European Radiology (2019) 29:439–449
https://doi.org/10.1007/s00330-018-5539-3

ONCOLOGY

A clinical-radiomics nomogram for the preoperative prediction of lung metastasis in colorectal cancer patients with indeterminate pulmonary nodules

TingDan Hu[1] · ShengPing Wang[1] · Lv Huang[2] · JiaZhou Wang[2] · DeBing Shi[3] · Yuan Li[4] · Tong Tong[1] · Weijun Peng[1]

Received: 17 January 2018 / Revised: 1 May 2018 / Accepted: 14 May 2018 / Published online: 12 June 2018

Abstract

Objectives To develop and validate a clinical-radiomics nomogram for preoperative prediction of lung metastasis for colorectal cancer (CRC) patients with indeterminate pulmonary nodules (IPN).

Methods 194 CRC patients with lung nodules were enrolled in this study (136 in the training cohort and 58 in the validation cohort). To evaluate the probability of lung metastasis, we developed three models, the clinical model with significant clinical risk factors, the radiomics model with radiomics features constructed by the least absolute shrinkage and selection operator algorithm, and the clinical-radiomics model with significant variables selected by the stepwise logistic regression. The Akaike information criterion (AIC) was used to compare the relative strength of different models, and the area under the curve (AUC) was used to quantify the predictive accuracy. The nomogram was developed based on the most appropriate model. Decision-curve analysis was applied to assess the clinical usefulness.

Results The clinical-radiomics model (AIC = 98.893) with the lowest AIC value compared with that of the clinical-only model (AIC = 138.502) or the radiomics-only model (AIC = 116.146) was identified as the best model. The clinical-radiomics nomogram was also successfully developed with favourable discrimination in both training cohort (AUC = 0.929, 95% CI: 0.885–0.974) and validation cohort (AUC = 0.922, 95% CI: 0.857–0.986), and good calibration. Decision-curve analysis confirmed the clinical utility of the clinical-radiomics nomogram.

Conclusions In CRC patients with IPNs, the clinical-radiomics nomogram created by the radiomics signature and clinical risk factors exhibited favourable discriminatory ability and accuracy for a metastasis prediction.

Key Points

- *Clinical features can predict lung metastasis of colorectal cancer patients.*
- *Radiomics analysis outperformed clinical features in assessing the risk of pulmonary metastasis.*
- *A clinical-radiomics nomogram can help clinicians predict lung metastasis in colorectal cancer patients.*

Keywords Colorectal neoplasms · Nomograms · Decision making

TingDan Hu and ShengPing Wang contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5539-3) contains supplementary material, which is available to authorized users.

✉ Tong Tong
t983352@126.com

✉ Weijun Peng
cjr.pengweijun@vip.163.com

[1] Department of Radiology, Fudan University Shanghai Cancer Center, Department of Oncology, Shanghai Medical College, Fudan University, No. 270, Dongan Rd, Shanghai 200032, People's Republic of China

[2] Department of Radiotherapy, Fudan University Shanghai Cancer Center, Fudan University, Shanghai, People's Republic of China

[3] Department of Colorectal Surgery, Fudan University Shanghai Cancer Center, Fudan University, Shanghai, People's Republic of China

[4] Department of Pathology, Fudan University Shanghai Cancer Center, Fudan University, Shanghai, People's Republic of China

Abbreviations

AIC	Akaike information criterion
AUC	Area under the curve
CA19-9	Carbohydrate antigen 19-9
CEA	Carcinoembryonic antigen
CI	Confidence interval
CRC	Colorectal cancer
DCA	Decision curve analysis
IPN	Indeterminate pulmonary nodules
ITT	Intravascular tumour thrombus
LASSO	Least absolute shrinkage and selection operator
LM	Lung metastasis
LR test	Likelihood-ratio test
NM	Non-metastasis
NPV	Negative predictive value
PNI	Perineural invasion
PPV	Positive predictive value
ROC	Receiver operating characteristic curve

Introduction

Colorectal cancer (CRC) is one of the leading causes of morbidity and mortality in the world [1]. The lung is the second most common organ for metastasis in those with CRC [2], as 5–15% of CRC patients reportedly develop lung metastasis (LM) [3]. As early detection and surgical treatment of pulmonary metastasis from colorectal cancer results in a 5-year survival rate of more than 50% in a selected patient population [4], a high accuracy of metastasis prediction is therefore important for clinical decision making. A surgical resection would be performed when there is a definite and clear diagnosis [5]; however, following routine CT scans, an increasing number of patients have indeterminate pulmonary nodules (IPN) that are unable to be characterised as malignant or benign and that have unremarkable morphological characteristics or further evidence of metastasis [6].

Since a chest CT has limited specificity in differentiating lung nodules effectively [7] and a PET-CT does not provide further clinical value due to a risk of false-negative results for lung lesions smaller than 10 mm [7], other studies are concentrating on relevant clinical risk factors to look for metastasis evidence. In addition, texture analysis is emerging as one of several 'radiomics' approaches for interpretation of medical imaging and has attracted attention in recent years as an alternative method to differentiate the IPNs without the need for an extra examine procedure. By extracting radiomics features from routinely acquired images and then quantifying them into high-dimensional data [8], we can observe that radiomics signature correlates closely with tumour staging and grading [9, 10]. To date, radiomics has been applied in many kinds of tumours as a prognostic, diagnostic and treatment response imaging biomarker [11–13].

To our knowledge, however, there has been no combination and comparison of texture features and clinical risk factors to predict LM in a large population with CRC. Therefore, we aim to compare the performance of clinical features and radiomics signature in predicting pulmonary metastasis, and also develop a clinical-radiomics nomogram as a useful clinical tool to improve LM prediction in CRC patients with suspicious lung nodules.

