frontiers in Oncology

ORIGINAL RESEARCH
published: 11 August 2020
doi: 10.3389/fonc.2020.00872

# Radiomic-Based Quantitative CT Analysis of Pure Ground-Glass Nodules to Predict the Invasiveness of Lung Adenocarcinoma

*Fangyi Xu[1], Wenchao Zhu[1], Yao Shen[1,2], Jian Wang[3], Rui Xu[4,5], Chooah Outesh[1], Lijiang Song[6], Yi Gan[7], Cailing Pu[1] and Hongjie Hu[1]**

[1] Department of Radiology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2] Department of Radiology, Yinzhou Hospital Affiliated With the School of Medicine of Ningbo University, Ningbo, China, [3] Department of Radiology, Tongde Hospital of Zhejiang Province, Hangzhou, China, [4] DUT-RU International School of Information Science & Engineering, Dalian University of Technology, Dalian, China, [5] DUT-RU Co-Research Center of Advanced ICT for Active Life, Dalian, China, [6] Department of Cardiothoracic Surgery, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China, [7] Department of Pathology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China

OPEN ACCESS

**Edited by:**
Marco Lucchi,
University of Pisa, Italy

**Reviewed by:**
Ahmad Chaddad,
Guilin University of Electronic Technology, China
Niha Beig,
Case Western Reserve University, United States
Shuling Chen,
First Affiliated Hospital of Sun Yat-sen University, China

***Correspondence:**
Hongjie Hu
hongjiehu@zju.edu.cn

**Specialty section:**
This article was submitted to Thoracic Oncology, a section of the journal Frontiers in Oncology

**Received:** 19 December 2019
**Accepted:** 04 May 2020
**Published:** 11 August 2020

**Citation:**
Xu F, Zhu W, Shen Y, Wang J, Xu R, Outesh C, Song L, Gan Y, Pu C and Hu H (2020) Radiomic-Based Quantitative CT Analysis of Pure Ground-Glass Nodules to Predict the Invasiveness of Lung Adenocarcinoma. Front. Oncol. 10:872. doi: 10.3389/fonc.2020.00872

**Objectives:** To investigate the performance of radiomic-based quantitative analysis on CT images in predicting invasiveness of lung adenocarcinoma manifesting as pure ground-glass nodules (pGGNs).

**Methods:** A total of 275 lung adenocarcinoma cases, with 322 pGGNs resected surgically and confirmed pathologically, from January 2015 to October 2017 were enrolled in this retrospective study. All nodules were split into training and test cohorts randomly with a ratio of 4:1 to establish models to predict between pGGN-like adenocarcinoma *in situ* (AIS)/minimally invasive adenocarcinoma (MIA) and invasive adenocarcinoma (IVA). Radiomic feature extraction was performed using Pyradiomics with semi-automatically segmented tumor regions on CT scans that were contoured with an in-house plugin for 3D-Slicer. Random forest (RF) and support vector machine (SVM) were used for feature selection and predictive model building in the training cohort. Three different predictive models containing conventional, radiomic, and combined models were built on the basis of the selected clinical, radiological, and radiomic features. The predictive performance of each model was evaluated through the receiver operating characteristic curve (ROC) and the area under the curve (AUC). The predictive performance of two radiologists (A and B) and our radiomic predictive model were further investigated in the test cohort to see if radiomic predictive model could improve radiologists' performance in prediction between pGGN-like AIS/MIA and IVA.

**Results:** Among 322 nodules, 48 (14.9%) were AIS and 102 (31.7%) were MIA with 172 (53.4%) for IVA. Age, diameter, density, and nine meaningful radiomic features were selected for model building in the training cohort. Three predictive models showed good performance in prediction between pGGN-like AIS/MIA and IVA (AUC $> 0.8$, $P < 0.05$) in both training and test cohorts. The AUC values in the test cohort were 0.824 (95% CI, 0.723–0.924), 0.833 (95% CI, 0.733–0.934), and 0.848 (95% CI, 0.750–0.946) for conventional, radiomic, and combined models, respectively. The predictive accuracy was

73.44 and 59.38% for radiologist A and radiologist B in the test cohort and was improved dramatically to 79.69 and 75.00% with the aid of our radiomic predictive model.

**Conclusion:** The predictive models built in our study showed good predictive power with good accuracy and sensitivity, which provided a non-invasive, convenient, economic, and repeatable way for the prediction between IVA and AIS/MIA representing as pGGNs. The radiomic predictive model outperformed two radiologists in predicting pGGN-like AIS/MIA and IVA, and could significantly improve the predictive performance of the two radiologists, especially radiologist B with less experience in medical imaging diagnosis. The selected radiomic features in our research did not provide more useful information to improve the combined predictive model's performance.

**Keywords: radiomics, lung cancer, adenocarcinoma, computed tomography, machine learning**

## BACKGROUND

A new classification for lung adenocarcinoma was proposed in 2011 by the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society (IASLC/ATS/ERS) (1), which was also issued as the 4th edition WHO lung cancer classification in 2015 (2). According to the new classification, lung adenocarcinoma can be divided into preinvasive lesion, minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (IVA), and preinvasive lesion includes atypical adenomatous hyperplasia (AAH) and adenocarcinoma *in situ* (AIS) (1). The improvement of medical technology and the generalization of lung cancer screening project have led more attention to pure ground-glass nodules (pGGNs) detected on computed tomography (CT) images (3, 4).

Approximately 20% of lung adenocarcinoma including AIS, MIA, and even some early-stage IVA could present as pGGNs on CT images (4), which makes it quite difficult for radiologists and clinicians to make a precise diagnosis with conventional radiological parameters like size, density, etc. Kakinuma et al. reported that growth was observed in approximately 10% of pGGNs $\leq$5 mm, of which 1% would develop into IVA or MIA in their study (5). In another study, 57.8% of pGGNs showed growth during follow-up and 26.3% of them were adenocarcinoma (6). Eguchi et al. examined 124 cases with pGGNs, and 64 pGGNs (51.6%) showed growth during their 2-year follow-up (7). Several previous research revealed that nearly 50% of pGGNs were invasive lesions (3, 8–10). In clinical practice, pGGNs are usually prescribed to be followed up but data above demonstrated that more detailed diagnosis and more individualized management should be made for pGGNs.

Compared with IVA, AIS, and MIA are considered as indolent lung adenocarcinoma because of the excellent prognosis (11, 12). AIS/MIA could be followed up or treated with sublobar resection while more aggressive surgical interventions should be taken for IVA (13, 14). Several previous studies revealed that the 5-year survival rates of AIS and MIA could be 100% and near 100% with a complete resection while that of IVA in stage Ia is no more than 75% (11, 13, 14). Thus, it might provide some guidance for clinical therapeutic decision-making if pGGN-like IVA could be figured out on preoperative CT images.

There were many studies investigating the difference in radiological features among lung adenocarcinoma subtypes. Wang et al. found that the mean CT attenuation and lesion size differed significantly between MIA and non-invasive lesions and internal air bronchograms were more often seen in adenocarcinoma (15). Several investigations reported that the notched signs, spiculations, bubbly lucencies, and rapid volume expansion were more common in IVA (15–17). However, those radiological features could be subtly different because of the small size of pGGN-like adenocarcinoma. Furthermore, the assessment of those parameters tends to be subjective, which could be influenced by radiologists' experience and diagnostic ability. Percutaneous biopsy is one method used in clinical practice to determine the nature of pulmonary nodules, which could provide relatively accurate pathological information. However, percutaneous biopsy is an invasive operation, and patients may have some operation-related complications (18, 19). Considering the heterogeneity in adenocarcinoma, small pieces of tissue obtained by biopsy cannot represent the characteristics of the whole lesion (20). What is more, in some cases, it is difficult to complete biopsy due to patient's physical condition and bad cooperation as well as the location and size of nodules (20, 21). Thus, the accurate diagnosis of pGGNs remains a key point and a challenge in the field of medical imaging diagnosis.

