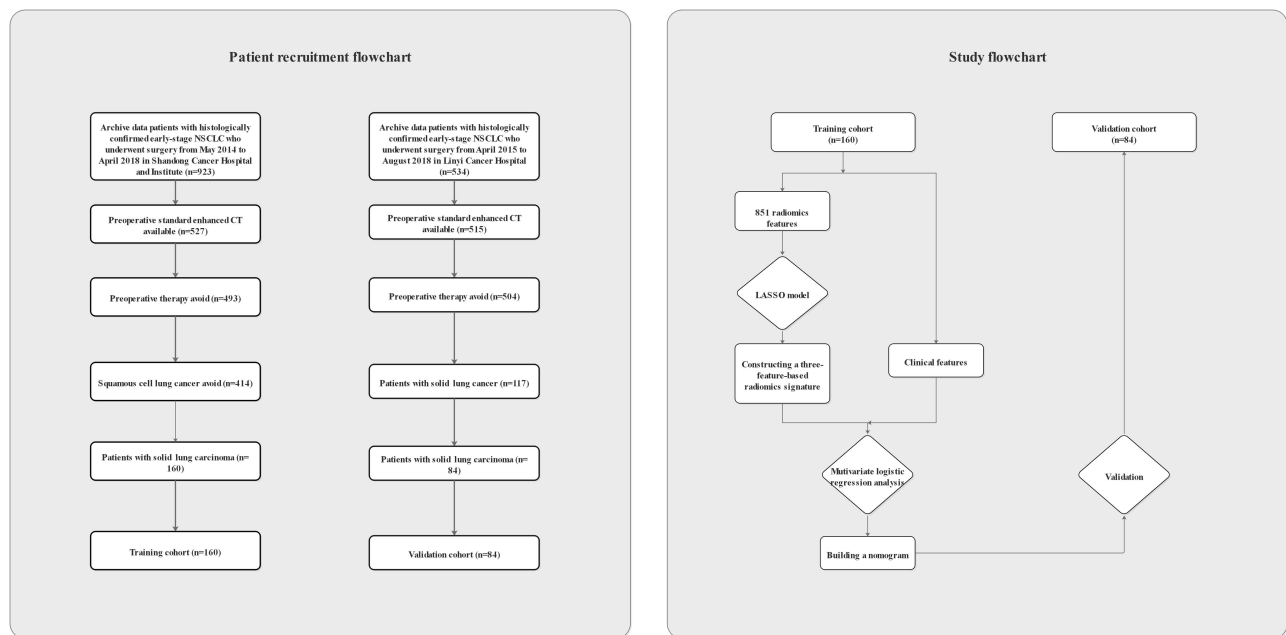
Radiomics, an emerging method to extract high-throughput data and analyze many quantitative features from medical images, has good prospects in precision medicine,<sup>13–16</sup> thus improving the detection of OLN. Recently, several studies have taken advantage of the radiomics approach to predict occult lymph nodes in lung cancer.<sup>17,18</sup> Zhong et al<sup>18</sup> found that the radiomics features extracted from unenhanced chest CT can predict OLN. Nevertheless, they brought approximately 49.7% ground-glass opacity (GGO) nodules covering minimally invasive adenocarcinoma and lepidic predominant adenocarcinoma into the including criteria, which had been demonstrated no tendency of lymph node metastasis.<sup>19</sup> It was unclear that the actually predictive value of radiomics features for patients with solid nodules who have the tendency of lymph node metastasis.<sup>19</sup> Yang et al<sup>17</sup> found that their radiomics nomogram derived from CT images of clinical stage I–III B NSCLC patients could predict lymph node metastasis, but it was unclear how many early-stage NSCLC cases were included. Accordingly, to detect the occult lymph nodes of early NSCLC patients before treatment, we only included cT1-T2N0M0 solid lung adenocarcinoma patients in this study to develop and validate a radiomics nomogram for predicting OLN.

## Materials and Methods

### Patients

Ethical approval was obtained from the institutional review board of Shandong Cancer Hospital and Institute (Jinan, China) and Linyi Cancer Hospital (Linyi, China) for this retrospective study, and written informed consent was waived. This study was conducted in accordance with the Declaration of Helsinki. Patients with histologically confirmed NSCLC who underwent surgery from October 2014 to April 2018 were reviewed. cOLNM was defined according to the fact that the short axis of the hilar or mediastinal lymph nodes was less than 10 mm on axial images of preoperative CT scans, but postoperative pathology results of lymph nodes demonstrated metastasis. The inclusion criteria were as follows: (1) patients with clinical T1-2N0M0 stage NSCLC diagnosed by preoperative CT scans (according to the 8th AJCC edition staging); (2) patients who underwent lobectomy, segmentectomy or wedge resection with systematic lymph node dissection; (3) patients who underwent standard contrast-enhanced CT less than 2 weeks before surgery; and (4) patients with postoperative biopsy-proven adenocarcinoma. The exclusion criteria were as follows: (1) preoperative therapy (chemotherapy, radiotherapy or chemoradiotherapy); (2) incomplete clinical data or postoperative pathology; (3) unsatisfactory image quality because of other diseases, such as pulmonary emphysema or inflammation; and (4) enlargement of lymph nodes on CT (identified as the short axis of a lymph node  $\geq 10$  mm on axial imaging<sup>9</sup>). Patients treated between April 2014 and May 2018 from Shandong Cancer Hospital and Institute comprised the training cohort, and patients treated between April 2015 and August 2018 from Linyi Cancer Hospital constituted the validation cohort. The patient recruitment pathways are shown in Figure 1.