Materials and methods

Patients

This retrospective study was approved by the ethics committee of Fudan University Shanghai Cancer Centre Institutional Review Board (Shanghai, China) and the requirement for informed consent was waived. Our study recruited 194 consecutive CRC patients (87 female (F)/107 male (M), mean age 58.6±10.73 years; age range, 32–84 years) with pulmonary lesions between January 2010 and October 2016 in our institution's database according to the following inclusion criteria: (i) histopathologically-confirmed colorectal cancer; (ii) with at least one pulmonary nodule measuring 5–20 mm in diameter detected by CT examination in our hospital, including synchronous nodules detected along with primary tumour and metachronous nodules detected at follow-up examination after CRC was diagnosed by at least 6 months; (iii) availability of clinical characteristics; (iv) with complete CT datasets and medical history at our institution. The exclusion criteria were as follows: (i) with pretreatment (including systemic chemotherapy or pneumonectomy) 6 months prior to initial CT examination; (ii) follow-up of no-progress lesions for less than 2 years after the initial examination; (iii) nodules detected 6 months before colorectal cancer was confirmed; (iv) obvious metastatic nodule or benign nodules (such as cyst, tuberculosis, inflammatory nodules, haemangioma) with typical imaging characteristics. The patients were randomly divided into a training cohort and a validation cohort by a computer algorithm in a ratio of 7:3. Figure 1 shows the patient recruitment pathway.

Grouping criteria and clinical data collection

Patients were confirmed with two outcomes: (i) pulmonary metastases (93 LMs, 41 F/52 M, mean age 58.39±10.26 years), which were confirmed by pathology (93 cases); and (ii) non-metastasis (101 NMs, 46 F/55 M, mean age 58.6 ±10.73 years), which included patients with benign pulmonary nodules with no progression in size (90 cases, confirmed by at least 2-year follow-up without treatment) and primary lung cancer cases (11 cases, confirmed by tissue samples).

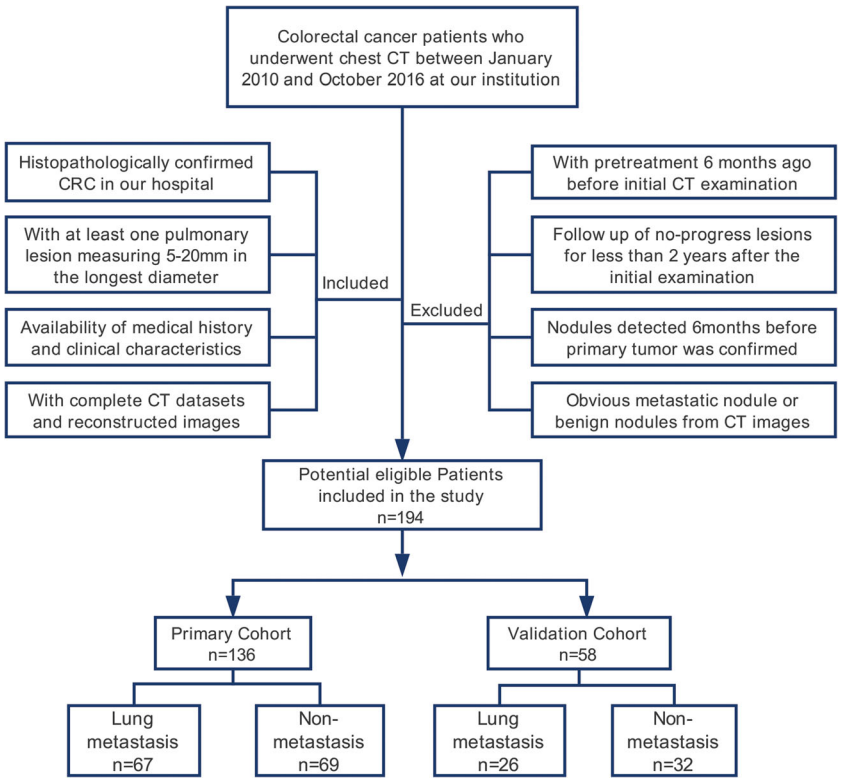

Fig. 1 Flow chart of patients' recruitment pathway

Image assessment and radiomics feature extraction

CT examinations were performed for the selected patients at our institution with the Somatom Definition AS (Siemens Healthcare, Erlangern, Germany), the Sensation 64 (Siemens Healthcare) and the Brilliance (Philips Healthcare, Best, The Netherlands) using a tube voltage of 120 kV and a current of 200 mA. The target lesion that was pathologically sampled or under at least 2-year follow-up surveillance was selected for reconstruction with the following standard reconstruction parameters: slice thickness, 1.0 mm; increment, 1 mm; pitch, 1.078; field of view, 15 cm; and a matrix of 512×512.

Reconstructed images were transferred from the hospital's picture archiving and communication system (PACS) to an off-line work station for texture analysis. Regions of interests (ROIs) covering the whole surface of the lesion on each consecutive slice were semi-automatically contoured in MIM software (v6.6.3; MIM Software Inc.), manually refined by an operator (T.D.H., graduate student) and then verified by an expert radiologist (S.P.W., with 20 years of experience in interpreting chest CT) to exclude the border of the lesion (to avoid partial volume effect) and any other irrelevant components such as air, peripheral vessels, normal tissue, pleura as well as surrounding organs. After a nodule was reconstructed and segmented, the volume of interest (VOI) images (DICOM format) were transferred to MATLAB (Math works Inc.) for feature extracting and analysis. From each segmented tumour, we extracted 203 texture features. More information about the radiomics feature extraction methodology can be found in the Electronic Supplementary Material.

Development of the clinical-only, radiomics-only and clinical-radiomics model

In the development of the clinical model, the chi-square test was performed to compare the differences in categorical variables, while a two-sample t-test was used to compare the differences in continuous variables between LM and NM. Significant clinical risk factors were introduced into the stepwise multivariate logistic regression analysis with the likelihood ratio test (LR test) and Akaike information criterion (AIC) employed as the stopping rule to build the clinical-only model. The diagnosis performance of the clinical model was then tested in the validation cohort with the multivariable regression formula derived from the training cohort applied to the patients in the validation cohort, and the probability of metastasis was calculated for each.

Next, the least absolute shrinkage and selection operator (LASSO) algorithm, which is suitable for the regression of high-dimensional data, was used to select the predictive radiomics feature in the training cohort. The tenfold cross-validation was implemented to avoid over-fitting. A radiomics score (rad-score) was calculated for each patient via a linear combination of selected texture features that were weighted by their respective coefficients. The formula was applied in the validation cohort to calculate the corresponding rad-score.

The univariate logistic regression analysis was performed to assess the association between the rad-score and lung metastasis.

Finally, clinical risk factors of the clinical-only model and rad-score were introduced to the stepwise multiple logistic regression to build the clinical-radiomics model, and the AIC and LR test were also employed as the stopping rule. A process of validation was performed.

