Am J Cancer Res 2025;15(2):797-810
www.ajcr.us /ISSN:2156-6976/ajcr0161941

Original Article

Nomogram integrating clinical-radiological and radiomics features for differentiating invasive from non-invasive pulmonary adenocarcinomas presenting as ground-glass nodules

Ning Dong[1*], Sirong Wei[2*], Lei Zheng[1], Delong Huang[1], Guowei Zhang[1], Yunxin Li[1], Hu Zhang[3], Aijie Wang[1], Ranran Huang[1], Xinyao Zhao[1], Peng Liang[1]

[1]Department of Radiology, Yantaishan Hospital, Yantai 264003, Shandong, China; [2]Department of Vascular Intervention, Yantai Qishan Hospital, Yantai 264001, Shandong, China; [3]Department of Radiology, Affiliated Hospital of Binzhou Medical College, Binzhou 256603, Shandong, China. [*]Equal contributors and co-first authors.

Received November 12, 2024; Accepted February 13, 2025; Epub February 15, 2025; Published February 28, 2025

Abstract: Objective: To construct a nomogram incorporating clinical-radiological and radiomics features from computed tomography (CT) for distinguishing invasive adenocarcinoma (IAC) from adenocarcinoma in situ (AIS) and minimally invasive adenocarcinoma (MIA) in ground-glass nodules (GGNs). Methods: This retrospective study included 473 GGN patients with postoperative pathological confirmation of AIS, MIA, or IAC. The training set comprised 257 patients from Yantaishan Hospital, while the test set, used for external validation, included 216 patients from the Affiliated Hospital of Binzhou Medical College. Radiomics features were selected, and a radiomics model was constructed using least absolute shrinkage and selection operator (LASSO) and minimum redundancy maximum relevance (mRMR) methods. A clinical-radiological model was developed using univariate and multivariate logistic regression. The nomogram was generated by combining the two models. Its performance was evaluated via receiver operating characteristic (ROC) curve analysis, calibration curve analysis, and decision curve analysis (DCA). Results: The radiomics model included 11 features, while the clinical-radiological model incorporated lobulation, age, and long diameter. The nomogram outperformed both individual models in terms of accuracy and area under the curve (AUC) in both the training and test sets. Calibration curve analysis confirmed good consistency between actual and predicted outcomes, and DCA indicated the nomogram's clinical utility. Conclusion: The nomogram is a non-invasive, accurate tool for preoperative differentiation of GGN types, providing valuable guidance for clinicians in treatment planning.

Keywords: Nomogram, radiomics, ground-glass nodules, lung adenocarcinoma

Introduction

Lung adenocarcinoma, one of the most common types of lung cancer, often presents as ground-glass nodules (GGNs) in its early stages [1-4]. GGNs are frequently the predominant manifestation in early-stage lung adenocarcinoma and may indicate different pathological subtypes, including adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (IAC) [5, 6]. Distinguishing between these subtypes is crucial, as each has distinct biological behaviors and prognostic implications [7]. AIS typically appears as pure GGNs with a lower risk of progression, while MIA may show partial solid components and a higher likelihood of local invasion [8]. IAC, with more extensive solid components, is more aggressive and may involve lymph nodes or metastasize to distant organs [9-11]. However, differentiating between AIS, MIA, and IAC based solely on imaging remains challenging due to overlapping radiological features. This highlights the need for advanced predictive tools, such as radiomics, to improve preoperative differentiation.

In routine clinical practice, radiological characteristics are often used to assess the invasiveness of GGNs. However, due to overlapping fea-

https://doi.org/10.62347/AOAN9966

tures [12-14], making a definitive diagnosis before surgery is difficult. This underscores the need for tools like radiomics to assist in differentiation. Radiomics, a quantitative approach to medical imaging analysis, has gained significant attention since its introduction in 2012 [15]. By extracting and analyzing large amounts of image data, radiomics identifies hidden patterns and associations within medical images that are not visible through conventional inspection [16].

The development of a nomogram that integrates radiomics with clinical-radiological features offers a promising approach to improving preoperative diagnostic accuracy [17, 18]. By providing a quantitative tool for individualized risk assessment, the nomogram can help clinicians tailor treatment strategies, optimize resource allocation, and ultimately improve patient outcomes [17]. Moreover, the nomogram's ability to predict the likelihood of specific pathological types can inform decisions about surgical intervention, adjuvant therapy, and follow-up management [19-21]. This study aims to construct such a nomogram to distinguish patients with GGNs, identifying whether they have AIS/MIA or IAC, thereby assisting clinicians in formulating appropriate treatment plans.

Additionally, the nomogram's utility extends to surgical planning and outcomes. For instance, distinguishing between AIS/MIA and IAC can influence the choice of surgical procedures - such as lobectomy or sublobar resection - and guide decisions on the extent of lymph node dissection [22, 23]. By accurately predicting the type of GGN, the nomogram can aid surgeons in planning less invasive surgeries, potentially reducing morbidity and improving recovery times [24, 25]. Moreover, the nomogram can help identify patients who may benefit from neoadjuvant or adjuvant therapies, further enhancing personalized care [26].

Radiomics is increasingly applied in various tasks, including tumor subtyping, clinical outcomes prediction, tumor recurrence, treatment response gene expression, and lymph node metastasis prediction [27-32]. Several studies have used radiomics to differentiate IAC from AIS/MIA [33-35], demonstrating its potential to improve diagnostic accuracy and support clinical decision-making. Differentiating IAC from AIS and MIA is particularly important, as MIA shares similar prognosis and treatment approaches with preinvasive lesions. Despite extensive research on combining radiomics with clinical features to predict the invasiveness of lung adenocarcinoma, our study offers several novel contributions. First, we propose a comprehensive nomogram integrating both radiomics and clinical-radiological features, providing a robust predictive tool that clinicians can easily use. Second, our study includes a larger, more diverse cohort, ensuring the generalizability and reliability of the findings. Third, we apply advanced statistical methods, such as least absolute shrinkage and selection operator (LASSO) and minimum redundancy maximum relevance (mRMR), to optimize feature selection and enhance the model's predictive performance. Finally, our decision curve analysis (DCA) demonstrates the clinical utility of the nomogram, contributing significantly to existing research.

In summary, this study aims to develop a nomogram that integrates radiomics with clinical-radiological features to distinguish patients with GGNs, identifying whether they have AIS/MIA or IAC. This model will assist clinicians in formulating appropriate treatment plans.

Materials and methods

This study was reviewed and approved by the ethics committee of Yantaishan Hospital. The requirement for informed consent was waived, as the study was retrospective. **Figure 1** illustrates the overall flow of the study.

Patients

Patients who underwent pulmonary nodule resection between February 2020 and November 2022 at Yantaishan Hospital and the Affiliated Hospital of Binzhou Medical College were retrospectively included. The inclusion criteria were: (1) chest CT performed within two weeks before surgery; (2) lesions presenting as subsolid nodules, including pure ground-glass nodules (pGGNs) and part-solid nodules (PSNs). A pGGN is defined as a lesion with no solid component, where the underlying lung parenchyma is visible through the opacity, whereas a PSN (or mixed ground-glass nodule, mGGN) includes both solid and ground-glass components. Both pGGNs and PSNs were considered sub-

