Clinical and Radiomic Factors for Predicting Invasiveness in Pulmonary Ground-Glass Opacity.

YT Dang (✉ dangyutao2003@163.com)

Shijingshan Hospital of Beijing City <https://orcid.org/0000-0002-5937-4109>

RT Wang

Xuanwu Hospital

K Qian

Xuanwu Hospital

J Lu

Xuanwu Hospital Department of Radiology

Y Zhang

Xuanwu Hospital

Research

Keywords: Ground-glass opacity; consolidation/tumour ratio; radiomics; prediction model; nomogram

Posted Date: July 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-42438/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Purpose: To identify the invasiveness of pulmonary ground-glass opacity (GGO) by analysing clinical and radiomic features.

Materials and methods: Patients with pulmonary GGOs between 2014 and 2019 were included. Clinical features were collected, and radiomic features were extracted from CT data by 3D Slicer software. Predictors of invasiveness of GGO were selected by least absolute shrinkage and selection operator (LASSO) logistic regression analysis, and receiver operating characteristic (ROC) curves were drawn for each prediction model.

Results: A total of 195 GGOs were included in this study. Maximum_diameter, CTR and 2 radiomic features were significant predictors of invasive GGOs. The area under the ROC curve (AUC) for prediction models of clinical and radiomic features were 0.84 and 0.809, respectively, and the AUC for the combined prediction model was 0.848. Finally, a nomogram was established for individualized invasiveness prediction.

Conclusions: The combination of radiomic features with clinical features can enable differentiation between preinvasive and invasive GGOs. Maximum_diameter>1.3 cm, CTR>0.27 and a higher waveletLHLglcmlmnm were significant factors for predicting invasiveness.

Introduction

In recent years, the detection rate of pulmonary ground-glass opacity (GGO) increased significantly with the application of high-resolution chest CT (HRCT) [1–2]. Studies have demonstrated that 59%-75% of persistent GGOs were precancerous lesions or early stage adenocarcinoma [3–4]. According to the Internal Association for the Study of Lung Cancer, American Thoracic Society and European Respiratory Society classification in 2011, adenocarcinoma was classified as atypical adenomatous hyperplasia (AAH), adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA) and invasive adenocarcinoma (IA) [5]. In general, preinvasive GGOs contain AAH and AIS, and MIA and IA are categorized as invasive lesions [6]. Up to now, the treatment of pulmonary GGO was often based on CT manifestations and clinical experience. AAH and AIS are pure GGO or mixed GGO with few solid components on chest CT [7–8], and these types of preinvasive nodules often needed close follow-up or limited resection[9–10]. The mixed GGO with more solid components tend to be invasiveness lesions, and it is often necessary to perform segmentectomy or lobectomy with lymph node dissection for this type of GGO. After appropriate operation, compared with the 100% 5-year disease-free survival with AAH and AIS and the almost 100% 5-year disease-free survival with MIA[11], the long-term survival rate of patients with invasive adenocarcinoma remains poor[12–16], with reports of 74.6%-92.4% in the literature. Therefore, the identification of invasiveness in pulmonary GGO is essential to the decision-making of clinical treatment and prognosis.

Radiomics refers to the quantitative image features extraction from radiology images [17]. In oncology, tumour radiomic features determined by imaging data can be used to reveal the correlation between diagnosis, prediction and prognosis of cancer patients, and those features often include nodal shape, nodal volume, intensity and a series of "texture" features [18–20]. Our purpose was to identify the invasiveness of GGOs by the clinical and radiomic features from the chest CT.

Patients And Methods

1. Patient selection and grouping

From January 2014 to February 2019, a total of 268 patients (301 GGOs, 2 of which have 3 GGOs, 29 of which have 2 GGO, and the rest were single) who were operated for pulmonary GGOs in the Department of thoracic surgery of Xuanwu Hospital were collected. Inclusion criteria: (1) complete clinical characters, including gender, age, smoking history, family history of lung cancer and pathological type; (2) complete non-enhanced chest CT data within two months before operation; and (3) final pathological results indicate malignant lesions (including AAH, AIS, Mia and IPS). Exclusion criteria: (1) before the chest CT examination, the patients had undergone puncture biopsy, radiotherapy, radiofrequency ablation or other treatment of GGOs; and (2) the maximum diameter of GGO on CT images was greater than 3 cm. According to the pathological results, the final cases were divided into two

