Volume 24, No 3 | International Journal of Radiation Research, July 2026

A logistic regression model utilizing computed tomography features for predicting the degree of malignancy in pure ground-glass nodule lung adenocarcinoma: A clinical investigation

W. Zhang[1], Y. Dong[2], X. Wei[3*]

[1]*Department of Radiology, The First People's Hospital of Pinghu, Pinghu, Zhejiang Province, China*
[2]*Department of Pathology, The First People's Hospital of Pinghu, Pinghu, Zhejiang Province, China*
[3]*Department of Radiology, Pinghu Hospiral of Traditional Chinese Medicine, Pinghu, Zhejiang Province, China*

► ***Original article***

Corresponding author:
Xu Wei, Ph.D.,
E-mail: *19957485499@163.com*

Received: November 2025
Final revised: February 2026
Accepted: February 2026

Int. J. Radiat. Res., July 2026; 24(3): 755-762

DOI: 10.61186/ijrr.24.3.24

Keywords: *Logistic regression, predictive model, pure ground-glass nodule, lung adenocarcinoma, CT, invasion degree.*

ABSTRACT

Background: Differentiating the benign or malignant nature and the degree of infiltration of pure ground-glass nodules (pGGNs) is challenging and represents a clinical difficulty in the precise management of pulmonary nodules. It was to fabricate and validate a logistic regression (LR) model under high-resolution computed tomography (HRCT) features for the non-invasive preoperative prediction of the malignant infiltration degree in pGGN-type lung adenocarcinoma. ***Materials and Methods:*** A total of 180 patients (with 279 lesions) presenting with pure ground-glass nodules (pGGNs), defined as homogeneous ground-glass opacities without solid components on preoperative high-resolution computed tomography (HRCT), were retrospectively enrolled and randomly assigned to a model development group and a model validation group. HRCT imaging features were analyzed, including lesion size, mean CT value, margin characteristics, vacuole or cavity sign, and the relationship between vessels and the lesion. LR analyses identified independent predictors of invasive adenocarcinoma (IAC), to build the prediction model. Model discriminatory ability was evaluated. ***Results:*** Multivariate analysis identified lesion size, margin characteristics (irregular), vacuole or cavity sign, and type III vascular relationship as independent risk factors for predicting IAC. The model demonstrated good predictive performance in the development group, with area under the curve (AUC) of 0.810 (95%CI: 0.728-0.952), sensitivity of 0.782, and specificity of 0.722. The model also had robust performance in the validation group. ***Conclusion:*** The LR model under CT features effectively differentiated between pre-invasive lesions and invasive lung adenocarcinoma in pGGNs, providing a valuable non-invasive quantitative tool for formulating individualized clinical management strategies.

INTRODUCTION

Lung cancer is a common malignancy worldwide regarding mortality, with lung adenocarcinoma being the most prevalent pathological type [1]. Massive pulmonary ground-glass nodules (GGNs) are being detected by CT [2]. A GGN is a hazy area of increased density on high-resolution CT (HRCT) that does not obscure bronchial structures or vascular margins. Based on their internal composition, GGNs are classified as pure GGNs (pGGNs) or mixed GGNs (mGGNs) [3]. pGGNs correspond to a continuous pathological spectrum, ranging from preinvasive lesions (atypical adenomatous hyperplasia (AAH)) and adenocarcinoma in situ (AIS) to minimally invasive adenocarcinoma (MIA) and invasive adenocarcinoma (IAC) [4]. Among these, AAH and AIS are considered preinvasive lesions, typically characterized by slow growth and very low metastatic potential; however, once they progress to IAC, their biological behavior becomes more aggressive.

This significant pathological difference leads to fundamental differences in clinical management strategies. For preinvasive lesions (AAH/AIS) and some MIAs, given their indolent biological behavior, current clinical guidelines mostly recommend active imaging surveillance to avoid unnecessary surgical trauma and long-term impacts on patient quality of life [5,6]. In contrast, for lesions that have progressed to IAC, surgical resection is the preferred and potentially curative treatment [7]. Therefore, noninvasive preoperative accurate discrimination of the invasion degree of pGGNs is crucial for formulating individualized and precise treatment decisions, which can avoid overtreatment of indolent lesions while ensuring timely intervention for invasive carcinomas. Currently, the "gold standard" for preoperative assessment of pGGN invasion degree is postoperative pathological diagnosis. There is a clinical urgent need for a reliable preoperative noninvasive prediction method. HRCT is the most

fundamental and important imaging tool for evaluating pulmonary nodules. Previous studies have indicated that multiple CT features, such as lesion size, density (CT value), margin characteristics, vacuole or cavity sign, and the relationship between the lesion and blood vessels, exhibit certain correlations with the malignant invasion degree of pGGNs [8, 9]. However, the diagnostic value of any single imaging feature is limited, as its sensitivity and specificity often fall short of meeting the demands for precise clinical decision-making. To integrate multiple valuable predictors, statistically based prediction models have emerged [10]. The logistic regression (LR) model can synthesize various risk factors and quantify their probabilistic impact on the outcome event, thereby providing a more objective and accurate individualized risk assessment. Although previous studies have attempted to develop such models, many are limited by lack of validation using an independent internal validation set, leaving their generalizability and clinical applicability requiring further confirmation [6].

