Original Article

Construction and validation of a risk score system for diagnosing invasive adenocarcinoma presenting as pulmonary pure ground-glass nodules: a multi-center cohort study in China

Qingcheng Meng[1#], Changbao Zheng[2#], Lanwei Guo[3], Pengrui Gao[1], Wentao Liu[1], Hong Ge[4], Tong Liu[1], Hui Peng[1], Jie Lu[5], Xuejun Chen[1]

[1]Department of Radiology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China; [2]Department of Radiology, The Hainan Cancer Hospital, Haikou, China; [3]Henan Office for Cancer Control and Research, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China; [4]Department of Radiotherapy, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, Zhengzhou, China; [5]Department of Radiology, The People's Hospital of Xingyang Country, Zhengzhou, China

Contributions: (I) Conception and design: Q Meng, C Zheng; (II) Administrative support: X Chen, H Ge; (III) Provision of study materials or patients: C Zheng, W Liu, T Liu; (IV) Collection and assembly of data: P Gao, H Peng, J Lu; (V) Data analysis and interpretation: L Guo, P Gao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xuejun Chen, MD. Department of Radiology, The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital, No. 127 Dongming Road, Jinshui District, Zhengzhou 450008, China. Email: zlyychenxuejun1943@zzu.edu.cn.

Background: Anxiety-driven clinical interventions have been queried due to the nondeterminacy of pure ground-glass nodules (pGGNs). Although radiomics and radiogenomics aid diagnosis, standardization and reproducibility challenges persist. We aimed to assess a risk score system for invasive adenocarcinoma in pGGNs.

Methods: In a retrospective, multi-center study, 772 pGGNs from 707 individuals in The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital were grouped into training (509 patients with 558 observations) and validation (198 patients with 214 observations) sets consecutively from January 2017 to November 2021. An additional test set of 143 observations in Hainan Cancer Hospital was analyzed in the same period. Computed tomography (CT) signs and clinical features were manually collected, and the quantitative parameters were achieved by artificial intelligence (AI). The positive cutoff score was ≥3. Risk scores system 3 combined carcinoma history, chronic obstructive pulmonary disease (COPD), maximum diameters, nodule volume, mean CT values, type II or III vascular supply signs, and other radiographic characteristics. The evaluation included the area under the curves (AUCs), accuracy, sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV) for both the risk score systems 1, 2, 3 and the AI model.

Results: The risk score system 3 [AUC, 0.840; 95% confidence interval (CI): 0.789–0.890] outperformed the AI model (AUC, 0.553; 95% CI: 0.487–0.619), risk score system 1 (AUC, 0.802; 95% CI: 0.754–0.851), and risk score system 2 (AUC, 0.816; 95% CI: 0.766–0.867), with 88.0% (0.850–0.904) accuracy, 95.6% (0.932–0.972) PPV, 0.620 (0.535–0.702) NPV, 89.6% (0.864–0.920) sensitivity, and 80.6% (0.717–0.872) specificity in the training sets. In the validation and test sets, risk score system 3 performed best with AUCs of 0.769 (0.678–0.860) and 0.801 (0.669–0.933).

Conclusions: An AI-based risk scoring system using quantitative image parameters, clinical features, and radiographic characteristics effectively predicts invasive adenocarcinoma in pulmonary pGGNs.

Keywords: Pure ground-glass nodule (pGGN); lung cancer; artificial intelligence (AI); X-ray; computed tomography (CT)

Submitted Jan 27, 2024. Accepted for publication May 29, 2024. Published online Jun 24, 2024.
doi: 10.21037/qims-24-170
View this article at: https://dx.doi.org/10.21037/qims-24-170

Introduction

The increasing detection rate of pulmonary ground-glass nodules (GGNs) has attracted unprecedented attention in clinical practice owing to the widespread use of low-dose computed tomography (CT) for screening programs of lung cancer or physical examinations (1). GGNs are defined as nodules with a slightly increased density without obscure underlying bronchial structures or vascular margins on high-resolution CT (2). GGNs can be classified as pure GGNs (pGGN) without any solid components under the mediastinal window [window width: 300–500 Hounsfield units (HU); window level: 30–50 HU] of CT images and part-solid nodules in benign or malignant conditions (3). Most international guidelines adopt conservative approaches to pGGN treatment, owing to the stable or inert growth of pGGNs on CT follow-up and favorable prognosis (4-7). However, more than 40–55% of pGGNs are found to be invasive or metastatic lesions during follow-up and about 20% of lung adenocarcinomas manifest as pGGNs on CT imaging (4,8-10), all of which may be explained by the imbalance of tumor cells with the rapid growth period. Therefore, as far as the pGGNs with or without invasive components are concerned, a similar strategy with CT follow-up or intervention is inappropriate.

Previous studies have suggested that the invasive lesions among pGGNs are related to certain CT imaging signs, such as the maximum nodule diameter, large volume, high voxel attenuation, vessel changes, history of extrapulmonary cancers, or chronic obstructive pulmonary disease (COPD) (10-15). However, using CT imaging or epidemiological characteristics for diagnosis is challenging mainly due to the complex mechanism underlying lung cancer, the overlap of features of invasive and non-invasive lesions, or observer experience. Recently, several studies about artificial intelligence (AI) with deep learning (DL) technology or radiomic features have reported that DL facilitates the detection, diagnosis, or follow-up of pGGNs with satisfactory results, which may be the best approach to reduce the bias of artificial measurement (16-18). However, the performance of previous studies using DL to differentiate between invasive and non-invasive lesions needs further improvement (19-24). Furthermore, Meng *et al.* (25) established a radiomics nomogram with good predictive ability with an area under the curve (AUC) of 0.940 and 0.946 in the training and validation sets, respectively, for evaluating the invasiveness of pulmonary adenocarcinomas manifesting as GGNs. However, the results are strongly limited by the lack of standardized acquisition parameters, inconsistent methods, and lack of reproducibility (26). Therefore, a simple, practical, and effective risk system for diagnosing the invasiveness of adenocarcinoma in pGGNs is needed. This study aimed to establish a novel risk-scoring system for diagnosing invasive adenocarcinoma in pGGNs based on traditional clinical features, CT signs, and AI-based quantitative image features and to provide a basis for rational clinical decision-making. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-24-170/rc).