2. Chest CT examination and general imaging feature acquisition

All preoperative chest CTs were non-enhanced and performed by one of the two machines (Sensation Cardiac 64, Siemens Medical Solutions, Forchheim, Germany; Somatom Definition Flash, Siemens Medical Solutions, Forchheim, Germany). All CT examinations were performed with the following parameters: 120 kVp; pitch, 1.2; 100–200 mAs; and collimation, 5.0 mm. The chest CT images of the patients were analysed by two radiologists. The largest diameter of the tumour (T, tumour) and the consolidation (C, consolidation component) were measured on the lung window and on the mediastinal window separately. The CTR (consolidation/tumour ratio) of pure GGO alone was 0, while that of solid nodules was 1. A nodule that is part-solid had a CTR between 0 and 1[21]. The final result is the average of the two radiologists.

3. CT texture analysis

(1) Radiomic feature extraction

CT data in DICOM format were imported into 3D-slicer software (Version 4.6.2; <http://www.slicer.org>). The volume of interest (VOI) was obtained by semiautomatic segmentation using the Segment Editor package. The radiomic software package was used for feature extraction. The characteristics can be divided into two groups: original features: (1) shape-based (14 features), (2) grey-level dependence matrix (14 features), (3) first-order statistics (18 features), (4) grey-level cooccurrence matrix (24 features), (5) grey-level run-length matrix (16 features), (6) grey-level size zone matrix (16 features), and (7) neighbouring grey tone difference matrix (5 features). Wavelet features: Features are calculated from the intensity and texture features of the original image using a wavelet filter. Therefore, the features are concentrated in different frequency ranges within the tumour volume.

(2) Stable radiomic feature selection

To obtain stable radiomic features, each image data point is subjected to VOI segmentation and radiomic feature extraction twice, the intraclass correlation coefficient (ICC) for each radiomic feature is calculated, and the stable radiomic features were selected as $ICC > 0.75$.

4. Selection of prediction factors and establishment of the prediction model

Patients enrolled in our study were divided into a training cohort and a validation cohort. The multivariate logistic regression analysis with backward method was used to select independent predictors from clinical features (including CTR, Maximum_diameter of GGO, age, family history, gender and smoking history) in the training cohort, and receiver operating characteristic (ROC) curves was plotted, with area under curve (AUC) values representing the predictive ability of the clinical prediction model. For radiomic features, the LASSO algorithm and 10-fold cross-validation were used to obtain independent predictors in distinguishing the two pathologic subtypes in the training cohort, ROC curves representing the radiomic prediction model were plotted, and AUC values were calculated. Finally, all meaningful predictors were used to build a combined prediction model, which was compared with the clinical and radiomic prediction models. We also used the validation cohort to validate the predictive ability of the prediction models. A nomogram was constructed to predict invasiveness for individual GGO.

5. Statistical analysis

The means of continuous variables were compared using the Mann-Whitney U test, and Pearson's chi-square test was used for categorical variables between the two groups by SPSS (version 22.0, Inc., IBM Company, Chicago, Illinois, USA). R software (version 3.5.2, <http://www.R-project.org>) was also used for data analysis. ICC was calculated using the "psych" package in R. The "MASS" package was used for logistic regression in the clinical features group. The LASSO regression analysis was performed for radiomic features and combined predictor selection by the "ncvreg" package in R. The ROC curve was built by the "pROC" and "ggplot2" packages in R. The "OptimalCutpoints" package was used for cut-off calculation in R. A nomogram was formulated by

using the package “rms” in R. The concordance index (C-index), which represents the performance of the nomogram, was calculated with the “rccorrens” package in “Hmisc” in R. $P < 0.05$ was set as statistically significant. The related computerized programs with R are listed in the Appendix.

6. Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of both institutional and national research committees. Written informed consent from either the patients or their representatives was obtained before surgery.

