

Predicting the Ki-67 proliferation index in pulmonary adenocarcinoma patients presenting with subsolid nodules: construction of a nomogram based on CT images

Jing Yan^{1,2#}, Xing Xue^{3#}, Chen Gao^{1,2}, Yifan Guo^{1,2}, Linyu Wu^{1,2}, Changyu Zhou^{1,2}, Feng Chen³, Maosheng Xu^{1,2}^

¹The First Clinical Medical College of Zhejiang Chinese Medical University, Hangzhou, China; ²Department of Radiology, The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China; ³Department of Radiology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

Contributions: (I) Conception and design: M Xu; (II) Administrative support: M Xu; (III) Provision of study materials or patients: M Xu, F Chen; (IV) Collection and assembly of data: J Yan, X Xue, C Gao, L Wu, C Zhou; (V) Data analysis and interpretation: J Yan, X Xue, Y Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

"These authors contributed equally to this work.

Correspondence to: Maosheng Xu. Department of Radiology, The First Affiliated Hospital of Zhejiang Chinese Medical University, 54 Youdian Road, Hangzhou, China. Email: xums166@zcmu.edu.cn; Feng Chen. Department of Radiology, The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou, China. Email: chenfenghz@zju.edu.cn.

Background: The Ki-67 proliferation index (PI) reflects the proliferation of cells. However, the conventional methods for the acquisition of the Ki-67 PI, such as surgery and biopsy, are generally invasive. This study investigated a potential noninvasive method of predicting the Ki-67 PI in patients with lung adenocarcinoma presenting with subsolid nodules.

Methods: This retrospective study enrolled 153 patients who presented with pulmonary adenocarcinoma appearing as subsolid nodules (SSNs) on computed tomography (CT) images between January 2015 and December 2018. Presence of LUAD with SSNs was confirmed by histopathology. Of these participants, 107 patients were from institution 1 and were divided into a training cohort and an internal validation cohort in a 7:3 ratio. The other 46 patients were from institution 2 and were enrolled as an external validation cohort. All patients underwent conventional CT scans with thin-slice (≤ 1.25 mm) reconstruction, and 1,316 quantitative radiomic features were extracted from the CT images for each nodule. The minimum redundancy maximum relevance and the least absolute shrinkage and selection operator were used for feature selection, and the radiomics signature was constructed based on these selected features. Clinical features were examined using univariate logistic regression analysis. The nomogram was developed based on the radiomics signature and the independent clinical risk factors. The Delong test and *t* test were employed for statistical analysis. The performance of different models was assessed by the receiver operating characteristic (ROC) curve.

Results: The diameter of the nodules [odds ratio (OR) =1.17; P=0.003] was identified as an independent predictive parameter. Both the radiomics signature and the nomogram suggested a good predictive probability for Ki-67 expression. For the radiomics signature, the area under the ROC curve (AUC) for the training cohort, the internal validation cohort, and the external validation cohort was 0.86 [95% confidence interval (CI): 0.77 to 0.95], 0.81 (95% CI: 0.64 to 0.98), and 0.77 (95% CI: 0.62 to 0.91), respectively. For

^ ORCID: 0000-0002-2396-1600.

the nomogram, the AUC for the training cohort, the internal validation cohort, and the external validation cohort was 0.86 (95% CI: 0.77 to 0.95), 0.80 (95% CI: 0.64 to 0.97), and 0.79 (95% CI: 0.65 to 0.94), respectively. There were no statistical differences in the AUCs between the radiomics signature and the radiomic nomogram in the training cohort or the validation cohorts (all P>0.05).

Conclusions: The nomogram provides a novel strategy for determining the Ki-67 PI in predicting the proliferation of subsolid nodules, which may be beneficial for the management of patients with SSNs.

Keywords: Ground glass nodule; subsolid nodule; Ki-67; radiomics; nomogram.

Submitted Dec 22, 2020. Accepted for publication Jul 29, 2021. doi: 10.21037/qims-20-1385 View this article at: https://dx.doi.org/10.21037/qims-20-1385

Introduction

The introduction of low-dose computed tomography (CT) has facilitated the detection of pulmonary subsolid nodules (SSNs) in population (1,2). Many SSNs follow a general pattern of inert growth, with minimal changes that may be difficult for the naked eye to discern during annual follow-up (3). However, a recent study showed that the malignant rate of persistent pulmonary SSNs can be quite high, with 3–30% becoming malignant (4).