Radiomics is an emerging subject that could extract a large amount of invisible features from medical images for clinical decision-making (20, 22). Radiomics has had remarkable progress in central nervous system malignancies, thoracic imaging diagnosis, discrimination of hepatic mass, and some other diseases (23–26). Chaddad et al. performed retrospective analysis involving 315 patients diagnosed as non-small cell lung cancer (NSCLC) and significant correlation was observed between radiomic features and survival (27). Also, radiomics' promising performance in the distinction of benign and malignant pulmonary nodules and the discrimination of adenocarcinoma subtypes had been validated in several researches (11, 28, 29). However, in most previous radiomic studies, all types of pulmonary nodules including solid and subsolid nodules were recruited as the study population. Few studies focused on the use of machine learning in early-stage lung adenocarcinoma representing as pGGNs, which are usually very

difficult to manage. Since the diagnosis of solid components in pulmonary nodules on CT images is relatively uncomplicated while pGGNs remain a big challenge for medical imaging diagnosis, we aimed to explore the potential value of radiomic-based quantitative analysis to predict the invasiveness of pGGN-like adenocarcinoma to establish a comprehensive predictive model for clinical decision-making.

## MATERIALS AND METHODS

This retrospective study was approved by the institutional review committee of the Sir Run Run Shaw Hospital (No. 20190520-162) with an abstention of informed consents from all the patients involved according to the guidelines of the Council for International Organizations of Medical Science (CIOMS).

### Patients

We reviewed all the materials of 1,610 patients undergoing surgical resection for primary lung adenocarcinoma with complete clinical data and preoperative CT images from January 2015 to October 2017 in Sir Run Run Shaw Hospital, and reinterpreted the preoperative CT images from the Picture Archiving and Communication Systems (PACS) one by one. Clinical data like age, gender, and smoking status of all cases were collected from digital records. Patients who met any one of the following criteria were excluded: (1) nodules with solid components ($n = 776$), (2) nodule diameter >3 cm ($n = 292$), (3) patients with a history of other malignant diseases ($n = 87$), (4) CT images with bad quality ($n = 70$), and (5) patients who accepted thoracic surgical intervention, radiation, or any chemotherapeutics ($n = 110$).

Finally, 275 patients (72 men and 203 women, age range, 25∼78 years) with 322 pGGNs (82 men and 240 women, age range, 25∼78 years) were enrolled into this retrospective study (detailed in **Figure 1** and **Table 1**). The median time from the last preoperative CT scan to surgery was 6 (0–92) days.

### Histological Evaluation

All surgical specimens were fixed with formalin and stained with hematoxylin–eosin (HE). Two pathologists evaluated all slides using a multi-headed microscope and discussed about the diagnosis until a consensus was reached. According to the classification of lung adenocarcinoma issued in the 4th edition WHO lung cancer classification in 2015, each nodule was classified as AIS, MIA, and IVA (26, 30). Each histological pattern presented in targeted lesion including lepidic, acinar, solid, papillary, and micropapillary patterns was recorded in 5% increments (30, 31).

### Image Acquisition

All the plain CT images were obtained using multidetector computed tomography scanners (Siemens SOMATOM Definition Flash, Siemens FORCE CT, Siemens Sensation 16, Siemens Definition AS 40, and GE LightSpeed VCT). The protocol parameters for each scanning were detailed in **Supplemental Table 1**. Plain spiral acquisitions were obtained from thoracic inlet to lung bases on patients accepting breath-hold training. Images were reconstructed using a standard reconstruction kernels in lung window settings (mean, −500 HU; width, 1,500 HU). All images underwent multi-planar reconstruction (MPR) including coronal and sagittal reconstruction utilizing the post-processing station.

### Nodule Analysis and Segmentation

All transverse CT images were interpreted jointly by two radiologists who were both blind to the clinical and pathological information of all cases (radiologist A, with 10 years' experience in thoracic imaging diagnosis; radiologist B, with 2 years' experience in medical imaging diagnosis) on our professional reading screen. Conventional quantitative radiological features that were widely used in clinical diagnosis involving diameter (cm) and density (Hounsfield Unit, HU) were determined for each nodule. Nodule diameter was measured on the average of long- and short-axis diameters, both of which should be obtained on the same transverse revealing the greatest dimensions. The nodule density was measured at three different parts of each nodule avoiding vessels and bronchus and the mean value of the three results was calculated. The mean value of diameter and density measured by the two radiologists was calculated for our study.

Segmentation data consisted of all the 322 pGGNs. Plain CT images were loaded into 3D-slicer (http://www.slicer.org) (32), an open source image processing software, implemented with in-house algorithm for automatic nodule detection and segmentation. Radiologist B would verify the regions of interest (ROIs) of automatic segmentation and made some modifications when the ROIs were not satisfactory. Radiologist A would have a second review for the results of radiologist B's semi-automatic segmentation. A consensus would be achieved via negotiation between two radiologists for each case when meeting a collision on reviewing. While modifying ROIs, two radiologists would delineate manually around the nodule boundary on each section avoiding the bronchus and vessels as much as they could.

### Radiomic Feature Extraction and Predictive Models Building

Segmentation data were analyzed with Pyradiomics to extract radiomic features describing tumor phenotypes (33). All the segmentation data had a voxel resampling of $0.7 \times 0.7 \times 0.7$ mm$^3$ for standardization to reduce the impact from the heterogeneity of image acquisition. In the end, we obtained nine types totaling 960 radiomic features for each nodule, which have been listed in **Supplemental Table 2**. Features are commonly grouped as follows: (1) first-order statistical features: these describe the voxel intensity distribution in the delineated ROI. They are usually calculated on the basis of the intensity histogram, including energy, entropy, standard deviation, skewness, kurtosis, uniformity, mean, minimum, and maximum intensity values and so on (20, 26). (2) Shape-based features: descriptors of the two- and three-dimensional size and shape of the ROI. (3) Textural features: these contain gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), neighboring