Results

1. Clinical features

A total of 184 patients with 195 GGOs were included in this study. Among the 184 patients, 11 had two GGOs. Of the 195 GGOs, 73 (including 21 AAH and 52 AIS) were in the preinvasive GGO group, and 122 (including 31 MIA and 91 IA) were in the invasive GGO group. The clinical data of the two groups are shown in Table 1. The clinical features of the two group of patients were analysed, and there were no significant differences in gender ($P = 0.651$), age ($P = 0.382$), smoking history ($P = 0.685$), family history ($P = 0.183$) and nodule location ($P = 0.554$) between the preinvasive GGO and invasive GGO groups. However, there were significant differences in the Maximum_diameter ($P < 0.001$) and CTR ($P < 0.001$) between the two groups. All GGOs were randomly divided into two groups: a training cohort (including 144 GGOs) and a validation cohort (including 51 GGOs).

Table 1
Clinical Features of All Patients

	Preinvasive GGO	Invasive GGO	Total	p- value*
Number of patients	73	122	195	
Gender				0.651
Male	24 (35.3%)	44 (64.7%)	68	
Female	49 (38.6%)	78 (61.4%)	127	
Age [#]	59.15 ± 8.82	60.31 ± 9.04		0.382
Smoking history				0.685
Non-smokers	55 (36.7%)	95 (63.3%)	150	
Smokers	18 (40.0%)	27 (60.0%)	45	
Family history				0.183
No	62 (39.7%)	94 (60.3%)	156	
Yes	11 (28.2%)	28 (71.8%)	39	
Maximum_diameter(cm) [#]	0.97 ± 0.42	1.43 ± 0.63		< 0.001
CTR [#]	0.12 ± 0.24	0.44± 0.31		< 0.001
Location				0.554
Right upper	27 (37.5%)	45 (62.5%)	72	
Right middle	6 (54.5%)	5 (45.5%)	11	
Right low	9 (29.0%)	22 (71.0%)	31	
Left upper	22 (41.5%)	31 (58.5%)	52	
Left low	9 (32.1%)	19 (67.9%)	28	

[#] mean ± standard deviation.

* p-value was based on comparison between preinvasive group with invasive group.

GGO = ground glass opacity.

Table 2. Diagnostic Accuracy and AUC of Prediction Models.

Groups	Cohort	SEN (%)	SPE (%)	PPV (%)	NPV(%)	Accuracy (%)	AUC	95%CI	p-value*
Clinical	Training	70.5	89.8	87.5	93.1	77.1	0.84	0.703–0.936	0.764
	Validation	63.0	91.7	89.5	68.8	76.5	0.819	0.691–0.929	
Radiomic	Training	81.1	75.5	86.5	67.3	79.2	0.809	0.73–0.888	0.176
	Validation	91.7	62.3	86.7	61.1	68.6	0.693	0.547–0.839	

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

	Preinvasive GGO	Invasive GGO	Total	<i>p</i> - value*					
Combined	Training	70.5	89.8	93.1	61.1	77.1	0.848	0.781– 0.914	0.639
	Validation	66.7	91.7	90.0	71.0	78.4	0.815	0.696– 0.933	

* DeLong's test for ROC curves of training cohort and validation cohort.

SEN = Sensitivity, SPE = Specificity, PPV = Positive Predictive Value, NPV = Negative Predictive Value, AUC = Area Under Curve.

2. Radiomic feature selection

Through texture analysis of each patient's chest CT, 851 radiomic features were obtained, including 107 original features and 8 groups of wavelet features (each group contains 93 wavelet feature factors) obtained by decomposition of the original features (except 14 shape features). With ICC > 0.75 as the screening criterion, 567 stable radiomic features were obtained, including 511 wavelet features and 56 original features.

3. Clinical prediction model

The logistic regression analysis results revealed that maximum_diameter ($p < 0.001$) and CTR ($p < 0.001$) were independent predictors of invasiveness in the training cohort of 144 GGOs. The ROC curve based on these plots was used to represent the clinical prediction model (clinical_training) of clinical features for the invasiveness of GGOs, as shown in Fig. 1. The cut-off value of Maximum_diameter was 1.3 cm (Sensitivity = 0.526, Specificity = 0.918, Positive Predictive Value = 0.926, Negative Predictive Value = 0.5). The cut-off value of CTR was 0.27 (Sensitivity = 0.705, Specificity = 0.837, Positive Predictive Value = 0.893, Negative Predictive Value = 0.594). The formula for calculating the clinical prediction model score was as follows: Clinical-score = $-2.05 + 1.622 * \text{Maximum_diameter} + 3.442 * \text{CTR}$.