Uncontrolled cell proliferation is a hallmark of cancer, and the proliferation status is related to the tumor's biological behavior, which affects the treatment and prognosis of patients (5,6). The nuclear-associated antigen Ki-67 is associated with cell proliferation and has been widely used as a prognostic indicator of disease-free survival in many malignant tumor types including glioma, breast cancer, and hepatocellular carcinoma (6-8). A recent study reported that 81.8% of oncologists indicated that the Ki-67 proliferation index (PI) is useful for informing patients about their prognosis, and more than 90% of oncologists indicated that they would use the Ki-67 PI to suggest a more guarded prognosis and/or initiate therapy after surgical resection (9). Indeed, predicting the Ki-67 PI may indirectly reflect the aggressiveness of tumors. Traditional immunohistochemical staining for Ki-67 is a widely accepted method for evaluating proliferative activity and prognosis in various tumor types. However, this technique requires postoperative specimens or preoperative fineneedle biopsies, which are both invasive and may increase the rate of pneumothorax by approximately 15% (10). Therefore, an alternative, noninvasive assessment would be highly valuable in daily clinical practice.

The morphology of a tumor may be related to its malignancy (11). However, the accurate determination of the

malignancy of SSNs still remains a significant challenge (12). Radiomics is the process of extracting information from medical images with high-throughput methods and transforming them into mineable quantitative data, which provides more valuable information compared to traditional imaging analysis (13). Image-based biomarkers can noninvasively capture the radiographic phenotype and characterize the underlying histopathology of a tumor and are not subject to sampling bias. Several studies have shown the potential of radiomics in aiding cancer detection, diagnosis, tumor classification, and assessment of prognosis (8,14,15). In addition, the radiomics approach has been used for predicting the significance of the Ki-67 PI in lung cancer prognosis, disease-free survival, and lymph node metastasis (16-18). However, to the best of our knowledge, little work has been reported on using radiomic features for predicting the Ki-67 PI in lung cancer patients with SSNs (19).

Therefore, the objective of this study was to build and validate a nomogram for predicting the expression status of Ki-67 in lung adenocarcinoma (LUAD) patients with SNNs.

Methods

Patients

This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the institutional board of the First Affiliated Hospital of Zhejiang Chinese Medical University (no. 2019-KL-067-01), China. Individual patient consent was waived due to the retrospective nature of this study.

Between January 1, 2015, and December 31, 2018, 359 LUAD patients with confirmed histopathology results were retrospectively recruited from the First Affiliated



Figure 1 A flowchart showing the selection process of the study population.

Hospital of Zhejiang Chinese Medical University, China (institution 1) and the First Affiliated Hospital, Zhejiang University School of Medicine, China (institution 2). The diagnosis of LUAD satisfied the standards of the International Association for the Study of Lung Cancer, the American Thoracic Society, and the European Respiratory Society classifications (20). The inclusion criteria for this study were as follows: (I) a chest CT scan with thinthickness performed within 1 month before surgery; (II) LUAD presenting with a SSN, including pure groundglass nodules (pGGNs) and part solid ground-glass nodules (GGNs); (III) SSNs with a diameter ≤30 mm (window width 1,300 HU and window level -600 HU); and (IV) immunohistochemistry showing Ki-67-positive staining. The exclusion criteria were as follows: (I) patients who had any oncological treatment, biopsy, or noninvasive examination before surgery; (II) histopathological data of Ki-67 unavailable; (III) images that affected nodule sketching; or (IV) LUAD presenting as solid nodules. Finally, a total of 153 SSNs were enrolled (Figure 1). The 107 patients from institution 1 were randomly divided into a training cohort and an internal validation cohort in a 7:3 ratio. The 46 patients from institution 2 were used as the external validation cohort.

Immunohistochemical analysis

All specimens were surgically resected tumors to ensure that the Ki-67 PI represented the whole tumor. The specimens were evaluated by experienced pathologists who were blinded to the CT information. Two to three sections of lung cancer with the hottest spots were selected, and at least 5 fields of view (400× magnification) were examined in each section. The Ki-67 PI refers to the percentage of malignant cells with positive nuclear Ki-67 immunostaining. In accordance with previous studies (21,22), the median of the 5th percentile was selected as the Ki-67 PI cutoff value. Samples were classified into 2 groups: the high Ki-67 PI group (Ki-67 \geq 5%) and the low Ki-67 PI (Ki-67 <5%) group.

