



# Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer

Xinzhong Zhu<sup>1,2,3</sup> · Di Dong<sup>2,4</sup> · Zhendong Chen<sup>2,3</sup> · Mengjie Fang<sup>2,4</sup> · Liwen Zhang<sup>2</sup> · Jiangdian Song<sup>2</sup> · Dongdong Yu<sup>2,4</sup> · Yali Zang<sup>2,4</sup> · Zhenyu Liu<sup>2,4</sup> · Jingyun Shi<sup>5</sup> · Jie Tian<sup>1,2,4</sup> 

Received: 8 June 2017 / Revised: 7 November 2017 / Accepted: 28 November 2017  
© European Society of Radiology 2018

## Abstract

**Objectives** To distinguish squamous cell carcinoma (SCC) from lung adenocarcinoma (ADC) based on a radiomic signature  
**Methods** This study involved 129 patients with non-small cell lung cancer (NSCLC) (81 in the training cohort and 48 in the independent validation cohort). Approximately 485 features were extracted from a manually outlined tumor region. The LASSO logistic regression model selected the key features of a radiomic signature. Receiver operating characteristic curve and area under the curve (AUC) were used to evaluate the performance of the radiomic signature in the training and validation cohorts.  
**Results** Five features were selected to construct the radiomic signature for histologic subtype classification. The performance of the radiomic signature to distinguish between lung ADC and SCC in both training and validation cohorts was good, with an AUC of 0.905 (95% confidence interval [CI]: 0.838 to 0.971), sensitivity of 0.830, and specificity of 0.929. In the validation cohort, the radiomic signature showed an AUC of 0.893 (95% CI: 0.789 to 0.996), sensitivity of 0.828, and specificity of 0.900.  
**Conclusions** A unique radiomic signature was constructed for use as a diagnostic factor for discriminating lung ADC from SCC. Patients with NSCLC will benefit from the proposed radiomic signature.

## Key points

- Machine learning can be used for auxiliary distinguish in lung cancer.
- Radiomic signature can discriminate lung ADC from SCC.
- Radiomics can help to achieve precision medical treatment.

---

Xinzhong Zhu and Zhendong Chen contributed equally to this work.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00330-017-5221-1>) contains supplementary material, which is available to authorized users.

- ✉ Xinzhong Zhu  
zxz@zjnu.edu.cn
- ✉ Di Dong  
di.dong@ia.ac.cn
- ✉ Jingyun Shi  
shijingyun89179@126.com

- <sup>1</sup> School of Life Science and Technology, XIDIAN University, Xi'an, Shanxi, China
- <sup>2</sup> CAS Key Lab of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China
- <sup>3</sup> College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua, Zhengjiang, China
- <sup>4</sup> University of Chinese Academy of Sciences, Beijing, China
- <sup>5</sup> Department of Radiology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Tongji, China

**Keywords** Adenocarcinoma · Diagnostic imaging · Regression analysis · ROC curve · Squamous cell carcinoma

## Abbreviations

ADC	Adenocarcinoma
AUC	Area Under the Curve
CI	Confidence Interval
CT	Computed Tomography
LASSO	Least Absolute Shrinkage and Selection Operator
NSCLC	Non-small Cell Lung Cancer
PET	Positron Emission Tomography
ROC	Receiver Operating Characteristic
ROI	Region of Interest
SCC	Squamous Cell Carcinoma

## Introduction

Lung cancer has long been one of the most common cancers worldwide, and its incidence rates are second highest in countries such as China, European countries, and the United States

[1–5]. Non-small cell lung cancer (NSCLC) accounts for about 85% of all lung cancers.

Squamous cell carcinoma (SCC) and adenocarcinoma (ADC) are two major histologic subtypes of NSCLC. There is a significant difference in the prognosis and recurrence rate of lung ADC and SCC. Fukui et al. reported that among patients with stage IA and IB disease, those with SCC have significantly worse outcomes compared to those with ADC [6]. ADC is more associated with distant metastasis and vascular invasion than with local recurrence [7]. Usui et al. showed that patients with vascular invasion-positive ADC have significantly worse outcomes than patients with vascular invasion-negative ADC, whereas SCC shows no significant difference. However, the incidence of large vessel invasion is higher in SCC than in ADC [8]. The treatment of lung ADC and SCC is also very different. Therefore, neoadjuvant (chemo) radiotherapy should be performed along with resection for lung SCC. Chemotherapy with the antifolate pemetrexed is only effective for lung ADC, while bevacizumab is recommended for lung SCC [9, 10]. Hence, an accurate diagnosis would not only improve therapeutic efficacy, but would also avoid unnecessary side effects and/or severe unwarranted side effects. Therefore, it is important to distinguish between the two subtypes of NSCLC prior to initiating treatment.