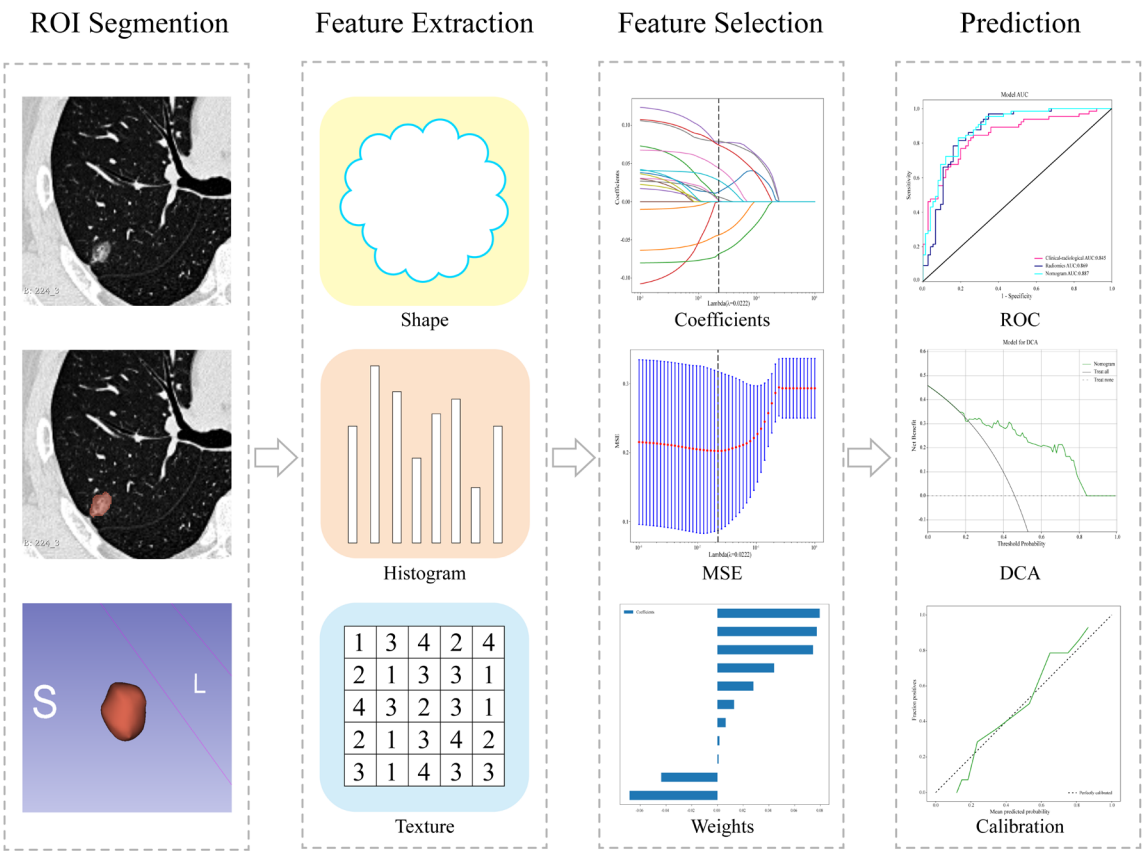

Figure 1. The workflow of this retrospective study. ROI, region of interest; ROC, receiver operating characteristic; DCA, decision curve analysis.

solid nodules for this study; (3) lesions imaged using lung window settings [width: 1300 Hounsfield units (HU), level: -650 HU]; (4) post-surgical diagnosis of AIS, MIA, or IAC.

The exclusion criteria were: (1) a history of chemotherapy, radiation therapy, or fine needle aspiration biopsy prior to CT examination; (2) multiple ground-glass nodules in the same lung lobe; (3) lesions > 30 mm in the longest diameter; (4) poor quality CT images, including those with severe motion artifacts.

To ensure adequate statistical power for detecting differences between the AIS, MIA, and IAC groups, the sample size was calculated based on previous studies with similar objectives [36, 37]. Power calculations were performed aiming for 80% power at an alpha level of 0.05 to detect a meaningful difference in AUC values. Considering the complexity of the radiomics model, we adhered to the rule of including at least 10 events per variable for model stability and accuracy. This study ultimately included 473 patients with 473 GGNs. The training set consisted of 257 patients from Yantaishan Hospital, while the test set included 216 patients from the Affiliated Hospital of Binzhou Medical College (**Figure 2**). The test set served as an external validation cohort, providing an independent evaluation of the model's performance.

CT protocols

CT scans were performed using one of three Philips Medical Systems CT scanners: Incisive 64, Brilliance 128, and Brilliance 64 (Philips Medical Systems, Netherlands). To minimize potential batch effects from using different scanners, all CT images were acquired using standardized protocols. Specifically, non-enhanced CT images were obtained with the parameters listed in **Table 1**.

Evaluation of clinical-radiological features

Clinical data, including sex, age, and smoking history, were extracted from electronic medical

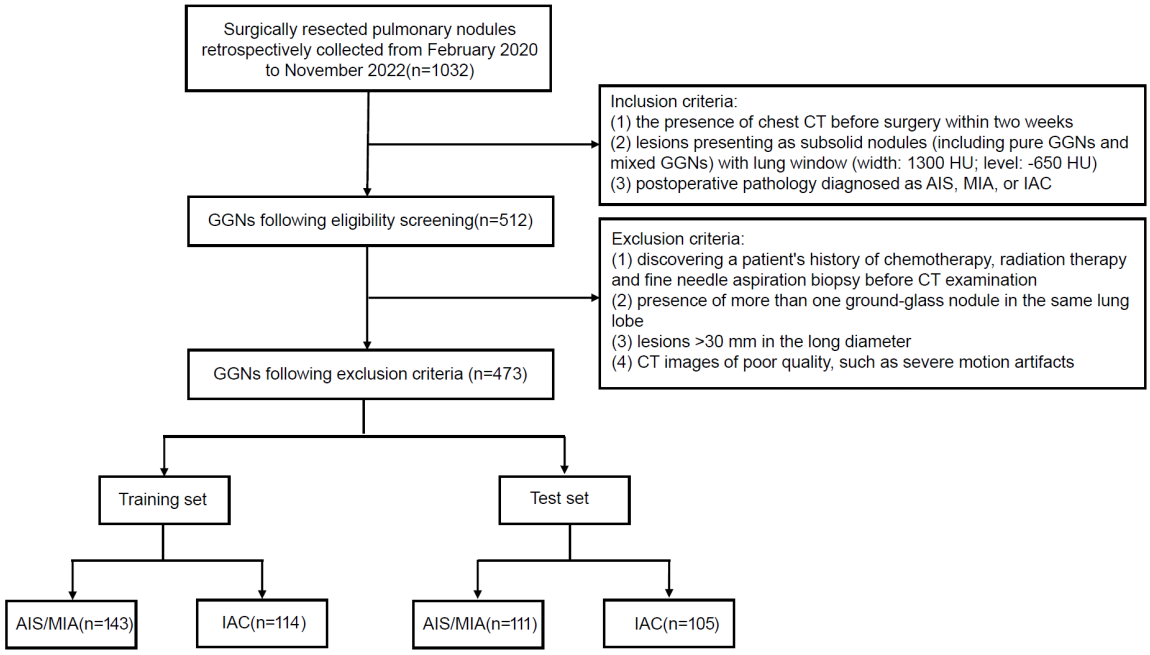

Figure 2. Flowchart of study population.

Table 1. The CT scanning parameters of different scanners

Settings	Brilliance 64	Brilliance 128	Incisive 64
Tube voltage	140 kVp	140 kVp	140 kVp
Tube current	300 mAs	300 mAs	300 mAs
Pitch	0.578	0.578	0.6
Collimation	0.625 mm ×64	0.625 mm ×64	0.625 mm ×32
Matrix size	1024×1024	1024×1024	1024×1024
Field of view	200 mm	200 mm	200 mm
Slice thickness of reconstruction	0.670 mm	0.670 mm	0.670 mm
Slice interval of reconstruction	0.340 mm	0.340 mm	0.335 mm
Reconstruction algorithm	Standard	Standard	Standard
Scanning methods	Plain scan	Plain scan	Plain scan

records. Radiological feature review was conducted by two independent chest radiologists, each with at least 15 years of experience in lung imaging, and with no prior knowledge of the study's details. Any disagreements were resolved through consensus. Thirteen radiological features were analyzed from the CT images, including lesion location, short and long diameters, lobulation, spiculation, margin, vascular change, air bronchus sign, bubble lucency, pleural indentation, shape, nodule type, and CT value. Both long and short diameters were included because they provide complementary information about the morphology and potential invasiveness of the nodules. The long diameter represents the largest dimension, indicating invasion extent, while the short diameter provides insight into the nodule's compactness or sphericity.

