European Radiology (2018) 28:5121–5128
https://doi.org/10.1007/s00330-018-5509-9

CHEST

The predictive value of CT-based radiomics in differentiating indolent from invasive lung adenocarcinoma in patients with pulmonary nodules

Yunlang She[1] · Lei Zhang[1] · Huiyuan Zhu[1] · Chenyang Dai[1] · Dong Xie[1] · Huikang Xie[2] · Wei Zhang[2] · Lilan Zhao[3] · Liling Zou[4,5] · Ke Fei[1] · Xiwen Sun[6] · Chang Chen[1]

Received: 16 October 2017 / Revised: 19 March 2018 / Accepted: 20 April 2018 / Published online: 4 June 2018

Abstract

Objectives Adenocarcinoma in situ (AIS) and minimally invasive adenocarcinoma (MIA) are assumed to be indolent lung adenocarcinoma with excellent prognosis. We aim to identify these lesions from invasive adenocarcinoma (IA) by a radiomics approach.

Methods This retrospective study was approved by institutional review board with a waiver of informed consent. Pathologically confirmed lung adenocarcinomas manifested as lung nodules less than 3 cm were retrospectively identified. In-house software was used to quantitatively extract 60 CT-based radiomics features quantifying nodule's volume, intensity and texture property through manual segmentation. In order to differentiate AIS/MIA from IA, least absolute shrinkage and selection operator (LASSO) logistic regression was used for feature selection and developing radiomics signatures. The predictive performance of the signature was evaluated via receiver operating curve (ROC) and calibration curve, and validated using an independent cohort.

Results 402 eligible patients were included and divided into the primary cohort (n = 207) and the validation cohort (n = 195). Using the primary cohort, we developed a radiomics signature based on five radiomics features. The signature showed good discrimination between MIA/AIS and IA in both the primary and validation cohort, with AUCs of 0.95 (95% CI, 0.91–0.98) and 0.89 (95% CI, 0.84–0.93), respectively. Multivariate logistic analysis revealed that the signature (OR, 13.3; 95% CI, 6.2–28.5; p < 0.001) and gender (OR, 3.5; 95% CI, 1.2–10.9; p = 0.03) were independent predictors of indolent lung adenocarcinoma.

Conclusion The signature based on radiomics features helps to differentiate indolent from invasive lung adenocarcinoma, which might be useful in guiding the intervention choice for patients with pulmonary nodules.

Key points

• *Based on radiomics features, a signature is established to differentiate adenocarcinoma in situ and minimally invasive adenocarcinoma from invasive lung adenocarcinoma.*

Keywords Lung neoplasms · Tomography, spiral computed · Radiomics · Multivariate analysis · Forecasting

Yunlang She and Lei Zhang contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5509-9) contains supplementary material, which is available to authorized users.

✉ Xiwen Sun
479082599@qq.com

✉ Chang Chen
chenthoracic@163.com

[1] Department of Thoracic Surgery, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Zhengmin Road 507, Shanghai 200433, People's Republic of China

[2] Department of Pathology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, People's Republic of China

[3] Department of General Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[4] Department of Medical Statistics, Tongji university School of Medicine, Shanghai, People's Republic of China

[5] Clinical and Translational Science Institute, University of Rochester Medical Center, Rochester, NY, USA

[6] Department of Radiology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Zhengmin Road 507, Shanghai 200433, People's Republic of China

Abbreviations

AIS	Adenocarcinoma in situ
AUC	Area under the curve
CT	Computed tomography
IA	Invasive adenocarcinoma
LASSO	Least absolute shrinkage and selection operator
MIA	Minimally invasive adenocarcinoma

Introduction

The prognostic and predictive value of the new classification system for lung adenocarcinoma proposed by the International Association for Study of Lung Cancer, American Thoracic Society and the European Respiratory Society has been validated [1], and it formed the basis of the fourth edition of the World Health Organization classification of lung cancer [2]. In particular, adenocarcinoma in situ (AIS) and minimally invasive adenocarcinoma (MIA) were newly defined and assumed to be indolent lesions because of the excellent prognosis compared with invasive adenocarcinomas (IA), and published evidence supports management of these lesions with sublobar resection [3–6]. Thus, it is of utmost importance to accurately diagnose these indolent lesions from invasive pulmonary adenocarcinoma prior to or during surgery. Unfortunately, even with intraoperative frozen section, it is still challenging to accurately diagnose these lesions because of the requirement to evaluate the entire tumour to rule out the existence of an invasive component [6, 7].