Pathological diagnosis is the gold standard for distinguishing lung ADC and SCC. However, pathological diagnosis requires invasive biopsy or the preparation of pathological tissue sections after surgery. In some cases, CT-guided needle biopsy could not be performed or not be suitable. For example, some small lesions are difficult to target and also cannot provide enough tissues for pathological diagnosis. What is more, for deep-located lesions, or lesions close to airways or blood vessels, the operation of CT-guided needle biopsy is challenging. In other patients with bad situations, CT-guided biopsy would not be recommended. Moreover, the tumour is often heterogeneous, which may affect the biopsy results. Pathological tissue sections provide more accurate results, but they can only be prepared after surgery, which delays the diagnosis. A non-invasive method for pathological classification prior to biopsy or surgery has not been developed yet.

Computed tomography (CT) images have been used to assess the relationship between imaging characteristics and pathological information in tumours because some pathological information, such as tumour enhancement characteristics and internal components (e.g., necrosis, calcification), can be obtained from CT images [11, 12]. However, radiologists have difficulty in distinguishing ADC from SCC based on morphological CT images, and inter-observer agreement is usually low. Recent studies have found that texture analysis can provide additional useful information based on widely available CT images, reflecting the potential of biological heterogeneity [13, 14]. CT-based texture analysis has recently been shown to predict tumour stage in oesophageal cancer and NSCLC [15, 16].

Radiomics is a novel technique that employs high-throughput quantitative image features for diagnosis and prognosis [17–21]. Radiomics regards images as data and performs data mining to predict clinical phenotype and even gene information [22–25]. Researchers also use quantitative CT images to characterize gene expression data and further predict the survival of NSCLC [26, 27]. Recently, radiomic signature has proven to be a significant classification biomarker for lung cancer and head/neck cancer staging [28]. However, there is still no quantitative method for non-invasive distinguishing of lung ADC and SCC. So far, scarcely any research has been done about the use of radiomic signatures to predict lung ADC and SCC.

Therefore, in this study, a CT-based radiomic signature was constructed for use as a diagnostic factor for discriminating lung ADC from SCC. The experimental results showed that the constructed radiomic signature has very good classification performance for lung ADC and SCC, indicating its promising clinical value.

## Materials and methods

### Patients

This retrospective study was approved by our institutional review board approval, and the need for informed patient consent was waived. About 441 consecutive patients diagnosed with lung cancer were recruited retrospectively between September 2010 and November 2013. The inclusion criteria were as follows: (1) pathologically confirmed lung cancer; (2) available CT images before treatment. The exclusion criteria were as follows: (1) patients with large-cell carcinoma, small-cell carcinoma, and sarcoma-like carcinoma ( $n=112$ ); (2) patients with preoperative chemotherapy ( $n=108$ ); (3) preoperative biopsy-proven histological grade unavailable ( $n=92$ ). Finally, 129 patients were selected for this study (Fig. 1). In this study, we wanted to use all the preoperative information to distinguish lung ADC from SCC so that our method will have the potential for assisting preoperative treatment decisions. Therefore, we used preoperative histological grade from a biopsy rather than grade from a surgical specimen.

### Image acquisition

All patients underwent pre-contrast CT imaging of the lungs with one of two multi-detector row CT systems (the GE Lightspeed Ultra 8, GE Healthcare, Hino, Japan; or the 64-slice LightSpeed VCT, GE Medical Systems, Milwaukee, WI, USA). The scanning parameters were as follows: 120 kV; 160 mAs; detector collimation:  $64 \times 0.625$  mm; rotation time: 0.5 s; and matrix size,  $512 \times 512$ . Each patient was scanned for the whole lung, the slice number varied from 100 to 600.