4. Radiomic prediction model

After MCP-penalized LASSO regression analysis and 10-fold cross-validation of 567 radiomic features in the training cohort of 144 GGOs, 2 radiomic features were identified as independent predictors of invasiveness: waveletLHLglcmldmn ($p < 0.001$) and waveletLLLglcmInverseVariance ($p < 0.001$). ROC curves were drawn based on these radiomic features. In the prediction model, the AUCs of the texture features ranged from 0.678 to 0.724. The predictive ability of a single texture feature for EGFR mutation was poor. The combined predictive ability of all texture features, radiomic_training, was 0.809, indicating improved predictive ability (Fig. 2). The formula for calculating the radiomic prediction model score was as follows: Radiomic-score = $-47.088 + 51.501 * \text{waveletLHLglcmldmn} - 9.871 * \text{waveletLLLglcmInverseVariance}$.

5. Combined prediction model

Maximum_diameter, CTR and waveletLHLglcmldmn were selected from all the clinical and radiomic features by MCP-penalized LASSO regression analysis and 10-fold cross-validation for building the combined prediction model. The ROC curves are shown in Fig. 3. The combined prediction model score was calculated as follows: Combined-score = $-23.036 + 1.196 * \text{Maximum_diameter} + 3.31 * \text{CTR} + 22.383 * \text{waveletLHLglcmldmn}$.

The sensitivity, specificity, positive predictive value, negative predictive value, accuracy, AUC and 95% CI of each prediction model in the training and validation cohorts were calculated to show the predictive ability (Table 2). No significant difference in AUC values was found between the training and validation cohorts for any of the three prediction models.

6. Nomogram establishment and validation

Based on the 3 predictors selected in the combined model, a nomogram was constructed to predict individual invasiveness of GGOs, as shown in Fig. 4. The C-indexes of the invasion prediction nomogram were 0.848 (95% CI, 0.814–0.881) in the training cohort and 0.815 (95% CI, 0.755–0.874) in the verification cohort. The nomogram was subjected to 1,000 bootstrap resamples for

internal validation, and the calibration curve was plotted (Fig. 5). The mean absolute errors of the calibration curves were 0.019 in the training cohort and 0.05 in the validation cohort.

Discussion

Determining the best treatment for pulmonary GGOs has remained a challenge for thoracic surgeons. The main reason is because it is difficult to classify GGOs before surgery, though a new classification of pulmonary adenocarcinoma was defined in 2011. Preoperative percutaneous CT-guided fine needle aspiration biopsy, endobronchial ultrasonography images and virtual bronchoscopy have been used for the pathological diagnosis of GGOs, but the diagnostic yield remains lower for smaller pGGOs[22–23]. In recent years, an increasing number of studies have been devoted to finding imaging biomarkers for GGO classification through chest CT, especially for the identification of invasive GGO lesions.

On chest CT, preinvasive GGO often appears as a pure GGO, and invasive GGO is more often a larger, mixed GGO [23–26]. Eguchi et al. reported that if the diameter of pure GGO is larger than 11 mm, it is more likely to be invasive[27]. Recently, Li et al. reported a cut-off diameter of 13.5 mm for evaluating the invasiveness in GGO nodules [28]. In 2013, Lee SM et al reported that the cut-off diameter for invasive GGO was 14 mm [16]. A study in 2019 reported that in the part-solid group of GGOs, a diameter larger than 1 cm was a significant factor for predicting invasiveness [29]. In our study, Maximum_diameter was an independent predictor of invasiveness for GGO, and the mean Maximum_diameter values were 0.97 ± 0.42 cm in the preinvasive group and 1.43 ± 0.63 cm in the invasive group ($p < 0.01$); the cut-off diameter for invasive GGO was 1.3 cm. The proportion of solid components of GGOs is often evaluated by CTR. Studies from 2001 to 2006 reported that $CTR > 0.5$ has been suggested as a predictor for pathological invasiveness [14, 30–35]. A multi-institutional prospective study concluded that an adenocarcinoma ≤ 2 cm with a $CTR \leq 0.25$ could be considered as a radiologically non-invasive lung adenocarcinoma[36]. In these studies [14, 30–36], based on the CTR range (0-0.25, 0.25–0.5, 0.5–0.75, 0.75-1) for group statistics, the obtained CTR value was not an accurate cut-off value. In our study, we included CTR as an independent variable into the multivariate regression analysis. In the clinical prediction model, it was determined that $CTR > 0.27$ was a significant predictor of invasive GGO. Zhou et al reported that the combination of $CTR > 0.38$ and $SUV_{max} > 1.46$ seems to be a better predictor of invasive GGO[37], but their study only enrolled patients with mixed GGOs. In our study, the AUCs of Maximum_diameter and CTR from ROC curves were 0.739 and 0.781, respectively, in the prediction model of invasiveness established by clinical features. The combined predictive ability of Maximum_diameter and CTR was found to be improved (AUC=0.84).