Segmentation and radiomics feature extraction

Manual segmentation of the nodules was performed by a junior chest radiologist with five years of experience using the open-source software 3D Slicer (Version 4.11; https://www.slicer.org) to define the region of interest (ROI) along the lesion's boundaries. The segmentation was then confirmed by a senior radiologist

with 20 years of chest CT experience. Any discrepancies were resolved through discussion between the two radiologists. During the segmentation process, the radiologists were instructed to carefully delineate the nodule, avoiding adjacent tissue structures such as blood vessels and the pleura. Radiomics features were extracted using Pyradiamics (http://pyradiomics.readthedocs.io).

Repeatability assessment

Two weeks after the initial analysis, 20 samples were selected for repeat testing. Both the senior and junior radiologists independently segmented the nodules and extracted features using the same methods. The repeatability of the segmentation and feature extraction was assessed using intra- and intergroup correlation coefficients (ICCs). Features with an ICC less than 0.75 were considered poorly repeatable and excluded from subsequent analysis.

Feature selection and radiomics model building

Before feature selection, the feature values of patients were normalized using the Z-score formula: $(x-\mu)/\sigma$, where x is the individual feature value, μ is the mean value of the feature across all patients, and σ is the standard deviation. For feature selection, two commonly used methods in radiomics studies were employed: (1) the mRMR method was used to remove redundant features and identify those with the highest correlation, and (2) an optimized subset of predictive features was selected, and a radiomics model was constructed using LASSO with 10-fold cross-validation. After feature selection, the corresponding coefficients for the selected radiomics features were calculated, and the Rad-score for each sample was derived by performing a weighted summation based on these feature coefficients.

Clinical-radiological model and nomogram building

Clinical and radiological features were selected using univariate and multivariate logistic regression methods to build a clinical-radiological model. Univariate analysis identified several potential risk factors for IAC, including age, lobulation, spiculation, vascular change, air bronchus sign, shape, nodule type, CT value, short diameter, and long diameter. Multivariate analysis revealed that age, lobulation, and long diameter were statistically significant and were used to construct the clinical-radiological model.

Model evaluation

The predictive performance of the test and training sets was evaluated using receiver operating characteristic (ROC) curve indicators, including negative predictive value (NPV), positive predictive value (PPV), area under the curve (AUC), sensitivity, specificity, and accuracy. The DeLong test was employed to assess the statistical significance of differences between the nomogram model developed in this study and other models. Calibration curves were used to evaluate the relationship between predicted and actual results of the nomogram model. The Hosmer-Lemeshow test assessed the model's goodness of fit, and DCA evaluated the clinical applicability of the nomogram model.

Statistical analysis

Statistical analysis was performed using SPSS 26.0 and the "One-key AI" platform (https://www.medai.icu). For data conforming to a normal distribution, continuous variables were presented as mean ± standard deviation ($\overline{x}\pm SD$), and group comparisons were made using the t-test. Categorical data were presented as frequencies (%), and the chi-square test was used for comparisons between two groups. A two-tailed *p*-value of less than 0.05 was considered statistically significant for all tests.

Results

Comparison of general data

A total of 473 patients were included in the study, with a mean age of 56.2±11.0 years (range: 24-77 years). The cohort comprised 333 females (70.40%) and 140 males (29.60%). Among the GGNs, there were 219 cases of IAC (46.30%), 165 cases of MIA (34.88%), and 89 cases of AIS (18.82%). Statistical analysis of patients' characteristics, including sex, age, smoking history, and radiological features, revealed no significant differences between the test and training sets (all > 0.05, Table S1). **Table 2** compares the clinical-radiological characteristics of the IAC group

Table 2. Comparison of clinical-radiological features between AIS/MIA and IAC groups

Items	Training set (n = 257)		P	Test set (n = 216)		P
Features	AIS/MIA (n = 143)	IAC (n = 114)		AIS/MIA (n = 111)	IAC (n = 105)	
Age (years)	53.32±10.61	61.03±9.87	<0.001	50.67±10.45	59.29±10.26	<0.001
Sex [n (%)]			0344			0.336
Male	48 (33.57%)	32 (27.98%)		34 (30.63%)	26 (24.76%)	
Female	95 (66.43%)	82 (72.02%)		77 (69.37%)	79 (75.24%)	
Smoking history [n (%)]			0.440			0.160
Absent	133 (92.99%)	103 (89.47%)		101 (%)	89 (84.76%)	
Present	10 (7.01%)	11 (10.53%)		10	16 (15.24%)	
Location [n (%)]			0.118			0.401
Right upper lobe	34 (23.78%)	38 (33.33%)		30 (27.03%)	30 (28.57%)	
Right middle lobe	15 (10.49%)	10 (8.77%)		12 (10.81%)	11 (10.48%)	
Right lower lobe	32 (22.38%)	26 (22.81%)		27 (24.32%)	21 (20.00%)	
Left upper lobe	42 (29.37%)	19 (16.67%)		25 (22.52%)	17 (16.19%)	
Left lower lobe	20 (13.99%)	21 (18.42%)		17 (15.32%)	26 (24.76%)	
Lobulation [n (%)]			<0.001			<0.001
Absent	108 (75.52%)	35 (30.70%)		87 (78.38%)	48 (45.71%)	
Present	35 (24.48%)	79 (69.30%)		24 (21.62%)	57 (54.29%)	
Spiculation [n (%)]			<0.001			0.074
Absent	128 (89.51%)	82 (71.93%)		96 (86.49%)	81 (77.14%)	
Present	15 (10.49%)	32 (28.07%)		15 (13.51%)	24 (22.86%)	
Margin [n (%)]			0.175			0.450
Clear	111 (77.62%)	80 (70.18%)		74 (66.67%)	75 (71.43%)	
Unclear	32 (22.38%)	34 (29.82%)		37 (33.33%)	30 (28.57%)	
Vascular change [n (%)]			<0.001			0.006
Absent	44 (30.77%)	7 (6.14%)		37 (33.33%)	18 (17.14%)	
Present	99 (69.23%)	107 (93.86%)		74 (66.67%)	87 (82.86%)	
Air bronchus sign [n (%)]			<0.001			<0.001
Absent	110 (76.92%)	50 (43.86%)		105 (94.59%)	45 (42.86%)	
Present	33 (23.08%)	64 (56.14%)		6 (5.41%)	60 (57.14%)	
Bubble lucency [n (%)]			0.102			0.213
Absent	129 (90.21%)	95 (83.33%)		94 (84.69%)	82 (78.09%)	
Present	14 (9.79%)	19 (16.67%)		17 (15.31%)	23 (21.91%)	
Pleural indentation [n (%)]			0.372			0.211
Absent	77 (53.85%)	55 (48.25%)		57 (51.35%)	45 (42.86%)	
Present	66 (46.15%)	59 (51.75%)		54 (48.65%)	60 (57.14%)	
Shape [n (%)]			<0.001			0.148
Round	105 (73.43%)	38 (33.33%)		71 (63.96%)	57 (54.29%)	
Irregular	38 (26.57%)	76 (66.67%)		40 (36.04%)	48 (45.71%)	
Nodule type [n (%)]			0.002			<0.001
pGGN	92 (64.34%)	51 (44.74%)		77 (69.37%)	46 (43.81%)	
mGGN	51 (35.66%)	63 (55.26%)		34 (30.63%)	59 (56.19%)	
CT value (HU)	-449.81±181.12	-349.80±163.23	<0.001	-495.21±152.80	-349.86±137.37	<0.001
Short diameter (mm)	5.94±2.62	8.11±3.08	<0.001	5.46±1.63	8.43±2.70	<0.001
Long diameter (mm)	8.66±3.68	14.20±5.73	<0.001	7.97±2.37	13.39±4.65	<0.001

CT, computed tomography; AIS, adenocarcinoma in situ; MIA, minimally invasive adenocarcinoma; IAC, invasive adenocarcinoma; pGGN, pure ground-glass nodule; mGGN, mixed ground-glass nodule; HU, Hounsfield units.

and the AIS/MIA group in both the training and test sets. Significant differences were observed in CT values, vascular changes, age, short diameter, long diameter, lobulation, air bronchus sign, and nodule type across patient groups in both sets.

