ORIGINAL ARTICLE

Computed tomography radiomics-based distinction of invasive adenocarcinoma from minimally invasive adenocarcinoma manifesting as pure ground-glass nodules with bubble-like signs

Yining Jiang[1] · Ziqi Xiong[1] · Wenjing Zhao[1] · Jingyu Zhang[1] · Yan Guo[2] · Guosheng Li[3] · Zhiyong Li[1,4]

Abstract

Background To explore an effective model based on radiomics features extracted from nonenhanced computed tomography (CT) images to distinguish invasive adenocarcinoma (IAC) from minimally invasive adenocarcinoma (MIA) presenting as pure ground-glass nodules (pGGNs) with bubble-like (B-pGGNs) signs.

Patients and methods We retrospectively reviewed 511 nodules (MIA, $n = 288$; IAC, $n = 223$) between November 2012 and June 2018 from almost all pGGNs pathologically confirmed MIA or IAC. Eventually, a total of 109 B-pGGNs (MIA, $n = 55$; IAC, $n = 54$) from 109 patients fulfilling the criteria were randomly assigned to the training and test cluster at a ratio of 7:3. The gradient boosting decision tree (GBDT) method and logistic regression (LR) analysis were applied to feature selection (radiomics, semantic, and conventional CT features). LR was performed to construct three models (the conventional, radiomics and combined model). The performance of the predictive models was evaluated using the area under the curve (AUC).

Results The radiomics model had good AUCs of 0.947 in the training cluster and of 0.945 in the test cluster. The combined model produced an AUC of 0.953 in the training cluster and of 0.945 in the test cluster. The combined model yielded no performance improvement (vs. the radiomics model). The rad_score was the only independent predictor of invasiveness.

Conclusion The radiomics model showed excellent predictive performance in discriminating IAC from MIA presenting as B-pGGNs and may provide a necessary reference for extending clinical practice.

Keywords Pure ground-glass nodule (pGGN) · Bubble-like sign · Invasive adenocarcinoma (IAC) · Minimally invasive adenocarcinoma (MIA) · Radiomics

Introduction

Pulmonary pure ground-glass nodules (pGGNs) on high-resolution computed tomography (HRCT) appearing as an isolated lesion of hazy increased does not obliterate the fundamental bronchial and vascular configuration [1]. About 75% of persistent pGGNs were attributed to lung adenocarcinomas (LACs) or their precursors [2]. The new concept of adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA) and invasive adenocarcinoma

✉ Zhiyong Li
zjy_lzy@126.com

Yining Jiang
yiningj7@163.com

Ziqi Xiong
xiongziqi17@163.com

Wenjing Zhao
zhaowenjing97@163.com

Jingyu Zhang
Jingyu_z33@163.com

Yan Guo
guoyan0112@126.com

Guosheng Li
guoshengli998@163.com

1. Department of Radiology, The First Affiliated Hospital of Dalian Medical University, Dalian, China

2. GE Healthcare, Beijing, China

3. Department of Pathology, The First Affiliated Hospital of Dalian Medical University, Dalian, China

4. Dalian Engineering Research Centre for Artificial Intelligence in Medical Imaging, Dalian, China

(IAC) was proposed by the International Association for Study of Lung Cancer, the American Thoracic Society and the European Respiratory Society (IASLC, ATS, and ERS) in 2011 [3] and adopted by the World Health Organization (WHO) in 2015 [4]. MIA refers to a solitary LAC (≤ 3 cm in diameter) with predominantly lepidic growth and a ≤ 0.5 cm invasive component pathologically. IACs are mainly categorised based on invasive components > 0.5 cm [4]. MIA has been shown to correlate with low metastatic potential and is rarely associated with postoperative recurrence. In contrast, IACs show aggressive biological behaviour associated with pleural invasion, lymph node metastasis and postoperative recurrence [5].

The relationships between intra-tumour gene heterogeneity and pathologic classification have also been demonstrated [6]. Sakamoto et al. [7] showed that KRAS mutations were detectable in 7% of MIA cases and 0% of IAC cases. Zhang et al. [8] showed that TP53 was detected more frequently in IAC than in AIS/MIA (5/14 vs. 1/16, $P = 0.04$). An accurate description of early pathological types is necessary to determine the tumour state. Although the criteria for resection range remain controversial, the appearance of pGGNs on CT is strongly suspected as MIA pathologically; close follow-up scanning is recommended, and sublobar resection is not performed until suspicious morphology or other risk factors are observed [9, 10]. Lobectomy is still considered the standard surgical treatment for IAC [10]. Therefore, more precise preoperative diagnostic and individualized management should be established for pGGNs.

The bubble-like sign is defined as small spots of round or ovoid air attenuation [11, 12]. Saito et al. [13] reported that bubble-like within GGNs were a histological characteristic of collapse and dilated bronchioles and were strongly associated with LAC. Takahashi et al. [14] and Qi et al. [15] suggested pGGNs with bubble-like signs (B-pGGNs) were significantly associated with progressive adenocarcinoma. Zhang et al. [16] suggested B-pGGNs were significantly more common in the IAC than in the AIS-MIA (14/63, 22.22% vs. 16/163, 9.82%, $P = 0.014$). Qi et al. [15] also concluded that B-pGGNs were more likely reminiscent of IAC (IAC: 8/24, 33.33% vs. MIA: 12/108, 11.11%, $P = 0.001$).

The bubble-like sign strongly suggests the invasiveness and progression of pGGN on imaging, it is a risk factor for the invasiveness of the IAC. Moreover, bubble-like as a risk factor is not uncommon in MIA. In previous studies, the probability of B-pGGN in MIA was 22.06–22.92%, the probability of B-pGGN in IAC was 20.96–33.83% [17–20]. The morphological features of MIA and IAC presented as B-pGGN have greater overlap and similarity. Therefore, it is of more radiological significance to identify and how to identify such MIA and IAC. It's may offer some guidance

for personalized medical decision-making if B-pGGN-like IAC can be identified on preoperative CT images.

Preoperative HRCT images predicting the invasiveness of pGGNs were based on larger nodule size, higher entropy, and the bubble-like sign, etc., had been widely recognized in previous studies [21–25]. However, the conventional assessments were usually based on the radiologist's direct view and experience. By comparison, in radiomics, extraction of numerous features from the entire tumour volume is an objective and highly reproducible process, including complex patterns that are impossible for the human visual system to recognize or quantify [26]. Previous studies suggested that radiomics feature decoding was strongly correlated with invasiveness indices at the histological level [27, 28]. A nomogram is a graphical representation to depict a statistical model that generates a probability of a clinical event, for a given individual [29].