### Region of interest (ROI) segmentation

The segmentation of a ROI is essential for the extraction of quantitative features. ITK-SNAP software (version 3.4.0; [www.itk-snap.org](http://www.itk-snap.org)) was used for three-dimensional manual segmentation by a radiologist with 10 years of experience. To test intra-class reproducibility, 20 cases were selected randomly and segmented twice by one radiologist in two weeks. To test inter-class reproducibility, the 20 cases were segmented by two radiologists. The Kruskal–Wallis H test or an independent samples *t*-test, where appropriate, was used to assess the differences between the features generated at different times and by different radiologists, as well as between the twice-generated features by the same radiologist. Inter- and intra-class correlation coefficients (ICCs) were used to evaluate the intra- and inter-observer agreement of

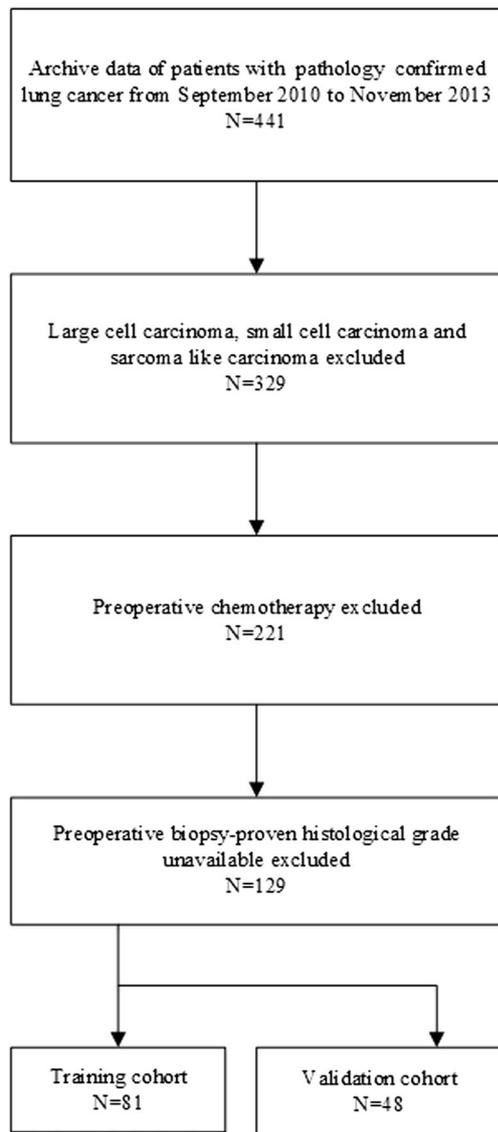


Fig. 1 Patient recruitment procedure

Table 1 Demographic differences in the training and validation cohorts

Characteristics	Training cohort	Validation cohort	<i>p</i> value
Gender(No [%])			
Male	42 (51.9)	27 (56.3)	0.628
Female	39 (48.1)	21 (43.7)	
Age(yr,m [r])	55 (41-78)	53 (43-75)	0.926
Subtype(No [%])			
SCC	34 (42)	19 (39.6)	0.790
ADC	47 (58)	29 (60.4)	

No, number; yr, year; m, median; r, range.

*p* value<0.05 indicates a significant difference in patients' characteristics between the training cohort and independent validation cohort

feature extraction. An ICC greater than 0.75 indicated good agreement. Approximately 485 features with an ICC higher than 0.75 were selected for further feature extraction. The dice coefficient of the intrastudy segmentations varied from 0.79 to 0.96. The dice coefficient of the interstudy segmentations varied from 0.75 to 0.94. Therefore, all outcomes were based on the features extracted by the same radiologist.