Baseline data, such as age, gender, smoking status, CT reported tumor size, T stage, histological grade, tumor location, tumor type, pleural traction, visceral pleural invasion, carcinoembryonic antigen (CEA) and lymph node dissection results were collected from medical records. For smoking status, patients were divided into nonsmokers and smokers, the latter including former smokers and current smokers. The CT-reported tumor size was defined as the longest diameter of the primary tumor. Tumor location was categorized as upper location or middle-lower location, and tumor type was categorized as central or peripheral. Laboratory analysis of CEA was done by blood tests within 2 weeks before surgery. The positive threshold value for CEA was  $>5$  ng/mL, according to the ranges routinely utilized at our institute.



**Figure 1** Recruitment procedure and study flowchart.

## CT Image Acquisition

Contrast-enhanced CT was performed with a 64-MDCT scanner (Definition AS+, Siemens SOMATOM) in the training cohort and a 64-MDCT scanner (Definition AS 64-Slice; Siemens, Germany) in the validation cohort. Generally, the enhancement method is taking 60–100 mL (1.5–2 mL/kg) of injected iopromide (300 mg/mL), and it is quickly injected with a high-pressure syringe (injection speed: 2 mL/s). The acquisition parameters were as follows: slice thickness, 5.00 mm; tube voltage, 120 kVp; tube current, 160 mAs for the training cohort and 220mAs for the validation cohort; field of view, 385 × 385 mm; and matrix, 512 × 512. All images were displayed with standard lung (width 1600 HU; level, -600 HU) and mediastinal window settings (width, 400 HU; level, 50 HU).

## Segmentation and Radiomics Feature Extraction

Segmentation of the entire region of interest (ROI) of the tumor was contoured independently and manually by two radiologists with 20 and 6 years of experience in venous-phase CT images, respectively. They reached a consensus by discussion. A total of 851 2 dimensional (2D) radiomics features were extracted from images using pyRadiomics software, a flexible open-source platform for extracting large amounts of features from medical images.<sup>20</sup> These features were categorized into groups of shape, first-order features, grey-level co-occurrence

matrix (GLCM), grey-level run-length matrix (GLRLM), grey-level size zone matrix (GLSZM), neighbor grey-tone difference matrix (NGTDM) of the original, and wavelet-based features. All the features are listed in [Supplementary Table 1](#). Then, 30 patients were randomly chosen for the assessment of interobserver reproducibility of the extracted radiomics features using the intraclass correlation coefficient (ICC), and the independent-samples *t*-test was utilized to assess the differences between the features generated by the different radiologists. Finally, 490 features with satisfactory consistency (ICC > 0.75, P > 0.05) were selected.

## Selecting Features and Building the Radiomics Score (Rad-Score)

The least absolute shrinkage and selection operator (LASSO) approach, which reduces the data dimensionality, was utilized to select the potential features that would predict OLN in the training cohort. A radiomics score was obtained for each patient through a linear combination of selected features weighed by their respective LASSO coefficients.

## Development of Radiomics Nomogram

Multivariable logistic regression analysis was utilized to identify the predictive factors for OLN among smoking (yes or no), Karnofsky performance status (KPS), CT-reported tumor size, T stage (T1 or T2), tumor location (upper or middle-lower), tumor type (central or peripheral),

wedge resection (yes or no), histology grade (poor or moderate/high), pleural traction (yes or no), visceral pleural invasion (yes or no), CEA (normal or abnormal) and Rad-score (per 0.1 increase). Backward step-wise selection was then done by using the likelihood ratio test with Akaike's information criterion as the stopping rule.<sup>21</sup> On the basis of the results of multivariable logistic analysis in the training cohort, we built a radiomics nomogram to quantitatively predict the individual probability of OLNLM.

## Performance, Validation and Clinical Use of the Nomogram

The calibration of the nomogram was evaluated by constructing calibration curves, and the Hosmer–Lemeshow test<sup>22</sup> was used to determine how perfectly the models were calibrated. ROC (receiver operating characteristic) was calculated to assess the discriminatory performance of tumor size, radiomics score and the nomogram.

The validation model was performed in the validation cohort. The logistic regression formula established in the training cohort was applied to all patients of the validation cohort. Then, the ROC analysis was performed and calibration curve was plotted.

Decision curve analysis (DCA) was performed to qualify the net benefits for different threshold probabilities and thereby assess the clinical usefulness of the nomogram in both the training cohort and the validation cohort.

## Statistical Analysis

Statistical analysis was conducted with R software (version 3.5.2), IBM SPSS (version 24.0) and MedCalc (version 15.6.1). Continuous variables were presented as mean  $\pm$  SD. The Mann–Whitney *U*-test or Student's *t*-test was used to assess the differences of continuous variables, such as age and tumor size. Fisher's exact or chi-squared test was used to compare the differences in categorical variables, such as gender, smoking status, and histological grade. The chi-squared test was also used to analyze the difference in OLNLM prevalence between the primary cohort and validation cohort. The LASSO model was used to reduce dimensionality of the data. Univariable and multivariable analyses were performed by binary logistic regression analysis. The LASSO model was established by the “glmnet” package. Multivariable binary logistic regression, nomograms and calibration plots were handled with the “rms” package. DCA was done using the “rmda” package. ROC analysis was performed by MedCalc software.