Therefore, this study aimed to systematically analyze HRCT imaging features in a relatively large cohort of pGGN patients, identify independent risk factors for invasive lung adenocarcinoma through multivariate LR analysis, and subsequently develop a comprehensive prediction model. Compared with complex machine-learning algorithms, logistic regression offers superior interpretability, transparent quantification of individual risk factors, and easier clinical acceptance, making it particularly suitable for decision support in routine radiological and surgical practice.

The novelty of the present study lies in several aspects. First, this work focuses exclusively on strictly defined pure ground-glass nodules without solid components, reducing heterogeneity compared with studies that combine pure and part-solid nodules. Second, beyond conventional CT features, we systematically incorporated the vessel–lesion relationship and further stratified it into three distinct types, highlighting the added predictive value of type III vascular alteration for invasive adenocarcinoma. Third, based on multivariate logistic regression, we developed and internally validated a quantitative and visually interpretable nomogram using an independent validation cohort, enhancing the model's clinical applicability and robustness. Collectively, this study provides a practical, noninvasive, and internally validated CT-based prediction tool to support individualized preoperative decision-making for patients with pGGNs.

MATERIALS AND METHDS

Study design

This single-center retrospective diagnostic study was approved by the institutional ethics review board and was conducted in accordance with the principles of the Declaration of Helsinki; written informed consent was obtained from all participants. Patients with lung adenocarcinoma confirmed by surgical resection or biopsy pathology were consecutively enrolled between January 2021 and December 2023. All patients underwent preoperative thoracic HRCT showing pGGNs.

Inclusion criteria were: (1) postoperative pathological diagnosis consistent with the 2015 WHO classification of lung adenocarcinoma, including AAH, AIS, MIA, and IAC; (2) clear CT images without significant motion or metal artifacts, meeting diagnostic and analytical requirements; (3) lesions identified as pGGNs, defined as homogeneous ground -glass opacities on HRCT without solid components and with a size ≤30 mm.

Exclusion criteria were: (1) concurrent other malignant tumors or active pulmonary infection; (2) previous ipsilateral lung surgery or history of radiotherapy/chemotherapy; (3) incomplete imaging data.

A total of 180 patients (comprising 279 pGGN lesions) were ultimately included for analysis. All patients were randomly assigned via computer-generated random numbers in a 7:3 ratio to either the model development group (n=126) or the model validation group (n=54).

CT image acquisition

All CT images were scanned via one of the following two 64-slice or higher spiral CT scanners: a Canon Prime 64-slice spiral CT scanner (Canon Medical Systems Corporation, Japan) and a Siemens Somatom Perspective 32-slice CT scanner (Siemens Healthineers AG, Germany). Scanning was performed with the patient in supine position and completed during a single breath-hold. Scan range extended from lung apex to lung base. Parameters include tube voltage=120 kV, automatic tube current modulation (noise index SD=10), and matrix 512×512. All data were reconstructed with 1.5 mm slice thickness via bone algorithm and were reviewed with both lung window (width=1500 HU, level=-400 HU) and mediastinal window (width=350 HU, level=40 HU) settings. All CT examinations were performed without intravenous contrast enhancement, as the assessment of pure ground-glass nodules and related morphological features was based exclusively on non –contrast-enhanced high-resolution CT images. Sample CT images are shown in figure 1.

Imaging analysis

Two associate experienced chief physicians in thoracic imaging diagnosis independently evaluated CT features of all pGGNs using three-dimensional pulmonary nodule analysis software (syngo.via; Siemens Healthineers, Erlangen, Germany), blinded to pathological results. Diagnostic disagreements

were resolved via consensus. The analyzed CT features were defined as follows: lesion size was determined by measuring the maximum diameter and its perpendicular diameter on the lung window and averaging the values; mean CT value was automatically calculated by the software after manually delineating a region of interest as large as possible on the largest cross-section of the lesion, avoiding vessels, bronchi, and vacuoles; margin characteristics were classified as "smooth" or "non-smooth"; the vacuole or cavity sign was defined as presence of round or oval air-density low-attenuation areas within lesion; the relationship between vessels and the lesion was categorized into three types: Type I, II, and III. Other signs, such as pleural indentation and air bronchogram, were also recorded. To ensure measurement consistency, if the discrepancy in lesion size measurements between the two physicians exceeded 2 mm or the discrepancy in mean CT values exceeded 50 HU, remeasurement and confirmation were required.

Pathological reference standard

Using postoperative histopathological results as the gold standard [11], lesions were classified according to the WHO classification: AAH and AIS were categorized into the "pre-invasive lesion" group, while MIA and IAC were categorized into the "IAC" group, both serving as dependent variable in the LR analysis.

Statistical analysis

Employing *SPSS 26.0*, measurement data were denoted as mean±standard deviation, and independent samples t-test conducted intergroup comparisons. Count data were presented as frequency, and χ^2 test or Fisher's exact test conducted intergroup comparisons. In model development group, univariate analysis was first conducted to screen for variables associated with IAC (P<0.05). Statistically notable variables were included in a multivariate binary LR analysis (backward stepwise method) to identify independent risk factors for predicting IAC, and their odds ratios (OR) and 95% confidence intervals (CI) were calculated.

Regarding multivariate LR analysis results, a nomogram model was constructed. The points score for each variable was calculated using the following equation: Raw Score=β coefficient × variable value, subsequently standardized to a 0-100 points system via linear transformation. The total points were converted into the probability of invasion using the logistic function: P=1 / [1 + e^ (-(-3.42 + total score))].