Image qualitative evaluation and segmentation

All patients underwent a conventional CT scan. Of the 107 enrolled patients from institution 1, 47 were scanned by a Somatom Sensation 64 machine (Siemens Healthineers, Erlangen, Germany), and 60 were scanned by an Aquilion ONE scanner (Toshiba Medical Systems, Tokyo, Japan). The CT scan parameters were as follows: tube voltage 100– 120 kV, automatic tube current modulation, collimation 0.6 mm × 64 (Siemens) and 0.5 mm × 64 (Toshiba), matrix 512×512, and slice thickness 0.75 mm (Siemens) and 1.0 mm (Toshiba).

Of the 46 enrolled patients from institution 2, 15 were scanned by a Brilliance 64 scanner (Philips Medical Systems, Best, The Netherlands), 23 were scanned using a Lightspeed VCT, and 8 were scanned by an Optima CT 540 machine (General Electric Company, Boston, MA, USA). CT scan parameters were as follows: tube voltage 120 kV, automatic tube current modulation, collimation 0.625 mm, matrix 512×512 , and slice thickness 1.0-1.25 mm.

The nodules were identified by 2 cardiothoracic radiologists with 5 and 10 years of experience. Any disagreements were resolved via consultation with a third cardiothoracic radiologist with nearly 30 years of experience. The region of interest (ROI) was manually segmented in a 3D space across all of the 2-dimensional thin-slice CT images in DICOM format using open-source software ITK-SNAP (version 3.6.0; www.itksnap.org). The radiologists were blinded to the pathological diagnosis but were provided with clinical information, such as age, and were given the option to vary the window and level setting within the software to efficiently annotate the nodule. Interand intraclass correlation coefficients (ICCs) were used to assess the agreement of feature extraction.

CT image feature extraction and radiomics signature establishment

The segmented ROI was imported into the AK software (Analysis Kit, GE Healthcare, Chicago, IL, USA). A total of 1,316 radiomic features, including first order, shape, gray level co-occurrence matrix (GLCM), gray level run-length matrix (GLRLM), and neighborhood gray tone difference matrix (NGTDM) parameters were calculated. All extracted radiomic features were normalized with *z* scores to remove the limitation of each unit. In addition, the outlier values were replaced by the median value of the column features.

The minimum redundancy maximum relevance (mRMR) and the least absolute shrinkage and selection operator (LASSO) were used to select the most useful predictive features in the training cohort and radiomics signature (radscore) building. The mRMR was determined to eliminate the redundant and irrelevant features. Subsequently, LASSO (23) was conducted to select the optimized subset to construct the radiomics signature. The rad-score was generated by weighing the selected features according to their respective coefficients.

Building the nomogram and model validation

Univariable logistic regression analysis was performed to evaluate clinical variables. The significant clinical risk factors and the radiomics signature were then imported into a multivariable logistic regression analysis to develop the nomogram. The flowchart of the nomogram is shown in *Figure 2*. The nomo-score was generated by weighing the rad-score and the diameter {the formula is shown in the Appendix 1 [2]}. Subsequently, the internal validation cohort and the external validation cohort were used to test the nomogram model. A calibration curve was used to investigate the performance of the nomogram with a calibration plot using 1,000 bootstrap resamples and the Hosmer-Lemeshow test. The performance of different models was assessed using the receiver operator characteristic (ROC) curve, sensitivity, specificity, and accuracy.

Decision curve analysis

To assess the clinical value of the nomogram, the radiomics signature, and the clinical model for predicting the expression of Ki-67, decision curve analysis (DCA) was performed. DCA is a method to evaluate clinical predictive models, diagnostic tests, and molecular markers (24,25). The probability threshold is a certain value at which the diagnostic patient would choose intervention. The net benefit represents the difference between the expected benefits and the expected damages.

Statistical analysis

The chi-squared test was applied for categorical variables, such as gender. The *t* test was used between groups of normally distributed variables, and the Mann-Whitney test was used for nonnormally distributed variables. The "RMS" package of the R software was used to build a nomogram. The LASSO logistic regression model was also used with penalty parameter tuning that was conducted by 10-fold cross-validation. The DeLong test was used to compare the difference in ROC curves between different models. Statistical analysis was performed with R software (version 3.6.1; https://www.r-project.org) and SPSS version 25 (IBM Corp., Armonk, NY, USA). A 2-tailed P value <0.05 was considered statistically significant.

Results

Clinical characteristics

A total of 153 consecutive patients with LUAD were enrolled in this retrospective study. There were 101 females and 52 males, with a mean age of 56.2 ± 11.0 years (range, 26–79 years). Within the cohort, 87 patients had high Ki-



Figure 2 A diagram showing the workflow of the development of the nomogram.

67 PI and 66 patients showed low Ki-67 PI. There were 108 cases of pGGNs and 45 cases of part-solid GGNs (Table S1). The mean diameter of the nodules was 13.8±6.1 mm (range, 5–30 mm).