In contrast, high resolution thin-section computed tomography (CT) allows comprehensive and non-invasive characterisation of pulmonary nodules. However, by visually assessing the CT images, we may miss out some important information to define the nodule's type. In this context, radiomics, which aims to extract high-throughput data and analyse large amounts of advanced quantitative features from medical images, may play a valuable role [8–10]. Recent studies have revealed that radiomic analysis showed promising potential for lung nodule characterisation [11–14].

However, previous studies only focused on finding some predictive radiomics features to differentiate preinvasive from invasive lung adenocarcinoma without validation [11–13]. In our study, a radiomics-based predictive model was established using least absolute shrinkage and selection operator (LASSO) logistic analysis, which showed excellent performance in predicting AIS/MIA from IA. To our knowledge, this is the first validated radiomics model for predicting preinvasive from invasive lung adenocarcinoma.

Materials and methods

Patients

Our institutional review board approved this retrospective study with a waiver of informed consent. Using the descriptive terms "subsolid nodule", "solid nodule", "part-solid nodule", "ground-glass nodule" or "ground glass opacity", we retrieved CT examinations of patients who underwent lung resection between January 2016 and December 2016 because of lung adenocarcinoma. Two thoracic radiologists (H.Z. and X.S.) re-evaluated the CT scans in consensus and assessed the lesions for inclusion. The pulmonary nodule is defined as a single, well-circumscribed, radiographic opacity that measures less than 30 mm in diameter and is completely surrounded by pulmonary parenchyma without atelectasis, mediastinal lymphadenopathy or pleural effusion [15]. Those with multiple lesions, large lesions (greater than 3 cm), CT slice thickness greater than 1 mm or without a diagnosis confirmed by pathologic analysis were excluded. In addition, patients' clinical characteristics were collected from medical records. Eligible patients were divided into the primary cohort (January to June) and the validation cohort (July to December).

Image acquisition protocol and radiomic features extraction

Preoperative chest CT scans were obtained using a Somatom Definition AS (Siemens Medical Systems, Germany) or Brilliance 40 (Philips Medical Systems, Netherlands). For both manufacturers, CT scans were acquired at full inspiration without contrast medium at 120 kVp tube energy and 200 mAs effective dose. In the Siemens group, a Somatom Definition AS scanner (64 × 0.625 mm detector, 1.0 pitch) was used. All images were reconstructed at 1.0 mm slice thickness, with 0.7 mm increment and a standard soft kernel (Siemens B31 filter, Siemens Medical Solutions, Forchheim, Germany). In the Philips group, scans were taken using a Brilliance 40 scanner (40 × 0.625 mm detector configuration, 0.4 pitch). Images were reconstructed at 1.0 mm slice thickness, with 0.7 mm increment and a sharp reconstruction kernel (C filter, Philips, Cleveland, OH).

Radiomic features were extracted from nodules with in-house software implemented with Python Programming Language (http://www.python.org). Nodule segmentation was performed manually by a radiologist (H.Z.) with 2 years of experience in chest CT imaging and confirmed by another radiologist (X.S.) with 20 years of experience. Regions of interest were delineated around the nodule boundary for each section, and the process of manual segmentation per

nodule took about 5 min for experienced performers. After nodule segmentation, the software automatically calculated radiomic features for each included patient. Generally, CT-based radiomics features can be divided into four groups: (1) volumetric features, (2) histogram features, (3) textural features, (4) wavelet features. The wavelet features calculated by wavelet transformation were much more obscure and difficult to interpret [16]. In this study, we selected 60 radiomics features of the first three groups, which made intuitive sense and these were mostly adopted in other research [11–13]. Algorithms for feature calculation are described in the Supplementary material.

Histological evaluation

All specimens were formalin-fixed and stained with haematoxylin and eosin (HE). Two pathologists (Huikang Xie; Wei Zhang) re-evaluated and discussed all slides at a multi-headed microscope and discussed until consensus was achieved. According to the new classification of lung adenocarcinoma [1], each histological pattern present was recorded in 5% increments. On the basis of the types of invasive pattern, patients were categorised into an indolent (AIS/MIA) or IA group.