## Results

### Patient Characteristics

The patient recruitment workflow and study workflow are shown in [Figure 1](#). The characteristics of the patients in the training and validation cohorts are given in [Table 1](#). There was no difference in the OLNLM rate between the two cohorts: 21.25% (34/160) in the training cohort and 25% (21/84) in the validation cohort ( $P = 0.505$ ). In the training cohort, there was no significant difference in the clinical characteristics between the OLNLM-positive group and the OLNLM-negative group, except for CT-reported tumor size ( $P = 0.001$ ), T stage ( $P = 0.001$ ), tumor type ( $P = 0.014$ ) and CEA level ( $P = 0.003$ ).

### Feature Selection, Construction and Diagnostic Validation of the Radiomics Signature

A total of 490 2D features, extracted via pyRadiomics, were reduced to 3 potential predictors with nonzero coefficients using the LASSO model ([Figure 2A](#) and [B](#)). Then, the radiomics signature was constructed. The Rad-score formula was calculated as follows:

$$\begin{aligned} \text{Rad-score} = & 0.08474311 \times \text{wavelet-HHH. first-order.} \\ & \text{Kurtosis} \\ & + 0.01010229 \times \text{wavelet-LHH. gldm. DependenceVariance} \\ & + 0.00174351 \times \text{original. gldm. LargeDependenceEmphasis.} \end{aligned}$$

The ROC curve analysis for the predictive efficacy of the radiomics signature in the training cohort and validation cohort is shown in [Figure 3A](#) and [B](#) and [Table 2](#). For the rad-score, the values of area under the curve (AUC) were 0.767 (95% CI, 0.680–0.855) in the training cohort and 0.767 (95% CI, 0.662–0.853) in the validation cohort. For the tumor size, the values of AUC were 0.730 (95% CI, 0.634–0.827) in the training cohort and 0.807 (95% CI, 0.706–0.885) in the validation cohort.

### Development of the Predictive Nomogram

In univariable logistic regression analysis, CT-reported tumor size, T stage, tumor type, CEA level and Rad-score (per 0.1 increase) were significant predictors of OLNLM in the training cohort ([Table 3](#)). CT-reported tumor size and Rad-score were identified as independent predictors of OLNLM in multivariable analysis. A model that combined the above two factors was constructed and is presented as a predictive nomogram in [Figure 4](#).

**Table I** Characteristics of Patients in the Training and Validation Cohorts

Characteristics	Training Cohort			Validation Cohort		
	OLN (+)	OLN (-)	P	OLN (+)	OLN (-)	P
Age, mean $\pm$ SD (y)	59.94 $\pm$ 8.30	62.56 $\pm$ 7.78	0.109	60.42 $\pm$ 7.94	59.71 $\pm$ 9.49	0.741
Gender			0.999			0.69
Male	16	61		6	23	
Female	18	65		15	40	
Smoking status			0.375			0.276
Smoker	7	38		4	22	
Nonsmoker	27	88		17	41	
CT-reported tumor size, mean $\pm$ SD (cm)	3.04 $\pm$ 0.93	2.25 $\pm$ 0.81	<b>0.001</b>	3.36 $\pm$ 0.92	2.29 $\pm$ 0.82	<b>0.001</b>
Clinical T stage			<b>0.001</b>			<b>0.001</b>
cT1	16	99		4	53	
cT2	18	27		17	10	
Histological grade			0.999			<b>0.001</b>
Highly/Moderately differentiated	31	112		10	61	
Poorly differentiated	3	14		11	2	
Tumor location			0.507			0.999
Upper	18	77		11	31	
Middle- Lower	16	49		10	32	
Tumor type			<b>0.014</b>			<b>0.012</b>
Central	18	36		11	13	
Peripheral	16	90		10	50	
Pleural traction			0.873			0.999
Yes	18	71		17	51	
No	16	55		4	12	
Visceral pleural invasion			0.092			0.115
Yes	21	55		17	37	
No	13	71		4	26	
CEA level			<b>0.003</b>			0.112
Normal	16	95		16	58	
Abnormal	18	31		5	5	
Wedge resection			0.413			0.104
Yes	3	20		0	9	
No	31	106		21	54	
Rad score, median	8.91	5.94	<b>0.001</b>	6.12	4.96	<b>0.001</b>

**Note:** Significant values are printed in bold.

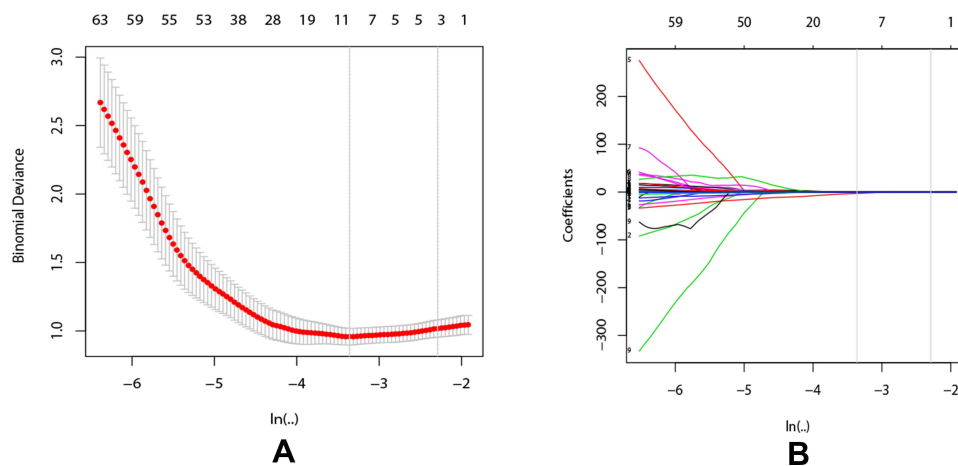
**Abbreviations:** SD, standard deviation; OLN, occult lymph node; CEA, carcinoembryonic antigen.