Methods

Study design and participants

This retrospective, multi-center cohort study was conducted in accordance with the Declaration of Helsinki (as revised in 2013), and the protocol was approved by The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital Ethics Committee (No. 2021-KY-0022). The requirement for individual consent for this retrospective analysis was waived. Medical records, including epidemiological characteristics, histopathological results, clinical characteristics, and serial chest CT scans of 2,410 patients with ≥1 pathologically GGNs were reviewed retrospectively and consecutively in the Henan Cancer Hospital from January 2017 to November 2021. We included patients (I) with pGGNs without solid components on CT imaging; (II) with stable or enlarged GGNs after ≥2 years of follow-up, no biopsy, or anti-tumor

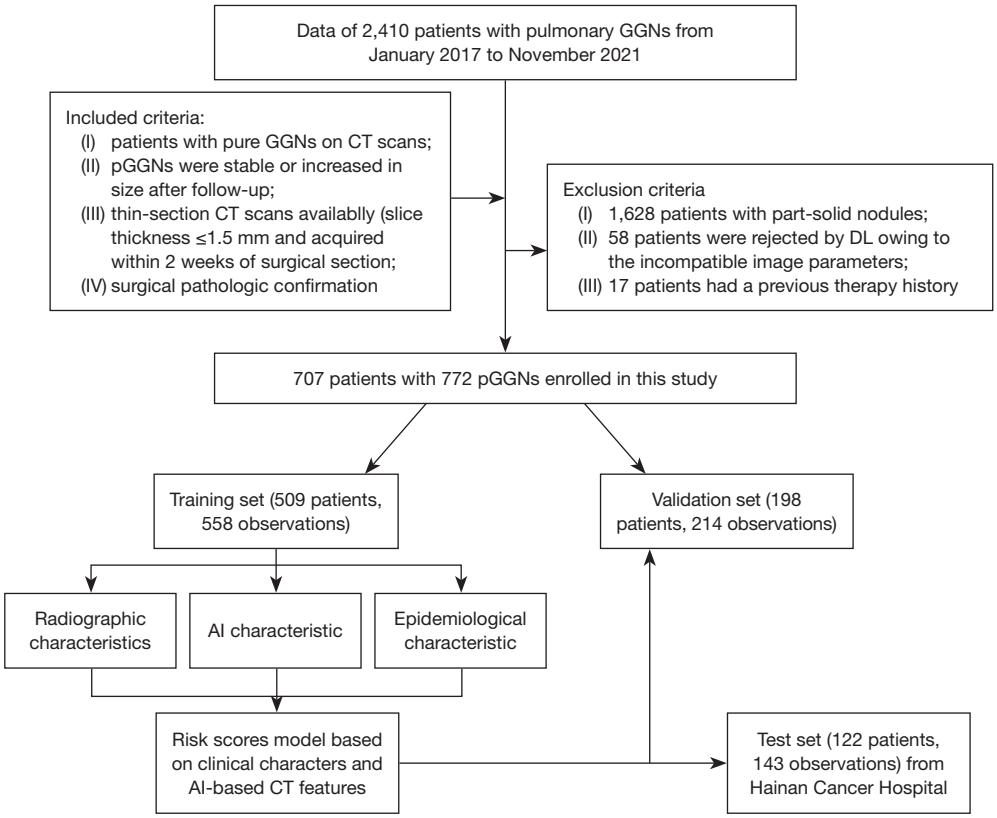

Figure 1 Flowchart of the patient selection procedure. GGN, ground-glass nodule; CT, computed tomography; pGGN, pure ground-glass nodule; DL, deep learning; AI, artificial intelligence.

treatment before CT imaging; (III) with thin-section CT scans (slice thickness ≤1.5 mm) acquired sequentially in a single examination within 2 weeks of pre-surgical resection; and (IV) pathologic confirmation by surgery. We excluded patients (I) without pathological diagnoses or follow-up; (II) with chest CT rejection by DL owing to incompatible scanning parameters (poor image quality or slice thickness of CT >1.5 mm); (III) with infective diseases of the lung or other organs. Finally, 707 patients (772 pGGNs) were included in this study, in which 506 patients with 561 GGNs from our prior study (10) were included and randomly assigned to training (509 patients, 558 observations) and validation (198 patients, 214 observations) sets. Data from 122 patients with 143 GGNs used as the test set were also operated and collected in The Hainan Cancer Hospital (*Figure 1*).

CT scanning

Chest CT of the apex to the base of the lung with the patient in a supine position was performed using 1 of 3 CT systems (GE Revolution, GE Healthcare, Chicago, IL, USA; Philips iCT256 or Icon, Philips Healthcare, Eindhoven, Netherlands). The scan parameters were as follows: tube voltage, 120 kV; tube current, 50–200 mA; rotation time, 0.5–1.0 s; pitch, 1.0–1.5; conventional layer thickness, 5.0 mm; reconstruction layer thickness, 1.25 mm. A nonionic contrast agent was used for the multiphase-enhanced scanning process.

AI algorithm and quantitative parameters

Commercially available AI software (approved by the China Food and Drug Administration), based on multi-stage 3-dimensional (3D) deep convolutional neural network (DCNN) algorithms (10) was used to process the arterial phase of enhanced CT with 1.25-mm datasets. The DL framework was based on ResNet101 combined with false positive reduction as the backbone. The fully convolutional one-stage object detection head and focal loss were used as the model to detect nodule boxes. Voxel segmentation