Construction of Rad-score using the LASSO regression model

As a result of the existence of multi-collinearity between radiomics features, the LASSO binary logistic regression model was used to select the optimal subset of radiomic features in order to develop the radiomic signature score (Rad-score). The mechanism of the LASSO regression analysis was to introduce a penalty parameter (also called as tuning parameter) to penalise the coefficient of variables entered into the LASSO regression model in order to avoid the overfitting problem. As the tuning parameter (λ) increases, more coefficients were set to zero (less variables are selected), and among the non-zero coefficients, more shrinkage was employed. The area under the receiver operating characteristic (AUC) curve was plotted versus log(λ) in order to identify the optimal value of log(λ). The optimal value was identified by the minimum criterion and the one standard error of the minimum criterion. The "glmnet" package of R software was used to perform the LASSO binary logistic regression analysis and the programing process is listed in the Supplementary material [17–19]

Statistical analysis

Continuous and categorical variables were compared using the *t* test and Fisher's exact test, respectively. Multivariable logistic regression analysis was used to select the independent prognostic factors. The performance of the model was assessed in the primary and validation cohorts. The discrimination of the signature was measured by the area under the curve (AUC). The apparent calibration curve was plotted using model-predicted probability against actual probability of invasive adenocarcinoma, and the bias-corrected curve was also produced with 1000 bootstrap resamples. Statistical analysis was performed with SPSS for Windows, version 20.0 (IBM, Armonk, NY, USA). A two-sided *p* value was always computed, and a difference was considered significant at $p < 0.05$.

Results

Clinicopathologic characteristics

In this study, most participants received lung resection without preoperative biopsy, except that 28 patients were diagnosed as IA by transthoracic needle aspiration. Also, 22 patients with lung nodules confirmed as benign or atypical adenomatous hyperplasia were excluded. In total, 402 eligible patients were included with 207 in the primary cohort and 195 in the validation cohort. Patients' baseline characteristics in the primary and validation cohorts are listed in Table 1. In the primary cohort, 61.8% of the patients were female with 29% having a smoking history, and the median age was 58 years (range 21–84 years). In the validation cohort, the majority of patients were female (64.1%) and never smokers (79.5%) with a median age of 57.5 years (range 24–81 years). Most of the patients had a normal carcinoembryonic antigen level.

There were 30.0% (62/207) of patients with AIS/MIA in the primary cohort and 32.8% (64/195) in the validation cohort. Gender was significantly different between the AIS/MIA and IA groups both in both the primary ($p = 0.007$) and validation ($p = 0.011$) cohorts. Age showed marginal significance between these two groups in the primary cohort ($p = 0.052$), but evident difference ($p = 0.014$) in the validation cohort. Carcinoembryonic antigen level and smoking history were not significantly different in either the primary or validation cohort.

Feature selection of the radiomic signature

λ in the LASSO model was selected using 10-fold cross-validation and log(λ) of − 2.216 was chosen for the optimal subset of radiomics features. At this value, 60 radiomic features were reduced to five potential predictors with non-zero coefficients in the LASSO logistic regression model in the primary cohort (Fig. 1). Figure 1a, b respectively show that, as the log(λ) changes from − 10 to 0, the number of variables that entered into the model is reduced, and the absolute values of the coefficients of the variables also shrink toward zero.

Table 1 Characteristics of patients in the primary and validation cohorts

Characteristic	Primary cohort (n = 207)			Validation cohort (n = 195)		
	AIS/MIA	IA	p	AIS/MIA	IA	p
Age (years)			0.052			0.014
< 65	44	82		46	70	
≥ 65	18	63		18	61	
Gender			0.007			0.011
Male	15	64		15	55	
Female	47	81		49	76	
Smoking history			0.242			0.263
Never	48	99		54	101	
Current or ever	14	46		10	30	
CEA level			0.325			0.150
Normal (< 5 ng/ml)	62	140		62	119	
Abnormal (≥ 5 ng/m)	0	5		2	12	
CT scanner			0.73			0.36
Somatom Definition AS	14	36		18	29	
Brilliance 40	48	109		46	102	
Rad-score	0.33 (− 1.21 to 2.52)	− 1.86 (− 3.86 to 1.20)	< 0.001	0.23 (− 1.35 to 2.51)	− 1.53 (− 4.89 to 1.27)	< 0.001
Maximum 3D diameter (mm)	16.11 ± 5.26	22.49 ± 4.74	< 0.001	16.29 ± 5.55	22.30 ± 5.42	< 0.001
Root mean squared	1383.35 ± 88.78	1609.83 ± 160.04	< 0.001	1375.02 ± 93.14	1562.33 ± 149.09	< 0.001
Standard deviation	129.60 ± 34.04	215.34 ± 48.17	< 0.001	129.40 ± 35.75	210.62 ± 63.60	< 0.001
Entropy	4.24 ± 0.38	4.83 ± 0.29	< 0.001	4.23 ± 0.39	4.79 ± 0.34	< 0.001
Low grey level run emphasis	0.02 ± 0.01	0.01 ± 0.006	< 0.001	0.02 ± 0.01	0.01 ± 0.008	< 0.001