Model comparison and nomogram development

The AIC and the LR test of the three models were applied to identify the most appropriate model. AIC and chi-square were calculated to compare the relative strength of different models, and the area under the curve (AUC) was used to quantify the predictive accuracy of the three models in both training and validation cohorts. We also calculated the probability of pulmonary metastasis for each patient with logistic regression analysis, and divided patients into metastasis and non-metastasis groups based on the probability of corresponding to the cut-off value with the highest Youden index. Compared with the actual metastasis results, we calculated the sensitivity, specificity, accuracy, positive-predictive value (PPV) and negative-predictive value (NPV) for the three models in both the primary and the validation cohort.

Finally, we built a nomogram based on the most appropriate model, and the calibration plots were used to graphically investigate the performance characteristics of the nomogram. Internal validation of the nomogram was performed in the validation cohort, and a calibration curve was performed.

Development of decision-curve analyses

To evaluate the added value of radiomics signature to clinical features in individually predicting pulmonary metastasis for CRC patients, we developed three decision curves based on the clinical risk factors, rad-score and the combined clinical-radiomics model, and the clinical utility could be demonstrated by calculating the net benefits for a range of threshold probabilities.

Statistical analysis was conducted with R software (version 3.3.3; http://www.Rproject.org). A two-sided $p < 0.05$ was considered significant. The statistical analysis packages are listed in the Electronic Supplementary Material.

Results

Clinical characteristics and development of the clinical-only model

Patients characteristics are shown in Table 1 and Supplementary Table S1. There were no significant differences between the training and validation cohorts (Supplementary Table S1). In addition, the chi-square test and t-test demonstrated the primary tumour site, N stage, diameter and chronicity as risk factors of developing lung metastasis. After multivariate analysis, N stage (odds ratio (OR) = 3.197; 95% CI: 1.352–7.560; $p = 0.008$), chronicity (OR = 9.013; 95% CI: 3.752–21.654; $p < 0.001$) and diameter (OR = 1.228; 95% CI: 1.104–1.367; $p < 0.001$) remained independent predictors (Table 2) in the clinical-only model as shown in Table 2.

Radiomics signature building and diagnostic validation

By using the LASSO regression model, 203 texture features were reduced to five potential predictors (40.6:1 ratio; Fig. 2A and B). These features were presented in the rad-score calculated by using the following formula: rad-score = -0.119404424+[W_GLRMS]-LL-LRE-Average*0.001169257+[W_GLCM]-HH-Information_Measures_I-Average*3.997517837+[W_GLCM]-LH-Difference variance-Average*(-0.041391190)+[GLCM]-Maximal_Correlation_Coefficient-Average*0.533372862+[GLCM]-Mean-Average*0.039599734. Figure 2C presents the distribution of the texture features and the combined rad-score in the LM and NM groups; we found that LM patients had higher rad-scores compared to NM patients. There was a significant difference in rad-score between the LM and NM groups in the training cohort ($p < 0.001$) using univariate logistic regression analysis (Table 2).

Development of the clinical-radiomics model

As seen in Table 2, the stepwise logistic regression model selected the rad-score (OR = 8.820; 95% CI: 4.149–18.747; $p < 0.001$), N stage (OR = 3.683; 95% CI: 1.251–10.846; $p = 0.018$) and chronicity (OR = 7.943; 95% CI: 2.648–23.830; $p < 0.001$) as significant predictors for pulmonary metastasis. Moreover, the rad-score was the dominant factor impacting the prediction of lung metastasis in the clinical-radiomics model.

Model comparison and nomogram apparent performance

The clinical-radiomics model (AIC = 98.893, χ^2 = 106.490) with the highest chi-square value and the lowest AIC value compared with the clinical-only model (AIC = 138.502, χ^2 = 66.884) or the radiomics-only model (AIC = 116.146, χ^2 = 85.240) was identified as the best model. The clinical-radiomics model also showed the best discrimination, with the AUC reaching 0.929 (95% CI: 0.885–0.974) in the training cohort and 0.922 (95% CI:

Table 1 Comparison of lung metastasis (LM) and non-metastasis (NM) in the training and validation cohorts

Characteristic	Training cohort			Validation cohort		
	LM	NM	*p*	LM	NM	*p*
Rad-score, mean SD	0.6± 0.56	(-0.7) ± 0.85	<0.001*	0.6±0.69	(-0.6) ±0.72	<0.001*
Age, mean SD	60.6± 10.54	57± 10.73	0.050	59.6±9.61	57±11.63	0.359
Gender (%)			0.747			0.412
Male	37 (55.2%)	40 (58%)		15 (57.7%)	15 (46.9%)	
Female	30 (44.8%)	29 (42%)		11 (42.3%)	17 (53.1%)	
Primary tumour (%)			0.033*			0.239
Rectum	46 (68.7%)	35 (50.7%)		17 (65.4%)	16 (50%)	
Colon	21 (31.3%)	34 (49.3%)		9 (34.6%)	16 (50%)	
CEA level (%)			0.626			0.280
0~5 ng/ml	51 (76.1%)	50 (72.5%)		17 (65.4%)	25 (78.1%)	
≥ 5 ng/ml	16 (23.9%)	19 (27.5%)		9 (34.6%)	7 (21.9%)	
CA19-9 level (%)			0.762			0.458
0~27 U/ml	54 (80.6%)	57 (82.6%)		19 (73.1%)	26 (81.2%)	
≥ 27 U/ml	13 (19.4%)	12 (17.4%)		7 (26.9%)	6 (18.8%)	
T stage (%)			0.636			0.080
T1-2	17 (25.4%)	20 (29%)		5 (19.2%)	13 (40.6%)	
T3-4	50 (74.6%)	49 (71%)		21 (80.8%)	19 (59.4%)	
N stage (%)			0.001*			0.771
N0	23 (34.3%)	43 (62.3%)		12 (46.2%)	16 (50%)	
N1-2	44 (65.7%)	26 (37.7%)		14 (53.8%)	16 (50%)	
ITT (%)			0.756			0.083
(-)	56 (83.6%)	59 (85.5%)		25 (96.2%)	26 (81.2%)	
(+)	11 (16.4%)	10 (14.5%)		1 (3.8%)	6 (18.8%)	
PNI (%)			0.399			0.980
(-)	57 (85.1%)	62 (89.9%)		22 (84.6%)	27 (84.4%)	
(+)	10 (14.9%)	7 (10.1%)		4 (15.4%)	5 (15.6%)	
Nodules number (%)			0.766			0.548
Single	37 (55.2%)	26 (37.7%)		9 (34.6%)	13 (40.6%)	
Double	15 (22.4%)	18 (26.1%)		6 (23.1%)	8 (25%)	
Multiple	27 (40.3%)	25 (36.2%)		11 (42.3%)	11 (34.4%)	
Diameter (%)			<0.001*			0.007*
5–10mm	10 (14.9%)	43 (62.3%)		8 (30.8%)	22 (68.8%)	
10–15mm	30 (44.8%)	18 (26.1%)		8 (30.8%)	6 (18.7%)	
15–20mm	27 (40.3%)	8 (11.6%)		10 (38.4%)	4 (12.5%)	
Laterality			0.624			0.786
Unilateral	39 (58.2%)	43 (62.3%)		17 (65.4%)	22 (68.8%)	
Bilateral	28 (41.8%)	26 (37.7%)		9 (34.6%)	10 (31.2%)	
Chronicity			<0.001*			<0.001*
Metachronous	51 (76.1%)	19 (27.5%)		21 (80.8%)	11 (34.4%)	
Synchronous	16 (23.9%)	50 (72.5%)		5 (19.2%)	21 (65.6%)	