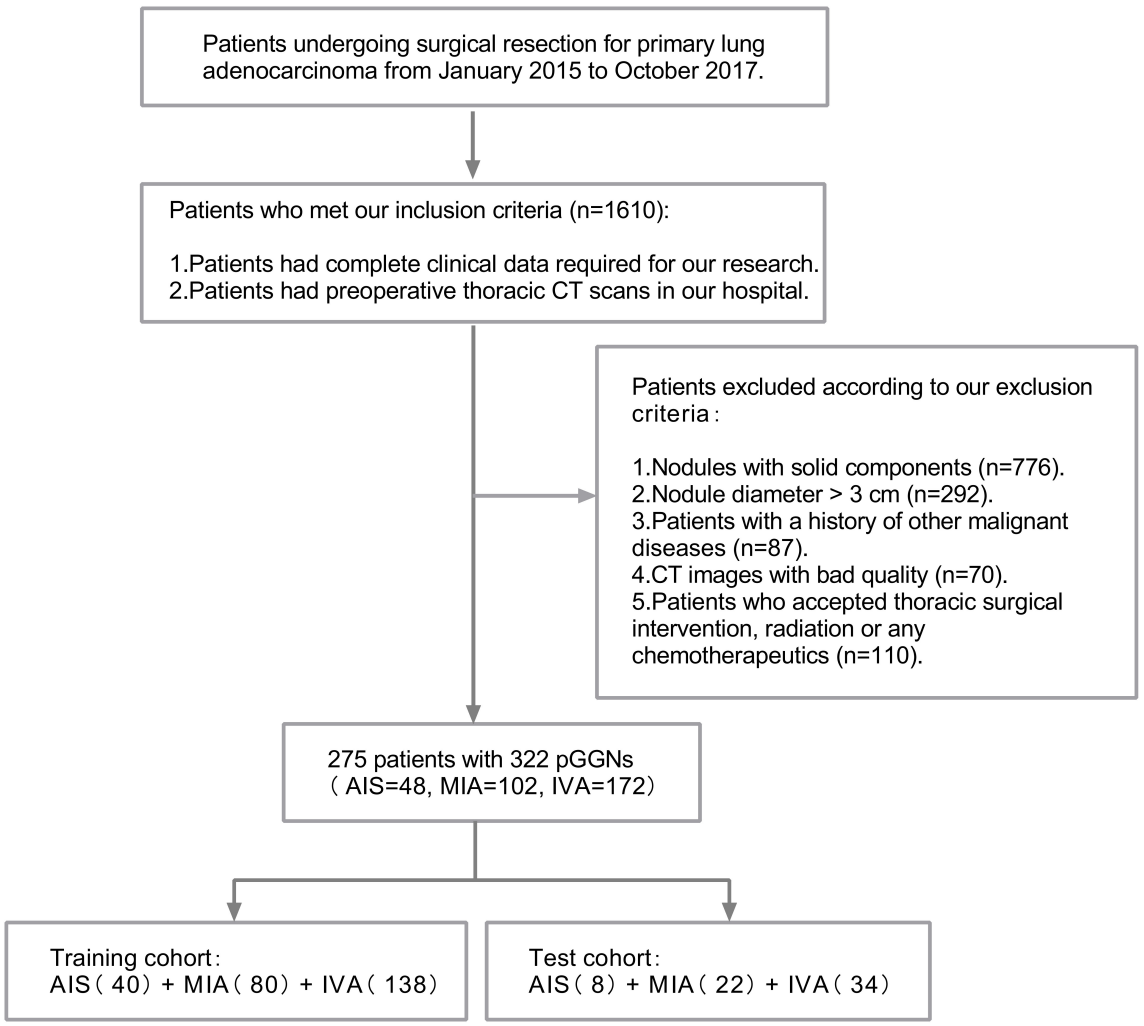


**FIGURE 1 |** The flowchart of patient selection.

gray size zone matrix (NGZDM), and gray-level dependence matrix (GLDM). They are computed on the analysis of the three-dimensional directions within the tumor and the consideration of the spatial location of each voxel in the ROI (26). (4) Transformed features: features in the first and third groups extracted from images applied with a series of wavelet or Laplacian-of-Gaussian filtration.

Random forest (RF) and support vector machine (SVM) with cross-validation (CV) was used for radiomic feature selection and predictive model building to distinct pGGN-like IVA from indolent adenocarcinoma (AIS/MIA). Multivariate models were made in training cohort and were tested in a separate test cohort. All pGGNs were split into training and test cohorts randomly by a ratio of 4:1. Three predictive models were created in our research: (1) Conventional (selected clinical and radiological quantitative features), (2) Radiomic (selected radiomic features), and (3) Combined (selected conventional and radiomic features) predictive models. Subsequently, a binary analysis in which pGGN-like IVA was set as positive while AIS/MIA was thought as negative was applied to compare the predictive performance between radiomic predictive model and two radiologists (A and B) in the test cohort. Two weeks later, two radiologists, knowing the performance of our radiomic predictive model and its diagnosis for each pGGN in the test cohort, reevaluated all pGGNs in the test cohort.

## Statistical Analysis

All the statistical analysis was applied using SPSS 25.0 (IBM, Armonk, NY, USA) and MedCalc 15.8 (MedCalc Software, Acacialaan 22, Ostend, Belgium). Tables and figures in our study were made with GraphPad Prism 5 (GraphPad Software Inc., San Diego, CA, USA) and Microsoft Office 2019 (Microsoft, Redmond, WDC, USA).

Thirty pGGNs were selected randomly to test the repeatability of nodule diameter and density measurement. Radiologist A and B did the measurement work of those 30 pGGNs, respectively. Two weeks later, radiologist B measured the diameter and density of these 30 pGGNs, again according to the same measurement criteria. Inter-/intra-observer correlation coefficient (ICC) was calculated for repeatability assessment.

**TABLE 1** | Analysis of clinical and radiological features of pGGNs in the training and test cohort.

| | Training cohort | | | | Test cohort | | | |
|---|---|---|---|---|---|---|---|---|
| | **Total** | **AIS/MIA** | **IVA** | ***P*** | **Total** | **AIS/MIA** | **IVA** | ***P*** |
| | 258 | 120 (46.5%) | 138 (53.5%) | | 64 | 30 (46.9%) | 34 (53.1%) | |
| **Gender** | | | | | | | | |
| Male | 70 (27.1%) | 31 (25.8%) | 39 (28.3%) | 0.662 | 12 (18.8%) | 3 (10.0%) | 9 (26.5%) | 0.092 |
| Female | 188 (72.9%) | 89 (74.2%) | 99 (71.7%) | | 52 (81.3%) | 27 (90.0%) | 25 (73.5%) | |
| Age (years) | 53 (25, 78) | 51 (25, 69) | 55 (26, 78) | 0.001 | 54 (30, 72) | 51 (31, 72) | 55 (30, 72) | 0.106# |
| **Smoking history** | | | | | | | | |
| Never | 230 (89.1%) | 109 (90.8%) | 121 (87.7%) | 0.417 | 59 (92.2%) | 29 (96.7%) | 30 (88.2%) | 0.360* |
| Current or ever | 28 (10.9%) | 11 (9.2%) | 17 (12.3%) | | 5 (7.8%) | 1 (3.3%) | 4 (11.8%) | |
| Diameter (cm) | 0.94 (0.24, 2.94) | 0.76 (0.24, 1.92) | 1.13 (0.51, 2.94) | <0.001 | 0.93 (0.34, 2.04) | 0.70 (0.34, 1.41) | 1.09 (0.61, 2.04) | <0.001# |
| Density (HU) | −583.0 (−829.2, −122.5) | −615.0 (−801.7, −122.5) | −529.2 (−829.2, −176.3) | <0.001 | −525.8 (−763.2, −179.8) | −544.7 (−763.2, −298.9) | −509.0 (−739.1, −179.8) | 0.088# |

*# Two-sample t-test.*

** Fisher exact probability test.*

*Continuous variables were represented in median (minimum, maximum) without any special statement.*

*pGGNs, pure ground-glass nodules; AIS, adenocarcinoma in situ; MIA, minimally invasive adenocarcinoma; IVA, invasive adenocarcinoma.*

For the assessment of clinical, quantitative radiological, and selected radiomic features, chi-square test or Fisher exact probability test was utilized for categorical variables. Two-sample *t*-test was adopted, if the continuous variables met the normal test and variance homogeneity test; otherwise, Wilcoxon signed-rank test was used. Predictive power of each predictive model was evaluated using receiver operating characteristic (ROC) curve and area under the curve (AUC). Models with an AUC $> 0.50$ and a $P < 0.05$ were thought to be predictive. McNemar's test and Kappa analysis were used to compare the binary diagnosis of two radiologists and radiomic predictive model.

# RESULTS

The measurement of nodule diameter and density between senior radiologist A and junior radiologist B was highly consistent (ICC $> 0.9$, $P < 0.05$). Two weeks later, radiologist B did the measurement for the 30 selected pGGNs, and the ICC values were up to 0.955 ($P < 0.05$) and 0.984 ($P < 0.05$) for diameter and density measurement.

## Clinical Data and Conventional Image Features

A total of 322 pGGNs were recruited into this study with 80% in the training cohort and 20% in the test cohort. The analysis of clinical and quantitative radiological features in the training and test cohort were listed in **Table 1**. In the training cohort, the median age was 53 years (age range, 25–78 years) and the majority of cases were female (72.9%) with 28 (10.9%) having a smoking history. In the test cohort, 52 (81.3%) were female with a median age of 54 years (age range, 30–72 years) and 59 (92.2%) never smoke. Diameter showed statistical discrepancy between AIS/MIA and IVA in both training and test cohort ($P < 0.001$) while nothing significantly different existed in gender and smoking status between the two groups. Age and density exhibited evident difference in training cohort between AIS/MIA and IVA ($P = 0.01$ and $P < 0.001$), but no significant difference ($P > 0.05$) appeared between the two groups in test cohort.