AIS adenocarcinoma in situ, *MIA* minimally invasive adenocarcinoma, *IA* invasive adenocarcinoma, *CEA* carcinoembryonic antigen

Patients with indolent lung adenocarcinoma generally had lower values of maximum 3D diameter, root mean squared, entropy and standard deviation, but higher values of long grey level run emphasis compared with those with IA (Table 1). In order to validate the stability and generalizability of the radiomics analysis by in-house software, all five selected features were re-extracted using 3D slicer (www.slicer.org), which demonstrated no significant difference between these two times of feature extraction (Supplementary material).

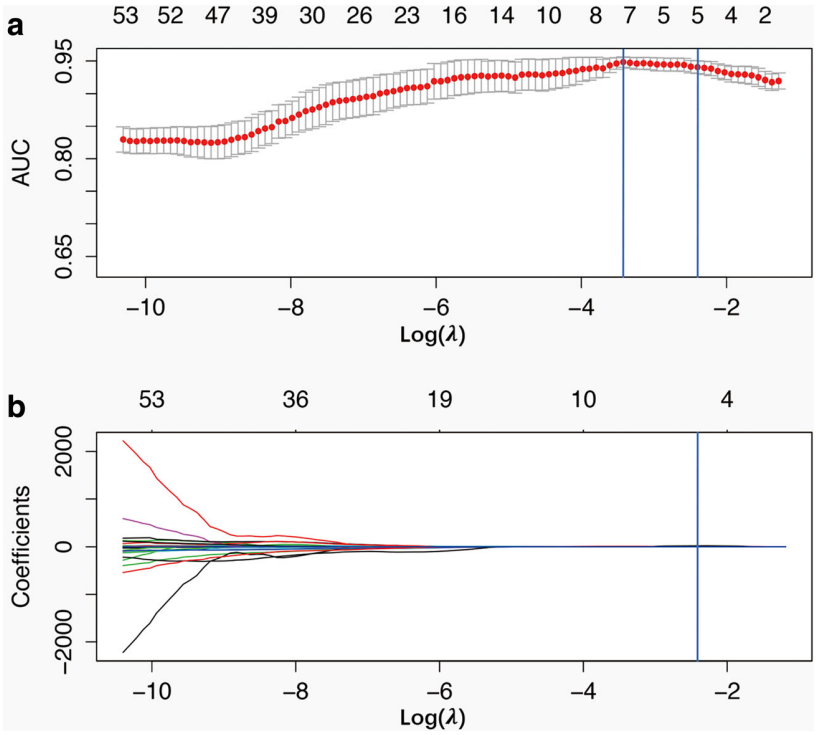

Fig. 1 Radiomic feature selection using LASSO regression model. **a** Optimal feature selection according to AUC value; **b** LASSO coefficient profiles of the 60 radiomic features. Vertical line was drawn at the selected value using 10-fold cross-validation, where optimal λ resulted in 5 non-zero coefficients