However, to the best of our knowledge, no studies have investigated identification of IAC and MIA in B-pGGNs. Therefore, the objective of the present retrospective study was to construct a radiomics nomogram for more precise prediction of IAC and MIA presenting as B-pGGNs on preoperative HRCT.

Patients and methods

Patients

With approval from the Ethical Committee and the Institutional Review Board, this retrospective study was completed with an exemption for informed consent. We retrospectively reviewed 511 nodules (MIA, $n = 288$; IAC, $n = 223$) between November 2012 and June 2018 who were pathologically confirmed after thoracic surgical resection. Inclusion criteria: (A) HRCT scans performed within two weeks before lesion resection; (B) complete specimens of pGGNs on lung window images (window width, WW, 1500 HU; window level, WL, − 500 HU) histologically confirmed to be MIA or IAC; (C) available picture archiving and communication systems (PACS, CT imaging thickness ≤ 1.5 mm); and (D) B-pGGNs characterized as small spots of round or ovoid air attenuation within the lesions. Eventually, after applying the exclusion criteria, a total of 109 B-pGGNs (MIA, $n = 55$; IAC, $n = 54$) from 109 patients fulfilled the inclusion criteria. The detailed process of case selection is shown in Fig. 1. Demographic variables (e.g., age, sex, smoking history, underlying diseases) were recorded.

CT scan parameters

Noncontrast chest CT examinations were performed with scanning machines (Optima CT660, Discovery CT750 HD,
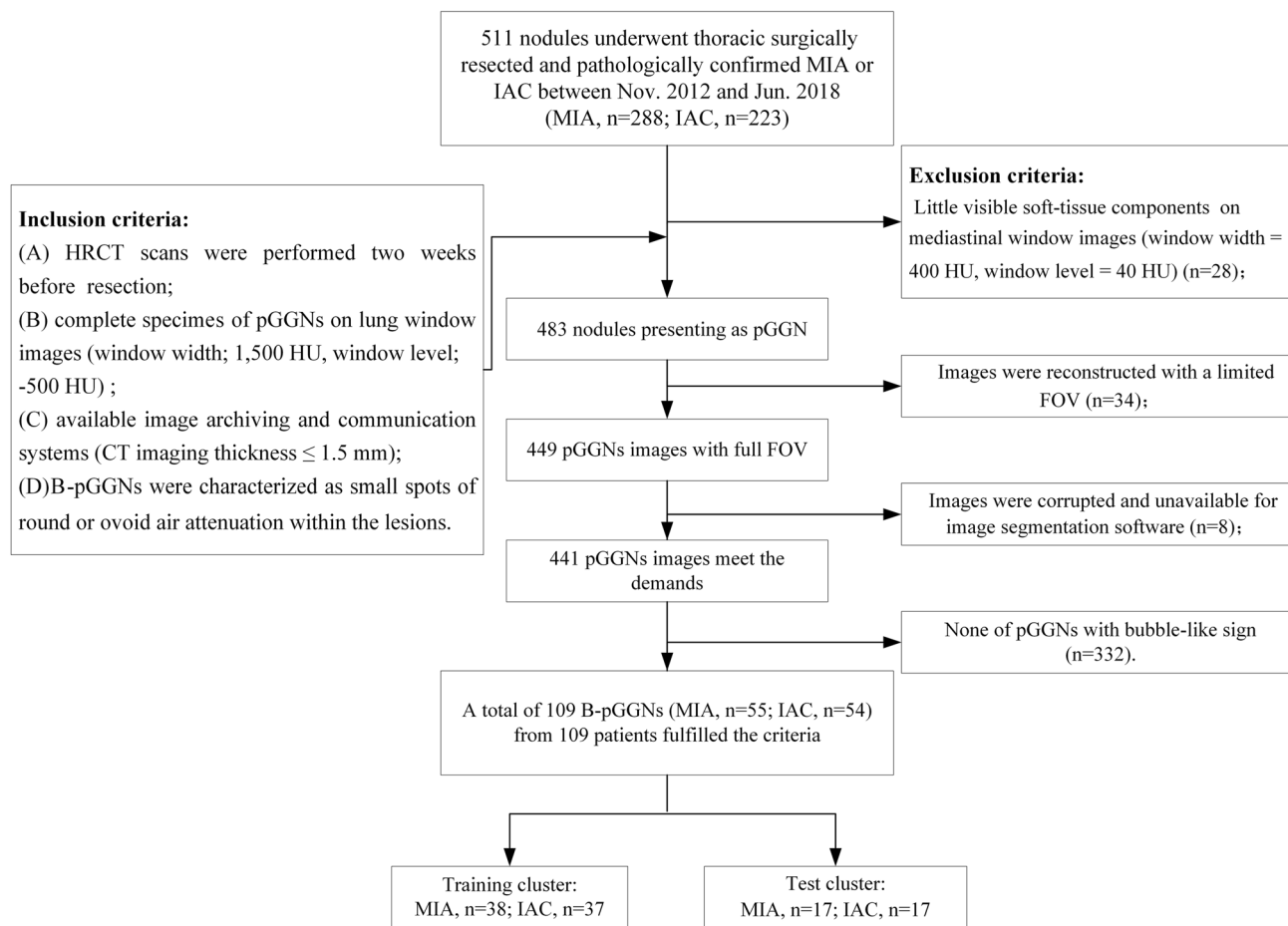
Fig. 1 The study flow chart shows the inclusion criteria and exclusion criteria for the final selected cases

Revolution CT or LightSpeed16 from General Electric, SOMATOM Perspective or Emotion 16 from Siemens, Brilliance 16P from Philips). Although various chest CT imaging protocols were used in this study, all examinations were performed with contiguous 1.00–1.50-mm axial sections and 1.00–1.50-mm slice intervals and then reconstructed at 1.00–1.50-mm intervals. Imaging parameters: a matrix of 512×512; tube current of 170–200 mA; a tube voltage of 120 kVp; rotation times of 0.5–0.6 s; a full field of view (FOV). Data were reconstructed with a lung kernel algorithm. Chest scans were performed with the patients in a supine position from the lung apex to the lung base area at the end of inspiration. A breath-hold exercise was implemented before each examination.

Histological type diagnosis

All pathological type diagnoses results were based on clinicopathological reports. During pathological diagnosis, all histological specimens were formalin fixed, paraffin embedded, haematoxylin–eosin stained and examined

by an experienced pathologist, then reviewed by a senior pathologist. Any disagreements were resolved by consensus, finally the pathological results were obtained. The histological types were recorded in accordance with the new classification suggested by the IASLC/ATS/ERS in 2011. We reviewed the diameter of invasion in all tumour recorded according to WHO in 2015.