## Calibration, Discrimination and Validation of the Nomogram

The calibration curve of the nomogram for the prediction rate of OLN showed good agreement in the training cohort (Figure 3C), with a nonsignificant result ( $P = 0.492$ ) of the Hosmer–Lemeshow test. The AUC of the model in the training cohort was 0.782 (95% CI,

0.691–0.874) (Figure 3A and Table 2). In the validation cohort, the Hosmer–Lemeshow test between prediction line and observation line of the calibration curve was not statistically significant ( $P = 0.128$ ) (Figure 3D). The AUC of the nomogram in the validation group was 0.813 (95% CI, 0.713–0.890) (Figure 3B and Table 2).





**Figure 2** Radiomics feature reduction and selection using the least absolute shrinkage and selection operator (LASSO) binary model. **(A)** Selection by the LASSO model utilized 10-fold cross-validation with the minimum criteria. The binomial deviances were plotted against log (lambda). Dotted vertical lines mark the optimal value applying the minimum criteria with 1 standard error (the 1-SE criteria). The optimal  $\lambda$  value of 0.101 with  $\log(\lambda) = -2.291$  was selected. **(B)** LASSO coefficient profiles of the 357 radiomics features. The vertical line was delineated at the optimal value by 10-fold cross-validation, and 3 features were chosen with nonzero coefficients in the plot.

## Clinical Use

DCA for the radiomics model applied to the validation cohort is presented in Figure 5. In the decision curve, if the threshold probability of a patient was between 0.22 and 0.83, the nomogram to predict OLN showed more benefit than the treat-all therapeutic plan or treat-none plan.

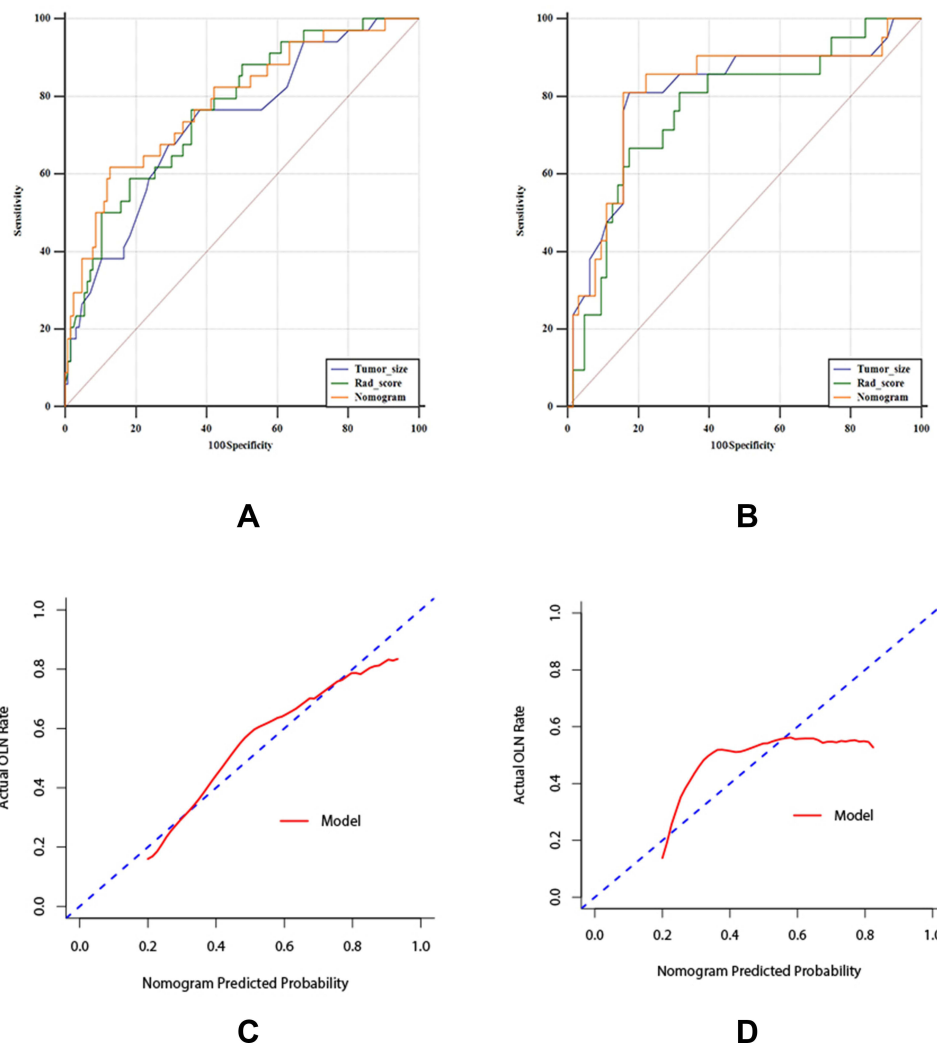
## Discussion

To develop a preoperative predictive model of clinical OLN for patients with early-stage solid lung adenocarcinoma, a nomogram combined with a radiomics signature was established and validated in this retrospective study. Consisting of both radiomics and clinical factors, the nomogram in our study was tailored to stratify patients on the basis of the risk of clinical OLN before surgery.