Model discriminatory performance was assessed via receiver operating characteristic (ROC) curve and area under the curve (AUC). Model sensitivity and specificity were calculated in development and validation groups, respectively. Hosmer-Lemeshow test assessed the model's calibration. P<0.05 meant statistically significant.

RESULTS

Patient and lesion characteristics

In table 1, 180 patients (85 males and 104 females, 66.34 ± 11.53 years old) were ultimately recruited, encompassing 279 pathologically confirmed pGGN lesions. All patients were randomly assigned to model development group (126 patients with 189 lesions) or model validation group (54 patients with 90 lesions). Negligible differences were observed in baseline characteristics such as gender, age, smoking history, or clinical symptoms between groups (P>0.05), indicating balanced and comparable grouping.

Table 1. Baseline clinical characteristics.

Characteristic	Overall (N=180)	Model development group (n=126)	Model validation group (n=54)	*P*
Sex, n (%)				0.852
Male	85(47.2)	59 (46.8)	26 (48.1)	
Female	95(52.8)	67 (53.2)	28 (51.9)	
Age (years), Mean ± SD	66.34± 11.53	65.89 ± 11.71	67.41 ± 11.15	0.401
Smoking history, n (%)	70(38.9)	48 (38.1)	22 (40.7)	0.736
Number of lesions per patient, n (%)				0.455
Single	175(97.2)	122 (96.8)	53 (98.1)	
Multiple	5(2.8)	4 (3.2)	1 (1.9)	

Univariate analysis of CT signs

In the model development group, a comprehensive comparison of CT features was conducted between pre-invasive lesions (AAH/AIS) and IACs (MIA/IAC). In table 2, versus pre-invasive lesion group, IAC group demonstrated notably larger lesion diameter (12.65 ± 0.21 mm *vs.* 7.81 ± 1.82 mm, P<0.05) and higher mean CT value (-720 ± 73 HU *vs.* -750 ± 98 HU, P=0.028). Regarding morphological characteristics, the IAC group more frequently exhibited irregular shape (68.6% *vs.* 31.4%, P<0.05), non-smooth margins (67.1% *vs.* 12.3%, P<0.001), and the presence of vacuole or cavity signs (22.6% *vs.* 15.7%, P=0.003). In vascular analysis, a Type III vascular relationship (vessel distortion/dilation) was observed significantly more often in IACs than in pre-invasive lesions (27.1% *vs.* 45.3%, P=0.005). Insignificant differences were found in the distribution of pleural indentation or air bronchogram between the two groups (P>0.05).

Multi-factor LR analysis and establishment of prediction model

Variables with P<0.05 in univariate analysis were analyzed in multivariate binary LR analysis. Ultimately, four indicators were identified as independent risk factors for predicting IAC: lesion size (OR=1.32, 95% CI: 1.05-1.67), non-smooth

margin (OR=4.54, 95% CI: 2.12-9.71), presence of vacuole or cavity sign (OR=2.51, 95% CI: 1.47-4.30), and type III vascular relationship (OR=2.94, 95% CI: 1.38-6.28). Details are presented in table 3.

Table 2. Comparison of CT features in the model development group.

CT feature	Pre-invasive lesions(n=106)	IAC(n=70)	Statistic	*P*
Mean diameter (mm), Mean ± SD	7.81 ± 1.82	12.65±0.21	t=-15.32	0.001
Mean CT value (HU), Mean ± SD	-750 ± 98	-720±73	t=-2.22	0.028
Shape, n (%)			χ^2=25.41	0.001
Round/Oval	88 (83.0)	37(52.9)		
Polygonal/Irregular	18 (17.0)	33(47.1)		
Edge feature, n (%)			χ^2=52.17	0.001
Smooth	93 (87.7)	23(32.9)		
Non-smooth	13 (12.3)	47(67.1)		
Vacuole/Cavity sign, n (%)	24 (22.6)	11(15.7)	χ^2=8.92	0.003
Vascular relationship, n (%)			χ^2=12.35	0.002
Type I	67 (63.2)	28(40.0)		
Type II	31 (29.2)	26(37.1)		
Type III	8 (7.5)	16(22.9)		
Pleural indentation, n (%)	15 (14.2)	13(18.6)	χ^2=0.65	0.420
Air bronchogram, n (%)	22 (20.8)	18(25.7)	χ^2=0.60	0.439

Table 3. Multivariate LR analysis for predicting IAC.

Variable	β (regression coefficient)	Standard error	Wald χ^2	*P*	OR	95% *CI* for OR
Lesion size (per 1 mm increase)	0.28	0.12	5.42	0.020	1.32	1.05 - 1.67
Margin feature (Non-smooth *vs.* Smooth)	1.52	0.41	13.78	0.001	4.54	2.12 - 9.71
Mean CT value	-0.03	0.06	0.26	0.610	0.97	0.86-1.09
Vacuole/Cavity sign (Present *vs.* Absent)	0.92	0.29	10.62	0.001	2.51	1.47 - 4.30
Vascular relationship (Type III *vs.* Type I)	1.08	0.38	7.97	0.005	2.94	1.38 - 6.28
Constant	-3.42	0.33	11.43	0.001	0.03	

Using independent predictors, the final LR prediction model was established with the following equation:

Logit(P)=-3.42 + (0.28 × Lesion size) + (1.52 × Margin feature) + (0.92 × Vacuole/Cavity sign) + (1.08 × Type III vascular relationship), where *P* represents the predicted probability of the lesion being IAC. Margin feature: (smooth=0, non-smooth=1), vacuole/cavity sign: (absent=0, present=1), and vascular relationship: (type I/II=0, type III=1).