CT feature analysis

CT features were observed by two thoracic radiologists (a junior and a senior with 20 years of experience) blinded to the pathological results in the lung window (WW, 1500 HU; WL, − 500 HU). B-pGGN images were evaluated, and the following imaging features were recorded: tumour location, shape (irregular or round and oval), the tumour-lung interface (clear or unclear), lobulation, and pleural indentation in the B-pGGNs. The maximum axial layer (MAL) was selected for all measurement from the lung window. MD-MAL: the maximum diameter; MVD-MAL: the maximum vertical diameter of the MD. CTv-MAL: the CT value. The

measurement details: (1) An irregular curve depicted the edge of the region of interest (ROI) of the B-pGGN as much as possible. (2) Large vessels, bronchi and bubble-like signs should be eliminated when present in the measurement layer. Submaximum layers should be measured when they cannot be completely removed.

Manual nodule segmentation

All anonymized DICOM-format image sequences were manually segmented with axial lung window settings (WW, 1500 HU; WL, -500 HU) by outlining the ROI layer-by-layer on B-pGGN images and eliminating large vessels, bronchi and bubble-like signs. Then, a three-dimensional ROI of the lesions was obtained.

Radiomics feature extraction and selection process

A total of 106 radiomics features were extracted by 3D slicer (version 4.8.1, https://www.slicer.org/). Seven categories included Shape ($n = 13$), Grey Level Dependence Matrix (GLDM; $n = 14$), Grey Level Co-occurrence Matrix (GLCM; $n = 24$), First-order ($n = 18$), Grey Level Run Length Matrix (GLRLM; $n = 16$), Grey Level Size Zone Matrix (GLSZM; $n = 16$) and Neighbouring Grey Tone Difference Matrix (NGTDM; $n = 5$).

The reproducibility and robustness of extracted features were calculated by intra- and inter-reviewer correlation coefficients (ICCs). One radiologist performed all nodule segmentations on the HRCT images. Then, the same radiologist and another radiologist independently resegmented the same 50 B-pGGNs image sequences selected randomly after an interval of no less than 30 days. Intra- and inter-reviewer agreement were calculated.

Our dataset was randomly assigned to two clusters at a 7:3 ratio, feature selection and model establishment were carried out for the larger cluster, and model performance was verified by the smaller clusters and further independently verified in cases grouped by the brands of CT scanning machines. Before processing, normalization was conducted. The features were interpreted as showing good agreement if ICCs > 0.75 were retained. The feature selection method based on the gradient boosting decision tree (GBDT) [30].

Six kinds of machine-learning algorithms constructed various radiomics models (rad_score) based on the optimal feature subsets of the training cluster: logistic regression (LR), support vector machine (SVM), naive Bayesian classifier (NBC), K-nearest neighbour (KNN), decision tree (DT), and random forest (RF). The performances of different models were comprehensively evaluated by receiver operator characteristic (ROC) curves and selected the optimal stable model.

Multivariable LR analysis was performed with clinical, semantic and conventional quantitative CT candidate predictors of nodules to constitute the conventional LR predictive model. In addition, the conventional model was combined with a radiomics model to construct a combined model. The tenfold cross-validation were also performed to evaluate the performance and reliability of our model.

The Kolmogorov–Smirnov test was performed to determine normal distribution. T tests were used for normally distributed data expressed as the mean ± standard deviation (SD), and the Mann–Whitney U test was used for nonnormally distributed data expressed as the median and quartile. A chi-squared test or Fisher's test was used to examine semantic signs. ROC curves were plotted to assess the performance of the three models. The accuracy (ACC), sensitivity (SEN), specificity (SPE) and AUC were also calculated. Calibration curves accompanied by the Hosmer–Lemeshow test (a nonsignificant test statistic indicates that the model is perfectly calibrated) were plotted to evaluate the predictive accuracy of the models. Decision curve analysis (DCA) was conducted to evaluate whether the models were sufficiently robust. Significance in Delong test suggested differences and stability between models.

All statistical analyses for the present study were executed by SPSS (version 26.0, IBM, Armonk, NY, USA), R (version 3.5.1) and Python (version 3.5.6). A two-tailed P value < 0.05 indicated statistical significance.

Results

A total of 109 patients (50 males and 59 females) with 109 nodules who were randomly allocated into the training cluster ($n = 75$) and the test cluster ($n = 34$), including 55 MIA (54.46%) and 54 IAC (49.54%), were enrolled in this study. No B-pGGNs were correlated with any mediastinal lymph node metastases. Our data showed that diameter of invasion in IAC group is 8.00 (6.25, 8.75) mm, the diameter of invasion in MIA group is 3.00 (2.00, 4.00) mm, the diameter of invasion in all B-pGGN is 5.00 (3.00, 8.00) mm. Few cases were associated with limited and mild underlying diseases, and more underlying diseases were noted in the IAC than in the MIA group ($P = 0.004$). The comparisons between clinical information and CT findings of all populations are shown in Supplementary Material Table S1. Representative axial CT images and corresponding HE-stained sections in B-pGGNs are shown in Fig. 2.