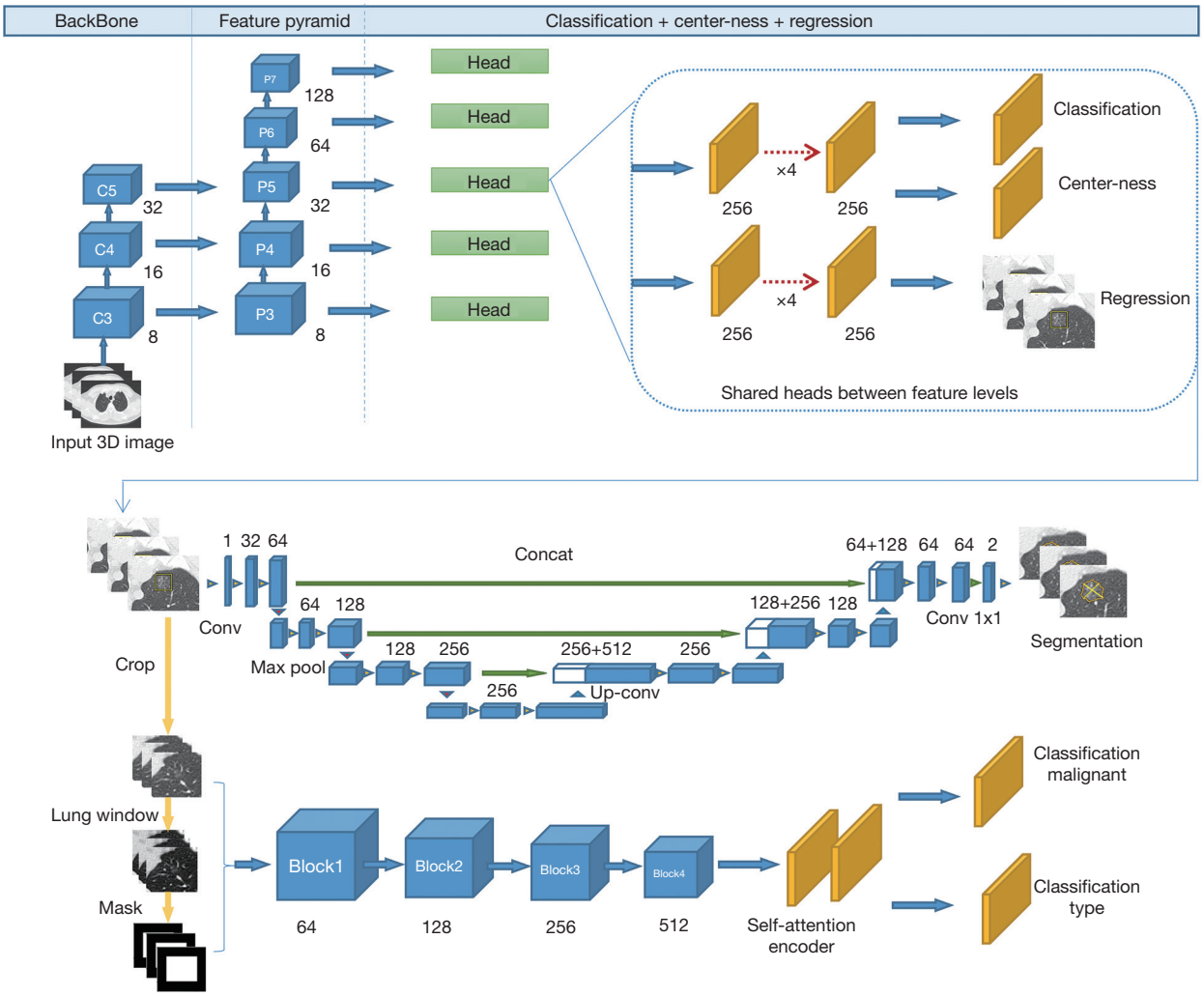

Figure 2 Frame structure of AI based on 3D-DCNN algorithms. 3D, 3-dimensional; Conv, convolution; Max, maximum; AI, artificial intelligence; 3D-DCNN, 3-dimensional deep convolutional neural network.

was performed using a 3D U-Net model with small images segmented by the nodule box. After windowing and mask processing, we used ResNet18 and self-attention as classification models (*Figure 2*). AI was used to detect nodules with boundary contours and venous phase CT values (in HU), minimum and maximum dimensions, and volume. A risk stratification model with low, medium, and high risk was scored for GGN malignancy by using the commercial DL approach, and the risk scores of positive nodules were those lesions with medium or high malignancy risk stratification of DL.

pGGN risk scoring system with AI-based quantitative parameters, clinical features, and radiographic characteristics

pGGN radiographic characteristics on lung windows (width, 1,500 HU; level, −500 HU) of venous phase-enhanced CT images were assessed independently by 3 experienced thoracic radiologists (7–17 years of chest CT interpretation experience; blinded to the pathological results); interpretation discrepancies were resolved through consensus. Radiographic characteristics included spiculation, bubble-like lucency, pleural indentation, and

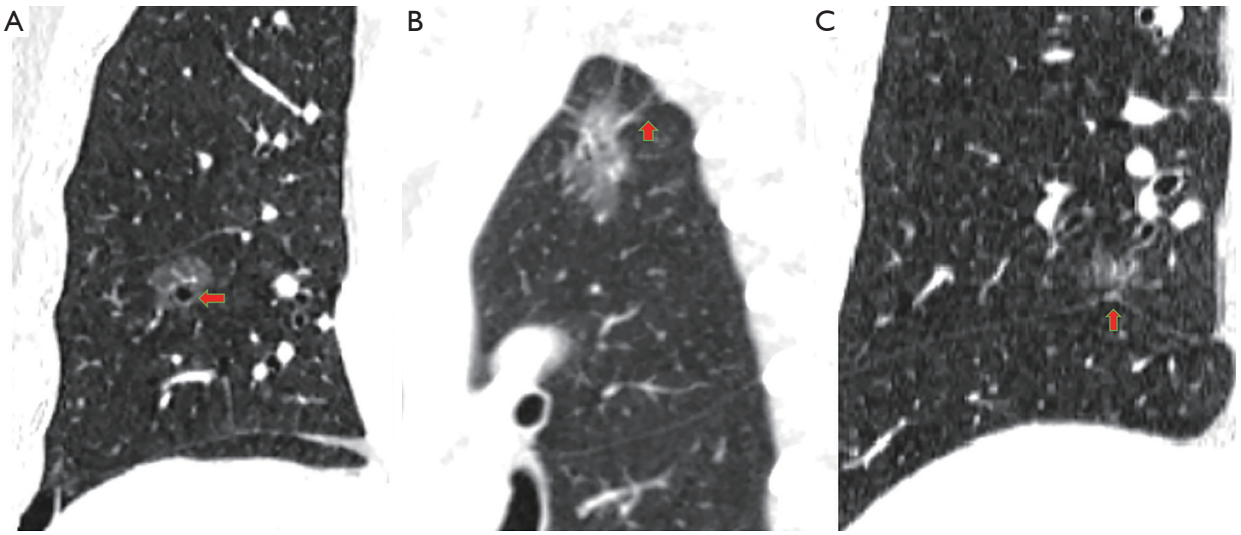

Figure 3 CT signs of pGGN. (A) Bubble-like appearance in the right upper lobe of the lung (red arrow). This appearance is characterized by air-attenuated, vesicle-like lucency within the pGGN. (B) Spiculation in the left upper lobe of the lung (red arrow). Spiculation is defined as the presence of strands extending from the margin of the pGGN into the lung parenchyma without reaching the pleural surface. (C) Pleural indentation along the horizontal fissure (red arrow). Pleural indentation is characterized by linear areas of high attenuation originating from the pGGN, extending peripherally to contact the horizontal fissure. CT, computed tomography; pGGN, pure ground-glass nodule.

the vascular supply sign. The spiculation was defined as the presence of the margin of the nodule stranding extending into the lung parenchyma without reaching the pleural surface. The bubble-like appearance was defined as vesicle-like lucency, air-attenuated within the nodule. Pleural indentation was defined as the high-attenuation linear areas extending peripherally to contact the pleura, which originated from the nodule (*Figure 3*). The vascular supply sign was grouped into 3 types (Type I, normal traveling path or size vessels passed through the lesions; Type II, dilated/distorted vessels within lesions; Type III, more complicated vasculature than that of Types I and II (*Figure 4*). According to the International Association for the Study of Lung Cancer (IASLC)/American Thoracic Society (ATS)/European Respiratory Society (ERS) (3) and the fifth edition of the classification of thoracic tumors issued by World Health Organization (WHO) in May 2021 (27), atypical adenomatous hyperplasia and adenocarcinoma in situ were categorized as adenomatous precursor lesions and all pGGNs were divided into non-invasive (benign, atypical adenomatous hyperplasia, and adenocarcinoma in situ) and invasive (minimal invasive adenocarcinoma and invasive adenocarcinoma) groups.