Texture analysis (TA) is an important means of medical image processing, as it can measure the tissue heterogeneity characteristics that cannot be observed by human eyes and can quantitatively display the subtle changes of image pixel value and their arrangement. In recent years, only a few studies had introduced TA and radiomic features of chest CT into the study of differentiating invasive pulmonary GGOs. Hee-Dong Chae et al [38] concluded that in part-solid GGOs, higher kurtosis and smaller mass could significantly differentiate preinvasive lesions from invasive pulmonary adenocarcinoma (IPA). Wei li et al [39] found that the voxel count and correlation features were significant differentiators of preinvasive lesions from IPAs and MIAs. Another 2018 study found that pure or mixed lesions (PM) and fractal dimension (FD) were predictors of invasive adenocarcinomas appearing as GGOs [40]. In a recent study [41], a support vector machine (SVM) trained on all the heterogeneity indicators showed higher accuracy in differentiating (88.1%) between indolent and invasive lesions. In the above studies, only two- or three-dimensional original texture features were used, and certain amounts of wavelet transform features were excluded. To the best of our knowledge, this is the first study to incorporate the 107 original features and 8 groups of wavelet features (each group contains 93 wavelet feature factors) into radiomic predictor selection for differentiating the invasiveness of GGO. For those high-dimensional data, to avoid overfitting in the prediction, we used MCP-penalized LASSO regression and 10-fold cross-validation analysis to identify relevant variables for subsequent establishment of the radiomic prediction model. Finally, waveletLHLglcldmn and waveletLLLglcmlInverseVariance were selected, and their AUCs were 0.724 and 0.678, respectively. However, the combination of the two radiomic predictors showed a better predictive ability of invasiveness (AUC = 0.809).

In the combined prediction model for differentiating the invasiveness of GGO, predictors were re-selected from all 6 clinical features and 567 extracted radiomic features. Maximum_diameter, CTR and waveletLHLglcldmn were enrolled. The predictive

ability of the combined prediction model (AUC = 0.848) was better than that of any single prediction model developed by clinical features (AUC = 0.84) or radiomic features (AUC = 0.809). The nomogram for the individual prediction model was constructed by the three predictors. The C-index (0.848 and 0.815 for the training and validation cohorts, respectively) and the calibration curve showed that the nomogram had good predictive ability.

Our study has some limitations. Due to the small sample size, we did not analyse the pure GGO (pGGO) group alone. Although we speculated that Maximum_diameter and some radiomic features might show a good predictive ability for invasiveness in the pGGO group, further confirmation by a large sample study is needed in the future.

Conclusion

The combined prediction model constructed by clinical and radiomic predictors showed a good predictive ability for invasiveness in GGOs. Maximum_diameter > 1.3 cm, CTR > 0.27 and a higher waveletLHLgldm may indicate a higher probability of invasive GGO. Our study may help thoracic surgeons better determine the best treatment strategy for patients with pulmonary GGOs.

Declarations

Funding

This study was supported by the Beijing Municipal Administration of Hospital Clinical Medicine Development of Special Funding Support (XMLX201702) awarded to Dr Zhang Yi.

Competing interests

There are no conflicts of interest.