For operable early-stage patients, the finding of cOLNM can determine the choice of lymph node dissection or systematic sampling, because lymph node dissection may cause an unnecessary trauma. The finding of cOLNM can also offer the option to use neoadjuvant therapy for high-risk patients with cOLNM, which could benefit the long-term survival outcomes of these patients. For inoperable early-stage patients, the finding of cOLNM is especially important, as the standard treatment pattern of inoperable early-stage patients is SBRT,<sup>23</sup> but lymph nodes are not delineated in the target volume of SBRT.<sup>24,25</sup> Thus, this Radiomics model could increase the accuracy of the current lymph node staging system, and identify patients a high risk of cOLNM who should therefore be offered additional

treatment, such as radiation of the cOLNM region, chemotherapy and even immunotherapy.

Clinical factors, such as CT-reported tumor size and tumor markers in laboratory analysis, have been associated with the existence of OLN in patients with early-stage NSCLC.<sup>17,18,26–28</sup> We found that CT-reported tumor size, T stage, tumor type (central or peripheral) and CEA level were predictors of OLN in patients with early-stage lung adenocarcinoma, though only CT-reported tumor size was still an independent predictor of OLN in the multivariate analysis. A similar result for T stage was seen in a retrospective study: Kim et al<sup>29</sup> investigated the predictability of OLN using metabolic parameters of pretreatment positron emission tomography (PET)/CT in patients with NSCLC and found that T stage was a significant factor in the univariate analysis but not in the multivariate analysis. Several previous studies<sup>27,30–32</sup> have found that CT-reported tumor size and central lung cancer were independent predictors of lymph node metastasis in patients with NSCLC. In terms of tumor markers, it was reported<sup>33,34</sup> in the past decades that tumor-associated markers from plasma or serum could facilitate early screening and treatment monitoring for NSCLC. Among the tumor markers, CEA has proven to be a sensitive tumor marker to predict occult lymph nodes in patients with NSCLC.<sup>35</sup> Other studies<sup>26,36</sup> have incorporated CEA level into models for lymph node metastasis prediction in NSCLC patients. However, CEA level, central location and T stage were not independent predictors of OLN in our study. The main reason that these variables were not independently predictive factors



**Figure 3** ROC curves and calibration curves. **(A)** ROC curves for training cohort. **(B)** ROC curve for radiomics signature of validation cohort. **(C)** Calibration curve of the radiomics model in the training group. **(D)** Calibration curve of the radiomics model in the validation group. Calibration curves described the calibration of each predictive model as a measure of the agreement between the predicted probabilities of occult lymph node metastasis (OLNM) and the observed result. The y-axis shows the actual OLN rate, and the x-axis shows the predicted OLN risks. The diagonal dotted blue line denotes a perfect prediction (ideal model). The red line reveals the actual performance of the nomogram.

in the multivariate analysis may be the relatively small sample size of the training cohort. Given the above, it was reasonable that we identified CT-reported tumor size as an element of the predictive model for occult lymph node metastasis in patients with early-stage solid lung adenocarcinoma.

**Table 2** The Results of ROC Analysis

Training Cohort	AUC (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	FI Score (%)
Tumor size	0.730 (0.634–0.827)	76.5	61.9	35.1	90.7	48.1
Rad score	0.767 (0.680–0.855)	76.5	64.3	36.6	91.0	49.5
Nomogram	0.782 (0.691–0.874)	61.8	87.3	56.8	89.4	59.2
Validation cohort						
Tumor size	0.807 (0.706–0.885)	80.9	82.5	60.7	92.9	69.4
Rad score	0.767 (0.662–0.853)	80.9	68.3	45.9	91.5	58.6
Nomogram	0.813 (0.713–0.890)	80.9	84.1	63.0	93.0	70.9

**Abbreviations:** ROC, receiver operating characteristic; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

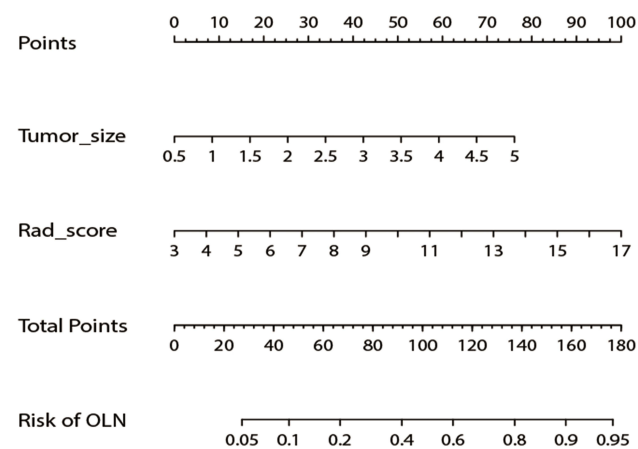
**Table 3** Univariable and Multivariable Logistic Regression Analyses for OLN

Variable and Intercept	Univariable Logistic Regression		Multivariable Logistic Regression	
	OR (95% CI)	P value	OR (95% CI)	P value
Smoking (yes or no)	0.600 (0.241–1.498)	0.274	NA	NA
KPS score	0.886 (0.361–2.179)	0.793	NA	NA
CT-reported tumor size	2.743 (1.706–4.411)	<b>0.001</b>	1.921 (1.114–3.311)	<b>0.019</b>
T stage (T1 or T2)	4.125 (1.860–9.149)	<b>0.001</b>	0.732 (0.171–3.138)	0.675
Tumor location (upper or middle-lower)	0.716 (0.334–1.535)	0.39	NA	NA
Tumor type (central or peripheral)	2.812 (1.294–6.114)	<b>0.009</b>	1.603 (0.657–3.910)	0.3
Wedge resection (yes or no)	0.513 (0.143–1.841)	0.306	NA	NA
Histologic grade (poor or moderate/high)	1.292 (0.349–4.782)	0.702	NA	NA
Histologic subtype (acinar or not)	0.869 (0.385–1.963)	0.736		
Pleural traction (yes or no)	0.871 (0.408–1.863)	0.723	NA	NA
Visceral pleural invasion (yes or no)	2.085 (0.960–4.531)	0.063	NA	NA
CEA level (normal or abnormal)	3.085 (1.401–6.790)	<b>0.005</b>	1.966 (0.801–4.825)	0.14
Rad score (per 0.1 increase)	1.438 (1.221–1.694)	<b>0.001</b>	1.318 (1.107–1.571)	<b>0.002</b>