Statistical analysis

Statistical analysis was performed using SPSS 19.0 (IBM Corp., Armonk, NY, USA) and MedCalc 15.2.2 (MedCalc Software, Ostend, Belgium). Inter-reader reliability of the types of vascular supply sign, pleural attachment, spiculation, or bubble lucency among the 3 readers was calculated using the intraclass correlation coefficient (ICC). For continuous predictors of pleural attachment, bubble lucency, spiculation, maximum diameter, volume, CT value, and vascular supply sign for pGGN differentiation were determined by optimal cutoff values of receiver operating characteristic (ROC) curves. Pearson's correlation analysis was conducted between the type of vascular supply sign and invasive adenocarcinoma of the lung. Performances of different risk scoring systems were assessed using AUCs and accuracy ($ACC = \frac{TP+TN}{TP+FP+TN+FN}$), sensitivity ($Sensitivity = \frac{TP}{TP+FN}$), specificity ($Specificity = \frac{TN}{TN+FP}$), positive predictive value ($PPV = \frac{TP}{TP+FP}$), and negative predictive value ($NPV = \frac{TN}{TN+FN}$), (where ACC is accuracy, TP is true positive, FP is false positive, TN is true negative, and FN is false negative). We considered P values with 2-sided <0.05 statistically significant.

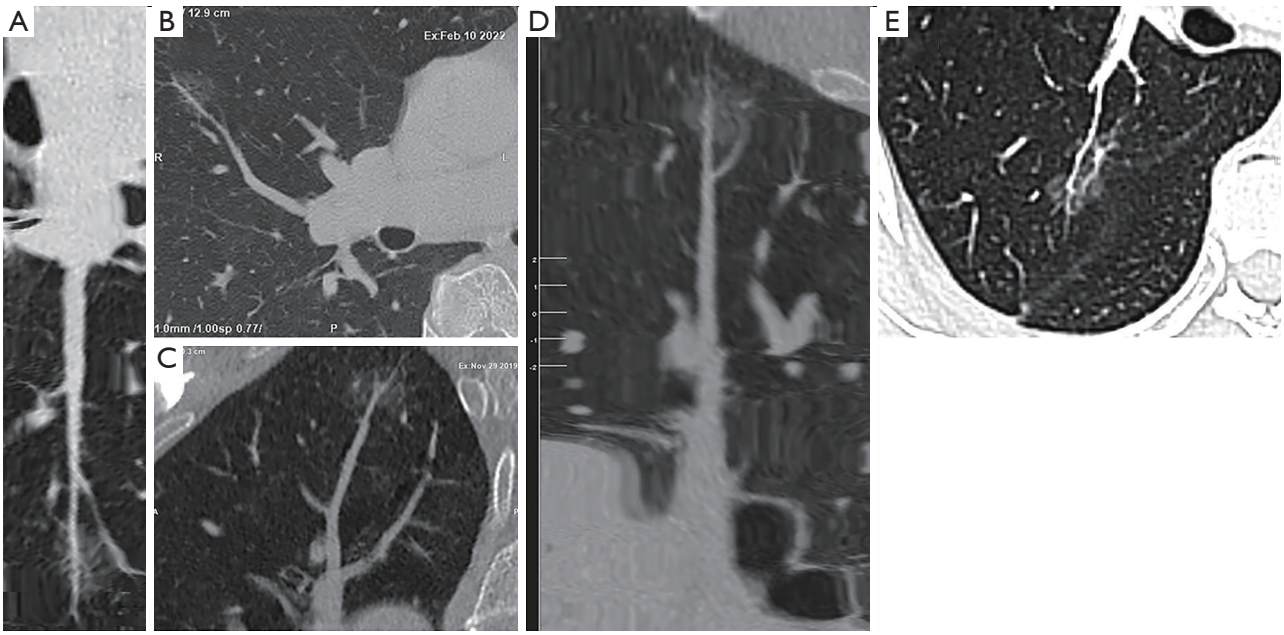

Figure 4 Types of vascular supply sign in pulmonary lesions. (A,B) Type I: pass-through, vessels passed through lesions with normal traveling path or size. (C,D) Type II: distorted/dilated vessels within lesions. (E) Type III: complicated, more complicated vasculature than type I and II.

Results

Dataset characteristics

Table 1 shows the baseline data of the patients. There were 13, 59, and 757 patients with 3, 2, and 1 nodule(s), respectively. Eventually among all the lesions in our study, 155 (16.94%) had non-invasive lesions [including 36 (3.93%) of atypical adenomatous hyperplasia, 66 (7.21%) of adenocarcinoma in situ, and 53 (5.79) of other benign lesions such as interstitial fibrosis, inflammation, and so on], and 760 (83.06%) had lung invasive adenocarcinomas (including minimally invasive and invasive adenocarcinomas). Some 58 participants had a family history of carcinoma, and 86 had COPD. There were 72 observations of Type I vascular supply signs, 145 of Type II, and 639 of Type III. Family history of carcinoma, COPD, bubble lucency, or spiculation on CT images, and volume significantly differed between the training, validation, and test sets ($P<0.05$ in all cases), and the sex, age, distribution of nodules in patients, pleural attachment, mean CT value, lung disease spectrum, long diameter, and types of vascular supply sign showed no statistical differences among the 3 sets ($P>0.05$ in all variables).

Radiographic characteristics based on CT scan analysis of pGGNs

To evaluate the inter-reader reliability of the types of vascular supply signs, 150 pGGNs were selected randomly and pleural attachment, spiculation, and bubble lucency were assessed independently by 3 experienced thoracic radiologists, and their ICCs were analyzed. The interobserver agreement for the measurements is summarized in *Table 2*. The ICCs for pleural attachment, bubble lucency, and types of vascular supply signs were 0.915, 0.807, and 0.807, respectively; however, the ICC for spiculation was 0.596. The performance of pleural attachment, spiculation, or bubble lucency in diagnosing pGGN lung cancer was evaluated using AUCs, and they achieved AUCs of 0.615 [95% confidence interval (CI): 0.560–0.670], 0.612 (95% CI: 0.558–0.667), and 0.563 (95% CI: 0.504–0.622), respectively. Under the cutoff value for Type II, the vascular supply sign yielded the highest AUC of 0.744 (95% CI: 0.681–0.808) among the radiographic characteristics, as presented in *Table 3*. A positive correlation was observed between the type of vascular supply sign and pulmonary invasive adenocarcinoma ($r=0.442$, P value <0.001).