## Radiomic Feature Selection and Predictive Model Building

A RF algorithm with 4-fold cross-validation was taken to calculate the contribution value of each radiomic feature in the training cohort for the prediction of pGGN-like IVA from AIS/MIA. Predictive model building was performed using SVM also combined with 4-fold cross-validation. All the extracted radiomic features were listed in descending order by the contribution value for the classifier and were added one by one as the input for the SVM model training in each iteration of cross-validation calculation (detailed in **Figure 2**). In the process of gradual accumulation of model training, the overall accuracy of the training cohort was recorded. When sequencing extracted radiomic features by their contribution value for the classifier in each iteration, we found that the first 20 radiomic features contributed much more than other features whose contribution values were <0.01 and even close to 0. Meanwhile, the first

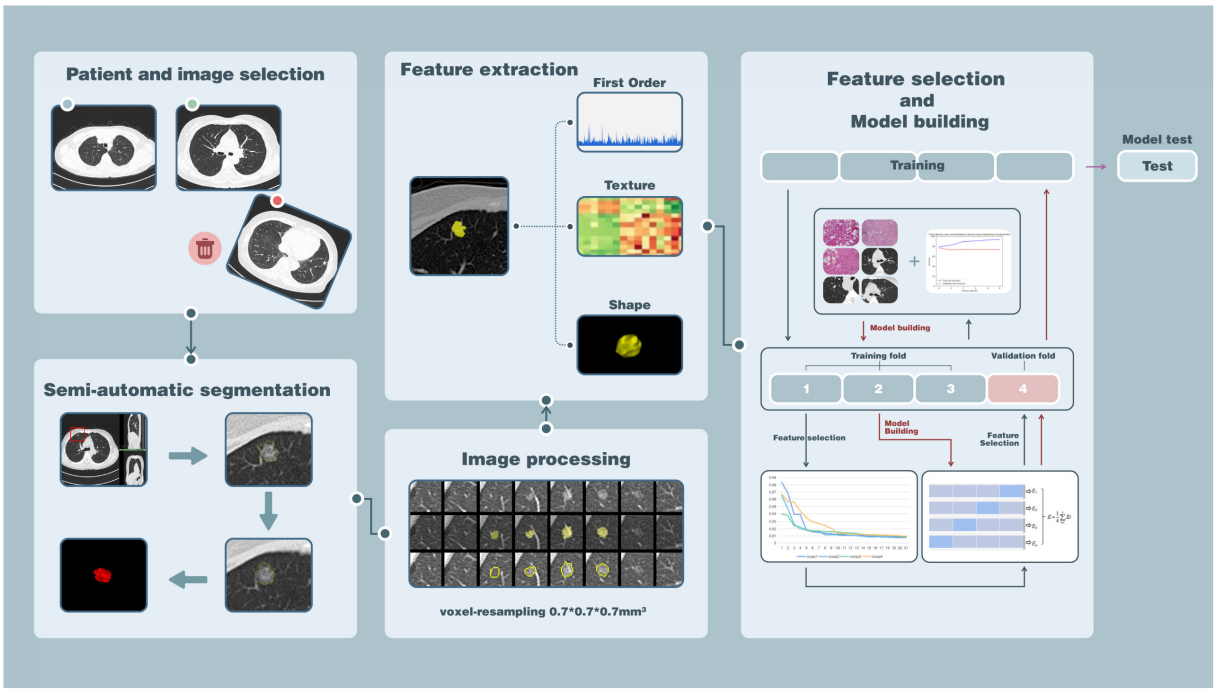


**FIGURE 2 |** The flowchart of the whole study.

**TABLE 2 |** Nine selected radiomic features.

| | Selected radiomic feature | Radiomic group | Filter associated |
|---|---|---|---|
| R1 | Maximum probability | GLCM | wavelet-LLL |
| R2 | Joint entropy | GLCM | None |
| R3 | Sum entropy | GLCM | wavelet-LLL |
| R4 | Joint energy | GLCM | None |
| R5 | Gray level Non-uniformity | GLSZM | log-sigma-1-0-mm-3D |
| R6 | Gray level Non-uniformity | GLSZM | wavelet-LLH |
| R7 | Gray level Non-uniformity | GLSZM | log-sigma-3-0-mm-3D |
| R8 | Size zone Non-uniformity | GLSZM | None |
| R9 | Low gray level run emphasis | GLRLM | log-sigma-3-0-mm-3D |

*GLCM, gray level co-occurrence matrix; GLSZM, gray level size zone matrix; GLRLM, gray level run length matrix.*

20 radiomic features could increase the classifying performance dramatically while inclusion of extra features did not make a big difference to the performance of the SVM classifier in each iteration. To improve the generalization of our predictive model, only radiomic features appearing more than three times in the first 20 of contribution value rank in four iterations of the 4-fold cross validation were selected for final model building. Those radiomic features that were used for final predictive model building incorporated four from GLCM, four from GLSZM, and one from GLRLM. All the nine features are detailed in **Table 2**.

We first investigated whether the nine advanced radiomic features could discriminate IVA from indolent adenocarcinoma (AIS/MIA) representing as pGGNs. **Figure 3** shows the comparison of those nine radiomic features for the distinction of pGGN-like AIS/MIA and IVA in training and test cohorts. All the selected radiomic features revealed significant difference between AIS/MIA and IVA in both training and test cohorts ($P < 0.001$). Among the nine radiomic features, three (Maximum Probability, Joint Energy, and Low Gray Level Run Emphasis) had larger median values for AIS/MIA while the median values of six other features (Joint Entropy, Sum Entropy, Gray Level Non-Uniformity, and Three Different Filtered Size Zone Non-Uniformity) were higher for IVA than that for AIS/MIA in both training and test cohorts (detailed in **Supplemental Table 3**).

Multivariate predictive models were created for each set of features, involving conventional (age, diameter, and density), radiomic (nine predictive features), and the combined (conventional and selected radiomic features) sets. The three different models presented good predictive power (AUC > 0.8, $P < 0.05$) in both training and test cohorts as shown in **Figure 4**, **Supplemental Table 4**. Then, DeLong's test (34, 35) was applied to complete the pairwise predictive performance comparison among the three models, in which no significant difference was observed ($P > 0.05$). **Figure 5** showed comprehensive parameters including accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), misdiagnosis rate (MR), and missed diagnosis rate (MDR) of the three predictive models' and two radiologists' binary diagnosis in the test cohort (detailed in **Supplemental Table 5**). The accuracy was 76.56, 71.88, and 78.13% for radiomic, conventional, and combined predictive models but no big difference was noted in

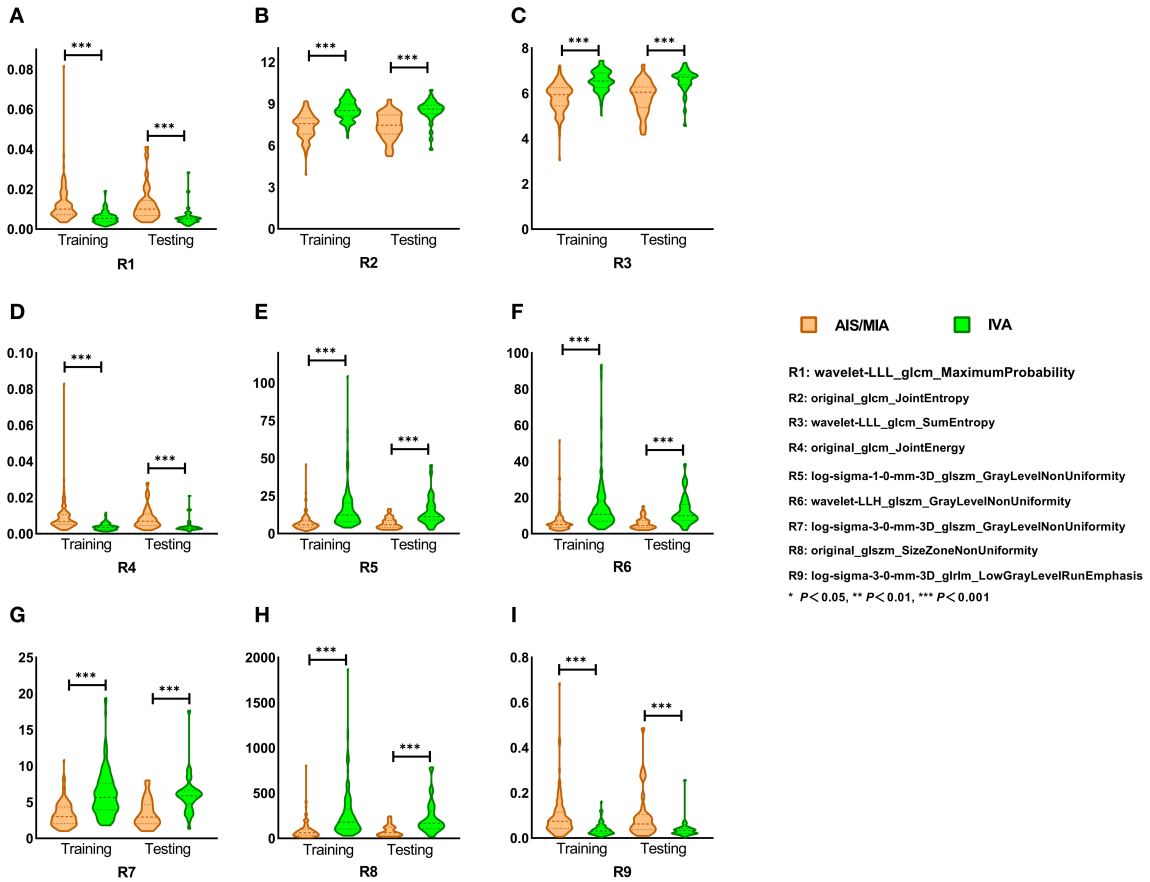


**FIGURE 3 | (A–I)** Selected radiomic features in training and test cohort. Nine selected radiomic features showed significant difference between AIS/MIA and IVA in both training and test cohorts. Maximum Probability, Joint Energy, and Low Gray Level Run Emphasis had larger median values for AIS/MIA, while the median values of Joint Entropy, Sum Entropy, Gray Level Non-Uniformity, and Size Zone Non-uniformity were higher for IVA in both training and test cohorts.