Nomogram of clinical prediction model establishment

Regarding independent predictors identified by multivariate LR analysis, the study fabricated a visual nomogram model (figure 2) to provide an individualized tool for predicting the risk of invasion.

During the variable scoring stage, the actual lesion

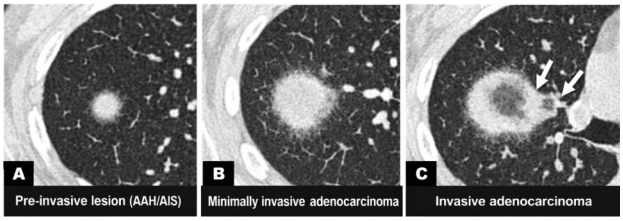


Figure 1. Representativeness high resolution Ct-images of pure ground-glass nodules with different invasion degrees. **A)** Pre-invasion lesion (AAH/AIS). **B)** Minimally invasive adenocarcinoma. **C)** Invasive adenocarcinoma with vacuole sign and vascular distortion (arrows).

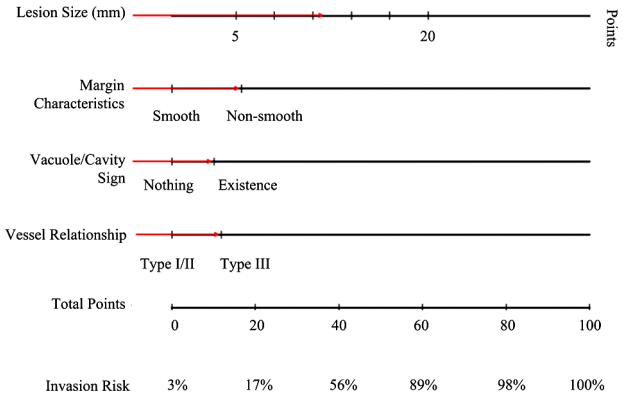


Figure 2. Nomogram for predicting invasion risk in pGGNs. (Note: This nomogram was built under multivariate LR model and incorporated four independent CT predictors. Instructions for use: (1) locate the patient's characteristics on respective variable axes; (2) draw lines upward to points axis to set each variable score; (3) sum all scores fortotal points; (4) draw a line downward from the total points to the risk of invasion axis to read predicted probability. The red arrow demonstrates an example calculation: a 12 mm lesion with a non-smooth margin, presence of a vacuole/cavity sign, and a type III vascular relationship, yielding a predicted invasion risk of 83.2%.).

diameter (range: 5-20 mm) is located on the "Lesion size" axis, and a line is plotted upward to "Points" axis to obtain corresponding score. On the "Margin characteristic" axis, if the lesion margin is non-smooth, the corresponding position is located and a line is plotted upward to obtain the score. On the "Vacuole/Cavity sign" axis, if this sign is present, the corresponding position is located and a line is drawn upward to obtain the score. On the "Vascular relationship" axis, if a Type III vascular relationship is present, the corresponding position is located and a line is plotted upward to obtain the score. The points obtained from the four variables are summed to yield total points, located on "Total points" axis. In the risk prediction stage, a line is plotted downward from these total points position to "Risk of invasion" axis to read corresponding predicted probability of invasion. The point allocation for each variable is based on their regression coefficient weights: lesion size contributes 0.28 points per 1 mm increase (linearly

converted to the points system); margin characteristic (non-smooth *vs.* smooth) corresponds to 1.52 points; vacuole/cavity sign (present *vs.* absent) corresponds to 0.92 points; and vascular relationship (type III *vs.* type I/II) corresponds to 1.08 points.

In figure 1, for a patient with a lesion size of 12 mm, non-smooth margins, presence of a vacuole or cavity sign, and a Type III vascular relationship, the total score is 68.5 points, corresponding to a predicted invasion risk of 83.2%. The prediction model had ideal calibration in validation cohort (Hosmer-Lemeshow test, P=0.325).

Prediction efficiency and verification of the model

The prediction model demonstrated excellent discrimination in the development cohort, with AUC value of 0.810 (95%CI: 0.728-0.952) (figure 2). The ROC curves for both cohorts are shown in figure 3. Hosmer-Lemeshow test indicated no significant difference between predicted and observed probabilities (development cohort P=0.325, validation cohort P=0.418), demonstrating good model calibration. At optimal cutoff value, model exhibited sensitivity of 78.2%, specificity of 72.2%, and accuracy of 75.4% (figure 4). The positive likelihood ratio (LR+) was 2.81, and the negative likelihood ratio (LR-) was 0.30.

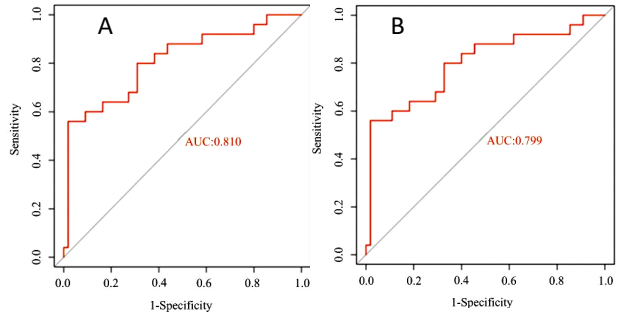


Figure 3. ROC curves of the prediction model. (**A:** model development cohort; **B:** model validation cohort).