Table 1 Comparison of clinical characteristics of patients between training and validation sets

Characteristics	Training set	Validation set	Test set	P value
Sex				0.241
Male	191	63	34	
Female	367	151	88	
Age (years)	56.09±10.07	56.31±9.05	57.11±10.32	0.550
Family history of carcinoma				<0.001
Yes	27	8	23	
No	531	206	120	
Chronic obstructive pulmonary disease				<0.001
Yes	46	12	28	
No	512	202	115	
Distribution of nodules in patients				0.315
One	467	184	106	
Two	35	12	12	
Three	7	2	4	
Pleural indentation				0.749
Yes	141	51	32	
No	417	163	111	
Bubble-like lucency				0.012
Yes	109	41	13	
No	449	173	130	
Spiculation				<0.001
Yes	126	49	7	
No	432	165	136	
Mean CT value (HU)	–(450.3±156.2)	–(463.3±154.4)	–(467.9±133.3)	0.062
Volume (mm^3)	1,608±2,243	1,761±2,976	1,035±1,728	0.012
Maximum diameter (mm)	13.55±6.06	13.84±6.53	13.11±6.39	0.555
Vascular supply sign				0.437
Yes				
Type I	44	21	7	
Type II	81	39	25	
Type III	399	138	102	
No	34	16	9	
Disease spectrum of lung				0.320
Non-invasive lesions				
AAH	25	7	4	
AIS	42	20	4	
Benign lesions	31	13	9	
Invasive adenocarcinoma	460	174	126	

Data are represented as means ± standard deviations or n. Type I, vessels passed through lesions with normal traveling path or size; Type II, distorted/dilated vessels within lesions; Type III, more complicated vasculature than Types I and II. CT, computed tomography; HU, Hounsfield unit; AAH, atypical adenomatous hyperplasia; AIS, adenocarcinoma in situ.

Table 2 Radiographic characteristics for intra- and inter-readers reliability assessment

Variables	Numbers	Reader-1	Reader-2	Reader-3	ICC	P value
Radiographic characteristics						
Pleural indentation	45	47	46	44	0.915	<0.001
Bubble lucency	16	17	17	18	0.807	<0.001
Spiculation	35	39	36	35	0.596	<0.001
Types of vascular supply sign					0.807	<0.001
Type I	13	11	14	13		
Type II	17	18	18	19		
Type III	111	108	111	109		

Type I, vessels passed through lesions with normal traveling path or size; Type II, distorted/dilated vessels within lesions; Type III, more complicated vasculature than Types I and II. ICC, intraclass correlation coefficient.

Table 3 The diagnostic value of AI-based quantitative image features and CT features for malignant pGGNs in the training set

Test result variable(s)	AUC	Standard error	Cutoff value	Asymptotic 95% confidence interval	
				Lower bound	Upper bound
Maximum diameter (mm)	0.789	0.025	12.45	0.740	0.839
Volume (mm^3)	0.805	0.024	441.25	0.758	0.853
Mean CT value (HU)	0.653	0.032	–487.5	0.591	0.715
Bubble lucency	0.563	0.030	–	0.504	0.622
Vascular supply sign	0.744	0.032	Type II	0.681	0.808
Spiculation	0.612	0.028	–	0.558	0.667
Pleural indentation	0.615	0.028	–	0.560	0.670

Type II, distorted/dilated vessels within lesions. AI, artificial intelligence; CT, computed tomography; pGGNs, pure ground-glass nodules; AUC, area under the curve; HU, Hounsfield unit.

AI-based quantitative parameters and analysis of CT signs of pGGNs

All the pGGNs were labeled and stratified automatically by the AI platform based on the DL algorithm. Similarly, quantitative parameters, including the 2-dimensional (2D) maximum diameter, 3D volume, and mean CT value, were calculated. The maximum diameter of the pGGNs achieved an AUC of 0.789 (95% CI: 0.740–0.839) with a cutoff value of 12.45 mm; volume yielded an AUC of 0.805 (95% CI: 0.758–0.853) with a cutoff value of 441.25 mm^3; and the AUC of the mean CT value was 0.653 (95% CI: 0.591–0.715) according to the cutoff value of –487.5 HU. Their performances were better than those of radiographic characteristics, as shown in *Table 3*.

Performance of AI-based risk scoring systems with different combinations in discriminating invasive adenocarcinoma presenting as pGGNs and ROC curve analysis

The radiographic characteristics of the bubble-like lucency with the lowest AUCs of 0.563 and spiculation with the lowest ICC of 0.596 were excluded from the novel risk score model. To individualize precision therapy for pGGN and enhance its applicability for clinicians in China, we adopted a simple grading system in our risk score system, instead of using multi-factor risk models. Therefore, based on high-risk factors for lung cancer in epidemiology (1,3,11,15), as well as the results of the analysis of radiographic characteristics or AI-based quantitative parameters [including maximum diameter (≥12.45 mm),

Table 4 Chart of risk scores with variable parameters

Variable parameters	Yes/no	Scores
Epidemiological characteristics		
Family history of carcinoma	Yes/no	1/0
Chronic obstructive pulmonary disease	Yes/no	1/0
Quantitative image features by AI		
Maximum diameter (≥12.45 mm)	Yes/no	1/0
Volume (≥441.25 mm^3)	Yes/no	1/0
Mean CT value (≥−487.5 HU)	Yes/no	1/0
Radiographic characteristics		
Pleural indentation	Yes/no	1/0
Vascular supply sign		
Type II	Yes/no	1/0
Type III	Yes/no	2/0
Total scores		0–9

Type II, distorted/dilated vessels within lesions; Type III, more complicated vasculature than Types I and II; Type I, vessels passed through lesions with normal traveling path or size. AI, artificial intelligence; CT, computed tomography; HU, Hounsfield unit.

volume (≥441.25 mm), mean CT value (≥−487.5 HU)], the variables, including a history of carcinoma, COPD, long diameters, volume of lesions, mean CT values, and other radiographic characteristics (spiculation, bubble-like lucency, and pleural indentation), were each assigned 1 score in the risk score system for pGGN. However, type III vascular supply sign was assigned 2 scores due to its positive correlation with pulmonary invasive adenocarcinoma, as presented in *Table 3*. To validate the diagnostic value of long diameters and volumes of pGGN, all variables were grouped into 3 models: Risk scoring system 1 = history of carcinoma × 1 + COPD × 1 + maximum diameters (≥ cutoff value) × 1 + volume of lesions (≥ cutoff value) + mean CT values (≥ cutoff value) × 1 + type II vascular supply sign × 1 or type III × 2 + other variables of radiographic characteristics × 1; Risk scoring system 2 = history of carcinoma × 1 + COPD × 1 + volume of lesions (≥ cutoff value) × 1 + mean CT values (≥ cutoff value) × 1 + type II vascular supply sign × 1 or type III × 2 + other variables of radiographic characteristics × 1; and Risk scoring system 3 = history of carcinoma × 1 + COPD × 1 + maximum diameters (≥ cutoff value) × 1 + volume of nodule (≥ cutoff value) + mean CT values (≥ cutoff value) × 1 + type II vascular supply sign × 1 or type III × 2 + other variables of radiographic characteristics × 1, as shown in *Table 4*. The total risk scores for different risk models ranged from 0 to 9, and the risk score cutoff for diagnosing lung cancer presenting as pGGNs was set at ≥3.