Chi-square tests were used to compare the differences in categorical variables (Gender, Primary tumour, CEA level, CA19-9 level, T stage, N stage, ITT, PNI, Nodules number, Laterality and Chronicity), while a two-sample t-test was used to compare the differences in age, diameter and Rad-score

OR odds ratio, *CEA* carcinoembryonic antigen, *CA19-9* carbohydrate antigen 19-9, *ITT* intravascular tumour thrombus, *PNI* perineural invasion

0.857–0.986) in the validation cohort compared with the clinical-only model (AUC = 0.849 in the training cohort and AUC = 0.852 in the validation cohort) or the radiomics-only model (AUC = 0.887 in the training cohort and AUC = 0.881 in the validation cohort). The ROC curves of the three models were presented in Fig. 4 in both the training and the validation cohort. On the other hand, as shown in Table 3, the best accuracy for predicting pulmonary metastasis was the clinical-radiomics model (AUC = 0.929, 95% CI: 0.885–0.974; sensitivity: 84.9%; specificity: 91.1%; accuracy: 88.2%; PPV: 92.5%; NPV: 84.1%).

The clinical-radiomics nomogram was successfully developed based on the clinical-radiomics model (Fig. 3A). As the nomogram illustrates, the rad-score accounted for a vast majority of the proportion compared to the other clinical features, which made radiomics signature the cardinal biomarker for LM prediction. The calibration plots also indicated good agreement between the nomogram

Table 2 Comparison of three models by multivariate logistic regression analysis

Characteristics	OR (95% CI)	*p*	AIC	LR (χ2)
Clinical-only model			138.502	66.884
N stage (N0 vs. N1-2)	3.197 (1.352–7.560)	0.008*		
Diameter	1.228 (1.104–1.367)	<0.001*		
Chronicity (synchronous vs. metachronous)	9.013 (3.752–21.654)	<0.001*		
Radiomics-only model			116.146	85.240
Rad-score	9.517 (4.744–19.093)	<0.001*		
Clinical-radiomics model			98.893	106.490
Rad-score	8.820 (4.149–18.747)	<0.001*		
N stage (N0 vs. N1-2)	3.683 (1.251–10.846)	0.018*		
Chronicity (synchronous vs. metachronous)	7.943 (2.648–23.830)	<0.001*		

Regarding discriminatory ability, homogeneity and monotonicity of gradients, the model with a higher χ2 value by the likelihood ratio test was considered the better model. Furthermore, the lower value for AIC value represented the better model for discriminatory ability

CI confidence interval, *AIC* Akaike information criterion, *LR* likelihood ratio