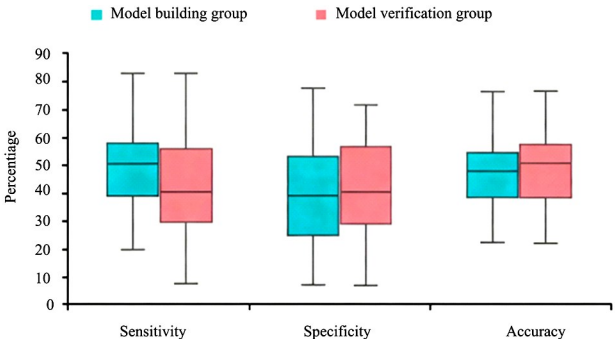


Figure 4. Contrast of prediction indicators between the two cohorts.

The model maintained robust predictive performance in validation cohort, with AUC value of 0.799 (95%CI: 0.753-0.899) and sensitivity, specificity, and accuracy of 72.6%, 64.8%, and 69.3%, respectively (figure 4).

To further demonstrate the model's clinical utility, lesions were categorized into low-risk (P<0.3), intermediate-risk (0.3≤P<0.7), and high-risk (P≥0.7) groups based on their predicted probabilities. In figure 5, in model development cohort, the proportion of pathologically confirmed IACs reached 88.9% in high-risk group, while this proportion was only 11.8% in low-risk group, demonstrating the model's excellent risk stratification capability.

In figure 6, the mean predicted probability was 0.340±0.138 in the pre-invasive lesion group and 0.749±0.145 in the IAC group. The clear distinction between these two groups indicates the model's strong discriminative power.

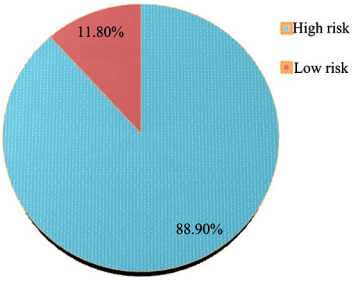


Figure 5. Proportion of cases in different risk groups.

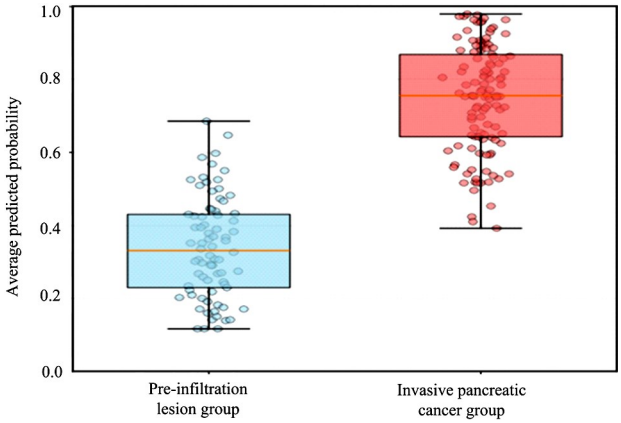


Figure 6. Distribution of predicted probabilities for pre-invasive lesions and IAC in the validation cohort. (Note: Each individual point in the figure represents a single pGGN lesion).

DISCUSSION

Based on HRCT imaging features, this work successfully built and validated a LR model for predicting the degree of malignant invasion in pGGN-type lung adenocarcinoma, providing an important auxiliary diagnostic tool for clinical practice. The model demonstrated good discriminatory ability in both development and validation cohorts (AUC=0.810 and 0.799, respectively), indicating high clinical applicability and robustness. Through comprehensive analysis of lesion size, margin characteristics (non-smooth), vacuole or cavity sign, and type III vascular relationship (vessel distortion/dilation), this study identified these as independent predictors for distinguishing between pre-invasive lesions (AAH and AIS) and IACs (including MIA and IAC). These findings are consistent with several recent studies and provide deeper imaging-based

evidence for the precise management of pGGNs [12].

The four independent predictors identified in this study are as follows. Lesion size, as the most intuitive imaging indicator, is closely associated with tumor proliferative capacity and invasiveness. This study revealed that the average diameter of IAC (12.65 mm) was markedly larger than that of pre-invasive lesions (7.81 mm), which is consistent with previous research findings and suggests that larger nodules are associated with a higher probability of malignant invasion. Liu *et al.* emphasized value of nodule diameter in predicting the invasiveness of pGGNs, while also noting inconsistencies in its diagnostic performance and optimal cutoff values [13]. Shi *et al.* also focused on identifying optimal CT thresholds to predict invasiveness in stage T1 lung adenocarcinoma [14]. Collectively, these studies indicated that lesion size is a critical indicator for assessing the malignancy degree of pGGNs. The irregularity of tumor margins, such as lobulation and spiculation, typically reflects the heterogeneous infiltration of tumor cells into surrounding tissues and an invasive growth pattern. Radiologically, these features are characteristic manifestations of malignant tumors, indicating that the tumor has breached normal tissue structures and is engaging in invasive interactions with the surrounding lung parenchyma [15]. The vacuole or cavity sign is more frequently observed in invasive lesions. This may be associated with internal tissue necrosis, destruction of alveolar structures, or mucus retention within the tumor, serving as one of the imaging markers of malignant transformation. These internal structural changes suggest the presence of more complex pathophysiological processes within the tumor, such as apoptosis, impaired angiogenesis, or enhanced secretory activity, all of which are features of malignant tumor progression [16]. Type III vascular relationships typically manifest as vessel distortion and dilation, indicating tumor-induced significant angiogenesis and blood supply remodeling. This abnormal vascular formation is a crucial characteristic of the malignant tumor microenvironment and is closely related to rapid tumor growth, high metabolic activity, and progression risk. Tumors utilize neovascularization to obtain nutrients and oxygen, as well as to serve as pathways for metastasis; therefore, the Type III vascular relationship is an important indicator for assessing tumor invasiveness and potential metastatic risk [17].