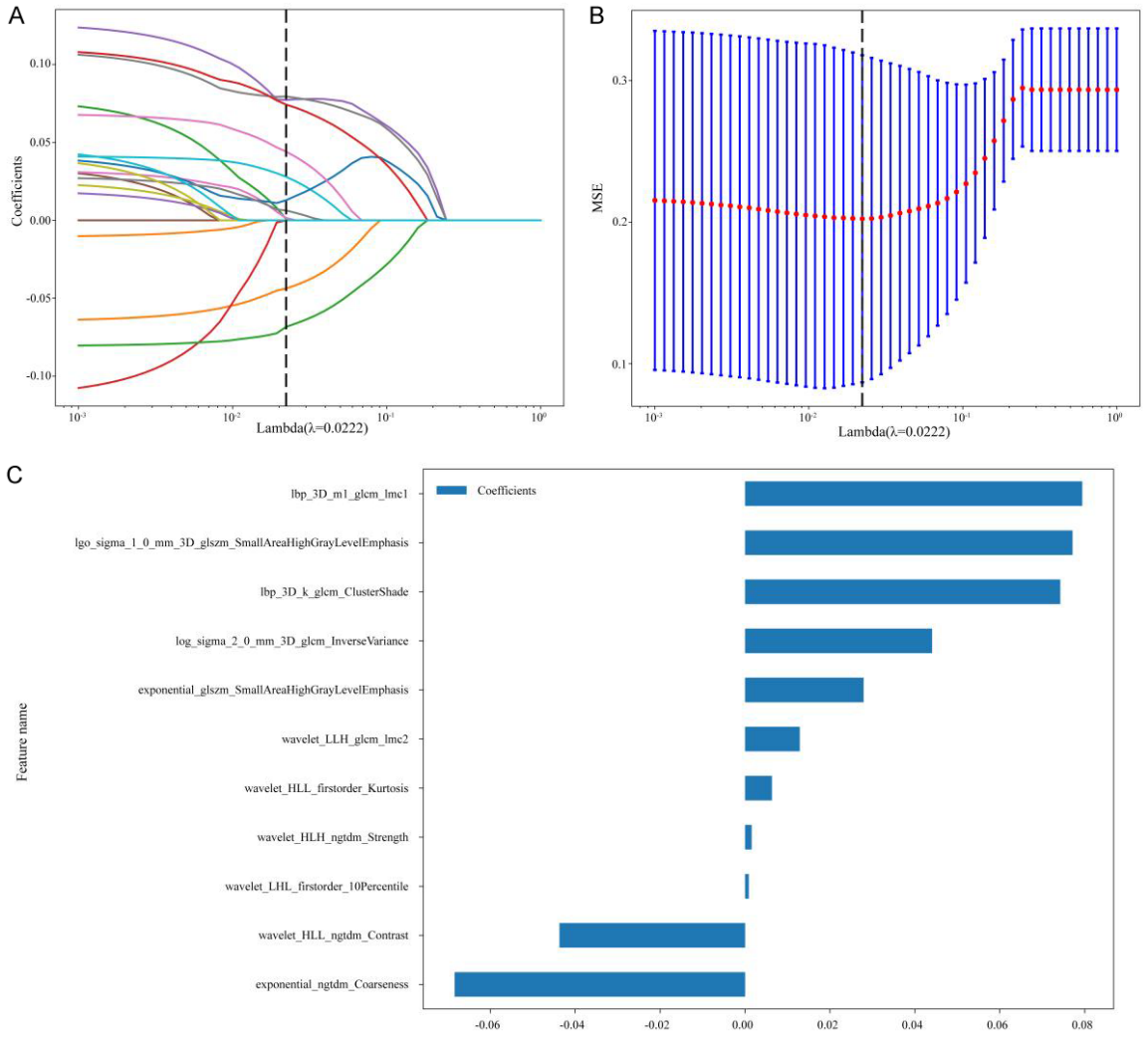

Figure 3. Feature selection using the least absolute shrinkage and selection operator (LASSO). A. Coefficients of 10-fold cross-validation; B. Mean squared error of 10-fold cross-validation; C. Selected features weight coefficients.

Radiomics model building and diagnostic validation

Each lesion's region of interest (ROI) was calculated with 1,834 features: 360 first-order features, 14 shape features, and 1,460 texture features. The texture features included 100 neighborhood gray tone difference matrix (NGTDM) features, 320 gray level size zone matrix (GLSZM) features, 440 gray level co-occurrence matrix (GLCM) features, 280 gray level dependence matrix (GLDM) features, and 320 gray level run length matrix (GLRLM) features. Features with an ICC greater than 0.75 were retained for further analysis, resulting in 1,463 radiomics features. The mRMR method was applied to remove irrelevant and redundant features, selecting the top 20 features based on correlation coefficients for LASSO screening. These 20 features were used to train the LASSO classifier on the training set. The LASSO logistic model contained 11 features with non-zero coefficients, forming the basis of the radiomics model (**Figure 3**). The Rad-score for each sample was calculated based on the selected features, with relevant formulas provided in the Supplementary Materials. In the training set, the AUC of the radiomics model for predicting IAC was 0.869 (95% CI: 0.841-0.897), while in the test set, it was 0.846 (95% CI: 0.818-0.874) (**Table 4**).

Development and predictive performance of the clinical-radiological model

Univariate logistic regression analysis identified age, lobulation, spiculation, vascular chan-

Table 3. Univariate and multivariate analysis of clinical-radiological features

Features	Univariate logistic regression		Multivariate logistic regression	
	OR (95% CI)	P	OR (95% CI)	P
Age	1.077 (1.049-1.108)	<0.001	1.071 (1.032-1.111)	<0.001
Lobulation	6.965 (4.056-12.234)	<0.001	6.640 (3.287-13.413)	<0.001
Spiculation	3.330 (1.725-6.684)	0.001		
Vascular change	6.794 (3.100-17.138)	<0.001		
Air bronchus sign	4.267 (2.513-7.368)	<0.001		
Shape	5.526 (3.255-9.562)	<0.001		
Nodule type	2.228 (1.351-3.703)	0.002		
CT value	1.001 (1.001-1.001)	<0.001		
Short diameter	1.074 (1.054-1.093)	<0.001		
Long diameter	1.289 (1.206-1.389)	<0.001	1.339 (1.226-1.462)	<0.001

OR, odds ratio; CI, confidence interval.

Table 4. Comparison of diagnosis efficiency of different models

	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy
Training set						
Clinical-radiological	0.845 (0.817-0.873)	0.846	0.733	0.773	0.846	0.786
Radiomics	0.869 (0.841-0.897)	0.815	0.813	0.791	0.836	0.814
Nomogram	0.887 (0.869-0.905)	0.831	0.813	0.794	0.847	0.821
Test set						
Clinical-radiological	0.823 (0.795-0.851)	0.571	0.970	0.941	0.727	0.787
Radiomics	0.846 (0.818-0.874)	0.857	0.758	0.750	0.862	0.803
Nomogram	0.865 (0.837-0.893)	0.857	0.818	0.800	0.871	0.836

AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value. The Clinical-radiological model includes age, lobulation, and long diameter as significant predictors identified through multivariate logistic regression analysis.

ge, air bronchus sign, shape, nodule type, CT value, short diameter, and long diameter as risk factors for IAC. Multivariate analysis revealed that age, lobulation, and long diameter were statistically significant and were included in the clinical-radiological model (**Table 3**). In the training set, the AUC of the clinical-radiological model was 0.845 (95% CI: 0.817-0.873), and 0.823 (95% CI: 0.795-0.851) in the test set (**Table 4**).

Radiomics nomogram building and evaluation

To enhance usability, a nomogram was constructed using multivariate logistic regression, incorporating age, lobulation, long diameter, and Rad-score (**Figure 4**). The AUC of the nomogram in the training set was 0.887 (95% CI: 0.869-0.905), with an accuracy of 0.821; in the test set, the AUC was 0.865 (95% CI: 0.837-0.893), with an accuracy of 0.836 (**Table 4**). The radiomics nomogram exhibited the highest accuracy and AUC in both the test and training sets. **Figure 5** displays the ROC curves for the clinical-radiological model, the radiomics model, and the nomogram. The DeLong test indicated that the nomogram outperformed the clinical-radiological model in the training set ($P<0.05$). **Figure 6** shows the calibration curve of the integrated nomogram. The Hosmer-Lemeshow test revealed no significant difference between the actual and predicted results in both the training set ($P = 0.620$) and the test set ($P = 0.096$), indicating good model fit. Specifically, DCA demonstrated that the nomogram provided a higher net benefit than both the radiomics and clinical-radiological models across threshold probabilities ranging from 0.1 to 0.9 (**Figure 7**). This suggests that the nomogram offers superior clinical utility compared to the other models, highlighting its potential for clinical application.