Univariate LR analysis showed statistically significant differences in shape, tumour-lung interface, the MD-MAL and the MVD-MAL between the MIA and IAC group in the training cluster ($P < 0.05$). Multivariate LR analysis showed that only tumour-lung interface and the MVD-MAL remained statistically significant. Comparisons of clinical
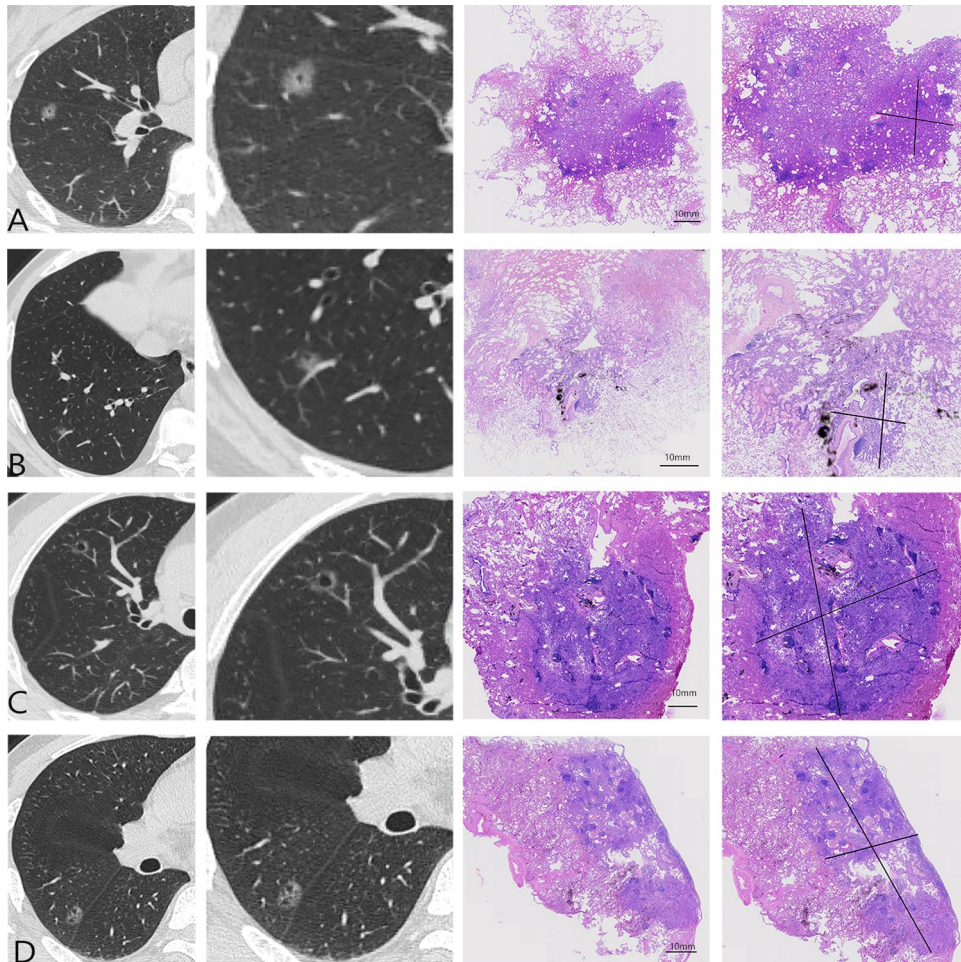
Fig. 2 Representative axial CT images and corresponding HE-stained sections in B-pGGNs. Row A: The axial CT image of a 45-year-old female's B-pGGN showed that CTv-MAL was − 553.49 HU, MD-MAL was 13.10 mm and MVD-MAL was 9.34 mm, with irregular shape, lobulation and plural indentation. The corresponding histological findings can be seen on the HE-stained section, under the microscope, the size of whole B-pGGN was 0.5×0.6 mm, the size of invasive component was 0.2×0.15 mm, the pathological diagnosis was MIA, and it was an atypical case clinically. Row B: The axial CT image of a 53-year-old male's B-pGGN showed that CTv-MAL was − 634.64 HU, MD-MAL was 9.99 mm and MVD-MAL was 7.60 mm, with lobulation. The corresponding histological findings can be seen on the HE-stained section, under the microscope, the size of whole B-pGGN was 0.4×0.5 mm, the size of invasive component was 0.1×0.1 mm, the pathological diagnosis was MIA, and it was a typical case clinically. Row C: The axial CT image of a 55-year-old male's B-pGGN showed that CTv-MAL was − 687.66 HU, MD-MAL was 13.40 mm and MVD-MAL was 11.80 mm. The corresponding histological findings can be seen on the HE-stained section, under the microscope, the size of whole B-pGGN was 0.5×0.6 mm, the size of invasive component was 0.5×0.6 mm, the pathological diagnosis was IAC, and it was an atypical case clinically. Row D: The axial CT image of a 50-year-old male's B-pGGN showed that CTv-MAL was − 603. 16 HU, MD-MAL was 14.30 mm and MVD-MAL was 10.90 mm, with pleural indentation. The corresponding histological findings can be seen on the HE-stained section, under the microscope, the size of whole B-pGGN was 0.3×0.7 mm, the size of invasive component was 0.3×0.7 mm, the pathological diagnosis was IAC, and it was a typical case clinically"

information and CT findings between the MIA and IAC group are listed in Table 1. The results of the LR analysis for conventional features are listed in Table 2.

In ICC tests, 100 features with ICCs > 0.75 were retained, and 6 features were eliminated (Supplementary Material Table S2). The GBDT method was applied to consolidate the 100 features into 16 potential predictors which were implemented to develop the radiomics model.

Based on the selected feature set, six machine-learning algorithms were adopted and the models were evaluated (Table 3). Considering the performance of the models between the training and the test cluster as well as the interpretability of the algorithms, the LR algorithm was finally selected to build the most appropriate model. The tenfold cross-validation curve for radiomics model and the combined model are presented in Supplementary Fig S1.

Table 1 Clinical information and CT findings of the training cluster and test cluster

Variables	Training cluster (n=75)			Test cluster (n=34)		
	MIA (n=38)	IAC (n=37)	P	MIA (n=17)	IAC (n=17)	P
Age (years)	56.71 ± 10.39	60.32 ± 10.33	0.135	57.00 ± 12.39	62.00 ± 9.62	0.198
Gender (%)			0.301			0.732
Female	24 (63.16%)	19 (51.35%)		9 (52.94%)	7 (41.18%)	
Male	14 (36.84%)	18 (48.65%)		8 (47.06%)	10 (58.82%)	
Smoking status			0.981			0.485
Never smoked	36 (94.74%)	36 (97.30%)		15 (88.24%)	17 (100.00%)	
Former or current smoker	2 (5.26%)	1 (2.70%)		2 (11.76%)	0 (0.00%)	
Tumour location			0.593			0.820
Right upper lobe	15 (39.47%)	14 (37.84%)		5 (29.41%)	6 (35.29%)	
Right middle lobe	5 (13.16%)	2 (5.41%)		1 (5.88%)	0 (0.00%)	
Right lower lobe	6 (15.79%)	4 (10.81%)		3 (17.65%)	5 (29.41%)	
Left upper lobe	7 (18.42%)	8 (21.62%)		4 (23.53%)	4 (23.53%)	
Left lower lobe	5 (13.16%)	9 (24.32%)		4 (23.53%)	2 (11.76%)	
Shape			0.015			1.000
Irregular	12 (31.58%)	22 (59.46%)		7 (41.18%)	8 (47.06%)	
Round and oval	26 (68.42%)	15 (40.54%)		10 (58.82%)	9 (52.94%)	
Tumour-lung interface			0.015			0.601
Clear	35 (92.11%)	26 (70.27%)		14 (82.35%)	16 (94.12%)	
Unclear	3 (7.89%)	11 (29.73%)		3 (17.65%)	1 (5.88%)	
Lobulation			0.744			0.018
Absent	13 (34.21%)	14 (37.84%)		0 (0.00%)	6 (35.29%)	
Presence	25 (65.79%)	23 (62.16%)		17 (100.00%)	11 (64.71%)	
Pleural indentation			0.300			0.296
Absent	25 (65.79%)	20 (54.05%)		12 (70.59%)	8 (47.06%)	
Presence	13 (34.21%)	17 (45.95%)		5 (29.41%)	9 (52.94%)	
CTv-MAL (HU)	− 686.29 (− 741.04, − 641.21)	− 627.77 (− 690.22, − 586.56)	0.007	− 670.03 ± 59.50	− 627.41 ± 79.84	0.087
MD-MAL (mm)	14.55 (11.40, 19.54)	18.20 (14.88, 21.56)	0.005	13.80 (9.79, 16.12)	20.40 (15.07, 26.31)	0.001
MVD-MAL (mm)	10.75 (8.85, 14.40)	13.50 (10.87, 17.33)	0.007	11.59 ± 4.14	15.35 ± 4.76	0.020