### Feature extraction

Approximately 485 candidate features with an ICC higher than 0.75 were extracted, including tumour intensity, shape and size, texture, and wavelet characteristics. Their extraction was performed in MATLAB 2015b (Mathworks, Natick, MA,

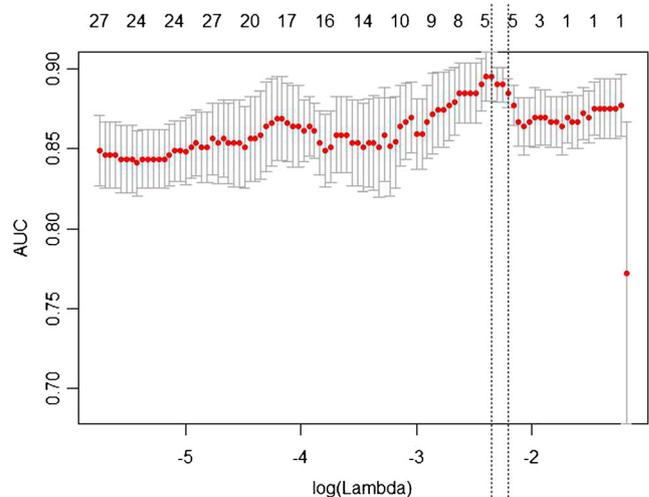
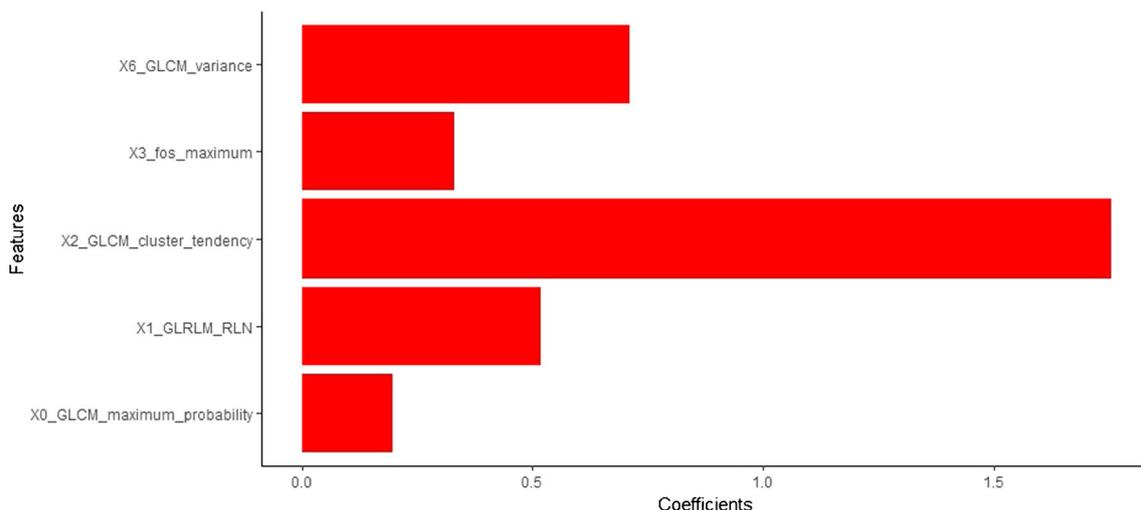


Fig. 2 Feature selection using the LASSO regression method. We used 5-fold cross validation in the LASSO model for the selection of the conditioning parameters (Lambda). The AUC was plotted versus log(Lambda) by using the minimum standard and the minimum standard of 1 standard error (1 - SE standard) to draw the vertical line with the best value. A Lambda value of 0.1098, with log(Lambda) - 2.2093 was chosen (1 - SE standard) according to the 5-fold cross validation



**Fig. 3** Histogram showing the contribution of each feature to the radiomic signature

USA). Detailed descriptions of these features are shown in [electronic supplementary materials](#).

### Feature selection

The least absolute shrinkage and selection operator (LASSO) logistic regression model was used for feature reduction and selection [29, 30]. Several features with non-zero coefficients were selected from the candidate features, and formed a radiomic signature. Radiomic score is obtained by computing the logistic regression product of these features.

### Predictive performance of the radiomic signature

In order to test the performance of the radiomic signature, a *T*-test was used to estimate the relationship between the radiomic signature and pathological types (lung ADC and SCC). Receiver operating characteristics (ROC) curves were plotted to show the performance of the signature [31]. The ROC curve is a comprehensive index reflecting specificity and sensitivity of continuous variables. The area under the curve (AUC) was calculated to evaluate the classification accuracy. The optimal cut-off threshold of the ROC curve was obtained in the