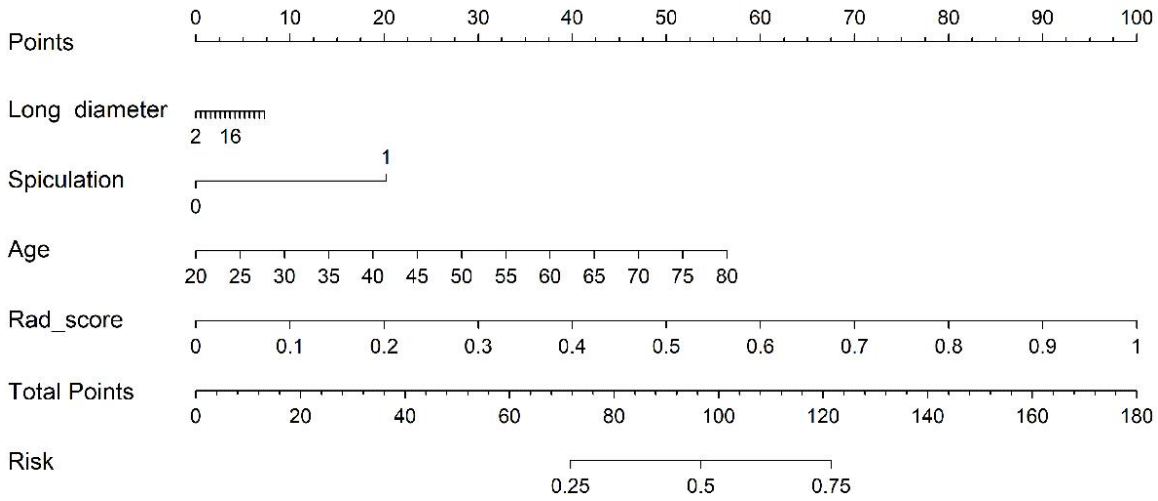

Figure 4. The integrated nomogram incorporating the age, lobulation, long diameter, and rad-score.

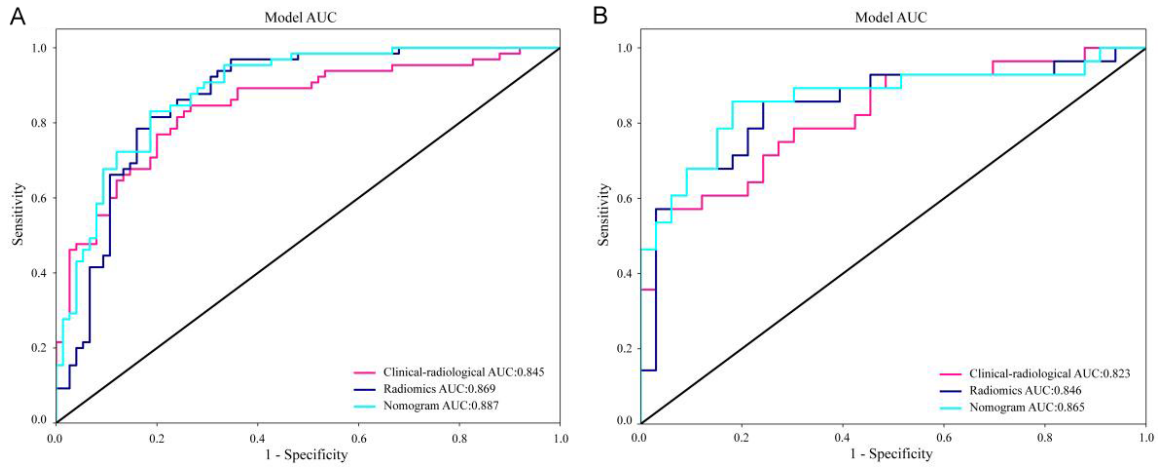

Figure 5. ROC curves of the clinical-radiological model, radiomics model, and nomogram in the training test and the test set. receiver operating characteristic (ROC). A. ROC curves for the clinical-radiological model, radiomics model, and nomogram in the training set. B. ROC curves for the clinical-radiological model, radiomics model, and nomogram in the test set.

Discussion

This study developed a nomogram model incorporating both the Rad-score and clinical-radiological features to differentiate patients with GGNs before surgery and determine whether they are AIS/MIA or IAC. The model demonstrated excellent discrimination, achieving an AUC of 0.887 in the training set and 0.865 in the test set. This model offers clinicians a valuable tool for designing tailored treatment plans, potentially improving patient outcomes.

Clinical-radiological features are commonly used to identify IAC in GGNs. Lee et al. [38] found that larger lesion size and lobulated borders are significant independent indicators of IAC. Another study [39] on pulmonary adenocarcinoma invasiveness reported that males over 60 have a significantly higher risk of IAC, and that nodule diameter significantly differs between the IAC cohort and the AAH/AIS/MIA cohort, with smaller diameters increasing the likelihood of an AAH/AIS/MIA diagnosis. In line with these findings, this study selected age, lobulation, and long diameter as predictive factors for IAC. The clinical-radiological model demonstrated an AUC of 0.845 in the training set and 0.823 in the test set, outperforming the model by Feng et al. [40], which reported

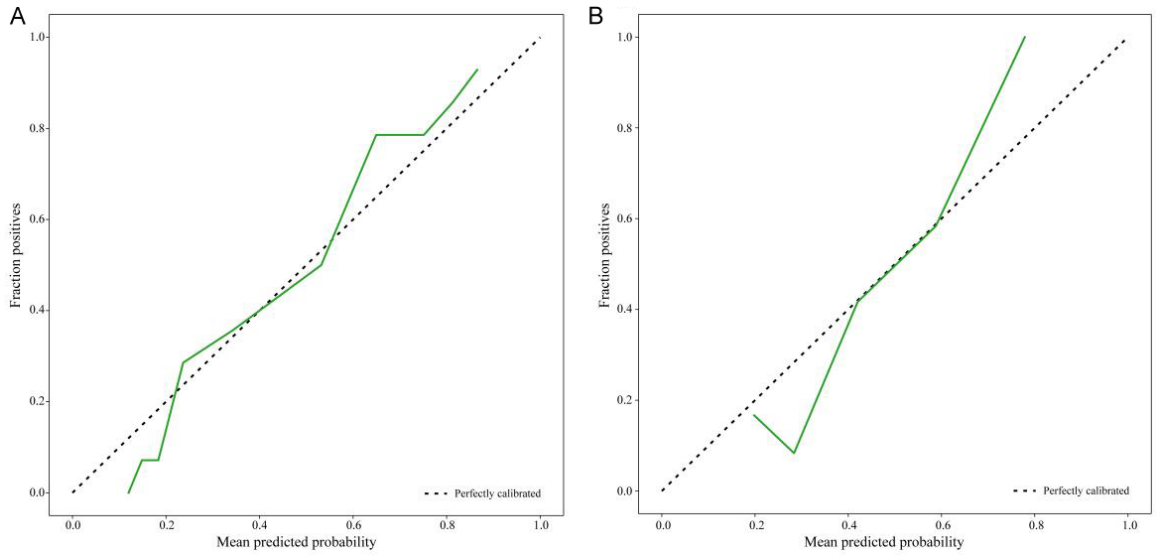

Figure 6. The calibration curve of the integrated nomogram in the training set and test set. A. Calibration curve for the integrated nomogram in the training set. B. Calibration curve for the integrated nomogram in the test set. The dashed line represents perfect calibration, and the solid line shows the actual performance of the nomogram.