The semantic features are represented as numbers (%). Normally distributed data are presented as the mean ± standard deviation (SD). Nonnormally distributed data are presented as the median (interquartile range). *IAC* invasive adenocarcinoma, *MIA* minimally invasive adenocarcinoma. CTv-MAL refers to the computed tomography attenuation value on the maximum axial layer. MD-MAL refers to the maximum diameter on the maximum axial layer. MVD-MAL refers to the maximum vertical diameter of the maximum diameter on the maximum axial layer

Table 2 Results of the logistic regression analysis for conventional features

Features	Univariate logistic regression			Multivariate logistic regression		
	OR	95% CI	P	OR	95% CI	P
Shape	3.178	1.231, 8.200	0.017	–		
Tumour–lung interface	4.936	1.250, 19.498	0.023	6.259	1.443, 27.147	0.014
MD-MAL (mm)	1.117	1.019, 1.223	0.018	-		
MVD-MAL (mm)	1.175	1.046, 1.320	0.007	1.198	1.057, 1.357	0.005

MD-MAL refers to the maximum diameter on the maximum axial layer. MVD-MAL refers to the maximum vertical diameter of the maximum diameter on the maximum axial layer. *OR* odds ratio, *CI* confidence interval

Table 3 Performance evaluation of all models established by the six machine-learning algorithms

Model	Train				Test			
	AUC (95%CI)	ACC	SPE	SEN	AUC (95%CI)	ACC	SPE	SEN
LR	0.947 (0.905, 0.980)	0.867	0.895	0.838	0.945 (0.875, 0.996)	0.853	0.882	0.824
SVM	0.979 (0.956, 0.995)	0.893	0.895	0.892	0.934 (0.857, 0.990)	0.853	0.824	0.882
NBC	0.879 (0.809, 0.936)	0.773	0.711	0.838	0.875 (0.758, 0.965)	0.853	0.824	0.882
KNN	0.860 (0.791, 0.922)	0.693	0.921	0.459	0.881 (0.773, 0.964)	0.794	0.882	0.706
DT	0.898 (0.841, 0.947)	0.827	0.711	0.946	0.799 (0.670, 0.912)	0.794	0.765	0.824
RF	0.970 (0.939, 0.991)	0.920	0.974	0.865	0.903 (0.803, 0.979)	0.794	0.882	0.706

LR logistic regression, *SVM* support vector machine, *NBC* naive Bayesian classifier, *KNN* K-nearest neighbour, *DT* decision tree, *RF* random forest, *AUC* area under the receiver operating characteristic curve, *CI* confidence interval, *ACC* accuracy, *SEN* sensitivity, *SPE* specificity
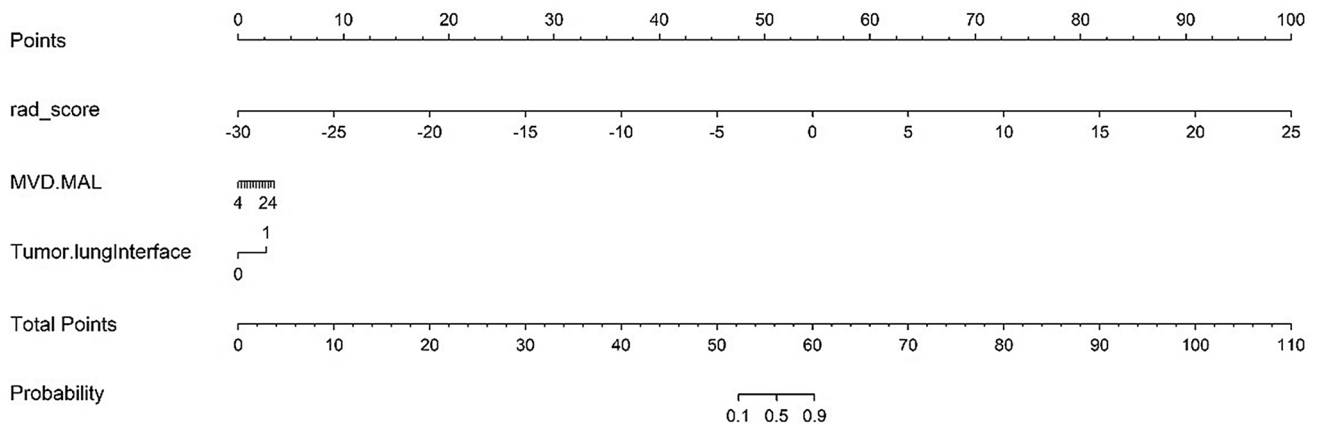

Fig. 3 The Nomogram of the combined model constructed by the rad_score, tumour-lung interface, and MVD-MAL to assess the probability of IAC in B-pGGNs. Among these, for tumour-lung interface, "1" indicates a pure ground-glass nodule with bubble-likes (B-pGGN) with an unclear edge appearance. A straight line was drawn perpendicularly from the marked position of each feature axis to the total points, and a corresponding point for each feature was obtained. All points were summed for all features. Next, the sum was marked on the total point axes, and a straight line was drawn perpendicular to the probability axis

Table 4 Evaluation of the three logistic regression prediction models

Item	Conventional model		Radiomics model		Combined model	
	Training	Test	Training	Test	Training	Test
AUC (95% CI)	0.750 (0.654,0.839)	0.692 (0.526, 0.840)	0.947 (0.905, 0.980)	0.945 (0.875, 0.996)	0.953 (0.915, 0.982)	0.945 (0.874,1.000)
ACC	0.653	0.618	0.867	0.853	0.867	0.882
SEN	0.784	0.824	0.838	0.824	0.892	0.882
SPE	0.526	0.412	0.895	0.882	0.842	0.882