**Note:** Significant values are printed in bold.

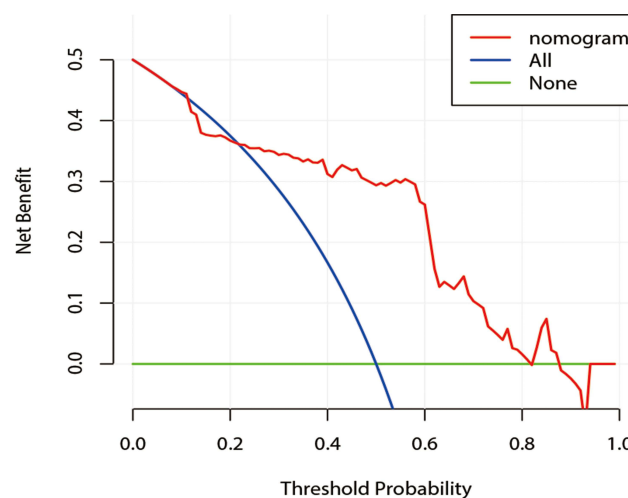
**Abbreviations:** OR, odds ratio; KPS, Karnofsky performance status; CT, computed tomography; CEA, carcinoembryonic antigen.

Radiomics, an emerging field in imaging analysis and processing, is able to convert imaging information into high dimensional mineable features<sup>37</sup> that could potentially be tailored to improve disease diagnosis, treatment planning and prognostic prediction.<sup>38</sup> Several studies<sup>17,18,26,39</sup> have identified radiomics features as part of their models for personalized prediction of postoperative occult lymph node metastasis in early-stage patients with NSCLC. However, none of the previous studies has built a radiomics model for OLN in patients with early-stage solid lung adenocarcinoma and validated this model in patients from another medical institution. Here, we developed a radiomics nomogram that incorporated CT-reported tumor size and a radiomics signature for the individualized preoperative



**Figure 4** Radiomics nomogram. The predictive nomogram was built in the training group. Its variables were CT-reported tumor size and Rad-score.

prediction of OLN in patients with solid early-stage lung adenocarcinoma and validated this predictive model in patients from another hospital. The major components of the radiomics signature in this study were wavelet features which can identify heterogeneity of primary tumor lesions at different scales and cannot be discerned by the naked eye. Wang et al<sup>36</sup> also found that peritumoral radiomics signatures made up mainly of wavelet features can improve the predictive abilities of lymph node metastasis in patients with



**Figure 5** Decision curve analysis for the radiomics nomogram in the validation cohort. The y-axis measures the net benefit, and the x-axis represents the threshold probability. The red line represents the radiomics nomogram. The blue line represents the assumption that all patients have OLN. The green line represents the assumption that no patients have OLN. The decision curve in the validation cohort indicates that if the threshold probability of a patient was between 0.22 and 0.83, the nomogram to predict OLN showed more benefit than the treat-all therapeutic plan or treat-none plan.



clinical stage T1 lung adenocarcinoma. In the training cohort, the AUCs of rad-score, tumor size and nomogram were 0.767 (0.680–0.855), 0.730 (0.634–0.827) and 0.782 (0.691–0.874), respectively. In the validation cohort, the AUCs of rad-score, tumor size and nomogram were 0.767 (0.662–0.853) and 0.807 (0.706–0.885) and 0.813 (0.713–0.890), respectively. The sensitivity analysis demonstrated that when excluding one factor from the predictive model, the remaining factor still showed a good performance of diagnosis. Thus, the results of model proved to be robust based on the AUCs.

Notably, the definition of OLNLM in our study was clinical OLNLM, which mainly depended on the imaging assessment by radiologists. It should also be considered that pathological OLNLM is an indispensable component in the study of OLNLM. Pathological OLNLM, also called occult metastasis (OM) in lymph nodes, means remaining undetected cancer cells in lymph nodes after the current pathological techniques for biopsy during pulmonary resection.<sup>40</sup> The standard technique in surgery to detect cancer cells in lymph nodes is to cut the lymph node serially into 3- to 4-mm-thick slices, which are embedded in paraffin, sectioned by a microtome, stained with hematoxylin and eosin (H&E), and tested for malignant cells by light microscopy. However, these techniques cannot identify OM, so immunohistochemistry (IHC) is currently used to find OM in lymph nodes. It was proved that<sup>40</sup> the presence of OM by IHC in lymph nodes of patients with NSCLC correlated with prognosis. In addition, none of the previous studies have reported the correlation between radiomics features and OMs in lymph nodes in patients with early-stage NSCLC. Consequently, further work should emphasize the preoperative prediction of both clinical and pathological OLNLM combined with radiomics imaging analysis.