*$p < 0.05$

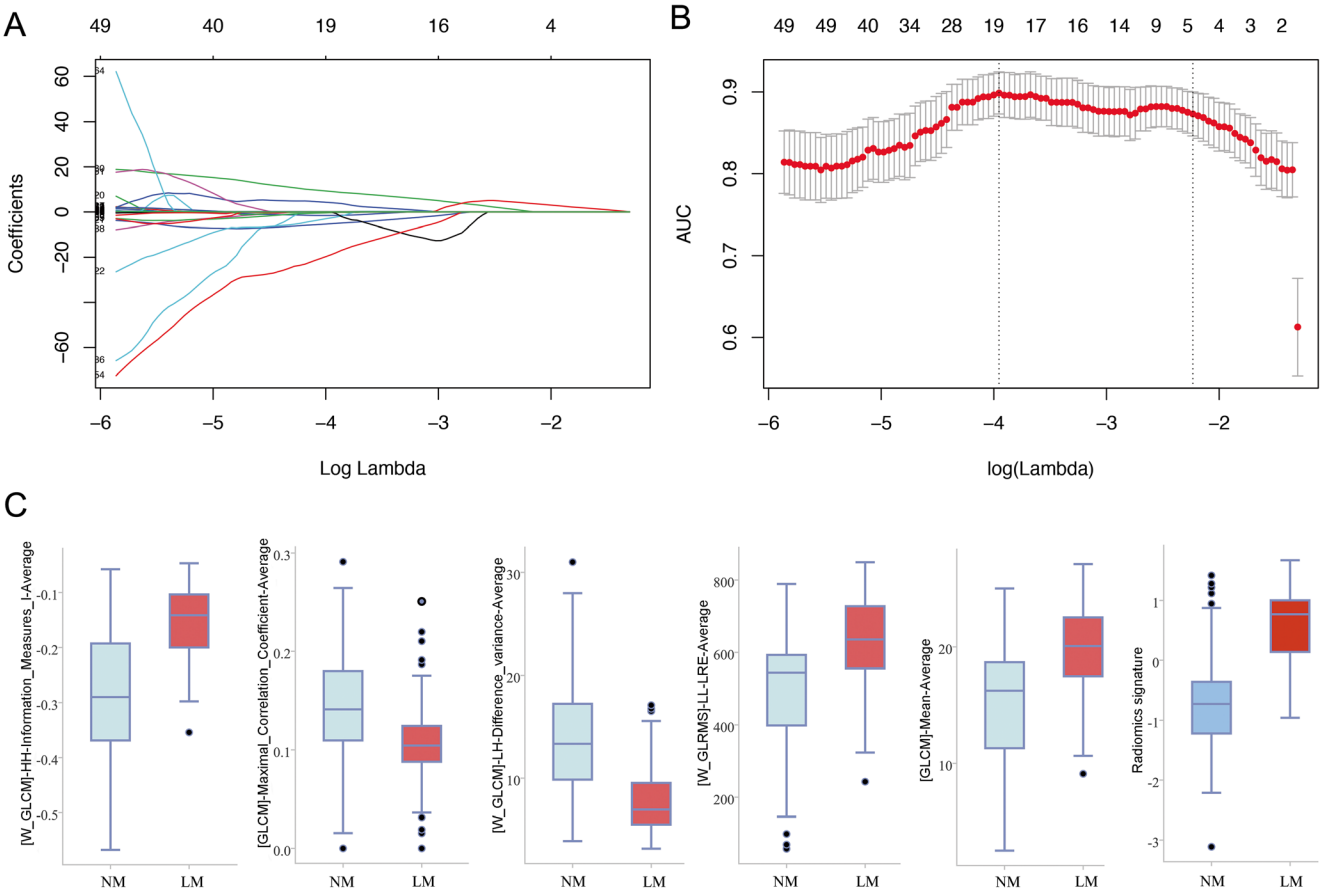

Fig. 2 (**A**) Texture feature selection by using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model. The differentiation performance of the radiomics signature was explored on the receiver operating characteristics (ROC) curve. Tuning parameter (λ) selection in the LASSO model used tenfold cross-validation via minimum criteria. Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria). A λ value of 0.21, with log (λ), -2.25 was chosen (1-SE criteria) according to tenfold cross-validation. (**B**) LASSO coefficient profiles of the 203 pulmonary-metastases-related texture features. A coefficient profile plot was produced against the log (λ) sequence. A vertical line is drawn at the value chosen by tenfold cross-validation. (**C**) Range of the selected five texture features ([GLCM]-LL-std-Average, [GLCM]-HH-Information_Measures_I-Average, [GLCM]-LH-Difference_variance-Average, [GLCM]-Maximal_Correlation_Coefficient-Average, [GLCM]-Mean-Average) and the Rad-score in the LM and NM group

Table 3 Accuracy and predictive value between three models

Training cohort	AUC	95% CI	Sensitivity	Specificity	Accuracy	PPV	NPV
Clinical features	0.849	0.784–0.915	81.3% (51/64)	77.8% (56/72)	78.7% (107/136)	76.1% (51/67)	81.2% (56/69)
Rad-score	0.887	0.826–0.948	82.2% (60/73)	88.9% (56/63)	85.3% (116/136)	89.6% (60/67)	81.2% (56/69)
Clinical-radiomics	0.929	0.885–0.974	84.9% (62/73)	91.1% (58/63)	88.2% (120/136)	92.5% (62/67)	84.1% (58/69)
Validation cohort	AUC	95% CI	Sensitivity	Specificity	Accuracy	PPV	NPV
Clinical features	0.852	0.756–0.947	71.4% (20/28)	80% (24/30)	75.9% (44/58)	76.9% (20/26)	75% (24/32)
Rad-score	0.881	0.796–0.966	70% (21/30)	82.1% (23/28)	75.9% (44/58)	80.8% (21/26)	71.9% (23/32)
Clinical-radiomics	0.922	0.857–0.986	75.9% (22/29)	86.2% (25/29)	81.0% (47/58)	84.6% (22/26)	78.1% (25/32)