Compared with the AI model, risk score system 1 and risk score system 2, risk score system 3 achieved the best performance with an AUC of 0.840 (0.789–0.890); 88.0% (0.850–0.904) accuracy; 95.6% (0.932–0.972) PPV; 62.0% (0.535–0.702) NPV; 89.6% (0.864–0.920) sensitivity; and 80.6% (0.717–0.872) specificity in the training and with an AUC of 0.769 (0.678–0.860); 83.2% (0.776–0.876) accuracy; 92.6% (0.875–0.957) PPV; 53.9% (0.405–0.667) NPV; 89.7% (0.842–0.936) sensitivity; and 72.5% (0.572–0.839) specificity in the validation sets. Furthermore, risk score system 3 yielded the best performance with an AUC of 0.801 (0.669–0.933) in the test set. However, of all the risk scoring systems, the AI model achieved the highest sensitivity of 94.1%, 94.3%, and 96.8% in the training, validation, and test sets, respectively, and the lowest specificity of 16.3%, 10.0%, and 17.7% in the training, validation, and test sets, respectively, as shown in *Table 5* and *Figure 5*.

Discussion

The results showed that risk score system 3 (including pleural attachment, vascular supply sign type II or III, maximum diameter, volume, and mean CT value; and family history of carcinoma and COPD) was superior to the AI model as well as risk score systems 1 and 2 in differentiating invasive adenocarcinoma that appeared as pGGNs, which can save the number of scans in the lung cancer screening, reduce the time to diagnosis for earlier resection, and therefore decrease the anxiety of patients.

Several studies have found that the long diameter, volume, and average CT value could be independent predictors of pGGN invasiveness (25,28,29). In this study, AI based on DL outperformed manual observation, required fewer labor hours, and decreased measurement error (30). The maximum diameter, volume, and mean CT value with cutoff values of 12.45 mm, 441.25 mm^3, and −487.5, respectively, performed best in diagnosing lung cancer that appeared as pGGNs with AUCs of 0.789, 0.805, and 0.653, respectively. The mean CT cutoff value in our study was similar to that of the study by Ikeda *et al.* (31) and higher than that of the prior study (29,32), which might be due to differences in the measurement techniques accompanying arterial phase-enhanced CT images, grouping, or other factors.

Table 5 Comparison of diagnostic values for benign and malignant pGGNs between AI and scores system-1, -2, and -3

Index	Training set				Validation set				Test set			
	AI	Scores system 1	Scores system 2	Scores system 3	AI	Scores system 1	Scores system 2	Scores system 3	AI	Scores system 1	Scores system 2	Scores system 3
TP	433	390	386	412	164	120	138	150	122	96	100	113
FP	82	16	20	19	36	12	11	12	14	5	5	5
FN	27	70	74	48	10	54	36	24	4	30	26	13
TN	16	82	78	79	4	28	29	28	3	12	12	12
Sensitivity, % (95% CI)	0.941 (0.916–0.959)	0.848 (0.812–0.878)	0.840 (0.803–0.870)	0.896 (0.864–0.920)	0.943 (0.897–0.969)	0.695 (0.623–0.759)	0.845 (0.784–0.891)	0.897 (0.842–0.936)	0.968 (0.921–0.988)	0.762 (0.681–0.828)	0.794 (0.715–0.855)	0.897 (0.831–0.939)
Specificity, % (95% CI)	0.163 (0.103–0.249)	0.837 (0.751–0.897)	0.796 (0.706–0.864)	0.806 (0.717–0.872)	0.1 (0.040–0.231)	0.775 (0.625–0.877)	0.7 (0.546–0.819)	0.725 (0.572–0.839)	0.177 (0.062–0.410)	0.706 (0.469–0.867)	0.706 (0.469–0.867)	0.706 (0.469–0.867)
PPV, % (95% CI)	0.841 (0.807–0.870)	0.961 (0.937–0.976)	0.951 (0.925–0.968)	0.956 (0.932–0.972)	0.82 (0.761–0.867)	0.909 (0.848–0.947)	0.926 (0.873–0.958)	0.926 (0.875–0.957)	0.897 (0.835–0.938)	0.951 (0.889–0.979)	0.952 (0.893–0.980)	0.958 (0.905–0.982)
NPV, % (95% CI)	0.372 (0.244–0.521)	0.540 (0.460–0.617)	0.513 (0.434–0.591)	0.620 (0.535–0.702)	0.286 (0.117–0.547)	0.244 (0.164–0.347)	0.446 (0.332–0.567)	0.539 (0.405–0.667)	0.429 (0.158–0.750)	0.286 (0.172–0.436)	0.316 (0.191–0.475)	0.480 (0.300–0.665)
Accuracy, % (95% CI)	0.805 (0.770–0.836)	0.846 (0.814–0.874)	0.832 (0.798–0.860)	0.880 (0.850–0.904)	0.785 (0.725–0.835)	0.692 (0.627–0.750)	0.780 (0.720–0.831)	0.832 (0.776–0.876)	0.874 (0.810–0.919)	0.755 (0.679–0.818)	0.783 (0.709–0.843)	0.874 (0.810–0.919)
AUC (95% CI)	0.553 (0.487–0.619)	0.802 (0.754–0.851)	0.816 (0.766–0.867)	0.840 (0.789–0.890)	0.521 (0.42–0.623)	0.695 (0.604–0.786)	0.759 (0.671–0.847)	0.769 (0.678–0.860)	0.556 (0.401–0.712)	0.730 (0.597–0.863)	0.746 (0.613–0.813)	0.801 (0.669–0.933)

pGGNs, pure ground-glass nodules; AI, artificial intelligence; TP, true-positive; FP, false-positive; FN, false-negative; TN, true-negative; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve.

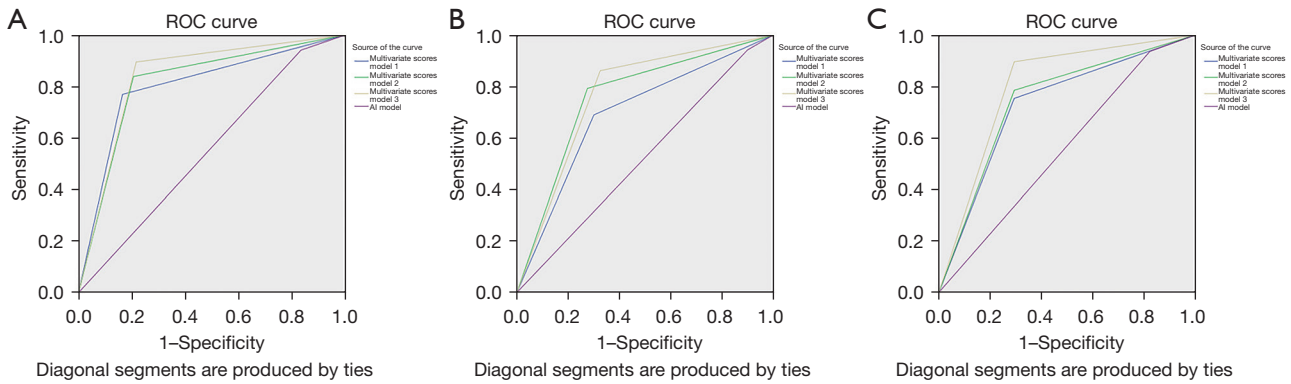

Figure 5 Diagnostic value of the AI model and the risk scores systems. The ROC curve corresponding to risk scores system 3 showed the highest performance among the four risk models for determining cases in which invasive adenocarcinoma appeared as pGGNs in the training set (A), validation set (B) and test set (C). ROC, receiver operating characteristic; AI, artificial intelligence; pGGNs, pure ground-glass nodules.