The association between a Type III vascular relationship and invasive adenocarcinoma is biologically plausible and can be explained by tumor-driven angiogenesis and tumor–stroma interactions. During the progression of lung adenocarcinoma from pre-invasive to invasive stages, tumor cells increasingly activate angiogenic signaling pathways, particularly vascular endothelial growth factor, resulting in abnormal neovascularization, vessel dilation, and architectural distortion [7, 17]. These angiogenesis-induced vascular alterations have been shown to correlate with aggressive tumor behavior and can be visualized radiologically as distorted or dilated vessels adjacent to or traversing the lesion, corresponding to Type III vessel–lesion relationships on CT imaging [12, 15].

In parallel, invasive tumor growth is accompanied by extensive remodeling of the tumor stroma, including activation of cancer-associated fibroblasts and degradation of the extracellular matrix, processes that facilitate tumor infiltration and mechanical deformation of surrounding vascular structures [18,19]. Histopathological studies have demonstrated that increased microvessel density and aberrant vascular architecture are significantly associated with invasive growth patterns, higher metastatic potential, and poorer prognosis in lung adenocarcinoma [16,20]. Therefore, the presence of a Type III vascular relationship on HRCT likely reflects an advanced tumor microenvironment characterized by active angiogenesis and stromal invasion, supporting its role as a meaningful imaging biomarker for malignant invasiveness.

This study possesses significant methodological advantages. It employed a relatively large sample size (180 patients with 279 lesions) and strictly adhered to the principle of random grouping for internal validation, effectively avoiding model overfitting and enhancing the generalizability of the results. Xue *et al.* also developed a multi-parameter prediction model for IAC in pGGNs smaller than 2 cm in diameter, and their study design similarly emphasized model accuracy and clinical application value [21]. This study not only constructed a LR model but also further developed a clinically practical nomogram tool. The nomogram transforms the complex statistical model into a visual risk assessment system, greatly facilitating practical application by clinicians [22, 23]. The study by Gao *et al.* also constructed a nomogram by combining subjective CT features and artificial intelligence-based histogram parameters to predict the invasiveness of pGGN-type lung adenocarcinoma and evaluated its diagnostic performance, further validating the utility of nomograms in clinical decision-making [24]. Compared with some previous studies, the present model maintains relatively high sensitivity (78.2%) while achieving acceptable specificity (72.2%), demonstrating a well-balanced performance in distinguishing between benign and malignant invasion. This indicates that the model can effectively identify malignant nodules requiring further intervention while avoiding overdiagnosis and unnecessary treatments. The precise management of pGGNs remains a challenge in current clinical practice for pulmonary nodules [25, 26]. The prediction model provided by this study offers a quantitative tool for the non-invasive preoperative

assessment of pGGN invasion extent, thereby aiding in guiding individualized treatment decisions. For lesions predicted by the model as low-risk (P<0.3), a more conservative follow-up strategy may be considered to avoid unnecessary surgical trauma; whereas for high-risk lesions (P≥0.7), active intervention is recommended to achieve early radical treatment [27, 28]. Introduction of the nomogram makes this model particularly suitable for primary hospitals or non-specialist radiologists, showing good potential for widespread adoption [22].

Despite the rigorous design and analysis of this study, several limitations remain. As a single-center retrospective study, potential selection bias may affect the generalizability of the results. Although the sample size is larger than that of some existing studies, multicenter, large-sample prospective studies are still needed to further validate the model's general applicability. The model did not incorporate clinical variables (such as tumor markers, patient age, smoking history) or more advanced imaging techniques (such as radiomics or artificial intelligence analysis). Future research could further integrate multimodal data on this basis to improve predictive accuracy. The microscopic views of the four stages of lung adenocarcinoma (LUAD) from normal to invasion and progression presented in the study by Liu *et al.* [13] align with our study's attempt to differentiate between AAH/AIS and MIA/IAC, providing a pathological basis for disease progression. At the biological level, genomic profiling studies of early-stage lung cancer have also revealed molecular differences among benign, pre-invasive, and invasive pulmonary nodules [18, 20]. Zhang *et al.* (2024) conducted a comprehensive proteomic analysis of 98 pre-invasive and 99 invasive lung adenocarcinoma cases, revealing evolutionary proteogenomic characteristics during the stepwise progression from pre-invasive to invasive lung adenocarcinoma [19]. These molecular-level findings can provide a deeper biological explanation for imaging features. Future research should focus on conducting multicenter prospective validation to assess its generalizability; simultaneously, integrating the CT features of this model with biomarkers such as radiomics, deep learning, and even liquid biopsy holds promise for constructing a more comprehensive risk assessment system, further enhancing the precision management level of pulmonary nodules [29, 30].