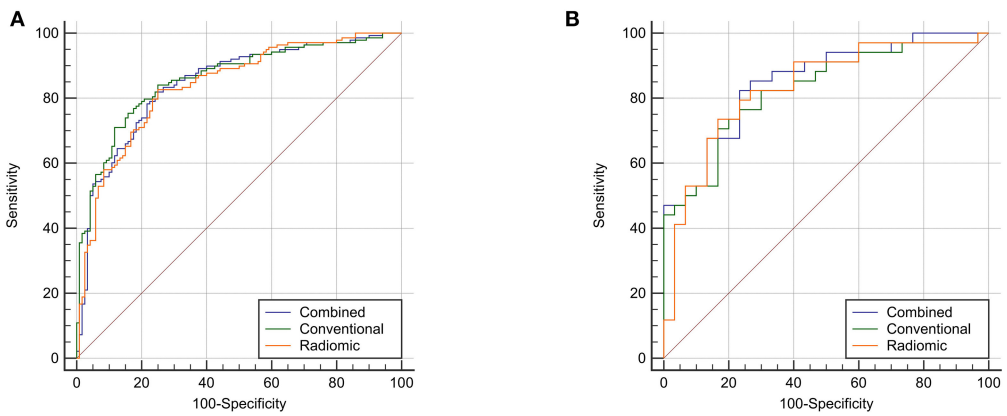


**FIGURE 4 |** The ROC analysis of the three different predictive model. Three predictive models presented good performance in discrimination between pGGN-like IVA and AIS/MIA. **(A)**, the predictive performance of three models in the 693 training cohort; **(B)**, the predictive performance of three models in the testing cohort.

comprehensive assessment among the three models (**Figure 5**), which was consistent with the results of DeLong's test above. In a word, no matter what features were used for model training (conventional or radiomic features), predictive models built with machine learning algorithm could predict pGGN-like IVA from AIS/MIA well. The combination of conventional and

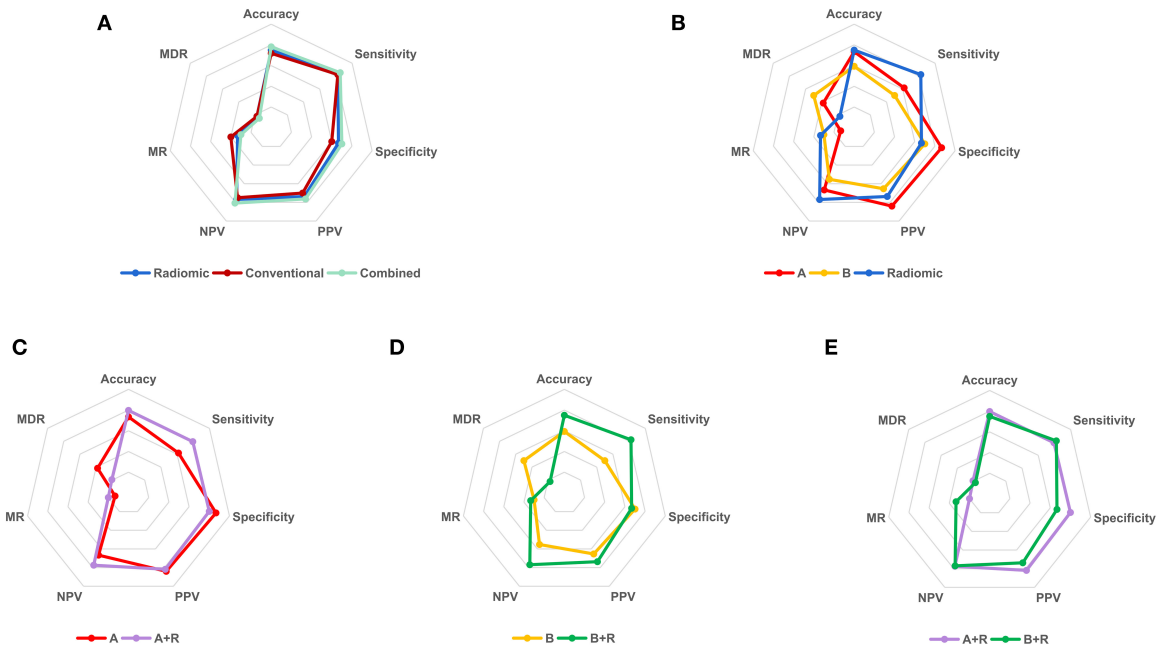


**FIGURE 5 |** The binary diagnosis of predictive models and two radiologists. (A), the first binary diagnosis of the senior radiologist A with 10-year experience in thoracic imaging diagnosis in test cohort; (B), the first binary diagnosis of the junior radiologist B with 2-year experience in medical imaging diagnosis in test cohort; R, the binary diagnosis of radiomic predictive model in test cohort; A+R, the second binary diagnosis of radiologist A in test cohort with the aid of radiomic predictive model; B+R, the second binary diagnosis of radiologist B in test cohort with the aid of radiomic predictive model. PPV, positive predictive value; NPV, negative predictive value; MR, misdiagnosis rate; MDR, missed diagnosis rate.

**TABLE 3 |** Performance comparison between radiomic predictive model and two radiologists in test cohort.

| | McNemar's test | | Kappa analysis | |
|---|---|---|---|---|
| | $\chi^2$ | *P* | $\kappa$ | *P* |
| A vs. B | 0 | 1.000 | 0.316 | 0.018 |
| A vs. R | 7.563 | 0.004 | 0.517 | <0.001 |
| B vs. R | 4 | 0.043 | 0.241 | 0.070 |
| A vs. A+R | 4.9 | 0.021 | 0.690 | <0.001 |
| B vs. B+R | 5.042 | 0.023 | 0.275 | 0.022 |
| A+R vs. B+R | 1.455 | 0.227 | 0.654 | <0.001 |

*A, the first binary diagnosis of the senior radiologist A with 10-year experience in thoracic imaging diagnosis in test cohort; B, the first binary diagnosis of the junior radiologist B with 2-year experience in medical imaging diagnosis in test cohort; R, the binary diagnosis of radiomic predictive model in test cohort; A+R, the second binary diagnosis of radiologist A in the test cohort with the aid of radiomic predictive model; B+R, the second binary diagnosis of radiologist B in the test cohort with the aid of radiomic predictive model.*

radiomic features could further improve the diagnosis accuracy of predictive model, but the improvement was not statistically significant in our study.

To investigate whether the radiomic predictive model could help radiologists improve their predictive performance, we then compared the diagnosis of radiomic predictive model and two radiologists (**Figure 5**, **Table 3**). Radiologist A performed better than radiologist B with higher diagnostic accuracy, sensitivity, specificity, PPV, and NPV. Either accuracy or sensitivity, radiomic predictive model outperformed radiologist A with the cost of decreased specificity. Significant difference was observed between the binary diagnosis of radiomic predictive model and that of two radiologists (A vs. R, $\chi^2 = 7.563$, $P = 0.004$, B vs. R, $\chi^2 = 4$, $P = 0.043$). Generally speaking, Radiomic predictive model showed better performance than two radiologists in the prediction between pGGN-like IVA and AIS/MIA. Two radiologists dramatically improved their diagnostic accuracy to 79.69 and 75.00% with the aid of radiomic predictive model (A vs. A+R, $\chi^2 = 4.9$, $P = 0.021$, B vs. B+R, $\chi^2 = 5.042$, $P = 0.023$). The comparison of the second diagnosis of two radiologists revealed that when having the guidance from radiomic predictive model, no statistical difference existed between two radiologists in prediction of pGGN-like IVA and AIS/MIA (A+R vs. B+R, $\chi^2 = 1.455$, $P = 0.227$).