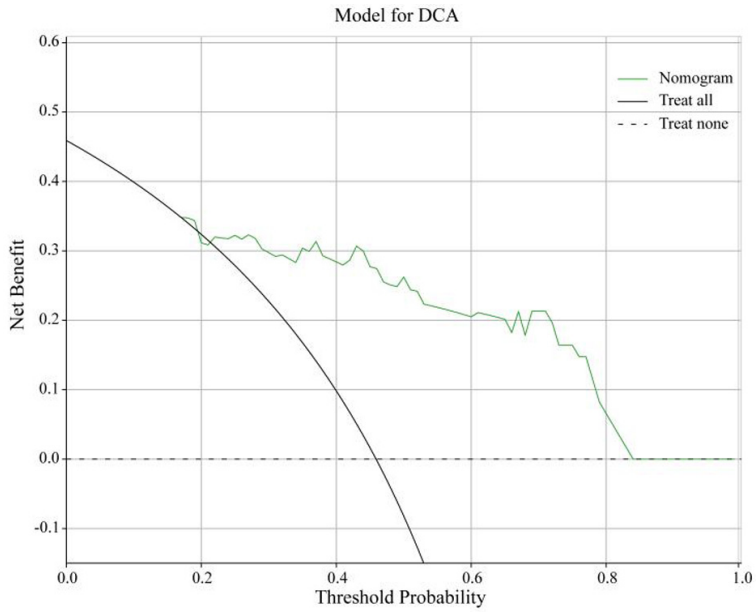

Figure 7. The decision curve analysis (DCA) of the nomogram model in the test set.

AUCs of 0.729 and 0.652, respectively. This difference may stem from our inclusion of both pure and mixed GGNs, while Feng et al. [40] focused solely on pure GGNs, which lack a solid component indicative of invasion [41-43], making them more challenging to assess based on clinician experience alone. Despite its overall efficiency, the model's sensitivity in the test set was only 0.571, which may result in IAC being misdiagnosed as AIS/MIA, potentially delaying treatment and reducing survival. Therefore, subjective methods remain challenging, as accurate evaluations require substantial experience and skill. Consequently, this study explored the use of radiomics features to develop a model with improved performance.

Radiomics is a rapidly advancing field that offers a non-invasive, cost-effective method for extracting high-throughput quantitative features from routine images. It captures intra-tumor heterogeneity, providing markers that aid clinical decision-making [44-46]. In a previous retrospective study, researchers developed a radiomics signature with five features to differentiate IAC from indolent lesions, demonstrating good efficacy (AUC = 0.89 in the validation cohort) [47]. Additionally, a model predicting invasiveness in lung adenocarcinoma based on seven features showed excellent performance (AUC = 0.924 in the external validation set)

[48]. In this study, 11 radiomics features were selected to construct the model, with texture features comprising the majority. This underscores the importance of texture features in distinguishing the invasiveness of GGNs, as they provide insights into the spatial relationships and heterogeneity within tumors [49, 50]. IAC, being more malignant than AIS/MIA, results in more complex and heterogeneous image textures. Notably, no shape features were included in the model, likely due to the morphological similarities between AIS/MIA and IAC, which complicate their differentiation. Overall, the findings of this study indicate that the radiomics model outperforms the clinical-radiological model in terms of accuracy and AUC.

To enhance the predictive performance of the model, we constructed a nomogram integrating the Rad-score and clinical-radiological features using multiple regression analysis. Combining multiple predictive factors allows for more accurate predictions of clinical outcomes, potentially accelerating the development of personalized medicine [51]. The nomogram in this study achieved the highest AUC and accuracy, demonstrating its effectiveness in distinguishing IAC from AIS/MIA. Additionally, the tool showed excellent calibration performance. We also evaluated its clinical practicality and found that, across most probability thresholds, the nomogram outperformed the treat-none or treat-all strategies. These results suggest that the nomogram could support clinicians in decision-making and help facilitate individualized, precision medicine.

However, several limitations need to be addressed. First, the retrospective design introduces the potential for selection bias. Despite rigorous methodology, retrospective studies are inherently prone to biases that may not be fully controlled. Second, the sample size in this study was limited to 473 patients. While this is adequate for initial model development and validation, a larger sample size would improve the robustness and generalizability of our findings. Future studies with larger datasets are necessary to confirm the reliability and applicability of the nomogram across diverse populations.

In conclusion, the radiomics nomogram is an accurate, non-invasive predictive tool that effectively differentiates between AIS/MIA and IAC. This capability allows clinical teams to develop more personalized treatment plans for patients.

Acknowledgements

I would like to thank my colleague Xuhong Pan for her support and help with the research. This study was supported by Yantai City Science and Technology Innovation Development Plan (No. 2024YD018).

Disclosure of conflict of interest

None.

Address correspondence to: Peng Liang, Department of Radiology, Yantaishan Hospital, Yantai 264003, Shandong, China. E-mail: liangpeng1122@126.com

References

[1] Chen X, Yao B, Li J, Liang C, Qi R and Yu J. Feasibility of using high-resolution computed tomography features for invasiveness differentiation of malignant nodules manifesting as ground-glass nodules. Can Respir J 2022; 2022: 2671772.

[2] Halpenny D, Das K, Ziv E, Plodkowski A, Zheng J, Capanu M, Rekhtman N, Montecalvo J, Solomon SB and Ginsberg MS. Percutaneous computed tomography guided biopsy of sub-solid pulmonary nodules: differentiating solid from ground glass components at the time of biopsy. Clin Imaging 2021; 69: 332-338.

[3] Han R, Wang LF, Teng F, Lin J, Xian YT, Lu Y and Wu AL. Presurgical computed tomography-guided localization of lung ground glass nodules: comparing hook-wire and indocyanine green. World J Surg Oncol 2024; 22: 51.

[4] He S, Chen C, Wang Z, Yu X, Liu S, Huang Z, Chen C, Liang Z and Chen C. The use of the mean computed-tomography value to predict the invasiveness of ground-glass nodules: a meta-analysis. Asian J Surg 2023; 46: 677-682.

[5] Koike H, Ashizawa K, Tsutsui S, Fukuda M, Okano S, Matsumoto K, Nagayasu T, Honda S and Uetani M. Surgically resected lung adenocarcinoma: do heterogeneous GGNs and part-solid nodules on thin-section CT show different prognosis? Jpn J Radiol 2023; 41: 164-171.

[6] Koratala A, Chandra NC, Balasubramanian P, Yu Lee-Mateus A, Barrios-Ruiz A, Garza-Salas A, Bowman A, Grage R, Fernandez-Bussy S and Abia-Trujillo D. Diagnostic accuracy of a com-