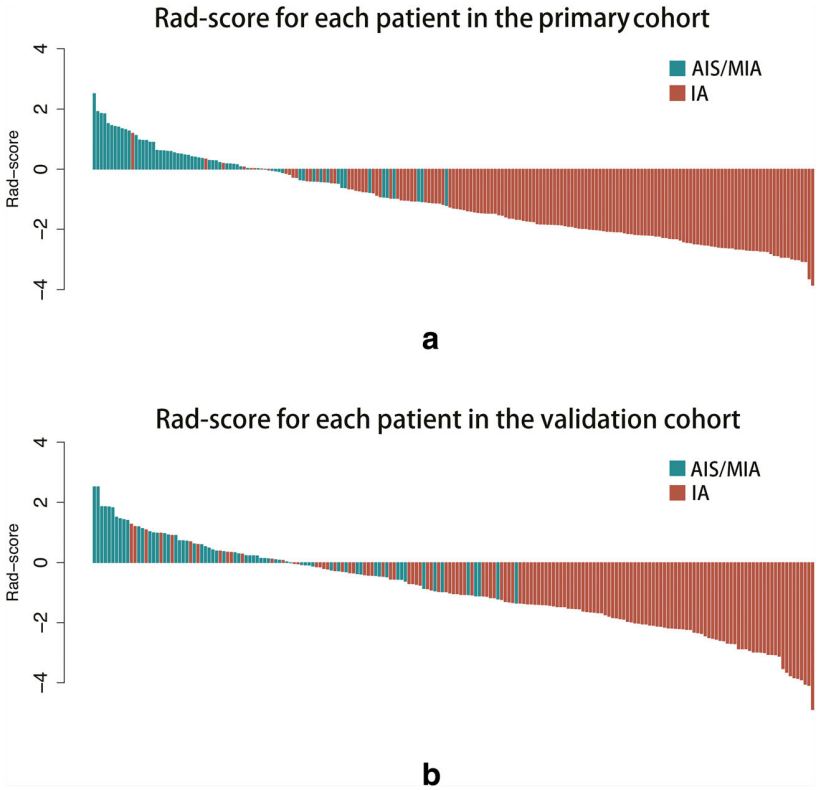

Fig. 2 Rad-score for patients in primary and validation cohort

Based on the five radiomic features, the radiomic signature score (Rad-score) was calculated for each patient (Supplementary material). The Rad-score for each patient is shown in a waterfall plot (Fig. 2). Rad-score was significantly different between AIS/MIA and IA group in both the primary and validation cohort ($p < 0.001$). The mean value of Rad-score for patients with indolent lung adenocarcinoma was significantly higher in both the primary and validation cohort (0.33 and 0.23, respectively) compared with those with IA (– 1.86 and – 1.53, respectively).

Performance of the radiomics signature

Multivariate logistic regression analysis identified the Rad-score (OR 13.29; $p < 0.001$) and gender (OR 3.55; $p = 0.027$) as independent factors (Table 2). On the basis of the discriminating analysis, the signature had AUCs of 0.95 (95% CI, 0.91–0.98) and 0.89 (95% CI, 0.84–0.93) (Fig. 3) in the primary and validation cohort, respectively. However, the AUC values of lesion size in differentiating AIS/MIA from IA were only 0.80 (95% CI, 0.73–0.87) and 0.77 (95% CI, 0.60–0.85) in primary and validation cohort, respectively. The complex model, which combines the signature and gender, had a higher AUC of 0.96 (95% CI, 0.92–0.98) and 0.90 (95% CI, 0.85–0.94) in the primary and validation cohort, respectively. However, the AUC difference between the Rad-score and complex model was not statistically significant (Table 3). The calibration curve of the signature is presented in Fig. 4.

Discussion

The largest study to date, which examined 440 radiomic features in 1019 patients with lung and head/neck cancer, substantiated that radiomic characteristics were correlated with both histology, tumour staging and overall survival [16]. A recent study also demonstrated an association between the imaging phenotype captured with radiomic signature and EGFR mutation in four independent cohorts of lung adenocarcinomas [20]. Furthermore, combining radiomic features with clinical features can provide added diagnostic value in identifying the persistent part-solid nodules, the presence of a

Table 2 Results of multivariate logistic regression analysis

Characteristic	β	Odds ratio (95% CI)	*p*
Intercept	− 0.05		0.119
Age (≥ 65 years)	0.02	1.02 (0.37–2.78)	0.971
Gender (female)	1.27	3.55 (1.16–10.89)	0.027
Rad-score	2.59	13.29 (6.2–28.45)	< 0.001