## DISCUSSION

When it comes to pure ground-glass pulmonary nodules, clinicians tend to choose follow-up as their first choice for management. However, according to previous studies (4, 5, 31), a certain proportion of IVA that needs surgical treatment could be pGGNs on CT scans. If a pGGN-like IVA was misdiagnosed as a benign lesion or indolent adenocarcinoma and was given a prescription of follow-up, it might progress during the interval or even metastasize and miss the optimal time for surgical intervention. Conventional radiological features like lobulated signs, spiculations, bubble lucencies, and pleura traction have been demonstrated to be helpful to differentiate

the malignancy of pulmonary nodules and the invasiveness of lung adenocarcinoma (36, 37). However, pGGN-like lung adenocarcinoma tends to be in small volume with a large similarity in morphological characteristics and the assessment of the conventional radiological features is easy to be affected by the subjectivity of doctors, it remains a challenge to make a precise judgement for pGGNs without surgical intervention. Percutaneous biopsies may be helpful in determining the nature of pulmonary nodules. Nevertheless, it is an invasive tissue extraction method with the possibility of postoperative complications, and due to the tumor heterogeneity, it is not persuasive to represent the characteristics of the entire lesion with only a tiny tissue (20, 38). In some cases, because of nodules' location and size, percutaneous biopsies may not be a good choice for diagnosis. Thus, the diagnosis of pGGNs remains a thorny point for clinical research.

Heidinger et al. reported that two-dimensional diameter could provide enough information for pulmonary nodule risk classification in their quantitative analysis based on CT images (4, 39). In our study, the diameter of pGGNs showed a significantly different distribution between IVA and indolent adenocarcinoma (AIS/MIA), which was consistent with previous reports. Kitami et al. acclaimed that almost all the pGGNs with a diameter $<10$ mm and a density of no more than $-600$ HU were demonstrated to be preinvasive lesions in their study (40). The medians of density in our training and test cohort were $-583.0$ and $-525.8$ HU, both slightly higher than $-600$ HU, which might have something to do with the difference in lung adenocarcinoma classification. We classified MIA into indolent adenocarcinoma, which might lead to the increase in density medians.

Radiomics is a new quantitative image analysis approach that allows thorough exploration in medical images and has attracted more and more attention in the field of medicine in recent years. In this study, we obtained plenty of quantitative radiomic features (960 for each nodule) from routine thoracic CT images using machine learning techniques and completed the quantitative analysis of pGGN-like adenocarcinoma classification on the basis of conventional clinical, quantitative radiological, and selected radiomic features. Nine selected radiomic features demonstrated good performance in distinction of the IVA and indolent adenocarcinoma representing pGGNs on unenhanced CT images. Nine selected radiomic features in our study consisted of four GLCM-based features, four GLSZM-based features, and one GLRLM-based feature, which were analyzed as follows:

(1) Four radiomic features from GLCM: Maximum Probability, Joint Entropy, Sum Entropy, and Joint Energy. GLCM was used to compare the gray level correlation between two points in a certain distance in a spatial position, which reflects the comprehensive information about pixel distribution including direction, distance, gray value, and the pattern of gray level arrangement (41, 42). Maximum Probability refers to a pair of pixels with the highest frequency in a GLCM (43, 44). Entropy is a parameter describing the complexity of an image, which means the larger the entropy value of an image is, the more complex the image is (43). Energy is related to the uniformity of gray level distribution and the roughness of image texture with a larger energy indicating a more regular and more stable texture (42, 43). In our study, Maximum Probability and Joint Energy got higher median values for AIS/MIA in the training as well as test cohort while Joint Entropy and Sum Entropy in IVA were higher than that in AIS/MIA in both training and test cohort. This might have something to do with that AIS/MIA tend to be homogeneous, which results in a higher probability of finding pixels with same distribution pattern in AIS/MIA and the different entropy and energy values between AIS/MIA and IVA.

(2) Four radiomic features from GLSZM and one from GLRLM: Gray Level Non-Uniformity filtered with LOG or wavelet algorithm, Size Zone Non-Uniformity, and Low Gray Level Run Emphasis. GLSZM refers to the number of pixels that share the same gray level intensity and the same arrangement pattern in an image while GLRLM calculates the number of pixels with the same gray level value and distribution pattern in a certain direction (45–47). Gray Level Non-Uniformity and Size Zone Non-Uniformity indicate the variability of gray level and size zone volumes in an image, with a higher value referring to more heterogeneity in ROIs (46, 47). In our study, the median values of Gray Level Non-Uniformity and Size Zone Non-Uniformity were higher for IVA, which might be related to the fact that IVA tends to be more heterogeneous. Low Gray Level Run Emphasis analyzes the distribution of low gray level values in an image (47). The homogeneity and relatively lower average density of AIS/MIA might lead to the higher median value of Low Gray Level Run Emphasis for AIS/MIA in our study.

Chen et al. picked 76 features meaningful for the distinction of malignancy of pulmonary nodules from 750 extracted radiomic features and built a predictive model whose accuracy was up to 84% using four selected advanced features (20). Yagi et al. carried out the texture analysis of high-resolution computed tomography (HRCT) and found that 90th percentile and entropy performed well in discrimination between AIS/MIA and IVA with an AUC value of 0.90 (95% CI: 0.84–0.95) (13). Three different predictive models set with clinical, radiological, and nine selected radiomic features from 960 features extracted from unenhanced CT images in our study all presented good predictive power in the discrimination between AIS/MIA (indolent adenocarcinoma) and IVA (AUC $> 0.8$, $P < 0.05$). She et al. extracted radiomic features from radiological data of 402 cases (207 for training and 195 for test) diagnosed with lung adenocarcinoma and selected five meaningful radiomic features to build a predictive model that outperformed significantly the predictive model only built with conventional radiological features including nodule diameter for the discrimination between IVA and AIS/MIA (11). However, no apparent difference existed between conventional and radiomic predictive models in our study ($P > 0.05$). Combining conventional and radiomic features could improve the AUC value of combined predictive model, but it was not statistically significant. This might be caused by the difference of our study population. All types of pulmonary nodules including solid and subsolid nodules, which have heterogeneous internal density in lesions, were used for She's study. Therefore, compared with conventional radiological features such as diameter and density, radiomic features could more thoroughly analyze the variability and distribution of gray level intensity in ROI, which

would provide more valuable information for improvement in diagnostic performance of predictive models. However, the relatively obscure variability of gray level intensity in pGGNs might result in the limited reference value for radiomic features in prediction between pGGN-like IVA and AIS/MIA, which potentially led to the similar diagnostic performance among our three predictive models. Nevertheless, predictive models built using either radiomic or conventional radiological features presented good performance in distinction between pGGN-like IVA and AIS/MIA, which further confirmed the possibility of machine learning methods for the differentiation of the invasion of pGGN-like lung adenocarcinoma. In conclusion, the predictive models established in our study could still provide certain guidance for clinicians to make accurate diagnosis.