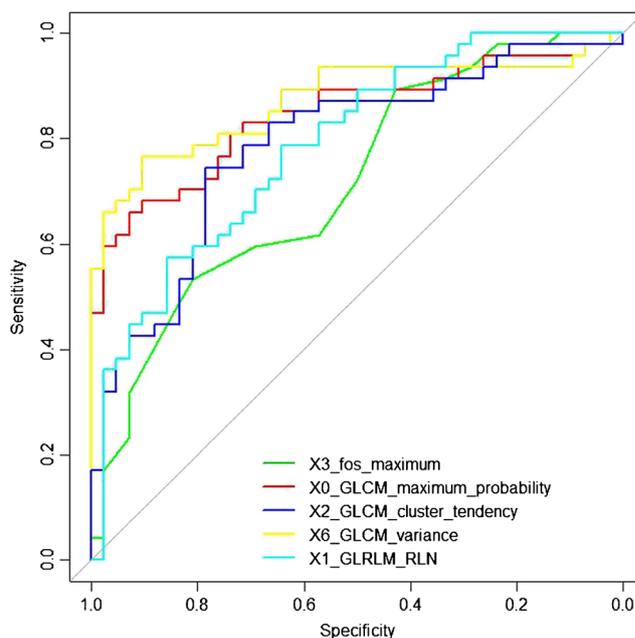
**Table 2** Classification of the five discriminating features selected by LASSO

Feature name	AUC	Sensitivity	Specificity
X3_fos_maximum	0.718	0.532	0.810
X0_GLCM_maximum_probability	0.845	0.660	0.929
X2_GLCM_cluster_tendency	0.872	0.766	0.905
X6_GLCM_variance	0.788	0.745	0.786
X1_GLRLM_RLN	0.787	0.574	0.857

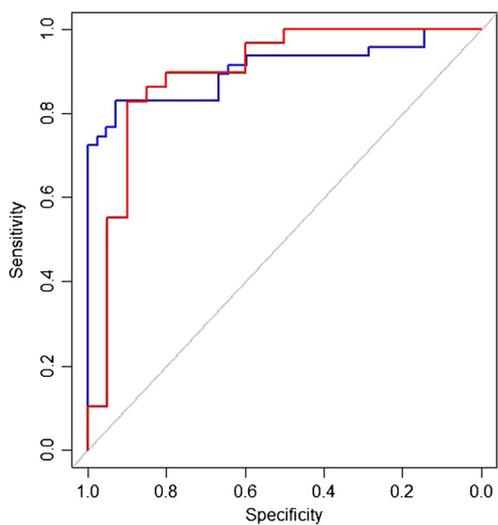
training cohort and then applied to the validation cohort for evaluation of sensitivity and specificity.

### Statistical analysis

Statistical analysis was conducted with R software (version 3.3.2; <http://www.Rproject.org>). The LASSO logistic regression was performed using the “glmnet” package. The reported statistical significance levels were all two-sided, with the statistical significance set at 0.05.



**Fig 4** A single characteristic ROC curve. ROC classification model of pathological classification based on the ROC threshold analysis for different features. Red, green, blue, yellow and cyan represent the model curves established using X0\_GLCM\_maximum\_probability, X3\_fos\_maximum, X2\_GLCM\_cluster\_tendency, X6\_GLCM\_variance and X1\_GLRLM\_RLN



**Fig. 5** ROC curves of the radiomic signature in the training (blue) and validation (red) cohorts. The AUC of the training cohort was 0.905 (95% CI: 0.838 to 0.971) and that of the validation cohort was 0.893 (95% CI: 0.789 to 0.996)

**Results**

**Demographic characteristics of the patients**

The study patients were divided into two groups: the training cohort with 81 patients (42 male and 39 female; median age, 55 years; range, 41 to 78 years); and the validation cohort with 48 patients (27 male and 21 female; median age, 53 years; range, 43 to 75 years). The demographic and tumour characteristics of all patients in the training and validation cohorts are listed in Table 1.