CI confidence interval, β regression coefficient

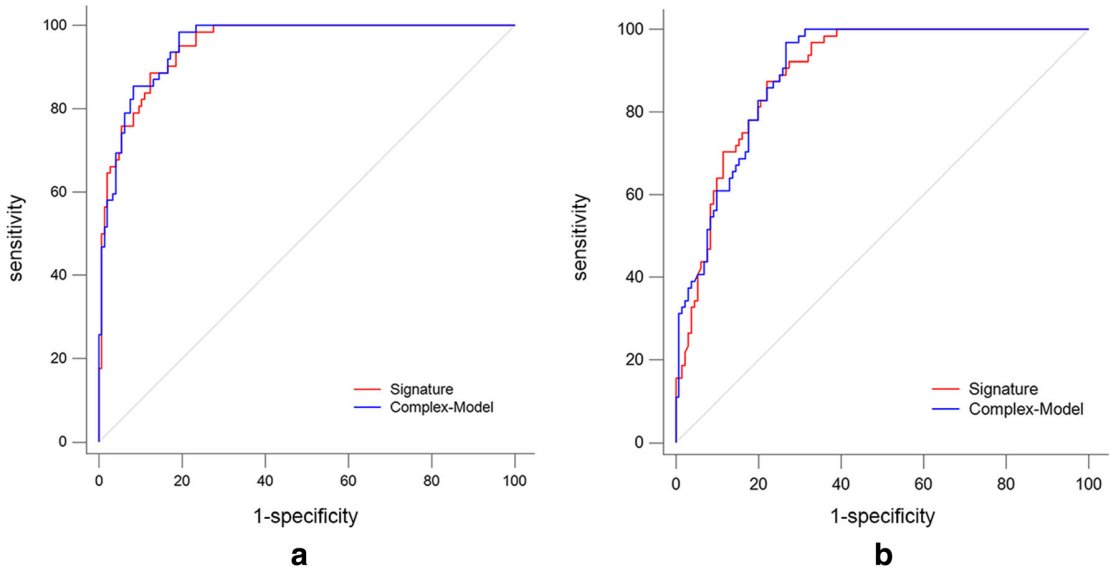

Fig. 3 Area under the curve (AUC) of the signature and the complex model in **a** the primary cohort and **b** the validation cohort

micropapillary component, lymph node phenotypic information and distant metastasis in lung adenocarcinomas [21–24].

In our study, we demonstrated that a novel radiomic signature based on five radiomic features (maximum 3D diameter, root mean squared, entropy, standard deviation and long grey level run emphasis) was an independent factor to discriminate indolent lung adenocarcinoma from IA in patients with solitary lung adenocarcinoma less than 3 cm. For the construction of the radiomic signature, 60 candidate radiomic features were reduced to five potential predictors by examining the predictor–outcome association, which was conducted by shrinking the regression coefficients with the LASSO method. This is a popular method for regression with high dimensional data, which has been extended and broadly applied to the Cox proportional hazard regression model for survival analysis, and to the logistic regression model for predicting outcome [18, 19, 25, 26]. This approach not only works better than the conventional method of choosing predictors on the basis of the intensity of their univariate association with outcome but it also allows researcher to combine the selected features into a single signature.

Lung nodule size was conventionally reported by using the long-axis diameter alone, and guidelines from the Fleischner Society recommended that lung nodule measurements should be based on the long- and short-axis diameters in the same plane [27]. Radiomics allows one to measure three-dimensional volume of lung nodules, which offers more information to precisely characterise the nodule size. In our study, maximum 3D diameter was significantly different between AIS/MIA and IA group in both the primary cohort and validation cohort.

Root mean squared, entropy and standard deviation were obtained from the histogram of voxel intensities and represented the heterogeneity of lung nodules [10, 28]. The root mean squared and standard deviation describe the histogram dispersion, which is a measure of how much the grey level differs from the mean. Similarly, entropy is an imaging feature able to represent heterogeneity of the nodule density, which is related to invasive tumour biology, such as the advanced stage [12, 13]. By quantitative analysis of CT image, radiomics could objectively reflect both the attenuation and dispersion of grey level intensity, which might not be evident on

Table 3 ROC analysis of the signature and the complex model

Model	Primary cohort			Validation cohort		
	AUC	95% CI	p^*	AUC	95% CI	p^*
Radiomics signature	0.95	0.91–0.98	0.421	0.89	0.84–0.93	0.437
Complex model	0.96	0.92–0.98		0.90	0.85–0.94	

ROC receiver operating characteristic, *AUC* area under the curve, *CI* confidence interval