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

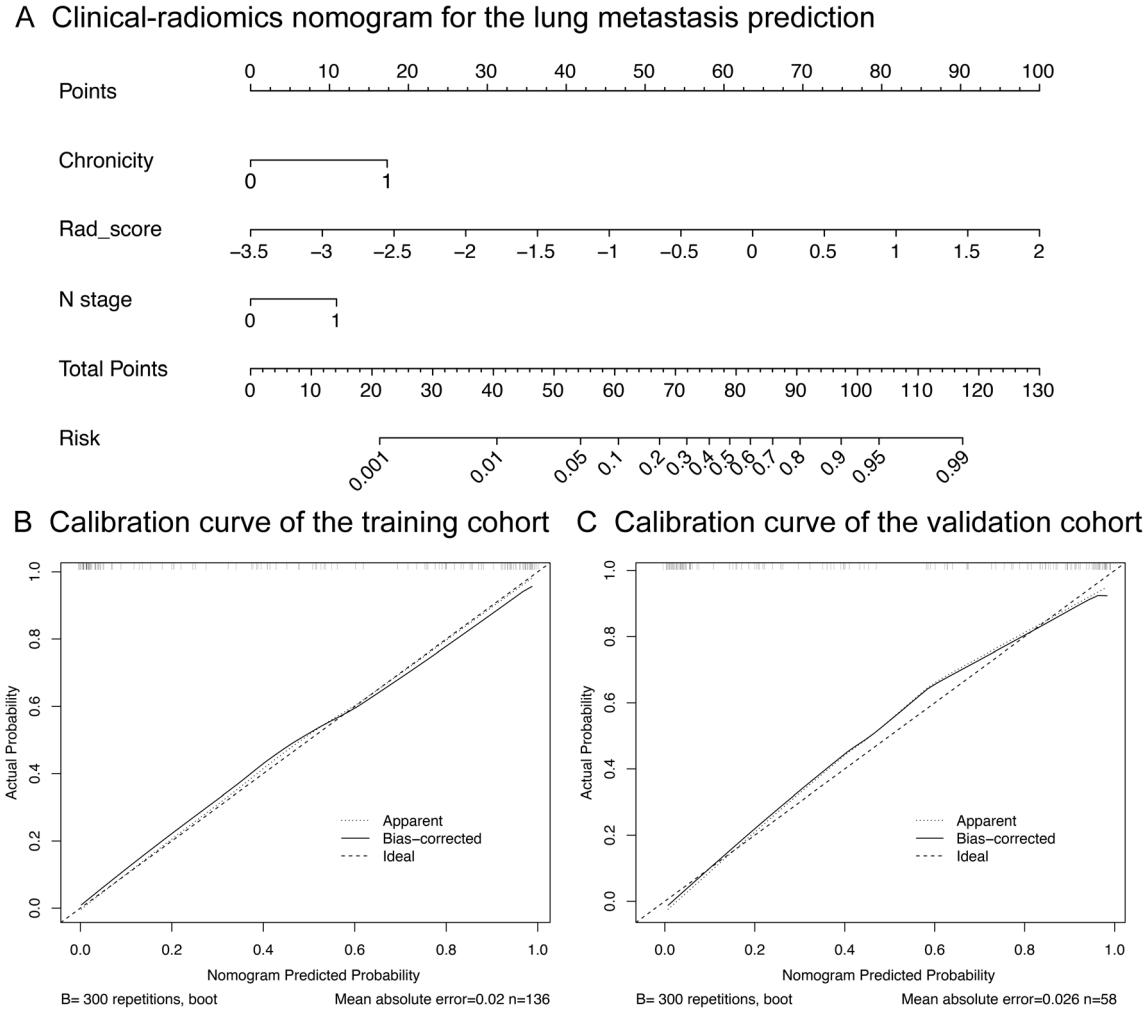

Fig. 3 (A) The developed clinical-radiomics nomogram for predicting the probability of pulmonary metastases. For the chronicity, 0 represents the synchronous nodule, while 1 represents metachronous lesion. For the N stage, 0 for the N0 and 1 for the N1-2. To use, locate the patient's N stage, draw a line straight up to the points axis to establish the score associated with that site. Repeat for the other covariates (chronicity and rad-score). By summing the scores of each point and locating it on the total score scale, the estimated probability of pulmonary metastases could be determined. (B and C) Calibration curves for predicting pulmonary metastases in the training and validation cohort. The 45° straight line represents the perfect match between the actual (Y-axis) and nomogram-predicted (X-axis) survival probabilities. A closer distance between two curves indicates higher accuracy

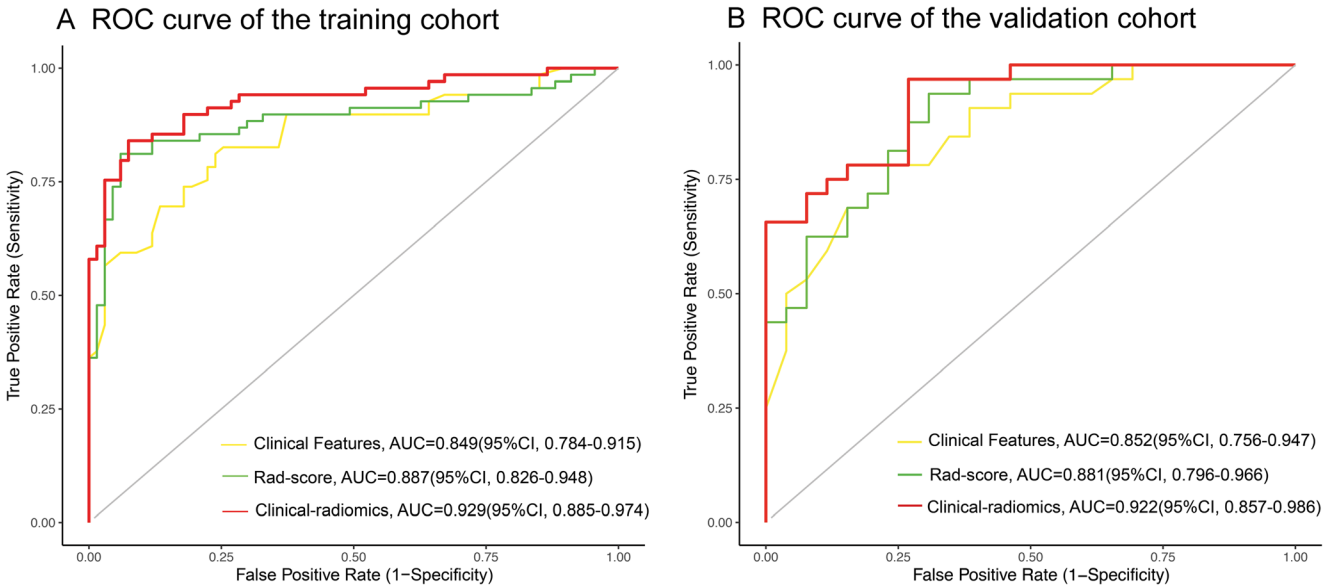

Fig. 4 The receiver operating characteristic (ROC) curves of the clinical features, rad-score and clinical-radiomics in the training (**A**) and validation sets (**B**), respectively

prediction and actual observation for LM and NM in both the training and the validation cohort (Fig. 3B and C).

Clinical use of DCA

The DCAs based on the clinical model, radiomics model and the combined clinical-radiomics model are shown in Fig. 5. The clinical DCA showed less benefit while the rad-score DCA showed more benefit in predicting the risk of metastases in 10–90% threshold probabilities, which indicates that within this range, rad-score outperformed the clinical features with more accuracy in LM prediction. With the adding of the radiomic signature, the clinical-radiomics nomogram achieved the most clinical utility with almost