AUC area under the receiver operating characteristic curve, *CI* confidence interval, *ACC* accuracy, *SEN* sensitivity, *SPE* specificity
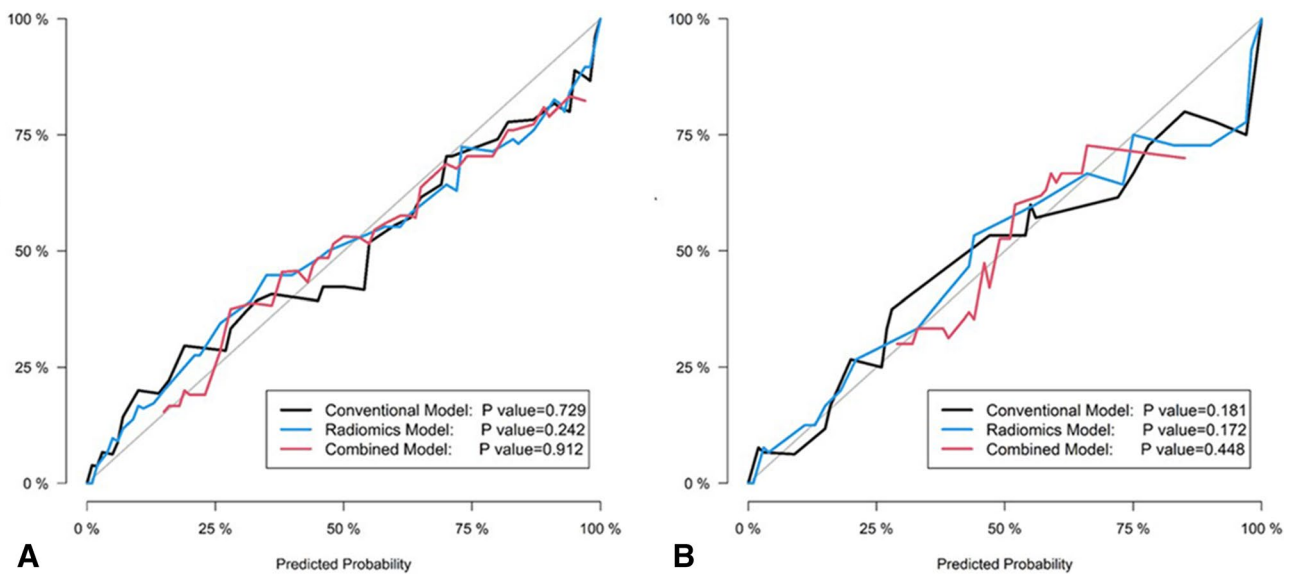
Fig. 4 The calibration curves of the three models for the training (**A**) and test clusters (**B**). Calibration curves evaluated the correspondence between predicted and observed probabilities. A colourful line closer to the grey line corresponds to a better prediction model

The rad_score was the only independent predictor in the combine model. The rad_score of 16 radiomics features was calculated as follows:

The Delong test showed that the combined model yielded no performance improvement (vs. the radiomics model) in the training cluster ($P = 0.485$). Regardless of cluster, both

$$
\begin{aligned}
\textbf{\textit{Rad\_score}} = {}& 0.725 - 2.036 \times shapeSphericity + 0.452 \times shapeMajorAxis + 4.958 \\
& \times shapeSurfaceArea - 1.757 \times firstorderSkewness - 2.793 \\
& \times firstorderMedian + 0.162 \times firstorderMaximum + 2.186 \\
& \times firstorder10Percentile + 0.080 \times firstorderKurtosis - 3.278 \\
& \times glrlmShortRunLowGrayLevelEmphasis - 0.029 \\
& \times glrlmLongRunLowGrayLevelEmphasis - 0.928 \\
& \times glszmSizeZoneNonUniformityNormalized + 2.385 \\
& \times glszmSmallAreaHighGrayLevelEmphasis + 2.646 \\
& \times glszmLowGrayLevelZoneEmphasis + 1.538 \\
& \times glszmSmallAreaLowGrayLevelEmphasis - 0.222 \\
& \times ngtdmCoarseness \\
& - 6.344 \times ngtdmBusyness
\end{aligned}
$$

In the present study, the conventional model composed of tumour-lung interface and the MVD-MAL produced medium AUCs of 0.750 in the training cluster and 0.692 in the test cluster. The radiomics model had better AUCs of 0.947 in the training cluster and 0.945 in the test cluster. The combined model produced AUCs of 0.953 in the training cluster and 0.945 in the test cluster, and was visualized by a nomogram (Fig. 3). The performance comparison details of the three models can be viewed in Table 4. The confusion matrixes of the three models are presented in Supplementary Tables S3–S11.

the radiomics model and the combined model were statistically better than the conventional model ($P < 0.05$). None of the three models showed significant differences between the training cluster and the test cluster, which further demonstrated the stability of our model.

The Hosmer–Lemeshow test with a nonsignificant test statistic ($P > 0.05$) indicated that our models had been well calibrated. The calibration curves are shown in Fig. 4. The decision curves showed that the radiomics model (blue line) and combined model (red line) generated a good net benefit in both clusters (Fig. 5).
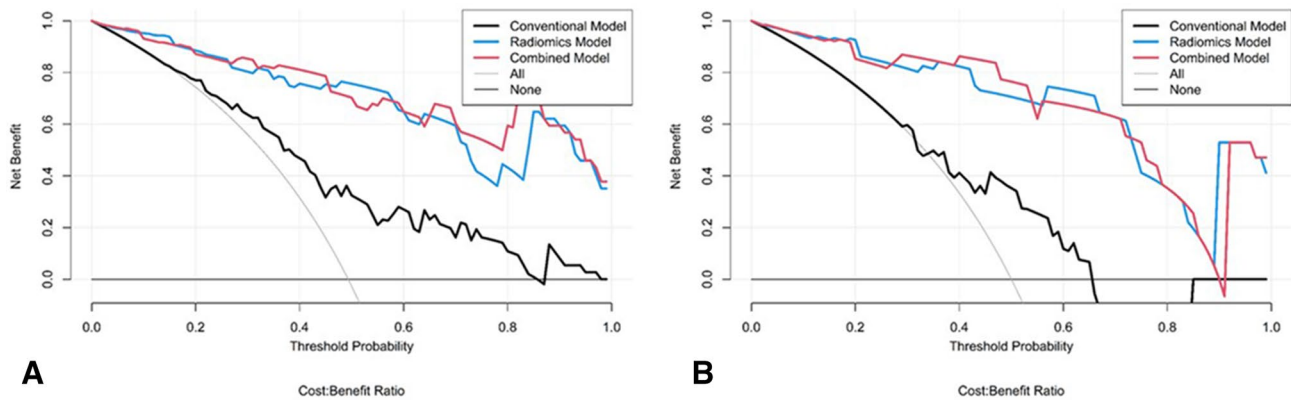
Fig. 5 The decision curves of the three models in the training (**A**) and test clusters (**B**). The net benefit is shown on the y-axis, and the probability threshold is shown on the x-axis. The radiomics model (blue line) and the combined model (red line) generated a good net benefit in both training and testing samples

Since only 5 nodules were scanned by the Philips machine, the independent validation groups were divided into a General Electric cluster (MIA, $n = 40$; IAC, $n = 45$) and a Siemens cluster (MIA, $n = 10$; IAC, $n = 9$). The performance of the radiomics model and the combined model was sufficiently verified in the General Electric cluster (AUC = 0.956 for the radiomics model; AUC = 0.956 for the combined model) and Siemens cluster (AUC = 0.897 for the radiomics model; AUC = 0.936 for the combined model). Furthermore, the Delong test demonstrated the appropriate stability and universality of the models.