Due to the differentiation of tumor margin cells, different growth rates or contraction of fibrous tissues are present within tumors, and the lesions including benign and malignant may grow with different patterns, resulting in irregular shapes, mixed ground-glass opacity, lobulation, higher frequency of pleural indentation, air bronchograms, spiculation, and abnormal dilation and distortion of the blood vessels. Therefore, it is worth questioning the role of CT signs in diagnosing malignant pGGNs, which might be attributable to the natural processes of pGGNs (12-14,29,31). Regarding the CT features, the vascular supply sign type has a positive relationship with pGGN malignancy (10). The AUC of the vascular supply sign type II was 0.744 in the training set; therefore, the type II and type III vascular supply signs were assigned scores of 1 and 2, respectively. The AUC of pleural attachment (0.615) was higher than that of bubble lucency (0.563) and spiculation (0.612). Notably, the interobserver correlation of vascular supply sign was lower than that of the other CT signs, which might due to the differentiation between type I and II vascular supply signs and that vascular post-processing technology was needed to improve it. Above all, the CT features of the type of vascular supply sign and pleural attachment were selected as variables of the risk score model. As far as the ICC among the different CT features was concerned, the spiculation with the lowest ICC of 0.596 in this study might be due to the lowest incidence rate of the malignant pGGNs and further study should be conducted for accurate cognition among different radiologists.

To improve individual stratification management of pGGNs in clinical practice, a total of 772 patients with adenocarcinomas appearing as pGGNs were investigated, which was more than in most recent studies, and the novel risk scoring system in our study was also designed based on pleural attachment, type II or III vascular supply signs, long diameter ≥12.45 mm, volume ≥441.25 mm^3, mean CT values ≥–487.5 HU, a family history of carcinoma, and COPD according to the risk degree of malignant pGGNs to validate the performance between the long diameter and volume of pGGNs. Risk score system 3, including the variables of long diameter and volume, had the best performance in the training (AUC =0.840), validation (AUC =0.769), and test (AUC =0.801) sets, which confirmed a synergistic effect of the long diameter and the volume owing to the different growth rates in different directions in the tumor cells.

This study has some limitations. First, as the study was retrospective, some selection biases, for example, the smoking status of patients was not recorded in this study due to insufficient data, although it is currently recognized as the most important risk factor for lung cancer. Second, the pGGNs number of benign lesions was small, which might be biased for the distribution of pathological subtypes of pGGNs. This was because benign lesions are mostly followed-up due to lack of vascular supply signs and they are rarely excised via surgery. We will continue to collect data on benign lesions with pathological confirmation in the future. Third, non-contrast CT scans are often used for lung cancer screening or follow-up, but only the enhanced CT scans were used for this retrospective study, and therefore the differentiation in diagnosing the types of vascular supply sign between the non-contrast

CT scan and the enhanced CT scan will be validated in future prospective research. Additionally, the improved AI algorithm or large language models in medical image processing may enhance the diagnosis of the invasiveness of pGGNs in clinical practice (33,34), and iterative upgrading of the algorithm is needed in the future.

In conclusion, various quantitative and qualitative variables showed different degrees of statistical difference in discriminating the malignant pGGNs. The predictive value of the risk score system based on AI parameters, clinical features, and radiographic characteristics of pGGNs was remarkably higher than the independent use of each variable; therefore, effectively improving the diagnostic accuracy of pGGNs in patients with lung adenocarcinoma. In future, a prospective clinical trial with more pGGN cases is warranted to validate the diagnostic value of the findings in this study.

Acknowledgments

We would like to thank Editage (www.editage.cn) for English language editing. Part of this study has been presented as an oral presentation at the section RPS 2104 of the European Congress of Radiology 2024 (details at: https://connect.myesr.org/course/pulmonary-nodules-and-lung-cancer-screening/?).

Funding: This work was supported by the Key Project of Medical Science and Technology of Henan Province in China (No. SBGJ202102057).

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-24-170/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-24-170/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This retrospective, multi-center cohort study was conducted in accordance with the Declaration of Helsinki (as revised in 2013), and the protocol was approved by The Affiliated Cancer Hospital of Zhengzhou University & Henan Cancer Hospital Ethics Committee (No. 2021-KY-0022), and individual consent for this retrospective analysis was waived.

References

1. Tammemägi MC, Katki HA, Hocking WG, Church TR, Caporaso N, Kvale PA, Chaturvedi AK, Silvestri GA, Riley TL, Commins J, Berg CD. Selection criteria for lung-cancer screening. N Engl J Med 2013;368:728-36.
2. Austin JH, Müller NL, Friedman PJ, Hansell DM, Naidich DP, Remy-Jardin M, Webb WR, Zerhouni EA. Glossary of terms for CT of the lungs: recommendations of the Nomenclature Committee of the Fleischner Society. Radiology 1996;200:327-31.
3. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, 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. Kodama K, Higashiyama M, Yokouchi H, Takami K, Kuriyama K, Kusunoki Y, Nakayama T, Imamura F. Natural history of pure ground-glass opacity after long-term follow-up of more than 2 years. Ann Thorac Surg 2002;73:386-92; discussion 392-3.
5. Wang J, Ma H, Ni CJ, He JK, Ma HT, Ge JF. Clinical characteristics and prognosis of ground-glass opacity nodules in young patients. J Thorac Dis 2019;11:557-63.
6. Haro A, Yano T, Kohno M, Yoshida T, Okamoto T, Maehara Y. Ground-glass opacity lesions on computed tomography during postoperative surveillance for primary non-small cell lung cancer. Lung Cancer 2012;76:56-60.
7. Chang B, Hwang JH, Choi YH, Chung MP, Kim H, Kwon OJ, Lee HY, Lee KS, Shim YM, Han J, Um SW. Natural history of pure ground-glass opacity lung nodules detected by low-dose CT scan. Chest 2013;143:172-8.
8. Eguchi T, Kondo R, Kawakami S, Matsushita M,