puted tomography-guided transthoracic needle biopsy for ground-glass opacities and subsolid pulmonary nodules. Cureus 2024; 16: e57414.
[7] Li J, Gao J, Wang G, Cui D and Deng K. A phantom study using dual-energy spectral computed tomography imaging: comparison of image quality between two adaptive statistical iterative reconstruction (ASiR, ASiR-V) algorithms for evaluating ground-glass nodules of the lung. J Cancer Res Ther 2021; 17: 1742-1747.
[8] Liu SZ, Yang SH, Ye M, Fu BJ, Lv FJ and Chu ZG. Bubble-like lucency in pulmonary ground glass nodules on computed tomography: a specific pattern of air-containing space for diagnosing neoplastic lesions. Cancer Imaging 2024; 24: 47.
[9] Niu R, Shao X, Shao X, Jiang Z, Wang J and Wang Y. Establishment and verification of a prediction model based on clinical characteristics and positron emission tomography/computed tomography (PET/CT) parameters for distinguishing malignant from benign ground-glass nodules. Quant Imaging Med Surg 2021; 11: 1710-1722.
[10] Oh J, Piao Z, Cho HJ, Chong Y, Kim SS, Kim JH and Kang MW. CT-based three-dimensional invasiveness analysis of adenocarcinoma presenting as pure ground-glass nodules. Transl Cancer Res 2023; 12: 765-773.
[11] Qiu Z, Wu Q, Wang S, Chen Z, Lin F, Zhou Y, Jin J, Xian J, Tian J and Li W. Development of a deep learning-based method to diagnose pulmonary ground-glass nodules by sequential computed tomography imaging. Thorac Cancer 2022; 13: 602-612.
[12] Sun SL, Yang ZJ, Sun HL, Huang ZG, Xu YY, Wang YL, Gao BX and Li CD. Air embolism after CT-guided localization of pulmonary ground-glass nodules. Br J Radiol 2023; 96: 20220583.
[13] Sun Y, Ma Z, Zhao W, Jin L, Gao P, Wang K, Huang X, Duan S and Li M. Computed tomography radiomics in growth prediction of pulmonary ground-glass nodules. Eur J Radiol 2023; 159: 110684.
[14] Wang C, Wu N, Zhang Z, Zhang LX and Yuan XD. Evaluation of the dual vascular supply patterns in ground-glass nodules with a dynamic volume computed tomography. World J Radiol 2022; 14: 155-164.
[15] Wu N, Liu S, Li J, Hu Z, Yan S, Duan H, Wu D, Ma Y, Li S, Wang X, Wang Y, Li X and Lu X. Deep sequencing reveals the genomic characteristics of lung adenocarcinoma presenting as ground-glass nodules (GGNs). Transl Lung Cancer Res 2021; 10: 1239-1255.
[16] Zhang H, Wang D, Li W, Tian Z, Ma L, Guo J, Wang Y, Sun X, Ma X, Ma L and Zhu L. Artificial intelligence system-based histogram analysis of computed tomography features to predict tumor invasiveness of ground-glass nodules. Quant Imaging Med Surg 2023; 13: 5783-5795.
[17] Meng F, Guo Y, Li M, Lu X, Wang S, Zhang L and Zhang H. Radiomics nomogram: a noninvasive tool for preoperative evaluation of the invasiveness of pulmonary adenocarcinomas manifesting as ground-glass nodules. Transl Oncol 2021; 14: 100936.
[18] Zheng H, Chen W, Qi W, Liu H and Zuo Z. Enhancing the prediction of the invasiveness of pulmonary adenocarcinomas presenting as pure ground-glass nodules: integrating intratumor heterogeneity score with clinical-radiological features via machine learning in a multicenter study. Digit Health 2024; 10: 20552076241289181.
[19] Zuo Z, Zeng W, Peng K, Mao Y, Wu Y, Zhou Y and Qi W. Development of a novel combined nomogram integrating deep-learning-assisted CT texture and clinical-radiological features to predict the invasiveness of clinical stage IA part-solid lung adenocarcinoma: a multicentre study. Clin Radiol 2023; 78: e698-e706.
[20] Liao RQ, Li AW, Yan HH, Lin JT, Liu SY, Wang JW, Fang JS, Liu HB, Hou YH, Song C, Yang HF, Li B, Jiang BY, Dong S, Nie Q, Zhong WZ, Wu YL and Yang XN. Deep learning-based growth prediction for sub-solid pulmonary nodules on CT images. Front Oncol 2022; 12: 1002953.
[21] Ma ZJ, Ma ZX, Sun YL, Li C, Jin L, Gao P, Li C and Li M. Prediction of subsolid pulmonary nodule growth rate using radiomics. BMC Med Imaging 2023; 23: 177.
[22] Wu FZ, Wu YJ, Chen CS and Tang EK. Prediction of interval growth of lung adenocarcinomas manifesting as persistent subsolid nodules ≤3 cm based on radiomic features. Acad Radiol 2023; 30: 2856-2869.
[23] Li Q, Song Z, Li X, Zhang D, Yu J, Li Z, Huang J, Su K, Liu Q, Zhang X and Tang Z. Development of a CT radiomics nomogram for preoperative prediction of Ki-67 index in pancreatic ductal adenocarcinoma: a two-center retrospective study. Eur Radiol 2024; 34: 2934-2943.
[24] Liu Q, Huang Y, Chen H, Liu Y, Liang R and Zeng Q. The development and validation of a radiomic nomogram for the preoperative prediction of lung adenocarcinoma. BMC Cancer 2020; 20: 533.
[25] Ren H, Xiao Z, Ling C, Wang J, Wu S, Zeng Y and Li P. Development of a novel nomogram-based model incorporating 3D radiomic signatures and lung CT radiological features for differentiating invasive adenocarcinoma from adenocarcinoma in situ and minimally invasive ade-

nocarcinoma. Quant Imaging Med Surg 2023; 13: 237-248.
[26] Shi L, Shi W, Peng X, Zhan Y, Zhou L, Wang Y, Feng M, Zhao J, Shan F and Liu L. Development and validation a nomogram incorporating CT radiomics signatures and radiological features for differentiating invasive adenocarcinoma from adenocarcinoma in situ and minimally invasive adenocarcinoma presenting as ground-glass nodules measuring 5-10 mm in diameter. Front Oncol 2021; 11: 618677.
[27] Han Y, Chai F, Wei J, Yue Y, Cheng J, Gu D, Zhang Y, Tong T, Sheng W, Hong N, Ye Y, Wang Y and Tian J. Identification of predominant histopathological growth patterns of colorectal liver metastasis by multi-habitat and multi-sequence based radiomics analysis. Front Oncol 2020; 10: 1363.
[28] Tong H, Sun J, Fang J, Zhang M, Liu H, Xia R, Zhou W, Liu K and Chen X. A machine learning model based on PET/CT radiomics and clinical characteristics predicts tumor immune profiles in non-small cell lung cancer: a retrospective multicohort study. Front Immunol 2022; 13: 859323.
[29] Xia TY, Zhou ZH, Meng XP, Zha JH, Yu Q, Wang WL, Song Y, Wang YC, Tang TY, Xu J, Zhang T, Long XY, Liang Y, Xiao WB and Ju SH. Predicting microvascular invasion in hepatocellular carcinoma using CT-based radiomics model. Radiology 2023; 307: e222729.
[30] Jiang Y, Zhou K, Sun Z, Wang H, Xie J, Zhang T, Sang S, Islam MT, Wang JY, Chen C, Yuan Q, Xi S, Li T, Xu Y, Xiong W, Wang W, Li G and Li R. Non-invasive tumor microenvironment evaluation and treatment response prediction in gastric cancer using deep learning radiomics. Cell Rep Med 2023; 4: 101146.
[31] Wang Y, Weng W, Liang R, Zhou Q, Hu H, Li M, Chen L, Chen S, Peng S, Kuang M, Xiao H and Wang W. Predicting T cell-inflamed gene expression profile in hepatocellular carcinoma based on dynamic contrast-enhanced ultrasound radiomics. J Hepatocell Carcinoma 2023; 10: 2291-2303.
[32] Fu N, Fu W, Chen H, Chai W, Qian X, Wang W, Jiang Y and Shen B. A deep-learning radiomics-based lymph node metastasis predictive model for pancreatic cancer: a diagnostic study. Int J Surg 2023; 109: 2196-2203.
[33] Cai J, Liu H, Yuan H, Wu Y, Xu Q, Lv Y, Li J, Fu J and Ye J. A radiomics study to predict invasive pulmonary adenocarcinoma appearing as pure ground-glass nodules. Clin Radiol 2021; 76: 143-151.
[34] Johnson SF, Tabatabaei SMH, Kim GHJ, Villegas BE, Brown M, Genshaft S, Suh RD, Barjaktarevic I, Wallace WD and Abtin F. Predicting invasiveness in lepidic pattern adenocarcinoma of lung: analysis of visual semantic and radiomic features. Med Sci (Basel) 2024; 12: 57.
[35] Kao TN, Hsieh MS, Chen LW, Yang CJ, Chuang CC, Chiang XH, Chen YC, Lee YH, Hsu HH, Chen CM, Lin MW and Chen JS. CT-based radiomic analysis for preoperative prediction of tumor invasiveness in lung adenocarcinoma presenting as pure ground-glass nodule. Cancers (Basel) 2022; 14: 5888.
[36] Zhao W, Xu Y, Yang Z, Sun Y, Li C, Jin L, Gao P, He W, Wang P, Shi H, Hua Y and Li M. Development and validation of a radiomics nomogram for identifying invasiveness of pulmonary adenocarcinomas appearing as subcentimeter ground-glass opacity nodules. Eur J Radiol 2019; 112: 161-168.
[37] Huang L, Lin W, Xie D, Yu Y, Cao H, Liao G, Wu S, Yao L, Wang Z, Wang M, Wang S, Wang G, Zhang D, Yao S, He Z, Cho WC, Chen D, Zhang Z, Li W, Qiao G, Chan LW and Zhou H. Development and validation of a preoperative CT-based radiomic nomogram to predict pathology invasiveness in patients with a solitary pulmonary nodule: a machine learning approach, multicenter, diagnostic study. Eur Radiol 2022; 32: 1983-1996.
[38] Lee SM, Park CM, Goo JM, Lee HJ, Wi JY and Kang CH. Invasive pulmonary adenocarcinomas versus preinvasive lesions appearing as ground-glass nodules: differentiation by using CT features. Radiology 2013; 268: 265-273.
[39] Wu G, Woodruff HC, Sanduleanu S, Refaee T, Jochems A, Leijenaar R, Gietema H, Shen J, Wang R, Xiong J, Bian J, Wu J and Lambin P. Preoperative CT-based radiomics combined with intraoperative frozen section is predictive of invasive adenocarcinoma in pulmonary nodules: a multicenter study. Eur Radiol 2020; 30: 2680-2691.
[40] Feng H, Shi G, Xu Q, Ren J, Wang L and Cai X. Radiomics-based analysis of CT imaging for the preoperative prediction of invasiveness in pure ground-glass nodule lung adenocarcinomas. Insights Imaging 2023; 14: 24.
[41] Chen W, Gong M, Zhou D, Zhang L, Kong J, Jiang F, Feng S and Yuan R. CT-based deep learning radiomics signature for the preoperative prediction of the muscle-invasive status of bladder cancer. Front Oncol 2022; 12: 1019749.
[42] Lin L, Li H, Wang X, Wang Z, Su G, Zhou J, Sun S, Ma X, Chen Y, You C and Gu Y. Components of the tumor immune microenvironment based on m-IHC correlate with prognosis and subtype of triple-negative breast cancer. Cancer Med 2023; 12: 21639-21650.
[43] Zhang T, Li X and Liu J. Prediction of the invasiveness of ground-glass nodules