There are several limitations to our study. First, although external validation was adopted, the sample size in the validation cohort was relatively small, so larger samples of patients with early-stage lung adenocarcinoma from hospitals other than ours and multi-center trials are required to evaluate the model we built in the future. Second, due to the retrospective design of our study, patient selection bias and even radiomics feature selection bias could exist. Prospective multi-center studies are needed to validate the preliminary findings of the current study. Third, the CT slices were relatively thick. We will validate our results by thin-slice CT when we can enroll enough patients with preoperative thin-slice CT scans in

the future. Fourth, recent studies<sup>41,42</sup> have revealed a correlation between gene expression and lymph node metastasis in NSCLC. Future studies should establish predictive models incorporating both radiomics features and genetic features. We manually delineated ROIs in our study, as most authors have done. Based on convolutional neural networks,<sup>43</sup> contouring the ROIs automatically by a deep learning method could be realized in the near future.

## Conclusion

This study developed and validated a nomogram that combines both a radiomics signature and clinical variables. It could be tailored to the individualized preoperative prediction of OLNLM in early-stage solid lung adenocarcinoma patients.

## Acknowledgments

Manuscript is approved by all authors for publication. This work was financially supported by the grants from the National Natural Science Foundation of China (82172866), the Shandong Natural Science Foundation (ZR2019LZL019), the Jinan Scientific and Technology Development Project (201805005), and the Start-up Fund for Talent Introduction of Shanghai Pulmonary Hospital (YX) (grant number: 201801).

## Disclosure

The authors declare no conflicts of interest in this work.

## References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin.* 2013;63(1):11–30. doi:10.3322/caac.21166
2. Aberle DR, Adams AM; The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med.* 2011;365(5):395–409.
3. Howington JA, Blum MG, Chang AC, Balekian AA, Murthy SC. Treatment of stage I and II non-small cell lung cancer: diagnosis and management of lung cancer, 3rd ed: American college of chest physicians evidence-based clinical practice guidelines. *Chest.* 2013;143(5):e278S–e313S.
4. Timmerman R, Paulus R, Galvin J, et al. Stereotactic body radiation therapy for inoperable early stage lung cancer. *JAMA.* 2010;303(11):1070–1076. doi:10.1001/jama.2010.261
5. Bilal H, Mahmood S, Rajashanker B, Shah R. Is radiofrequency ablation more effective than stereotactic ablative radiotherapy in patients with early stage medically inoperable non-small cell lung cancer? *Interact Cardiovasc Thorac Surg.* 2012;15(2):258–265. doi:10.1093/icvts/ivs179
6. Shirvani SM, Jiang J, Chang JY, et al. Comparative effectiveness of five treatment strategies for early-stage non-small cell lung cancer in the elderly. *Int J Radiat Oncol Biol Phys.* 2012;84(5):1060–1070. doi:10.1016/j.ijrobp.2012.07.2354