Consent for publication

Not applicable.

Availability of data and materials

On reasonable request, the datasets used during the current study are available from the corresponding author.

Code availability

All codes used with R are available in the Appendix.

Authors' contributions

Dang YT was responsible for conceiving and designing the study, data analysis, writing of the manuscript and all manuscript revisions. Wang RT and Qian K were responsible for patient data collection. Lu J was responsible for CT data collection and editing of the manuscript. Zhang Y was responsible for project conceptualization, manuscript revisions and editing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

[1] Tsutsui S, Ashizawa K, Minami K et al. Multiple focal pure ground-glass opacities on high-resolution CT images: Clinical significance in patients with lung cancer. *AJR Am J Roentgenol* 2010; 195: W131–8.

- [2] Miller A, Markowitz S, Manowitz A, Miller JA. Lung cancer screening using low-dose high-resolution CT scanning in a high-risk workforce: 3500 nuclear fuel workers in three US states. *Chest* 2004; 125 (5 Suppl): 152S–3S.
- [3] Kim HY, Shim YM, Lee KS, Han J, Yi CA, Kim YK. Persistent pulmonary nodular ground-glass opacity at thin-section CT: histopathologic comparisons. *Radiology* 2007;245(1):267-75.
- [4] Cho S, Yang H, Kim K, Jheon S. Pathology and prognosis of persistent stable pure ground-glass opacity nodules after surgical resection. *Ann Thorac Surg* 2013; 96: 1190–5.
- [5] Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 2011;6:244–85.
- [6] Dai J, Yu G, Yu J. Can CT imaging features of ground-glass opacity predict invasiveness? A meta-analysis. *Thorac Cancer* 2018;9(4):452-458.
- [7] Park CM, Goo JM, Lee HJ, et al. CT findings of atypical adenomatous hyperplasia in the lung. *Korean J Radiol* 2006;7(2):80–86.
- [8] Nagao M, Murase K, Yasuhara Y, et al. Measurement of localized ground-glass attenuation on thin-section computed tomography images: correlation with the progression of bronchioloalveolar carcinoma of the lung. *Invest Radiol* 2002;37(12):692–697.
- [9] Van Schil PE, Asamura H, Rusch VW, et al. Surgical implications of the new IASLC/ATS/ERS adenocarcinoma classification. *Eur Respir J* 2012;39(2):478–486.
- [10] Koike T, Togashi K, Shirato T, et al. Limited resection for noninvasive bronchioloalveolar carcinoma diagnosed by intraoperative pathologic examination. *Ann Thorac Surg* 2009;88(4):1106–1111.
- [11] Pedersen JH, Saghir Z, Wille MM, Thomsen LH, Skov BG, Ashraf H. Ground-glass opacity lung nodules in the era of lung cancer CT screening: Radiology, pathology, and clinical management. *Oncology (Williston Park)* 2016; 30: 266–74.
- [12] Asamura H, Hishida T, Suzuki K, et al. Radiographically determined noninvasive adenocarcinoma of the lung: survival outcomes of Japan Clinical Oncology Group 0201. *J Thorac Cardiovasc Surg* 2013;146:24-30.
- [13] Naidich DP, Bankier AA, MacMahon H, et al. Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the Fleischner Society. *Radiology* 2013;266(1):304–317.
- [14] Ohde Y, Nagai K, Yoshida J, et al. The proportion of consolidation to ground-glass opacity on high resolution CT is a good predictor for distinguishing the population of non-invasive peripheral adenocarcinoma. *Lung Cancer* 2003;42(3):303–310.
- [15] Kim EA, Johkoh T, Lee KS, et al. Quantification of ground-glass opacity on high-resolution CT of small peripheral adenocarcinoma of the lung: pathologic and prognostic implications. *AJR Am J Roentgenol* 2001;177(6):1417–1422.
- [16] Lee SM, Park CM, Goo JM, Lee HJ, Wi JY, Kang CH. Invasive pulmonary adenocarcinomas versus preinvasive lesions appearing as ground-glass nodules: Differentiation by using CT features. *Radiology* 2013; 268 (1): 265–73.
- [17] P. Lambin, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48(4):441-6.
- [18] R. Wilson, A. Devaraj. Radiomics of pulmonary nodules and lung cancer. *Transl Lung Cancer Res* 2017;6(1):86-91.
- [19] F. Han, et al. Texture feature analysis for computer-aided diagnosis on pulmonary nodules. *J Digit Imaging* 2015;28(1):99-115.
- [20] V. Kumar, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012;30(9):1234-48.