* AUC was compared between the signature and the complex model

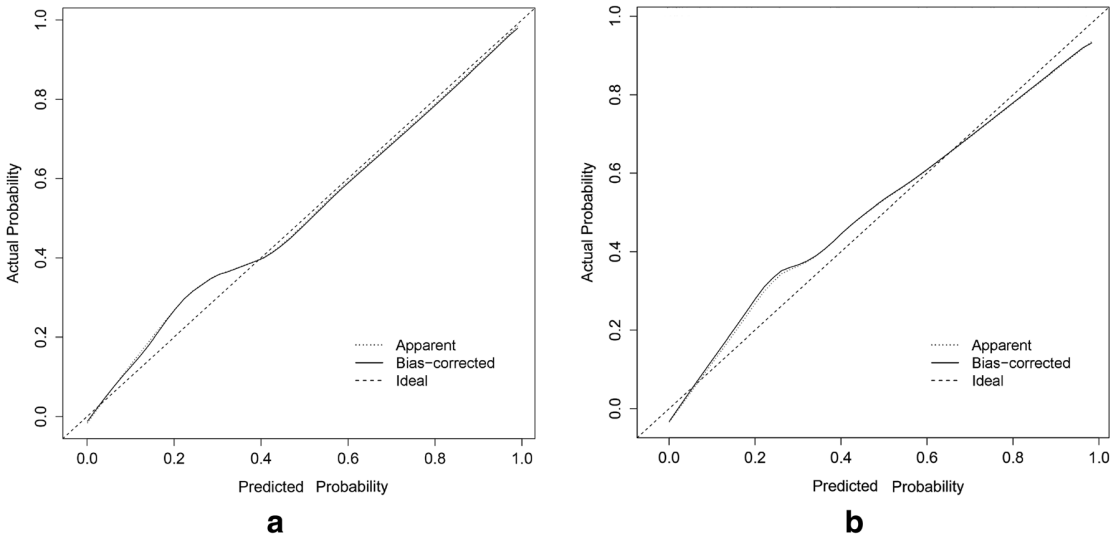

Fig. 4 Calibration curve showing the predicted versus actual probability for invasive adenocarcinoma. Calibration curve of radiomics signature in **a** the primary cohort and **b** the validation cohort

direct visual assessment. Because invasive nodules tended to be more heterogeneous on CT images [15], these parameters could provide valuable evidence to discriminate AIS/MIA from IA. In our study, these three parameters were significantly higher for patients with IA in both the primary cohort and validation cohort. Unlike histogram features, second-order statistics can retain spatial information among voxels, thus reflecting the texture characteristics of a lung nodule [28]. In our study, long grey level run emphasis was selected as one of the second-order statistics to quantify the nodule texture and showed significant difference between AIS/MIA and IA.

As indolent and invasive nodules have a large overlap in both the nodule size and visual morphology on CT images [13], it is highly challenging to differentiate them on the basis of visual assessment. In this regard, radiomics might provide added information to more accurately characterise these lesions. Recently several studies [11–13] have been conducted to utilise radiomics for lung nodule characterisation and they showed promising ability of differentiating invasive lung adenocarcinoma from pre-invasive lesions. In our study, we established an image signature based on five radiomic features, which showed excellent performance in differentiation of IA from AIS/MIA. When the signature is less than – 1.5, no patient was misclassified in both the primary and validation cohort. Therefore, it might serve as an important tool in determining the optimal management of patients with small lung nodules.

However, our study has several limitations. Firstly, patients diagnosed as having benign or atypical adenomatous hyperplasia were excluded, which may subject the study to selection bias. Secondly, the CT acquisition protocol was not standardised among patients. This may have resulted in the variability of CT attenuation values with resultant bias for estimation of radiomic features. Finally, the radiomic features in this study were derived from the results of manual segmentation by radiologists, and the segmentation variability was not evaluated. We believe that a reliable and robust automatic boundary extraction method should be further developed to address this issue.

In conclusion, the radiomics approach can be used to decode lung nodules in a non-invasive manner, thus enabling the identification of imaging phenotypes to characterise lung nodules. The developed radiomics signature provides added diagnostic value to differentiate IA from indolent lung adenocarcinoma in lung nodules less than 3 cm, which might offer useful information for the clinician to choose the optimal intervention.

Funding This study has received funding by Shanghai Hospital Development Center (16CR3116B).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Chang Chen.