in lung adenocarcinoma by radiomics analysis using high-resolution computed tomography imaging. Cancer Control 2022; 29: 10732748221089408.
[44] Jiang L, You C, Xiao Y, Wang H, Su GH, Xia BQ, Zheng RC, Zhang DD, Jiang YZ, Gu YJ and Shao ZM. Radiogenomic analysis reveals tumor heterogeneity of triple-negative breast cancer. Cell Rep Med 2022; 3: 100694.
[45] Su GH, Xiao Y, You C, Zheng RC, Zhao S, Sun SY, Zhou JY, Lin LY, Wang H, Shao ZM, Gu YJ and Jiang YZ. Radiogenomic-based multiomic analysis reveals imaging intratumor heterogeneity phenotypes and therapeutic targets. Sci Adv 2023; 9: eadf0837.
[46] Wang X, Xie T, Luo J, Zhou Z, Yu X and Guo X. Radiomics predicts the prognosis of patients with locally advanced breast cancer by reflecting the heterogeneity of tumor cells and the tumor microenvironment. Breast Cancer Res 2022; 24: 20.
[47] She Y, Zhang L, Zhu H, Dai C, Xie D, Xie H, Zhang W, Zhao L, Zou L, Fei K, Sun X and Chen C. The predictive value of CT-based radiomics in differentiating indolent from invasive lung adenocarcinoma in patients with pulmonary nodules. Eur Radiol 2018; 28: 5121-5128.
[48] Zhang T, Zhang C, Zhong Y, Sun Y, Wang H, Li H, Yang G, Zhu Q and Yuan M. A radiomics nomogram for invasiveness prediction in lung adenocarcinoma manifesting as part-solid nodules with solid components smaller than 6 mm. Front Oncol 2022; 12: 900049.
[49] Peng J, Zou D, Zhang X, Ma H, Han L and Yao B. A novel sub-regional radiomics model to predict immunotherapy response in non-small cell lung carcinoma. J Transl Med 2024; 22: 87.
[50] Zhang G, Man Q, Shang L, Zhang J, Cao Y, Li S, Qian R, Ren J, Pu H, Zhou J, Zhang Z and Kong W. Using multi-phase CT radiomics features to predict EGFR mutation status in lung adenocarcinoma patients. Acad Radiol 2024; 31: 2591-2600.
[51] Li S, Yang Z, Li Y, Zhao N, Yang Y, Zhang S, Jiang M, Wang J, Sun H and Xie Z. Preoperative prediction of vasculogenic mimicry in lung adenocarcinoma using a CT radiomics model. Clin Radiol 2024; 79: e164-e173.

Table S1. Comparison of clinical-radiological features between the training and test set

Features	Training set (n = 257)	Test set (n = 216)	P
Age (years)	56.90±10.94	55.62±11.15	0.213
Sex [n (%)]			0.427
Male	80 (31.13%)	60 (27.78%)	
Female	177 (68.87%)	156 (72.22%)	
Smoking history [n (%)]			0.162
Absent	236 (91.83%)	190 (87.96%)	
Present	21 (8.17%)	26 (12.04%)	
Location [n (%)]			0.711
Right upper lobe	72 (28.01%)	60 (27.78%)	
Right middle lobe	25 (9.73%)	23 (10.65%)	
Right lower lobe	58 (22.57%)	48 (22.22%)	
Left upper lobe	61 (23.74%)	42 (19.44%)	
Left lower lobe	41 (15.95%)	43 (19.91%)	
Lobulation [n (%)]			0.131
Absent	143 (55.64%)	135 (62.50%)	
Present	114 (44.36%)	81 (37.50%)	
Spiculation [n (%)]			0.948
Absent	210 (81.71%)	177 (81.94%)	
Present	47 (18.29%)	39 (18.06%)	
Margin [n (%)]			0.167
Clear	192 (74.71%)	149 (69.00%)	
Unclear	65 (25.29%)	67 (31.00%)	
Vascular change [n (%)]			0.144
Absent	51 (19.84%)	55 (25.46%)	
Present	206 (80.16%)	161 (74.54%)	
Air bronchus sign [n (%)]			0.101
Absent	160 (62.26%)	150 (69.44%)	
Present	97 (37.74%)	66 (30.56%)	
Bubble lucency [n (%)]			0.089
Absent	224 (87.16%)	176 (81.48%)	
Present	33 (12.84%)	40 (18.52%)	
Pleural indentation [n (%)]			0.370
Absent	132 (51.36%)	102 (47.22%)	
Present	125 (48.64%)	114 (52.78%)	
Shape [n (%)]			0.428
Round	143 (55.64%)	128 (59.26%)	
Irregular	114 (44.36%)	88 (40.74%)	
Nodule type [n (%)]			0.776
pGGN	143 (55.64%)	123 (56.94%)	
mGGN	114 (44.36%)	93 (43.06%)	
CT value (HU)	-428.49±162.10	-428.49±162.10	1.000
Short diameter (mm)	6.82±2.63	6.82±2.63	0.995
Long diameter (mm)	11.23±5.48	10.46±4.49	0.092

CT, computed tomography; pGGN, pure ground-glass nodule; mGGN, mixed ground-glass nodule; HU, Hounsfield units.

The formula of rad-score:
Rad-score = 0.464-0.043 * wavelet_HLL_ngtdm_Contrast
+0.875e-0.3 * wavelet_LHL_firstorder_10Percentile
+0.077 * log_sigma_1_0_mm_3D_glszm_SmallAreaHighGrayLevelEmphasis
+0.002 * wavelet_HLH_ngtdm_Strength
+0.079 * lbp_3D_m1_glcm_Imc1
+0.028 * exponential_glszm_SmallAreaHighGrayLevelEmphasis
+0.013 * wavelet_LLH_glcm_Imc2
-0.068 * exponential_ngtdm_Coarseness
+0.074 * lbp_3D_k_glcm_ClusterShade
+0.044 * log_sigma_2_0_mm_3D_glcm_InverseVariance
+0.006 * wavelet_HLL_firstorder_Kurtosis.