To assess whether the radiomic predictive model could improve radiologists' performance in diagnosis of pGGN-like lung adenocarcinoma, we further compared the dichotomous diagnosis of the radiomic predictive model and two radiologists. There was a dramatic difference between radiologist A and radiologist B in the values of diagnostic accuracy, sensitivity, and specificity. The Kappa analysis revealed bad consistency between the results of the two radiologists ($\kappa = 0.316$, $P = 0.018$) while McNemar's test between that of two radiologists showed no significant difference ($P < 0.05$). This was related to the mechanisms of two statistical methods. McNemar's test only compares results with collision in two diagnostic methods instead of using comprehensive data acquired in a study while Kappa analysis calculates the consistency in all data (48–50). **Figure 6** shows the mechanism of McNemar's test and the formula for calculating the value of $\chi^2$. A small ($b - c$) leads to a small $\chi^2$-value, which results in a $P$-value of more than 0.05 no matter whether the data have actual clinical significance or not. In our test cohort, 32.8% (21/64) of the pGGNs had diverse diagnoses from two radiologists, while the value of ($b - c$) is only 1, which resulted in a value of 0 for $\chi^2$. In this special situation, McNemar's test and Kappa analysis should be combined with actual data distribution to complete the comparison of two radiologists' diagnosis. The comprehensive analysis showed that the diagnostic ability of senior radiologist A was higher than that of junior radiologist B and the radiomic predictive model outperformed two radiologists. When having the diagnosis of the radiomic predictive model for reference, two radiologists could significantly improve their performance in prediction pGGN-like IVA from AIS/MIA. What is more, no significant difference existed between the second diagnosis from two radiologists with the aid of the radiomic predictive model. The predictive model built using selected radiomic features in our research could obviously improve the ability of radiologists in prediction between pGGN-like IVA and AIS/MIA, especially for the junior radiologist; it could help radiologist B reach the level of the senior radiologist A's diagnostic ability, which had certain potential clinical meaning.

There were also some limitations in our study. First, this was a retrospective study in which all the cases were sorted according to rigorous exclusion criteria. There was certain inherent selection bias in it. Meanwhile, 322 pGGNs were not large enough for this quantitative study, compared with 960 features for each nodule. Further data collection including data from other clinic centers would be done to evaluate these models' predictive reproducibility. Second, the feature selection driven by restrictive algorithm (<1% features remaining after) might lead to a certain loss of potential predictive features for distinction. Despite the significant feature reduction, we were still able to find predictive features with high robustness. Third, limitations of this trial included the lack of standardization in image acquisition completed on various CT scanners. A voxel-resampling was applied to reduce the influence from the variability in image acquisition protocols in our study, but the above problem may still have a certain impact on the feature selection and model building. Thus, the standardization of image acquisition and establishment of database with high quality are urgently required for radiomic research. Finally, we chose a semi-automatic method to complete the pGGN segmentation. Though consistent segmentation for each pGGN had been reached through negotiation by two radiologists, there was still some interobserver difference existing in this procedure. A reliable

**A**

| | | Method A | |
|---|---|---|---|
| | | Positive (+) | Negative (−) |
| Method B | Positive (+) | a | b |
| | Negative (−) | c | d |

**B**

$$\chi^2 = \frac{(b-c)^2}{b+c}, v = 1 \cdots\cdots (1)$$

$$\chi_c^2 = \frac{(|b-c|-1)^2}{b+c}, v = 1 \cdots\cdots (2)$$

**FIGURE 6 | (A)** The mechanism of McNemar's test. **(B)** The calculation formula for $\chi^2$-value. Formula (1) would be used if ($b + c$) ≥40; otherwise, formula (2) should be chosen.

automatic segmentation algorithm that can be applied in clinical practice is still to be developed.

## CONCLUSION

Nine selected radiomic features in our research showed different distribution between IVA and AIS/MIA, which could provide some guidance for clinical practice. The predictive models established using conventional clinical, radiological, or radiomic features could help to distinguish the invasiveness of pGGN-like lung adenocarcinoma, but radiomic features could not offer more meaningful information to improve the performance of the combined model created in our study. The diagnostic performance of the radiomic predictive model established in our study was better than that of the two radiologists, and the predictive model could provide auxiliary information for radiologists (especially for junior radiologists) to improve their diagnostic ability in discrimination between pGGN-like IVA and AIS/MIA.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

This study was approved by the Institutional Review Committee of the Sir Run Run Shaw Hospital.

## AUTHOR CONTRIBUTIONS

FX and WZ conceived and designed this study. FX, YS, and CP applied the data collection for this study. FX and YS did the data reproof work during the manscript review and now they are collecting new data for further exploration. FX and JW did the segmentation work for each pGGN recruited for this study. WZ and RX contributed to the algorithm design for this study. YG completed the pathological interpretation of all pGGNs in this study. FX drafted the manuscript. LS and CQ and HH contributed equally to the manuscript review.

## FUNDING

This research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ20F030018, Key Research and Development Program of Zhejiang Province under Grant No. 2019C03064, Program Co-sponsored by Province and Ministry under Grant No. WKJ-ZJ-1926 and National Natural Science Foundation of China (NSFC) under Grant No. 61772106.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.00872/full#supplementary-material

## REFERENCES

1. 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. doi: 10.1097/JTO.0b013e318206a221
2. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin J, Beasley MB, et al. The (2015). World health organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol.* (2015). 10:1243–60. doi: 10.1097/JTO.0000000000000630
3. Sun Y, Li C, Jin L, Gao P, Zhao W, Ma W, et al. Radiomics for lung adenocarcinoma manifesting as pure ground-glass nodules: invasive prediction. *Eur Radiol.* (2020) 30:3650–9. doi: 10.1007/s00330-020-06776-y
4. Heidinger BH, Anderson KR, Nemec U, Costa DB, Gangadharan SP, VanderLaan PA, et al. 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. doi: 10.1016/j.jtho.2017.05.017
5. Kakinuma R, Muramatsu Y, Kusumoto M, Tsuchida T, Tsuta K, Maeshima AM, et al. Solitary pure ground-glass nodules 5 mm or smaller: frequency of growth. *Radiology.* (2015) 276:873–82. doi: 10.1148/radiol.2015141071
6. Kodama K, Higashiyama M, Yokouchi H, Takami K, Kuriyama K, Kusunoki Y, et al. Natural history of pure ground-glass opacity after long-term follow-up of more than 2 years. *Ann Thorac Surg.* (2002) 73:386–92. doi: 10.1016/S0003-4975(01)03410-5
7. Eguchi T, Kondo R, Kawakami S, Matsushita M, Yoshizawa A, Hara D, et al. Computed tomography attenuation predicts the growth of pure ground-glass nodules. *Lung Cancer.* (2014) 84:242–7. doi: 10.1016/j.lungcan.2014.03.009
8. Yang Y, Li K, Sun D, Yu J, Cai Z, Cao Y, et al. Invasive pulmonary adenocarcinomas versus preinvasive lesions appearing as pure ground-glass nodules: differentiation using enhanced dual-source dual-energy CT. *AJR Am J Roentgenol.* (2019) 213:W114–22. doi: 10.2214/AJR.19.21245
9. Wu F, Tian S P, Jin X, Jing R, Yang Y Q, Jin M, et al. CT and histopathologic characteristics of lung adenocarcinoma with pure ground-glass nodules 10 mm or less in diameter. *Eur Radiol.* (2017) 27:4037–43. doi: 10.1007/s00330-017-4829-5
10. She Y, Zhao L, Dai C, Ren Y, Zha J, Xie H, et al. Preoperative nomogram for identifying invasive pulmonary adenocarcinoma in patients with pure ground-glass nodule: a multi-institutional study. *Oncotarget.* (2017) 8:17229–38. doi: 10.18632/oncotarget.11236
11. She Y, Zhang L, Zhu H, Dai C, Xie D, Xie H, et al. The predictive value of CT-based radiomics in differentiating indolent from invasive lung adenocarcinoma in patients with pulmonary nodules. *Eur Radiol.* (2018) 28:5121–8. doi: 10.1007/s00330-018-5509-9
12. Kadota K, Villena-Vargas J, Yoshizawa A, Motoi N, Sima CS, Riely GJ, et al. Prognostic significance of adenocarcinoma *in situ*, minimally invasive adenocarcinoma, and nonmucinous lepidic predominant invasive adenocarcinoma of the lung in patients with stage I disease. *Am J Surg Pathol.* (2014) 38:448–60. doi: 10.1097/PAS.0000000000000134
13. Yagi T, Yamazaki M, Ohashi R, Ogawa R, Ishikawa H, Yoshimura N, et al. HRCT texture analysis for pure or part-solid ground-glass nodules: distinguishability of adenocarcinoma *in situ* or minimally invasive adenocarcinoma from invasive adenocarcinoma. *Jpn J Radiol.* (2018) 36:113–21. doi: 10.1007/s11604-017-0711-2
14. Zhang J, Wu J, Tan Q, Zhu L, Gao W. Why do pathological stage IA lung adenocarcinomas vary from prognosis?: a clinicopathologic study of 176 patients with pathological stage IA lung adenocarcinoma based on the IASLC/ATS/ERS classification. *J Thorac Oncol.* (2013) 8:1196–202. doi: 10.1097/JTO.0b013e31829f09a7
15. Wang X, Wang L, Zhang W, Zhao H, Li F. Can we differentiate minimally invasive adenocarcinoma and non-invasive neoplasms based on