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

Statistics and biometry One of the authors 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. Travis WD, Brambilla E, Noguchi M et al (2011) International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society International Multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 6:244–285
2. Travis WD, Brambilla E, Nicholson AG et al (2015) The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. J Thorac Oncol 10:1243–1260
3. Yoshizawa A, Motoi N, Riely GJ et al (2011) Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. Mod Pathol 24:653–664
4. Woo T, Okudela K, Mitsui H et al (2012) Prognostic value of the IASLC/ATS/ERS classification of lung adenocarcinoma in stage I disease of Japanese cases. Pathol Int 62:785–791
5. Kadota K, Villena-Vargas J, Yoshizawa A et al (2014) Prognostic significance of adenocarcinoma in situ, minimally invasive adenocarcinoma, and nonmucinous lepidic predominant invasive adenocarcinoma of the lung in patients with stage I disease. Am J Surg Pathol 38:448–460
6. Liu S, Wang R, Zhang Y et al (2016) Precise diagnosis of intraoperative frozen section is an effective method to guide resection strategy for peripheral small-sized lung adenocarcinoma. J Clin Oncol 34:307–313
7. Yeh YC, Nitadori J, Kadota K et al (2015) Using frozen section to identify histological patterns in stage I lung adenocarcinoma of </= 3 cm: accuracy and interobserver agreement. Histopathology 66:922–938
8. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577
9. Verma V, Simone CB, Krishnan S, Lin SH, Yang J, Hahn SM (2017) The rise of radiomics and implications for oncologic management. J Natl Cancer Inst. https://doi.org/10.1093/jnci/djx055
10. 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
11. Son JY, Lee HY, Lee KS et al (2014) Quantitative CT analysis of pulmonary ground-glass opacity nodules for the distinction of invasive adenocarcinoma from pre-invasive or minimally invasive adenocarcinoma. PLoS One 9:e104066
12. Chae HD, Park CM, Park SJ, Lee SM, Kim KG, Goo JM (2014) Computerized texture analysis of persistent part-solid ground-glass nodules: differentiation of preinvasive lesions from invasive pulmonary adenocarcinomas. Radiology 273:285–293
13. Hwang IP, Park CM, Park SJ et al (2015) Persistent pure ground-glass nodules larger than 5 mm: differentiation of invasive pulmonary adenocarcinomas from preinvasive lesions or minimally invasive adenocarcinomas using texture analysis. Invest Radiol 50:798–804
14. Hawkins S, Wang H, Liu Y et al (2016) Predicting malignant nodules from screening CT scans. J Thorac Oncol 11:2120–2128
15. Ost DE, Gould MK (2012) Decision making in patients with pulmonary nodules. Am J Respir Crit Care Med 185:363–372
16. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006
17. McNeish DM (2015) Using Lasso for predictor selection and to assuage overfitting: a method long overlooked in behavioral sciences. Multivariate Behav Res 50:471–484
18. Jiang Y, Zhang Q, Hu Y et al (2016) ImmunoScore signature: a prognostic and predictive tool in gastric cancer. Ann Surg 267:504–513
19. Guo BL, Ouyang FS, Yang SM et al (2017) Development of a preprocedure nomogram for predicting contrast-induced acute kidney injury after coronary angiography or percutaneous coronary intervention. Oncotarget 8:75087–75093
20. Rios Velazquez E, Parmar C, Liu Y et al (2017) Somatic mutations drive distinct imaging phenotypes in lung cancer. Cancer Res 77:3922–3930
21. Lee SH, Lee SM, Goo JM et al (2014) Usefulness of texture analysis in differentiating transient from persistent part-solid nodules (PSNs): a retrospective study. PLoS One 9:e85167
22. Song SH, Park H, Lee G et al (2017) Imaging phenotyping using radiomics to predict micropapillary pattern within lung adenocarcinoma. J Thorac Oncol 12:624–632
23. Coroller TP, Agrawal V, Huynh E et al (2017) Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC. J Thorac Oncol 12:467–476
24. Coroller TP, Grossmann P, Hou Y et al (2015) CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol 114:345–350
25. Huang YQ, Liang CH, He L et al (2016) Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 34:2157–2164
26. Zhang JX, Song W, Chen ZH et al (2013) Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. Lancet Oncol 14:1295–1306
27. Bankier AA, MacMahon H, Goo JM, Rubin GD, Schaefer-Prokop CM, Naidich DP (2017) Recommendations for measuring pulmonary nodules at CT: a statement from the Fleischner Society. Radiology. https://doi.org/10.1148/radiol.2017162894
28. Davnall F, Yip CS, Ljungqvist G et al (2012) Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? Insights Imaging 3:573–589