Yoshizawa A, Hara D, Matsuoka S, Takeda T, Miura K, Agatsuma H, Sakaizawa T, Tominaga Y, Saito G, Toishi M, Hamanaka K, Hashizume M, Shiina T, Amano J, Koizumi T, Yoshida K. Computed tomography attenuation predicts the growth of pure ground-glass nodules. Lung Cancer 2014;84:242-7.
9. Zhao W, Xu Y, Yang Z, Sun Y, Li C, Jin L, Gao P, He W, Wang P, Shi H, Hua Y, 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;112:161-8.
10. Meng Q, Li B, Gao P, Liu W, Zhou P, Ding J, Zhang J, Ge H. Development and Validation of a Risk Stratification Model of Pulmonary Ground-Glass Nodules Based on Complementary Lung-RADS 1.1 and Deep Learning Scores. Front Public Health 2022;10:891306.
11. Park CM, Goo JM, Kim TJ, Lee HJ, Lee KW, Lee CH, Kim YT, Kim KG, Lee HY, Park EA, Im JG. Pulmonary nodular ground-glass opacities in patients with extrapulmonary cancers: what is their clinical significance and how can we determine whether they are malignant or benign lesions? Chest 2008;133:1402-9.
12. Li Q, Fan L, Cao ET, Li QC, Gu YF, Liu SY. Quantitative CT analysis of pulmonary pure ground-glass nodule predicts histological invasiveness. Eur J Radiol 2017;89:67-71.
13. Xiang W, Xing Y, Jiang S, Chen G, Mao H, Labh K, Jia X, Sun X. Morphological factors differentiating between early lung adenocarcinomas appearing as pure ground-glass nodules measuring ≤10 mm on thin-section computed tomography. Cancer Imaging 2014;14:33.
14. Heidinger BH, Anderson KR, Nemec U, Costa DB, Gangadharan SP, VanderLaan PA, Bankier AA. Lung Adenocarcinoma Manifesting as Pure Ground-Glass Nodules: Correlating CT Size, Volume, Density, and Roundness with Histopathologic Invasion and Size. J Thorac Oncol 2017;12:1288-98.
15. Qi C, Sun SW, Xiong XZ. From COPD to Lung Cancer: Mechanisms Linking, Diagnosis, Treatment, and Prognosis. Int J Chron Obstruct Pulmon Dis 2022;17:2603-21.
16. Kong B, Wang X, Bai J, Lu Y, Gao F, Cao K, Xia J, Song Q, Yin Y. Learning tree-structured representation for 3D coronary artery segmentation. Comput Med Imaging Graph 2020;80:101688.
17. Ye H, Gao F, Yin Y, Guo D, Zhao P, Lu Y, Wang X, Bai J, Cao K, Song Q, Zhang H, Chen W, Guo X, Xia J. Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. Eur Radiol 2019;29:6191-201.
18. Ohno Y, Aoyagi K, Yaguchi A, Seki S, Ueno Y, Kishida Y, Takenaka D, Yoshikawa T. Differentiation of Benign from Malignant Pulmonary Nodules by Using a Convolutional Neural Network to Determine Volume Change at Chest CT. Radiology 2020;296:432-43.
19. Hsu YC, Tsai YH, Weng HH, Hsu LS, Tsai YH, Lin YC, Hung MS, Fang YH, Chen CW. Artificial neural networks improve LDCT lung cancer screening: a comparative validation study. BMC Cancer 2020;20:1023.
20. Cai J, Liu H, Yuan H, Wu Y, Xu Q, Lv Y, Li J, Fu J, Ye J. A radiomics study to predict invasive pulmonary adenocarcinoma appearing as pure ground-glass nodules. Clin Radiol 2021;76:143-51.
21. Singh R, Kalra MK, Homayounieh F, Nitiwarangkul C, McDermott S, Little BP, Lennes IT, Shepard JO, Digumarthy SR. Artificial intelligence-based vessel suppression for detection of sub-solid nodules in lung cancer screening computed tomography. Quant Imaging Med Surg 2021;11:1134-43.
22. Qi LL, Wu BT, Tang W, Zhou LN, Huang Y, Zhao SJ, Liu L, Li M, Zhang L, Feng SC, Hou DH, Zhou Z, Li XL, Wang YZ, Wu N, Wang JW. Long-term follow-up of persistent pulmonary pure ground-glass nodules with deep learning-assisted nodule segmentation. Eur Radiol 2020;30:744-55.
23. Sun Y, Li C, Jin L, Gao P, Zhao W, Ma W, Tan M, Wu W, Duan S, Shan Y, Li M. Radiomics for lung adenocarcinoma manifesting as pure ground-glass nodules: invasive prediction. Eur Radiol 2020;30:3650-9.
24. Xia X, Gong J, Hao W, Yang T, Lin Y, Wang S, Peng W. Comparison and Fusion of Deep Learning and Radiomics Features of Ground-Glass Nodules to Predict the Invasiveness Risk of Stage-I Lung Adenocarcinomas in CT Scan. Front Oncol 2020;10:418.
25. Meng F, Guo Y, Li M, Lu X, Wang S, Zhang L, 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.
26. Thawani R, McLane M, Beig N, Ghose S, Prasanna P, Velcheti V, Madabhushi A. Radiomics and radiogenomics in lung cancer: A review for the clinician. Lung Cancer 2018;115:34-41.
27. WHO Classification of Tumours Editorial Board. WHO classification of tumours. Thoracic Tumours (5th edition). IARC: Lyon, 2021.

28. Hsu WC, Huang PC, Pan KT, Chuang WY, Wu CY, Wong HF, Yang CT, Wan YL. Predictors of Invasive Adenocarcinomas among Pure Ground-Glass Nodules Less Than 2 cm in Diameter. Cancers (Basel) 2021;13:3945.
29. Li X, Zhang W, Yu Y, Zhang G, Zhou L, Wu Z, Liu B. CT features and quantitative analysis of subsolid nodule lung adenocarcinoma for pathological classification prediction. BMC Cancer 2020;20:60.
30. Cao C, Liu F, Tan H, Song D, Shu W, Li W, Zhou Y, Bo X, Xie Z. Deep Learning and Its Applications in Biomedicine. Genomics Proteomics Bioinformatics 2018;16:17-32.
31. Ikeda K, Awai K, Mori T, Kawanaka K, Yamashita Y, Nomori H. Differential diagnosis of ground-glass opacity nodules: CT number analysis by three-dimensional computerized quantification. Chest 2007;132:984-90.
32. Yang HH, Lv YL, Fan XH, Ai ZY, Xu XC, Ye B, Hu DZ. Factors distinguishing invasive from pre-invasive adenocarcinoma presenting as pure ground glass pulmonary nodules. Radiat Oncol 2020;15:186.
33. Wang S, Tong X, Cheng Q, Xiao Q, Cui J, Li J, Liu Y, Fang X. Fully automated deep learning system for osteoporosis screening using chest computed tomography images. Quant Imaging Med Surg 2024;14:2816-27.
34. Tian D, Jiang S, Zhang L, Lu X, Xu Y. The role of large language models in medical image processing: a narrative review. Quant Imaging Med Surg 2024;14:1108-21.

Cite this article as: Meng Q, Zheng C, Guo L, Gao P, Liu W, Ge H, Liu T, Peng H, Lu J, Chen X. Construction and validation of a risk score system for diagnosing invasive adenocarcinoma presenting as pulmonary pure ground-glass nodules: a multi-center cohort study in China. Quant Imaging Med Surg 2024;14(7):4864-4877. doi: 10.21037/qims-24-170