**Radiomic signature construction**

The LASSO logistic regression model was used to reduce the 485 features to five features on the training cohort. The radiomic score (Rad-score) can be calculated from these selected features. Their coefficients are shown in Fig. 2. These features include X3\_fos\_maximum, X0\_GLCM\_maximum\_probability, X2\_GLCM\_cluster\_tendency, X6\_GLCM\_variance and X1\_GLRLM\_RLN. Rad-score =  $-0.1770059 + X3\_fos\_maximum \times -0.7681409 + X0\_GLCM\_maximum\_probability \times -0.9542600 + X2\_GLCM\_cluster\_tendency \times 3.7099175 + X6\_GLCM\_variance \times 0.2290265 + X1\_GLRLM\_RLN$

$\times -0.5012380$ . The contribution of the selected features and their corresponding regression coefficients are shown in Fig. 3. The performance of the ROC curve was calculated separately for each of the five features selected by LASSO. The classification threshold of each feature was determined by ROC curve threshold analysis. The AUC value, sensitivity, and specificity of classification using this threshold are shown in Table 2. The closer the AUC value is to 1, the better the prediction effect. We selected the five AUC features which constructed a radiomic signature. The five characteristics of the ROC curve are shown in Fig. 4.

**Predictive performance**

The ROC curves for training and validation cohorts are shown in Fig. 5. The radiomics model could differentiate lung ADC from SCC. The AUC of the training cohort was 0.905 (95%CI: 0.838 to 0.971, sensitivity 0.830, specificity 0.929) and that of the validation cohort was 0.893 (95% CI: 0.789 to 0.996, sensitivity 0.828, specificity 0.900). The average time required to analyse one case in the training cohort was approximately 3.25 min, and the average time required for the validation cohort was about 3.2 min. The diagnostic accuracy of the model in the training and validation cohorts is shown in Table 3.

**Discussion**

Medical imaging provides valuable information for the diagnosis and treatment of cancer patients. It is not unusual to extract some basic indicators from these images as prognostic factors or to evaluate treatment response [32]. However, these methods do not capture more information about tumours beyond the images. Radiomics uses an automated high-throughput data feature extraction algorithm to transform the image data into high-resolution extractable image feature data, which can describe the organizational properties of tumours.

In this study, a radiomic signature comprising five quantitative CT image features was built to distinguish lung ADC from SCC pre-biopsy and pre-operation. The experimental results showed that our method was effective in both the training cohort (AUC, 0.905) and independent validation cohort (AUC, 0.893). It is proven that radiomics could provide a powerful prediction value by employing large amounts of quantitative image features (485 features were originally

**Table 3** Diagnostic accuracy of the model in the training and validation cohorts

Data	Sensitivity	Specificity	AUC(95%CI)	Time
Training cohort	0.830	0.929	0.905 (0.838-0.971)	3.25min
Validation cohort	0.828	0.900	0.893 (0.789-0.996)	3.2min

extracted). Moreover, most of the selected features were texture features, which reflected the heterogeneity of the tumour ROI. Among these features, the most important value was X2\_GLCM\_cluster\_tendency, which reflected differences in information between images. Therefore, the X2\_GLCM\_cluster\_tendency could be a predicting factor for differentiating lung ADC and SCC.

Currently, there are several biomarkers to distinguish lung SCC from ADC, such as cytokeratin 7 (CK7), trefoil factor 3 (TFF3), and thyroid transcription factor 1 (TTF-1) for lung ADC, and CK5/6 and P63 for lung SCC [33–35]. The problem with the above biomarkers in the distinguishing of lung SCC and ADC is that their individual sensitivity and specificity are not sufficient for accurately identifying cancer subtypes. Diagnosing lung ADC and SCC by combining several biomarkers may be a trend in the future. Patnaik et al. used MiR-205 and MiR-375 microRNAs to distinguish between lung ADC and SCC with high sensitivity and specificity [36]. However, using biomarkers is invasive and costly. Ha et al. obtained texture features from positron emission tomography (PET) images to distinguish lung SCC from lung ADC. A total of 200 characteristic parameters were found. Among them, 15 characteristics were found to differ significantly between lung ADC and SCC [37]. However, the number of experimental cases was only 30, and there was no established prediction model between the characteristics and lung ADC and SCC.

Lung SCC is more sensitive to (chemo)radiotherapy, indicating the clinical value of the proposed radiomic signature. Our radiomic signature could be a low-cost, non-invasive diagnostic factor for pathological type classification and guided neoadjuvant therapy. Patients with lung SCC would benefit from this radiomic signature, since they could choose neoadjuvant (chemo)radiotherapy preoperatively to improve survival. Patients with lung ADC could also avoid unnecessary chemoradiotherapy according to the outcome of our radiomics model.