- [21] Moon Y, Park JK, Lee KY, Ahn S, Shin J. Predictive factors for invasive adenocarcinoma in patients with clinical non-invasive or minimally invasive lung cancer. *J Thorac Dis* 2018;10(11):6010-6019.
- [22] Shimizu K, Ikeda N, Tsuboi M, et al. Percutaneous CT-guided fine needle aspiration for lung cancer smaller than 2 cm and revealed by ground-glass opacity at CT. *Lung Cancer* 2006;51:173-9.
- [23] Ikezawa Y, Shinagawa N, Sukoh N, et al. Usefulness of Endobronchial Ultrasonography With a Guide Sheath and Virtual Bronchoscopic Navigation for Ground-Glass Opacity Lesions. *Ann Thorac Surg* 2017;103:470-5.
- [23] Lee HY, Lee KS, Han J, et al. Mucinous versus nonmucinous solitary pulmonary nodular bronchioloalveolar carcinoma: CT and FDG PET findings and pathologic comparisons. *Lung Cancer* 2009;65:170-5.
- [24] Zhang Y, Qiang JW, Ye JD, et al. High resolution CT in differentiating minimally invasive component in early lung adenocarcinoma. *Lung Cancer* 2014;84:236-41.
- [25] Lee HJ, Lee CH, Jeong YJ, et al. IASLC/ATS/ERS International Multidisciplinary Classification of Lung Adenocarcinoma: novel concepts and radiologic implications. *J Thorac Imaging* 2012;27:340-53.
- [26] Kobayashi Y, Ambrogio C, Mitsudomi T. Ground-glass nodules of the lung in never-smokers and smokers: clinical and genetic insights. *Transl Lung Cancer Res* 2018;7(4):487-497.
- [27] Eguchi T, Yoshizawa A, Kawakami S et al. Tumor size and computed tomography attenuation of pulmonary pure ground-glass nodules are useful for predicting pathological invasiveness. *PLOS One* 2014; 9 (5): e97867.
- [28] Li M, Wang Y, Chen Y, Zhang Z. Identification of preoperative prediction factors of tumor subtypes for patients with solitary ground-glass opacity pulmonary nodules. *J Cardiothorac Surg* 2018; 13 (1): 9.
- [29] Chen PH, Chang KM, Tseng WC, Chen CH, Chao JI. Invasiveness and surgical timing evaluation by clinical features of ground-glass opacity nodules in lung cancers. *Thorac Cancer* 2019;10(11):2133-2141.
- [30] Asamura H, Suzuki K, Watanabe S, et al. A clinicopathological study of resected subcentimeter lung cancers: a favorable prognosis for ground glass opacity lesions. *Ann Thorac Surg* 2003;76:1016-22.
- [31] Ikeda N, Maeda J, Yashima K, et al. A clinicopathological study of resected adenocarcinoma 2 cm or less in diameter. *Ann Thorac Surg* 2004;78:1011-6.
- [32] Suzuki K, Kusumoto M, Watanabe S, et al. Radiologic classification of small adenocarcinoma of the lung: radiologic-pathologic correlation and its prognostic impact. *Ann Thorac Surg* 2006;81:413-9.
- [33] Aoki T, Tomoda Y, Watanabe H, et al. Peripheral lung adenocarcinoma: correlation of thin-section CT findings with histologic prognostic factors and survival. *Radiology* 2001;220:803-9.
- [34] Matsuguma H, Yokoi K, Anraku M, et al. Proportion of ground-glass opacity on high-resolution computed tomography in clinical T1 N0 M0 adenocarcinoma of the lung: A predictor of lymph node metastasis. *J Thorac Cardiovasc Surg* 2002;124:278-84.
- [35] Nakata M, Sawada S, Yamashita M, et al. Objective radiologic analysis of ground-glass opacity aimed at curative limited resection for small peripheral non- small cell lung cancer. *J Thorac Cardiovasc Surg* 2005;129:1226-31.
- [36] Suzuki K, Koike T, Asakawa T, et al. A prospective radiological study of thin-section computed tomography to predict pathological noninvasiveness in peripheral clinical IA lung cancer (Japan Clinical Oncology Group 0201). *J Thorac Oncol* 2011;6:751-6.
- [37] Zhou J, Li Y, Zhang Y, et al. Solitary ground-glass opacity nodules of stage IA pulmonary adenocarcinoma: combination of 18F-FDG PET/CT and high-resolution computed tomography features to predict invasive adenocarcinoma. *Oncotarget* 2017;8:22212-21.