Despite the encouraging results, several limitations of this study should be acknowledged. First, this was a single-center retrospective study, which may introduce selection bias and limit the generalizability of the findings. Second, although a relatively large number of lesions were analyzed, some patients contributed more than one pGGN, which may have introduced lesion-level dependency despite random group allocation. Third, the model was internally validated using a split-sample approach, but external validation in independent, multicenter cohorts is still required to confirm its robustness and clinical applicability. Fourth, only conventional CT morphological features were included; potentially informative clinical variables (such as serum biomarkers or genetic data) and advanced imaging approaches (e.g., radiomics or deep learning–based features) were not incorporated. Finally, although logistic regression offers strong interpretability and clinical acceptance, it may not fully capture complex nonlinear relationships present in imaging data. Future prospective, multicenter studies integrating multimodal data and external validation are warranted to further improve and validate the predictive performance of this model.

It was confirmed that the conventional CT feature-based model possesses significant clinical application value and promotion potential. Future efforts should focus on further validating the model's generalizability through multicenter, prospective studies and exploring its integration with emerging technologies such as radiomics and artificial intelligence to continuously enhance the precision management of pulmonary nodules.

CONCLUSION

In this study, a logistic regression model based on high-resolution CT features was developed and internally validated to predict the degree of malignant invasion in pure ground-glass nodule lung adenocarcinoma. Lesion size, non-smooth margins, vacuole or cavity sign, and Type III vascular relationship were identified as independent predictors of invasiveness. The model demonstrated good discriminative performance and was translated into a clinically practical nomogram, providing a noninvasive and interpretable tool to support individualized preoperative management of patients with pGGNs.

***Acknowledgement*:** the Department of Pathology of The First People's Hospital of Pinghu for their support in data collection and pathological confirmation. We also appreciate the technical assistance provided by the imaging staff during CT scanning and image reconstruction.

***Ethical consideration*:** This retrospective study was approved by the Institutional Ethics Review Board of The First People's Hospital of Pinghu (Approval No. Pinghu First Hospital 2025021) and was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrollment.

***Conflicts of interest*:** The authors declare that they have no competing interests relevant to the content

of this manuscript. No financial or personal relationships with other people or organizations have inappropriately influenced this work.

***Fundings*:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

***Author contribution*:** W. Zhang: Conceptualization, data curation, formal analysis, investigation, methodology, writing – original draft, writing – review & editing. Y. Dong: Data curation, investigation, resources, validation, visualization. X. Wei: Conceptualization, methodology, project administration, supervision, writing – review & editing. All authors read and approved the final manuscript.