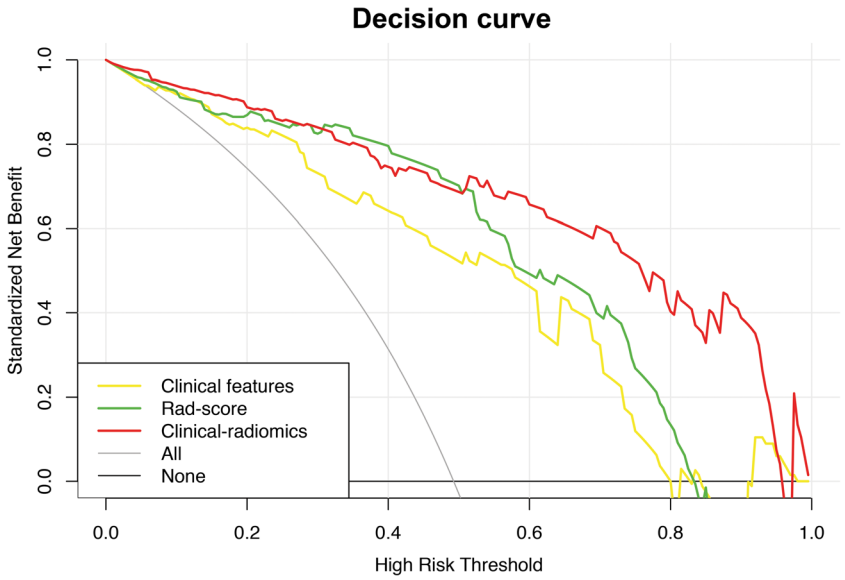

Fig. 5 The decision-curve analysis for the clinical features and the rad-score as well as the combined model. The y-axis measures the net benefit. The yellow line represents the model of clinical features, and the green line the rad-score and the red line the combination of clinical-radiomics. The grey line represents the assumption that all patients have LM metastases. The black line represents the assumption that no patients have LM metastases. Threshold probability refers to the point at which a patient considers the benefit of treatment for intermediate- to high-risk pulmonary metastases equivalent to the harm of over-treatment for low-risk disease and thus reflects how the patient weighs up the benefits and risks associated with the decision. The higher curve at any given threshold probability is the optimal prediction to maximise net benefit. Across the various threshold probabilities, the clinical-radiomics curve showed maximised net benefit compared with clinical features individually performance

all of the threshold probabilities, which indicates that the nomogram is a reliable clinical treatment tool to predict pulmonary metastases for CRC patients with lung nodules.

Discussion

This is the first study to develop a clinical-radiomics nomogram based on the radiomics analysis incorporating with key clinical risk factors to quantitatively assess the risk of LM for colorectal cancer patients with IPNs. Firstly, the clinical factors most associated with the metastasis are based on N stage, chronicity and size of the nodule. Secondly, these factors are not sufficient for decision making, as we found that the addition of a radiomic analysis would strongly and independently affect metastasis prediction. The performance of the clinical-radiomics nomogram is enhanced compared with the clinical features or radiomics signature with the AUC of 0.929 compared with 0.849 and 0.887, respectively, and this new model was capable of differentiating metastasis from non-metastasis nodules with a sensitivity of 84.9% and a specificity of 91.1% in our training cohort. Therefore, we considered that the clinical-radiomics nomogram will be helpful to differentiate metastatic nodules in clinical decision making.

Of the two clinical features used to predict metastasis in our study, a positive nodal status was consistently demonstrated in previous studies. There are many studies [14, 15] reporting nodal involvement as the risk factor for metastasis. Similarly, we found higher nodal stage to be a significant factor in metastasis in CRC patients with pulmonary nodules, which suggested that patients with a positive nodal disease were at increased risk of malignancy of coexisting IPNs and they may need close monitoring for early detection of the progression of lung lesions. On the other hand, metachronous nodules were closely correlated with the occurrence of LM and have an excess predictive impact compared with that of N stage, as shown in the nomograms. To date, only a few studies have researched the association between metachronous nodules and metastases. Consistent with our findings, Kim et al. [14] identified metachronous IPNs as a risk factor for progression to metastases. A reasonable explanation may be that compared with synchronous lesions, metachronous nodules could be associated with a relatively higher rate of IPNs [16], which may lead to a higher metastasis probability.

Compared with the published studies concentrating on relevant risk factors for metastasis evidence for IPNs, our study not only evaluated the clinical features but also tried to explore more information in CT images and use the radiomics process to quantify not only the morphological characteristics such as intensity, shape, size or volume but also the internal texture features that are unable to be captured by our human visual system but can be reflected on radiomics analysis. Radiomics has gained increasing attention due to its potential to build predictive models relating image features to phenotypes or gene signatures [17, 18]. As we know, benign and malignant lung nodules are completely different in cell morphology and biological behaviour [19], with different internal structures reflected through their diverse extent of hypoxia, angiogenesis, inflammation and glucose metabolism [20, 21]. The heterogeneity between benign nodules and metastatic ones therefore should be revealed by radiomics features. Several studies had demonstrated that feature extraction of lung nodules could detect small invasive components to differentiate between malignant and benign nodules with the AUC varying from 0.72 to 0.90 [22–25].

In the radiomics analysis, we identified five parameters: [W_GLRMS]-LL-std-Average, [W_GLCM]-HH-Information_Measures_I-Average, [W_GLCM]-LH-Difference_variance-Average, [GLCM]-Maximal_Correlation_Coefficient-Average and [GLCM]-Mean-Average (four in five textural within the grey-level co-occurrence matrix (GLCM)) and the combined radiomics signature showed superior performance in predicting LM (AUC = 0.887, 95% CI: 0.826–0.948). In addition, as has been previously investigated, Hanania et al. [26] identified 14 top-performing radiomic features (all GLCM parameters) that can differentiate between benign and malignant IPN pathology. In another study with a large cohort [8], they also found that GLCM texture features showed strong correlations with gene expression in both lung cancer and head and neck cancer. As the most commonly used texture parameters, the GLCM parameters are constructed by using the number, distance and angle of a combination of grey levels that is reflected in the image [27]. This may have an extraordinary performance in lung cancer diagnosis as there is a natural advantage with the high contrast between pulmonary nodules and lung parenchyma [19].

Finally, to extrapolate to clinical use, we developed and validated the clinical-radiomics nomogram for the clinician to individually predict the LM risk for each CRC patient. As the clinical-radiomics model outperformed the clinical features and radiomics signature with a higher AUC and more net benefits across the majority of the range of threshold probabilities in decision-curve analysis, it may be the most promising approach to guide clinical management. We recommend that patients with a higher N stage or metachronous nodules should be subjected to close supervision of follow-up computed tomographic examinations to assess lesion progression. For those with a higher risk of metastasis after calculating the total points, we suggest using them as potential operative candidates to achieve

prolonged survival as there exists a higher risk of LM. We believe that the clinical use of the nomogram can not only avoid unnecessary surgery in patients with benign nodules, but also reduce the burden of costs from follow-up diagnostic procedures and anxiety associated with false-positive tests, or exclude patients with a surgically curable disease from a potentially beneficial treatment.