high-resolution computed tomography features of pure ground glass nodules? *PLoS ONE.* (2017) 12:e180502. doi: 10.1371/journal.pone.0180502
16. Honda T, Kondo T, Murakami S, Saito H, Oshita F, Ito H, et al. Radiographic and pathological analysis of small lung adenocarcinoma using the new IASLC classification. *Clin Radiol.* (2013) 68:e21–6. doi: 10.1016/j.crad.2012.09.002
17. Lee S W, Leem CS, Kim TJ, Lee KW, Chung JH, Jheon S, et al. The long-term course of ground-glass opacities detected on thin-section computed tomography. *Respir Med.* (2013) 107:904–10. doi: 10.1016/j.rmed.2013.02.014
18. Choi CM, Um SW, Yoo CG, Kim YW, Han SK, Shim YS, et al. Incidence and risk factors of delayed pneumothorax after transthoracic needle biopsy of the lung. *Chest.* (2004) 126:1516–21. doi: 10.1378/chest.126.5.1516
19. Kazerooni EA, Lim FT, Mikhail A, Martinez FJ. Risk of pneumothorax in CT-guided transthoracic needle aspiration biopsy of the lung. *Radiology.* (1996) 198:371–5. doi: 10.1148/radiology.198.2.8596834
20. Chen CH, Chang CK, Tu CY, Liao WC, Wu BR, Chou KT, et al. Radiomic features analysis in computed tomography images of lung nodule classification. *PLoS ONE.* (2018) 13:e192002. doi: 10.1371/journal.pone.0192002
21. Priola AM, Priola SM, Cataldi A, Errico L, Di Franco M, Campisi P, et al. Accuracy of CT-guided transthoracic needle biopsy of lung lesions: factors affecting diagnostic yield. *Radiol Med.* (2007) 112:1142–59. doi: 10.1007/s11547-007-0212-y
22. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036
23. Xu X, Zhang H, Liu Q, Sun S, Zhang J, Zhu F, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J Hepatol.* (2019) 70:1133–44. doi: 10.1016/j.jhep.2019.02.023
24. Chaddad A, Kucharczyk MJ, Daniel P, Sabri S, Jean-Claude BJ, Niazi T, et al. Radiomics in glioblastoma: current status and challenges facing clinical implementation. *Front Oncol.* (2019) 9:374. doi: 10.3389/fonc.2019.00374
25. Caballo M, Pangallo DR, Mann RM, Sechopoulos I. Deep learning-based segmentation of breast masses in dedicated breast CT imaging: radiomic feature stability between radiologists and artificial intelligence. *Comput Biol Med.* (2020) 118:103629. doi: 10.1016/j.compbiomed.2020.103629
26. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res.* (2017) 77:3922–30. doi: 10.1158/0008-5472.CAN-17-0122
27. Chaddad A, Desrosiers C, Toews M, Abdulkarim B. Predicting survival time of lung cancer patients using radiomic analysis. *Oncotarget.* (2017) 8:104393. doi: 10.18632/oncotarget.22251
28. Hwang IP, Park CM, Park SJ, Lee SM, McAdams HP, Jeon YK, et al. Persistent pure ground-glass nodules larger than 5 mm: differentiation of invasive pulmonary adenocarcinomas from preinvasive lesions or minimally invasive adenocarcinomas using texture analysis. *Invest Radiol.* (2015) 50:798–804. doi: 10.1097/RLI.0000000000000186
29. Son JY, Lee HY, Lee KS, Kim JH, Han J, Jeong JY, et al. Quantitative CT analysis of pulmonary ground-glass opacity nodules for the distinction of invasive adenocarcinoma from pre-invasive or minimally invasive adenocarcinoma. *PLoS ONE.* (2014) 9:e104066. doi: 10.1371/journal.pone.0104066
30. Xiang W, Xing Y, Jiang S, Chen G, Mao H, Labh K, et al. 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. doi: 10.1186/s40644-014-0033-x
31. Milanese G, Sverzellati N, Pastorino U, Silva M. Adenocarcinoma in pure ground glass nodules: histological evidence of invasion and open debate on optimal management. *J Thorac Dis.* (2017) 9:2862–7. doi: 10.21037/jtd.2017.08.120
32. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D slicer as an image computing platform for the quantitative imaging network. *Magn Reson Imaging.* (2012) 30:1323–41. doi: 10.1016/j.mri.2012.05.001
33. van Griethuysen J, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339
34. Sun X, Xu W. Fast Implementation of DeLong's Algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Proce Lett.* (2014) 21:1389–93. doi: 10.1109/LSP.2014.2337313
35. Molodianovitch K, Faraggi D, Reiser B. Comparing the areas under two correlated ROC curves: parametric and non-parametric approaches. *Biometrical J.* (2006) 48:754–57. doi: 10.1002/bimj.200610223
36. Heidinger BH, Anderson KR, Nemec U, Costa DB, Gangadharan SP, VanderLaan PA, et al. Morphologic characteristics of pulmonary adenocarcinomas manifesting as pure ground-glass nodules on CT. *J Thorac Dis.* (2017) 9:E1148–50. doi: 10.21037/jtd.2017.11.25
37. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* (2014) 5:4006. doi: 10.1038/ncomms5644
38. Kitami A, Sano F, Hayashi S, Suzuki K, Uematsu S, Kamio Y, et al. Correlation between histological invasiveness and the computed tomography value in pure ground-glass nodules. *Surg Today.* (2016) 46:593–8. doi: 10.1007/s00595-015-1208-1
39. Zhang Y, Shen Y, Qiang JW, Ye J D, Zhang J, Zhao RY. HRCT features distinguishing pre-invasive from invasive pulmonary adenocarcinomas appearing as ground-glass nodules. *Eur Radiol.* (2016) 26:2921–8. doi: 10.1007/s00330-015-4131-3
40. Parmar C, Leijenaar RT, Grossmann P, Rios VE, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep.* (2015) 5:11044. doi: 10.1038/srep 11044
41. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybernetics.* (1973) SMC-3:610–21. doi: 10.1109/TSMC.1973.4309314
42. Iskyne. *GLCM in textural analysis[EB/OL].* (2013). Available online at: https://blog.csdn.net/lskyne/article/details/8659225?depth__1-utm__source=distribute.pc__relevant.none-task&utm__source=distribute.pc__relevant.none-task (accessed March 9, 2020).
43. Kanying. *Gray-Level Co-Occurence Matrix From an Image[EB/OL].* (2016). Available online at: https://www.cnblogs.com/8335IT/p/5648445.html (accessed April 6, 2020).
44. Liu B, Zhao Z. Texture feature fusion-based segmentation method of SAR images. *App. Res. Com.* (2011) 28:756–8. doi: 10.3969/j.issn.1001-3695.2011.02.099
45. Thibault G, Fertil B, Navarro C, Pereira S, Lévy N, Sequeira J, et al. Texture indexes and gray level size zone matrix application to cell nuclei classification. In: *10th International Conference on Pattern Recognition and Information Processing* (2009).
46. Tustison N, Gee J. Run-length matrices for texture analysis. *Insight J.* (2008) 2008.
47. Xu D, Kurani A, Furst J, Raicu D. Run-length encoding for volumetric texture. *The 4th IASTED International Conference on Visualization, Imaging, and Image Processing.* Marbella (2004).
48. Jiang F. Consistency and variability evaluation of two kinds of measurement results. *Clin Res Practice.* (2017) 2:116–7. doi: 10.19347/j.cnki.2096-1413.201728057
49. Ma B, Li K, He J, Yang T, Ma J. *Medical Statistics. Vol. 6.* Beijing: People's Medical Publishing House (2013).
50. Viera A J, Garrett JM. Understanding inter'observer agreement: the kappa statistic. *Fam Med.* (2005) 37:360–3.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.