Discussion

In this study, we aimed to analyse conventional CT and radiomics features to construct and verify a model that can accurately predict the pathological invasiveness of B-pGGNs before resection. To our knowledge, we are the first to focus on preoperative predictions of B-pGGNs. We discovered that the rad_score was significantly associated with the invasiveness of B-pGGNs and could predict IAC as an independent biomarker, with AUCs of 0.947 and 0.945 for the training and test clusters, respectively. The performance of the combined model was not significantly improved.

Suspicious morphology offered valuable information on pGGNs for managing [9, 31]. Bubble-like signs include a dilated bronchiole or residual air-filled space on pathology owing to partial alveolar filling by proliferative cells and/or thickening of alveolar septa and were valuable features for differentiating IAC in pGGNs [32]. Our results revealed that the probability of bubble-like signs in MIA reached 19.10% (55/288), which was similar to another report (22.1%) [20]. Considering the significance of precisely differentiation, our aim was to adopt a more comprehensive and objective model instead of a single morphology criterion.

Among conventional features, the tumour-lung interface and MVD-MAL were significantly correlated with the invasiveness of B-pGGNs. In terms of the marginal features of pGGNs, Chu et al. [20] found that invasive pGGNs (MIAs/IACs) usually had coarse margins ($P < 0.001$). Gao et al. [33] showed no difference in clear tumour-lung interfaces between preinvasive and invasive nodules ($P = 0.140$) but noted a difference between benign nodules and IAC ($P = 0.001$). Our study showed that IACs presenting B-pGGNs with an unclear tumour–lung interface were more common. The MVD-MAL was a quantitative index associated with nodule roundness [34]. Han et al. [23] showed that the MVD-MAL of pGGNs may be helpful for differentiating IAC from preinvasive/MIA lesions. Our study showed similar results: the MVD-MAL was a valuable predictor of B-pGGNs invasiveness.

Our research demonstrated a strong correlation between radiomics and tumour histological invasiveness. By using a combination of sixteen radiomics features, the B-pGGNs histological subtype could be reliably predicted, including three "Shape", five "First-order", two "GLRLM", four "GLSZM" and two "NGTDM" features. "Shape" defines the physical characteristics of the whole tumour in terms of shape and volume; for example, "Sphericity" determines how spherical (round) a tumour is, and "Surface major axis" and "Surface area" describe the 3D geometric properties of the tumour [35]. The "Shape" was also adopted by Xiong et al. [36] to discriminate between IAC from MIA presenting as pGGNs ($P = 0.028$). Using "First-order histogram" statistics of voxel intensity features, "Skewness" and "Median" quantified the density characteristics of tumour regions on CT images [37]. "Skewness" and "Median" had good performance in Zhang et al.'s [16] study distinguishing between IAC and AIS/MIA appearing as pGGNs. Texture features are derived from "GLRLM", "GLSZM" and "NGTDM", which reflect the textural characteristics of intra-tumour

heterogeneity [38]. As a non-invasive indicator, the high heterogeneity of lung adenocarcinoma by texture extraction was associated with a higher pathological grade [39].

Xu et al. [40] analysed 322 pGGNs and showed good predictive AUC for AIS/MIA from IAC in the test cohort of 0.824 for the conventional model, 0.833 for the radiomics model and 0.848 for the combined model. Compared with the features that they selected, "Shape" and "First-order" features were also included in ours, which yielded a better AUC. Sun et al. [41] predicted MIA/IAC from benign/AAH/AIS manifesting as pGGNs and indicated that the combined model (AUC = 0.80) performed better than the radiomics model (AUC = 0.73) in the training cluster. However, our results, the combined model did not improve diagnostic efficiency. In contrast to previous research, our study focused on B-pGGNs for the first time, showed excellent discriminative performance for the radiomics model (AUC, 0.947) in the training cluster and the superiority of the rad_score as an independent biomarker. The radiomics model had stable AUC, ACC, SPE and SEN values and was also proven to be excellent in the General Electric cluster (AUC = 0.956) and Siemens cluster (AUC = 0.897).

Nevertheless, our study had some limitations. First, multicentre B-pGGNs imaging data were not used for additional validation. We will conduct prospective experiments with a larger amount. Second, automatic generation of a three-dimensional mask of nodules is still an immature technology; thus, some subjective errors exist in the delineation of the ROIs. Remaining vessels and bronchi may still generate inaccurate predictions from some radiomics features. Because we focused on invasive lesions in case collection, AIS was not included in our study. In further refined experiments, comparative research on bubble-like signs in lung pGGNs will be performed between pathology and radiology settings, elastic fiber staining will be used for evaluation of alveolar rupture, collapse and invasion.

Conclusions

We established a model based on combining conventional and radiomics features to facilitate differentiation between IAC from MIA presenting as B-pGGNs on preoperative HRCT and demonstrated the excellent predictive performance. If these B-pGGNs are appropriately classified in a timely and reasonable manner, blind or extensive radical treatment will be avoided.

Declarations

References

1. Hansell DM, Bankier AA, MacMahon H, et al. Fleischner society: glossary of terms for thoracic imaging. Radiology. 2008;246:697–722.
2. Kim HY, Shim YM, Lee KS, et al. Persistent pulmonary nodular ground-glass opacity at thin-section CT: histopathologic comparisons. Radiology. 2007;245:267–75.
3. Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol. 2011;6:244–85.
4. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 Classification. J Thorac Oncol. 2015;10:1243–60.
5. Yeh YC, Kadota K, Nitadori J, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification predicts occult lymph node metastasis in clinically mediastinal node-negative lung adenocarcinoma. Eur J Cardiothorac Surg. 2016;49:e9–15.
6. Inamura K. Clinicopathological characteristics and mutations driving development of early lung adenocarcinoma: tumor initiation and progression. Int J Mol Sci. 2018;19:1259.
7. Sakamoto H, Shimizu J, Horio Y, et al. Disproportionate representation of KRAS gene mutation in atypical adenomatous hyperplasia, but even distribution of EGFR gene mutation from preinvasive to invasive adenocarcinomas. J Pathol. 2007;212:287–94.
8. Zhang C, Zhang JJ, Xu FP, et al. Genomic landscape and immune microenvironment features of preinvasive and early invasive lung adenocarcinoma. J Thorac Oncol. 2019;14:1912–23.
9. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. Radiology. 2017;284:228–43.
10. Van Schil PE, Asamura H, Rusch VW, et al. Surgical implications of the new IASLC/ATS/ERS adenocarcinoma classification. Eur Respir J. 2012;39:478–86.
11. Kim HY, Shim YM, Lee KS, et al. Persistent pulmonary nodular ground-glass opacity at thin-section CT: histopathologic comparisons. Radiology. 2017;245:267–75.
12. Wu F, Tian SP, Jin X, et al. CT and histopathologic characteristics of lung adenocarcinoma with pure ground-glass nodules 10 mm or less in diameter. Eur Radiol. 2017;27:4037–43.
13. Saito H, Yamada K, Hamanaka N, et al. Initial findings and progression of lung adenocarcinoma on serial computed tomography scans. J Comput Assist Tomogr. 2009;33:42–8.
14. Takahashi S, Tanaka N, Okimoto T, Tanaka T, et al. Long term follow-up for small pure ground-glass nodules: implications of determining an optimum follow-up period and high-resolution CT findings to predict the growth of nodules. Jpn J Radiol. 2012;30:206–17.