Our study had certain limitations. With the retrospective design, there was a potential for selection bias by excluding patients receiving systemic treatment; the selected cases without chemoradiotherapy might have had a lower histopathological grade and different distribution of targeted clinical characteristics. In addition, stringent external validation still needs to be performed in multicentre clinical trials for further generalisation. Finally, at present radiomics analysis seems to be time consuming; however, compared with the clinical dilemma of whether to receive chemoradiotherapy or surgery or imaging follow-up, we consider it more cost-efficient to get a relatively definite diagnosis and initiate timely treatment, thus maximising the patient's benefit. In addition, one promising direction that radiomics is moving toward is adopting the machine-learning method, which can automatically make predictions from complex data thus making end-to-end models possible. Therefore, radiomics analysis is still worthy and we think this study shows bright prospects.

In conclusion, the identified clinical features and radiomics signature have the potential to be used as non-invasive biomarkers for LM risk prediction for CRC patients with lung nodules. The clinical-radiomics nomogram described in this study demonstrated the incremental value of the radiomics signature for individualised metastasis estimation and may serve as a potential clinical treatment tool.

Funding This study has received funding by the National Science Foundation for Young Scientists of China (Grant No.81501437).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Tong Tong.

Conflict of interest The authors of this article declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry Shengping Wang has significant statistical expertise.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- observational
- performed at one institution

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. CA Cancer J Clin 65:87–108
2. Kobayashi H, Mochizuki H, Sugihara K et al (2007) Characteristics of recurrence and surveillance tools after curative resection for colorectal cancer: a multicenter study. Surgery 141:67–75
3. Desch CE, Benson AB 3rd, Somerfield MR et al (2005) Colorectal cancer surveillance: 2005 update of an American Society of Clinical Oncology practice guideline. J Clin Oncol 23:8512–8519
4. Vachani A, Tanner NT, Aggarwal J et al (2014) Factors that influence physician decision making for indeterminate pulmonary nodules. Ann Am Thorac Soc 11:1586–1591
5. NM A (2003) The solitary pulmonary nodule. N Engl J Med 349(16):1575
6. Inoue M, Ohta M, Iuchi K et al (2004) Benefits of surgery for patients with pulmonary metastases from colorectal carcinoma. Ann Thorac Surg 78:238–244
7. De Wever W, Meylaerts L, De Ceuninck L, Stroobants S, Verschakelen JA (2007) Additional value of integrated PET-CT in the detection and characterization of lung metastases: correlation with CT alone and PET alone. Eur Radiol 17:467–473
8. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006
9. Zhang Y, Moore GR, Laule C et al (2013) Pathological correlates of magnetic resonance imaging texture heterogeneity in multiple sclerosis. Ann Neurol 74:91–99
10. Yoon HJ, Kim Y, Kim BS (2015) Intratumoral metabolic heterogeneity predicts invasive components in breast ductal carcinoma in situ. Eur Radiol 25:3648–3658
11. Win T, Miles KA, Janes SM et al (2013) Tumor heterogeneity and permeability as measured on the CT component of PET/CT predict survival in patients with non-small cell lung cancer. Clin Cancer Res 19:3591–3599
12. Ganeshan B, Goh V, Mandeville HC (2013) Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. Radiology 266:326–336
13. Ahmed A, Gibbs P, Pickles M, Turnbull L (2013) Texture analysis in assessment and prediction of chemotherapy response in breast cancer. J Magn Reson Imaging 38:89–101
14. Baek SJ, Kim SH, Kwak JM et al (2012) Indeterminate pulmonary nodules in rectal cancer: a recommendation for follow-up guidelines. J Surg Oncol 106:481–485
15. Kim CH, Huh JW, Kim HR, Kim YJ (2015) Indeterminate pulmonary nodules in colorectal cancer: follow-up guidelines based on a risk predictive model. Ann Surg 261:1145–1152
16. Mitry E, Guiu B, Cosconea S, Jooste V, Faivre J, Bouvier AM (2010) Epidemiology, management and prognosis of colorectal cancer with lung metastases: a 30-year population-based study. Gut 59:1383–1388
17. Cufer T, Ovcaricek T, O'Brien ME (2013) Systemic therapy of advanced non-small cell lung cancer: major-developments of the last 5-years. Eur J Cancer 49:1216–1225
18. Kumar V, Gu Y, Basu S et al (2012) Radiomics: the process and the challenges. Magn Reson Imaging 30:1234–1248
19. Wei K, Su H, Zhou G et al (2016) Potential Application of Radiomics for Differentiating Solitary Pulmonary Nodules. OMICS J Radiol 5:1-5
20. Park J, Kobayashi Y, Urayama KY, Yamaura H, Yatabe Y, Hida T (2016) Imaging Characteristics of Driver Mutations in EGFR, KRAS, and ALK among Treatment-Naive Patients with Advanced Lung Adenocarcinoma. PLoS One 11:e0161081
21. Divine MR, Katiyar P, Kohlhofer U, Quintanilla-Martinez L, Pichler BJ, Disselhorst JA (2016) A Population-Based Gaussian Mixture

Model Incorporating 18F-FDG PET and Diffusion-Weighted MRI Quantifies Tumor Tissue Classes. J Nucl Med 57:473–479
22. Wu W, Parmar C, Grossmann P et al (2016) Exploratory study to identify radiomics classifiers for lung cancer histology. Front Oncol 6:71
23. Kamiya A, Murayama S, Kamiya H, Yamashiro T, Oshiro Y, Tanaka N (2014) Kurtosis and skewness assessments of solid lung nodule density histograms: differentiating malignant from benign nodules on CT. Jpn J Radiol 32:14–21
24. Ko JP, Suh J, Ibidapo O et al (2016) Lung adenocarcinoma: correlation of quantitative CT findings with pathologic findings. Radiology 280:931–939
25. Son JY, Lee HY, Kim JH et al (2016) Quantitative CT analysis of pulmonary ground-glass opacity nodules for distinguishing invasive adenocarcinoma from non-invasive or minimally invasive adenocarcinoma: the added value of using iodine mapping. Eur Radiol 26:43–54
26. Hanania AN, Bantis LE, Feng Z et al (2016) Quantitative imaging to evaluate malignant potential of IPMNs. Oncotarget 7:85776–85784
27. Lee G, Lee HY, Park H et al (2017) Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. Eur J Radiol 86:297–307