7. Judy GD, Kaidar-Person O, Deal A, et al. The persistent problem of local/regional failure after surgical intervention for early-stage lung cancer. *Ann Thorac Surg.* 2018;106(2):382–389. doi:10.1016/j.athoracsur.2018.03.062
8. van den Berg LL, Klinkenberg TJ, Groen HJ, Widder J. Patterns of recurrence and survival after surgery or stereotactic radiotherapy for early stage NSCLC. *J Thorac Oncol.* 2015;10(5):826–831. doi:10.1097/JTO.0000000000000483
9. Silvestri GA, Gould MK, Margolis ML, et al. Noninvasive staging of non-small cell lung cancer. *Chest.* 2007;132(3):178S–201S. doi:10.1378/chest.07-1360
10. Andersen MB, Harders SW, Ganeshan B, Thygesen J, Torp Madsen HH, Rasmussen F. CT texture analysis can help differentiate between malignant and benign lymph nodes in the mediastinum in patients suspected for lung cancer. *Acta Radiol.* 2016;57(6):669–676. doi:10.1177/0284185115598808
11. Rogasch JM, Apostolova I, Steffen IG, et al. Standardized visual reading of F18-FDG-PET in patients with non-small cell lung cancer scheduled for preoperative thoracic lymph node staging. *Eur J Radiol.* 2016;85(8):1345–1350. doi:10.1016/j.ejrad.2016.05.004
12. Feng M, Yang X, Ma Q, He Y. Retrospective analysis for the false positive diagnosis of PET-CT scan in lung cancer patients. *Medicine.* 2017;96(42):e7415. doi:10.1097/MD.00000000000007415
13. Huang Y, Liu Z, He L, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology.* 2016;281(3):947–957. doi:10.1148/radiol.2016152234
14. 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
15. Lee G, Lee HY, Park H, et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: state of the art. *Eur J Radiol.* 2017;86:297–307. doi:10.1016/j.ejrad.2016.09.005
16. Verma V, Simone CB 2nd, Krishnan S, Lin SH, Yang J, Hahn SM. The rise of radiomics and implications for oncologic management. *J Natl Cancer Inst.* 2017;109(7). doi:10.1093/jnci/djx055
17. Yang X, Pan X, Liu H, et al. A new approach to predict lymph node metastasis in solid lung adenocarcinoma: a radiomics nomogram. *J Thorac Dis.* 2018;10(Suppl 7):S807–S819. doi:10.21037/jtd.2018.03.126
18. Zhong Y, Yuan M, Zhang T, Zhang YD, Li H, Yu TF. Radiomics approach to prediction of occult mediastinal lymph node metastasis of lung adenocarcinoma. *Am J Roentgenol.* 2018;211(1):109–113. doi:10.2214/AJR.17.19074
19. Matsuguma H, Yokoi K, Anraku M, et al. Proportion of ground-glass opacity on high-resolution computed tomography in clinical T1 N0 M0 adenocarcinoma of the lung: a predictor of lymph node metastasis. *J Thorac Cardiovasc Surg.* 2002;124(2):278–284. doi:10.1067/mtc.2002.122298
20. van Griethuysen JJM, 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
21. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* 2015;102(3):148–158.
22. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. *Crit Care Med.* 2007;35(9):2052–2056. doi:10.1097/01.CCM.0000275267.64078.B0
23. Giaj-Levra N, Borghetti P, Bruni A, et al. Current radiotherapy techniques in NSCLC: challenges and potential solutions. *Expert Rev Anticancer Ther.* 2020;20(5):387–402. doi:10.1080/14737140.2020.1760094
24. Figlia V, Mazzola R, Cuccia F, et al. Hypo-fractionated stereotactic radiation therapy for lung malignancies by means of helical tomotherapy: report of feasibility by a single-center experience. *Radiol Med.* 2018;123(6):406–414. doi:10.1007/s11547-018-0858-7
25. Cuccia F, Mortellaro G, Mazzola R, et al. Prognostic value of two geriatric screening tools in a cohort of older patients with early stage non-small cell lung cancer treated with hypofractionated stereotactic radiotherapy. *J Geriatr Oncol.* 2020;11(3):475–481. doi:10.1016/j.jgo.2019.05.002
26. Gu Y, She Y, Xie D, et al. A texture analysis-based prediction model for lymph node metastasis in stage I a lung adenocarcinoma. *Ann Thorac Surg.* 2018;106(1):214–220. doi:10.1016/j.athoracsur.2018.02.026
27. Li L, Ren S, Zhang Y, et al. Risk factors for predicting the occult nodal metastasis in T1-2N0M0 NSCLC patients staged by PET/CT: potential value in the clinic. *Lung Cancer.* 2013;81(2):213–217. doi:10.1016/j.lungcan.2013.04.012
28. Park HK, Jeon K, Koh WJ, et al. Occult nodal metastasis in patients with non-small cell lung cancer at clinical stage IA by PET/CT. *Respirology.* 2010;15(8):1179–1184. doi:10.1111/j.1440-1843.2010.01793.x
29. Kim DH, Song BI, Hong CM, et al. Metabolic parameters using 18F-FDG PET/CT correlate with occult lymph node metastasis in squamous cell lung carcinoma. *Eur J Nucl Med Mol Imaging.* 2014;41(11):2051–2057. doi:10.1007/s00259-014-2831-6
30. Zhang Y, Sun Y, Xiang J, Zhang Y, Hu H, Chen H. A prediction model for N2 disease in T1 non-small cell lung cancer. *J Thorac Cardiovasc Surg.* 2012;144(6):1360–1364. doi:10.1016/j.jtcvs.2012.06.050
31. Shafazand S, Gould MK. A clinical prediction rule to estimate the probability of mediastinal metastasis in patients with non-small cell lung cancer. *J Thorac Oncol.* 2006;1(9):953–959. doi:10.1016/S1556-0864(15)31627-0
32. Chen K, Yang F, Jiang G, Li J, Wang J. Development and validation of a clinical prediction model for n2 lymph node metastasis in non-small cell lung cancer. *Ann Thorac Surg.* 2013;96(5):1761–1768. doi:10.1016/j.athoracsur.2013.06.038
33. Feng Q, Yu M, Kiviat NB. Molecular biomarkers for cancer detection in blood and bodily fluids. *Crit Rev Clin Lab Sci.* 2006;43(5–6):497–560.
34. Villalobos P, Wistuba II. Lung cancer biomarkers. *Hematol Oncol Clin North Am.* 2017;31(1):13–29. doi:10.1016/j.hoc.2016.08.006
35. Moon Y, Choi SY, Park JK, Lee KY. Risk factors for occult lymph node metastasis in peripheral non-small cell lung cancer with invasive component size 3 cm or less. *World J Surg.* 2020;44(5):1658–1665. doi:10.1007/s00268-019-05365-5
36. Wang X, Zhao X, Li Q, et al. Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? *Eur Radiol.* 2019;29(11):6049–6058. doi:10.1007/s00330-019-06084-0
37. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48(4):441–446. doi:10.1016/j.ejca.2011.11.036
38. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5(1):4006. doi:10.1038/ncomms5006
39. Liu Y, Kim J, Balagurunathan Y, et al. Prediction of pathological nodal involvement by CT-based Radiomic features of the primary tumor in patients with clinically node-negative peripheral lung adenocarcinomas. *Med Phys.* 2018;45(6):2518–2526. doi:10.1002/mp.12901
40. Rusch VW, Hawes D, Decker PA, et al. Occult metastases in lymph nodes predict survival in resectable non-small-cell lung cancer: report of the ACOSOG Z0040 trial. *J Clin Oncol.* 2011;29(32):4313–4319. doi:10.1200/JCO.2011.35.2500

41. Seto K, Kuroda H, Yoshida T, et al. Higher frequency of occult lymph node metastasis in clinical N0 pulmonary adenocarcinoma with ALK rearrangement. *Cancer Manag Res.* 2018;10:2117–2124. doi:10.2147/CMAR.S147569
42. Meng W, Ye Z, Cui R, et al. MicroRNA-31 predicts the presence of lymph node metastases and survival in patients with lung adenocarcinoma. *Clin Cancer Res.* 2013;19(19):5423–5433. doi:10.1158/1078-0432.CCR-13-0320
43. Jian J, Xiong F, Xia W, et al. Fully convolutional networks (FCNs)-based segmentation method for colorectal tumors on T2-weighted magnetic resonance images. *Australas Phys Eng Sci Med.* 2018;41(2):393–401. doi:10.1007/s13246-018-0636-9

### 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>