15. Qi L, Xue K, Li C, et al. Analysis of CT morphologic features and attenuation for differentiating among transient lesions, atypical adenomatous hyperplasia, adenocarcinoma in situ, minimally invasive and invasive adenocarcinoma presenting as pure ground-glass nodules. Sci Rep. 2019;9:14586.

16. Zhang T, Pu XH, Yuan M, et al. Histogram analysis combined with morphological characteristics to discriminate adenocarcinoma in situ or minimally invasive adenocarcinoma from invasive adenocarcinoma appearing as pure ground-glass nodule. Eur J Radiol. 2019;113:238–44.

17. Meng Y, Liu CL, Cai Q, et al. Contrast analysis of the relationship between the HRCT sign and new pathologic classification in small ground glass nodule-like lung adenocarcinoma. Radiol Med. 2019;124:8–13.

18. Liu LH, Liu M, Wei R, Jin E, et al. CT findings of persistent pure ground glass opacity: can we predict the invasiveness? Asian Pac J Cancer Prev. 2015;16:1925–8.

19. Meng FY, Guo Y, Li MY, Lu XQ, et al. Radiomics nomogram: a noninvasive tool for preoperative evaluation of the invasiveness of pulmonary adenocarcinomas manifesting as ground-glass nodules. Transl Oncol. 2021;14:100936.

20. Chu ZG, Li WJ, Fu BJ, et al. CT characteristics for predicting invasiveness in pulmonary pure ground-glass nodules. AJR Am J Roentgenol. 2020;215:351–8.

21. Lim HJ, Ahn S, Lee KS, Han J, et al. Persistent pure ground-glass opacity lung nodules >/= 10 mm in diameter at CT scan: histopathologic comparisons and prognostic implications. Chest. 2013;144:1291–9.

22. Hwang I-P, Park CM, Park SJ, et al. 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. 2015;50:798–804.

23 Han L, Zhang P, Wang Y, et al. CT quantitative parameters to predict the invasiveness of lung pure ground-glass nodules (pGGNs). Clin Radiol. 2018;73:504e1–7.

24. Zhang Y, Tang J, Xu JR, et al. Analysis of pulmonary pure ground-glass nodule in enhanced dual energy CT imaging for predicting invasive adenocarcinoma: comparing with conventional thin-section CT imaging. J Thorac Dis. 2017;9:4967–78.

25. Zhan Y, Peng XQ, Shan F, et al. Attenuation and morphologic characteristics distinguishing a ground-glass nodule measuring 5–10 mm in diameter as invasive lung adenocarcinoma on thin-slice CT. AJR Am J Roentgenol. 2019;213:W162–70.

26. Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. Transl Lung Cancer Res. 2017;6:86–91.

27. Wang LJ, Shen WY, Xi Y, et al. Nomogram for predicting the risk of invasive pulmonary adenocarcinoma for pure ground-glass nodules. Ann Thorac Surg. 2018;105:1058–64.

28. Yang B, Guo LL, Lu GM, et al. Radiomic signature: a non-invasive biomarker for discriminating invasive and non-invasive cases of lung adenocarcinoma. Cancer Manag Res. 2019;11:7825–34.

29. Balachandran VP, Gonen M, Smith JJ, et al. Nomograms in oncology: more than meets the eye. Lancet Oncol. 2015;16:e173–80.

30. Jiang J, Wang R, Wang M, et al. Boosting tree-assisted multitask deep learning for small scientific datasets. J Chem Inf Model. 2020;60:1235–44.

31. Cho J, Ko S-J, Kim SJ, et al. Surgical resection of nodular ground-glass opacities without percutaneous needle aspiration or biopsy. BMC Cancer. 2014;18:838.

32. Ding HD, Shi JY, Zhou X, et al. Value of CT characteristics in predicting invasiveness of adenocarcinoma presented as pulmonary ground-glass nodules. Thorac Cardiovasc Surg. 2017;65:136–41.

33. Gao F, Sun Y, Zhang GZ, et al. CT characterization of different pathological types of subcentimeter pulmonary ground-glass nodular lesions. Br J Radiol. 2019;92:20180204.

34. Henschke CI, Naidich DP, Yankelevitz DF, et al. Early lung cancer action project: initial findings on repeat screenings. Cancer. 2001;92:153–9.

35. Lee GW, Lee HY, Park HJ, et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. Eur J Radiol. 2017;86:297–307.

36. Xiong ZQ, Jiang YN, Che SY, et al. Use of CT radiomics to differentiate minimally invasive adenocarcinomas and invasive adenocarcinomas presenting as pure ground-glass nodules larger than 10 mm. Eur J Radiol. 2021;141:109772.

37. Kamiya A, Murayama S, Kamiya H, et al. Kurtosis and skewness assessments of solid lung nodule density histograms: differentiating malignant from benign nodules on CT. Jpn J Radiol. 2014;32:14–21.

38. Ondimu SN, Murase H. Effect of probability-distance based Markovian texture extraction on discrimination in biological imaging. Comput Electron Agric. 2008;63:2–12.

39. Liu Y, Liu SC, Qu FY, et al. Tumor heterogeneity assessed by texture analysis on contrast-enhanced CT in lung adenocarcinoma: association with pathologic grade. Oncotarget. 2017;8:53664–74.

40. Xu FY, Zhu WC, Shen Y, et al. Radiomic-based quantitative CT analysis of pure ground-glass nodules to predict the invasiveness of lung adenocarcinoma. Front Oncol. 2020;10:872.

41. Sun YL, Li C, Jin L, et al. Radiomics for lung adenocarcinoma manifesting as pure ground-glass nodules: invasive prediction. Eur Radiol. 2020;30:3650–9.