Our study has some limitations. First, our study used only CT image features, but PET or other modalities may improve performance. PET presents functional information of the tumour, and may perfect the radiomic signature. Second, the small sample size and single-centre cohort limit the expansion of the radiomic signature. In the future, large, multi-centre cohorts should be recruited for validation. Third, in this study, only patients with preoperative diagnosis from a biopsy were included. However, the biopsy information like histological grade was not significantly related to the discrimination of ADC and SCC, and therefore, not included in our prediction model. In the future, we will include patients without a preoperative biopsy and expand our cohort.

In conclusion, these selected features of the constructed radiomic signature can distinguish between lung ADC and SCC, thus assisting doctors in clinical diagnosis.

**Funding** This work was supported by the National Natural Science Foundation of China (81227901, 81771924, 81501616, 61231004, 81671851, and 81527805), National Key R&D Program of China (2017YFA0205200, 2017YFC1308700, 2017YFC1308701, 2017YFC1309100), the Science and Technology Service Network Initiative of the Chinese Academy of Sciences (KFJ-SW-STS-160), the Instrument Developing Project of the Chinese Academy of Sciences (YZ201502), the Beijing Municipal Science and Technology Commission (Z161100002616022), and the Youth Innovation Promotion Association CAS.

## Compliance with ethical standards

**Guarantor** The scientific guarantor of this publication is Jie Tian.

**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** No complex statistical methods were necessary for this paper.

**Informed consent** Written informed consent was obtained from all subjects (patients) in this study.

**Ethical approval** Institutional Review Board approval was obtained.

## Methodology

- retrospective
- diagnostic or prognostic study
- performed at one institution

## References

1. Printz C (2015) Lung cancer new leading cause of death for women in developed countries: Data reflects increased rates of smoking. *Cancer* 121:1911–1912
2. Brawley OW (2011) Avoidable cancer deaths globally. *CA Cancer J Clin* 61:67–68
3. Ferlay J, Steliarovaoufoucher E, Lortettieulent J et al (2013) Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 49:1374–1403
4. Nair VS, Gevaert O, Davidzon G et al (2012) Prognostic PET 18F-FDG uptake imaging features are associated with major oncogenomic alterations in patients with resected non-small cell lung cancer. *Cancer Res* 72:3725–3734
5. Siegel RL, Miller KD, Jemal A (2017) Cancer Statistics, 2017. *CA Cancer J Clin*. 67:7–30
6. Fukui T, Taniguchi T, Kawaguchi K et al (2015) Comparisons of the clinicopathological features and survival outcomes between lung cancer patients with adenocarcinoma and squamous cell carcinoma. *Gen Thorac Cardiovasc Surg* 63:507–513
7. Bodendorf MO, Haas V, Laberke HG (2009) Prognostic value and therapeutic consequences of vascular invasion in non-small cell lung carcinoma. *Lung Cancer* 64:71–78
8. Usui S, Minami Y, Shiozawa T et al (2013) Differences in the prognostic implications of vascular invasion between lung adenocarcinoma and squamous cell carcinoma. *Lung Cancer* 82:407–412
9. Herbst RS, O'Neill VJ, Fehrenbacher L et al (2007) Phase II study of efficacy and safety of bevacizumab in combination with chemotherapy or erlotinib compared with chemotherapy alone for