REFERENCES

1. Wu L-L, Lin W-K, Qian J-Y, Ma S-S, *et al.* (2023) Prognostic assessment of lung adenocarcinoma patients with early-staging diseases: a nomogram based on coagulation-related factors. *European Journal of Cardio-Thoracic Surgery, **64**(5): ezad313.*
2. Todoroki K, Kawakami S, Kiniwa Y, Asaka S, *et al.* (2025) Pulmonary metastases from malignant melanoma showing ground-glass opacity nodules. *Japanese Journal of Radiology **43**(6): 949-957.*
3. Feng H, Shi G, Xu Q, Ren J, *et al.* (2023) Radiomics-based analysis of CT imaging for the preoperative prediction of invasiveness in pure ground-glass nodule lung adenocarcinomas. *Insights into Imaging, **14**(1): 24.*
4. Mei K, Feng Z, Liu H, Wang M, *et al.* (2025) Preoperative prediction of pulmonary ground-glass nodule infiltration status by CT-based radiomics combined with neural networks. *BMC Cancer, **25**(1): 659.*
5. AlShammari A, Patel A, Boyle M, Proli C, *et al.* (2024) Prevalence of invasive lung cancer in pure ground glass nodules less than 30 mm: A systematic review. *European Journal of Cancer, **213**: 115116.*
6. Huang Z, Peng K, Hong Z, Zhang P, Kang M (2022) Nomogram for predicting recurrence and metastasis of stage IA lung adenocarcinoma treated by videoassisted thoracoscopic lobectomy. *Asian Journal of Surgery, **45**(12): 2691-2699.*
7. Ren Y-F, Ma Q, Zeng X, Huang C-X, *et al.* (2024) Single-cell RNA sequencing reveals immune microenvironment niche transitions during the invasive and metastatic processes of ground-glass nodules and part-solid nodules in lung adenocarcinoma. *Molecular Cancer, **23**(1): 263.*
8. Bin J, Wu M, Huang M, Liao Y, *et al.* (2024) Predicting invasion in early-stage ground-glass opacity pulmonary adenocarcinoma: a radiomics-based machine learning approach. *BMC Medical Imaging, **24**(1): 240.*
9. Jin Z, Zhang W, Liu H, Ding A, *et al.* (2022) Potential Therapeutic Application of Local Anesthetics in Cancer Treatment. *Recent Patents on Anti-Cancer Drug Discovery, **17**(4): 326-342.*
10. Chen W-c, Zhang L-h, Bai Y-y, Liu Y-b, *et al.* (2022) Nomogram prediction of chronic postsurgical pain in patients with lung adenocarcinoma after video-assisted thoracoscopic surgery: A prospective study. *Frontiers in Surgery, **9**: 1004205.*
11. Sun J, Yan Y, Meng Y, Ma Y, *et al.* (2022) An immune-related nomogram model that predicts the overall survival of patients with lung adenocarcinoma. *BMC Pulmonary Medicine, **22** (1): 114.*
12. Gao Y, Yan C, Cao Z, Xie Z, *et al.* (2025) A computed tomography (CT)-based ternary classification model for predicting the invasiveness of pure ground-glass nodules. *Clinical Radiology, **90**: 107081.*
13. Liu J, Yang X, Li Y, Xu H, *et al.* (2024) Predicting the Invasiveness of Pulmonary Adenocarcinomas in Pure Ground-Glass Nodules Using the Nodule Diameter: A Systematic Review, Meta-Analysis, and Validation in an Independent Cohort. *Diagnostics, **14**(2): 147.*
14. Shi W, Wei Z, Hu Y, Sun Y, *et al.* (2025) Identification of the optimal threshold for predicting the infiltration degree of T1-stage lung adenocarcinoma using solid component volume and three-dimensional consolidation-to-tumor ratio in threshold segmentation. *Quantitative Imaging in Medicine and Surgery, **15** (7): 6446-6456.*
15. Yang Y, Xu J, Wang W, Zhao J, *et al.* (2023) Meta-analysis of the correlation between CT-based features and invasive properties of pure ground-glass nodules. *Asian Journal of Surgery, **46**(9): 3405-3416.*
16. Lee JH, Choi Y, Hong H, Kim YT, *et al.* (2023) Prognostic value of CT-defined ground-glass opacity in early-stage lung adenocarcinomas: a single-center study and meta-analysis. *European Radiology, **34** (3):1905-1920.*
17. Song F, Fu B, Liu M, Liu X, *et al.* (2023) Proposal of modified lung-RADS in assessing pulmonary nodules of patients with previous malignancies: A Primary Study. *Diagnostics, **13**(13): 2210.*
18. De Zuani M, Xue H, Park JS, Dentro SC, *et al.* (2024) Single-cell and spatial transcriptomics analysis of non-small cell lung cancer. *Nature Communications, **15**(1): 4388.*
19. Zhang Y, Fu F, Zhang Q, Li L, *et al.* (2024) Evolutionary proteogenomic landscape from pre-invasive to invasive lung adenocarcinoma. *Cell Reports Medicine, **5**(1): 101358.*
20. Liang P, Peng M, Tao J, Wang B, *et al.* (2024) Development of a genome atlas for discriminating benign, preinvasive, and invasive lung nodules. *MedComm, **5**(8): e644.*
21. Xue M, Li R, Wang K, Liu W, *et al.* (2024) Construction and validation of a predictive model of invasive adenocarcinoma in pure ground-glass nodules less than 2 cm in diameter. *BMC Surgery, **24**(1): 56.*
22. Song W, Shi J, Zhou B, Meng X, *et al.* (2024) Nomogram predicting overall and cancer specific prognosis for poorly differentiated lung adenocarcinoma after resection based on SEER cohort analysis. *Scientific Reports, **14**(1): 22045.*
23. Xue M, Li R, Liu J, Lu M, *et al.* (2024) Nomogram for predicting invasive lung adenocarcinoma in small solitary pulmonary nodules. *Frontiers in Oncology **14**: 1334504.*
24. Gao R, Gao Y, Zhang J, Zhu C, *et al.* (2023) A nomogram for predicting invasiveness of lung adenocarcinoma manifesting as pure ground-glass nodules: incorporating subjective CT signs and histogram parameters based on artificial intelligence. *J Cancer Res and Clinical Oncology, **149**(17): 15323-15333.*
25. Sun JD, Sugarbaker E, Byrne SC, Gagné A, *et al.* (2024) Clinical outcomes of resected pure ground-glass, heterogeneous ground-glass, and part-solid pulmonary nodules. *American Journal of Roentgenology, **222**(5): e2330504.*
26. Ye W, Fu W, Li C, Li J, *et al.* (2024) Diameter thresholds for pure ground-glass pulmonary nodules at low-dose CT screening: Chinese experience. *Thorax, **80**(2): 76-85.*
27. DeSimone AK, Byrne SC, Hammer MM (2024) Comparison of Lung-RADS Version 1.1 and Lung-RADS Version 2022 in Classifying Airway Nodules Detected at Lung Cancer Screening CT. *Radiology: Cardiothoracic Imaging, **6**(1): e230149.*
28. Woo W, Kang D-Y, Cha YJ, Kipkorir V, *et al.* (2024) Histopathologic fate of resected pulmonary pure ground glass nodule: a systematic review and meta-analysis. *Journal of Thoracic Disease, **16** (2): 924-934.*
29. Liu W, You W, Lan Z, Ren Y, *et al.* (2024) An immune cell map of human lung adenocarcinoma development reveals an anti-tumoral role of the Tfh-dependent tertiary lymphoid structure. *Cell Reports Medicine, **5**(3): 101448.*
30. Zhu W, Jin Y, Ma G, Chen G, *et al.* (2024) Classification of lung cancer subtypes on CT images with synthetic pathological priors. *Medical Image Analysis **95**: 103199.*