[38] Hee-Dong Chae, Chang Min Park, Sang Joon Park, Sang Min Lee, Kwang Gi Kim, et al. Computerized Texture Analysis of Persistent Part-Solid Ground-Glass Nodules: Differentiation of Preinvasive Lesions From Invasive Pulmonary Adenocarcinomas. *Radiology* 2014;273(1):285-93.

[39] Wei Li, Xuexiang Wang, Yuwei Zhang, Xubin Li, Qian Li, et al. Radiomic Analysis of Pulmonary Ground-Glass Opacity Nodules for Distinction of Preinvasive Lesions, Invasive Pulmonary Adenocarcinoma and Minimally Invasive Adenocarcinoma Based on Quantitative Texture Analysis of CT. *Chin J Cancer Res* 2018;30(4):415-424.

[40] Xing Xue, Yong Yang, Qiang Huang, Feng Cui, Yuqing Lian, et al. Use of a Radiomics Model to Predict Tumor Invasiveness of Pulmonary Adenocarcinomas Appearing as Pulmonary Ground-Glass Nodules. *Biomed Res Int* 2018:6803971.

[41] Ming Li, Vivek Narayan, Ritu R Gill, Jyothi P Jagannathan, Maria F Barile, et al. Computer-Aided Diagnosis of Ground-Glass Opacity Nodules Using Open-Source Software for Quantifying Tumor Heterogeneity. *R Am J Roentgenol* 2017; 209(6):1216-1227.

Figures

Figure 1

Clinical prediction model of invasiveness. (A) ROC curve for invasiveness prediction of GGOs with clinical predictors separately and combined in the training cohort. (B) ROC curves for the training cohort (radiomic_training) and the validation cohort (radiomic_validation); the corresponding AUCs were 0.84 and 0.819 (p=0.764).

Figure 2

Radiomic feature selection and the development of the radiomic prediction model. (A) The LASSO algorithm and 10-fold cross-validation for radiomic predictor selection. The optimal lambda is 0.075 at the minimum cross-validation error (1.14), and the corresponding number of predictors is 2. (B) ROC curve for invasiveness prediction with radiomic predictors separately and combined in the training cohort. (C) ROC curves for the training cohort (radiomic_training) and the validation cohort (radiomic_validation); the corresponding AUCs were 0.809 and 0.693 ($p=0.176$).

Figure 3

Development of the combined prediction model. (A) The LASSO algorithm and 10-fold cross-validation for combined predictor selection. When the minimum cross validation error is 1.02, the optimal lambda value is 0.079, and the corresponding number of nonzero coefficients is 3. (B) ROC curves of the combined prediction model for the training cohort (combined_training AUC=0.848) and the validation cohort (combined_validation AUC=0.815). (C) ROC curves are depicted to describe the discrimination of the clinical prediction model (clinical_training), the general imaging prediction model (imaging_training), the radiomic prediction model (radiomic_training) and the combined prediction model (combined_training).

Figure 4

A nomogram incorporating all the significant predictors for invasiveness was constructed with the training cohort. The predictors include Maximum_diameter, CTR and radiomic (waveletLHLglcmldmn). The sum of points received for each variable value is located on the total points axis, and a line was drawn downward to the prediction axis to determine the invasiveness probability.

Figure 5

Calibration curve of the nomogram for predicting the probability of invasiveness in the training cohort (A) and the validation cohort (B). The actual invasiveness probability is plotted on the y-axis; the nomogram-predicted invasiveness probability is plotted on the x-axis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix.docx](#)