- treatment of recurrent or refractory non small-cell lung cancer. *J Clin Oncol* 25:4743–4750
10. Di Costanzo F, Mazzoni F, Micol Mela M et al (2008) Bevacizumab in non-small cell lung cancer. *Drugs* 68:737–746
  11. Nguyenkim TD, Frauenfelder T, Strobel K et al (2015) Assessment of Bronchial and Pulmonary Blood Supply in Non-Small Cell Lung Cancer Subtypes Using Computed Tomography Perfusion. *Investig Radiol* 50:179–186
  12. Yip C, Landau D, Kozarski R et al (2014) Primary esophageal cancer: heterogeneity as potential prognostic biomarker in patients treated with definitive chemotherapy and radiation therapy. *Radiology* 270:141–148
  13. Ng F, Ganeshan B, Kozarski R, Miles KA, Goh V (2013) Assessment of primary colorectal cancer heterogeneity by using whole-tumor texture analysis: contrast-enhanced CT texture as a biomarker of 5-year survival. *Radiology* 266:177–184
  14. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA (2013) Non-small cell lung cancer: histopathologic correlates for texture parameters at CT. *Radiology* 266:326–336
  15. Ganeshan B, Abaleke S, Young RC, Chatwin CR, Miles KA (2010) Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer Imaging* 10:137–143
  16. Ganeshan B, Skogen K, Pressney I, Coutroubis D and Miles K (2012) Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol* 67:157–164.
  17. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 48:441–446
  18. Kumar V, Gu Y, Basu S et al (2012) Radiomics: the process and the challenges. *Magn Reson Imaging* 30:1234–1248
  19. Choi ER, Lee HY, Jeong JY et al (2016) Quantitative image variables reflect the intratumoral pathologic heterogeneity of lung adenocarcinoma. *Oncotarget* 7:67302–67313
  20. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 278:563–577
  21. Aerts HJ, Velazquez ER, Leijenaar RT et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 5:4006
  22. Itakura H, Achrol AS, Mitchell LA et al (2015) Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities. *Sci Transl Med* 7:303ra138
  23. Li H, Zhu Y, Burnside ES et al (2016) MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays. *Radiology* 281:382–391
  24. Zinn PO, Singh SK, Kotrotsou A et al (2016) 139 Clinically Applicable and Biologically Validated MRI Radiomic Test Method Predicts Glioblastoma Genomic Landscape and Survival. *Neurosurgery* 63:156–157
  25. Huang Y, Liu Z, He L et al (2016) Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. *Radiology* 281:152234
  26. Yamamoto S, Korn RL, Oklu R et al (2014) ALK Molecular Phenotype in Non-Small Cell Lung Cancer: CT Radiogenomic Characterization. *Radiology* 272:568–576
  27. Gevaert O, Xu J, Hoang CD et al (2012) Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results. *Radiology* 264:387–396
  28. Parmar C, Lambin P, Aerts HJ et al (2015) Radiomic feature clusters and prognostic signatures specific for Lung and Head and Neck cancer. *Sci Rep* 5:11044
  29. Chi MA, Jian H (2016) Asymptotic properties of LASSO in high-dimensional partially linear models. *Science China Mathematics* 59: 1–20
  30. Rao SJ (2003) Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. *J Am Stat Assoc* 98:257–258
  31. Lloyd CJ (2015) Theory & Methods: Fitting Roc Curves Using Non-linear Binomial Regression. *Aust. N Z J Stat* 42:193–204
  32. Ball DL, Fisher RJ, Burmeister BH et al (2013) The complex relationship between lung tumor volume and survival in patients with non-small cell lung cancer treated by definitive radiotherapy: A prospective, observational prognostic factor study of the Trans-Tasman Radiation Oncology Group (TROG 99.05). *Radiother Oncol* 106:305–311
  33. Khayyata S, Yun S, Pasha T et al (2009) Value of P63 and CK5/6 in distinguishing squamous cell carcinoma from adenocarcinoma in lung fine-needle aspiration specimens. *Diagn Cytopathol* 37:178–183
  34. Montezuma D, Azevedo R, Lopes P et al (2013) A panel of four immunohistochemical markers (CK7, CK20, TTF-1, and p63) allows accurate diagnosis of primary and metastatic lung carcinoma on biopsy specimens. *Virchows Arch* 463:749–754
  35. Wang XN, Wang SJ, Pandey V et al (2015) Trefoil factor 3 as a novel biomarker to distinguish between adenocarcinoma and squamous cell carcinoma. *Medicine* 94:e860
  36. Patnaik S, Mallick R, Kannisto E et al (2015) MiR-205 and MiR-375 MicroRNA Assays to Distinguish Squamous Cell Carcinoma from Adenocarcinoma in Lung Cancer Biopsies. *J Thorac Oncol* 10:446–453
  37. Ha S, Choi H, Cheon GJ et al (2014) Autoclustering of Non-small Cell Lung Carcinoma Subtypes on 18F-FDG PET Using Texture Analysis: A Preliminary Result. *Nucl Med Mol Imaging* 48:278–286