The detailed clinical characteristics of the training cohort and the validation cohort are shown in Table 1. A high Ki-67 PI was often observed in patients with part-solid GGNs and higher-grade clinical tumor node metastasis (cTNM) staging.

Inter- and intraobserver reproducibility of radiomic feature extraction

The interobserver ICC for reader A and reader B ranged from 0.776 to 0.871. The intraobserver ICC between reader A and reader B ranged from 0.790 to 0.927. The results indicated a favorable inter-and intraobserver reproducibility for feature extraction.

Radiomic signature development and validation

A total of 20 features were retained by using the mRMR.

LASSO was then used to select the optimized features to construct the radiomics signature (rad-score). The detailed formula of the rad-score is shown in the Appendix 1 [1]. The value of the rad-score was higher in patients with a high Ki-67 PI compared to patients with a low Ki-67 PI. Compared with the clinical model, the radiomics signature showed a higher predictive ability for Ki-67 PI. The AUC of the training cohort, internal validation cohort, and external validation cohort for the radiomics signature was 0.86 (95% CI: 0.77 to 0.95), 0.81 (95% CI: 0.64 to 0.98), and 0.77 (95% CI: 0.62 to 0.91), respectively (Figure 3). The DeLong test between the radiomics signature and the clinical model showed statistical significance (P=0.02).

Building the nomogram and assessing its diagnostic value

Univariable logistic regression analysis in the training cohort revealed that the diameter of the nodule (OR 1.17; 95% CI: 1.06 to 1.31; P=0.003) and the clinical TNM staging (OR 3.77; 95% CI: 1.60 to 8.87; P=0.002) were independent predictors of high Ki-67 PI. Stepwise multivariable analysis was used to integrate the optimal predictor and develop the

Quantitative Imaging in Medicine and Surgery, Vol 12, No 1 January 2022

Table T The emiliar enalacteristics of the training conort, internal valuation conort, and external valuation conort									
Clinical characteristics	Training cohort, n=75		Durahua	Internal validation cohort, n=32		Dualua	External validation cohort, n=46		Dualua
	Low Ki-67	High Ki-67	· P value ·	Low Ki-67	High Ki-67	P value	Low Ki-67	High Ki-67	P value
Gender			0.359			0.503			0.039
Male	15 (42.9)	12 (30.0)		2 (13.3)	5 (29.4)		3 (18.8)	15 (50.0)	
Female	20 (57.1)	28 (70.0)		13 (86.7)	12 (70.6)		13 (81.3)	15 (50.0)	
Age (years)	56.5±13.7	55.6±9.8	0.758	59.0±10.5	54.6±9.1	0.202	56.1±12.7	56.1±9.7	0.872
Diameter (mm)	12.0±4.6	15.4±6.4	0.008	10.0±3.2	16.1±7.3	0.002	11.3±5.5	15.5±6.5	0.021
Nodule type			1.000			0.571			0.013
pGGN	28 (80.0)	31 (77.5)		12 (80.0)	11 (64.7)		13 (81.3)	13 (43.3)	
Part-solid GGN	7 (20.0)	9 (22.5)		3 (20.0)	6 (35.3)		3 (18.8)	17 (56.7)	
cTNM			0.034			0.012			0.015
IA1	15 (42.9)	11 (27.5)		10 (66.7)	4 (23.5)		10 (62.5)	6 (20.0)	
IA2	19 (54.3)	20 (50.0)		5 (33.3)	7 (41.2)		4 (25.0)	15 (50.0)	
IA3	1 (2.9)	9 (22.5)		0 (0.0)	6 (35.3)		2 (12.5)	9 (30.0)	

Table 1 The clinical characteristics of the training cohort, internal validation cohort, and external validation cohort

cTNM, clinical TNM stage according to the 8th Edition Lung Cancer Stage Classification; GGN, ground-glass nodule; pGGN, pure GGN.



Figure 3 Receiver operating characteristic (ROC) curves of predictive Ki-67 proliferation index (PI) of the clinical model, radiomics signature, and nomogram between the training cohort, internal validation cohort, and the external validation cohort.

nomogram (*Figure 4*). *Table 2* showed a superior predictive performance in Ki-67 expression compared to the clinical model in both the training cohort (AUC 0.86; 95% CI: 0.77 to 0.95), the internal validation cohort (AUC 0.80; 95% CI: 0.64 to 0.97), and the external validation cohort (AUC 0.79; 95% CI: 0.65 to 0.94). The Hosmer-Lemeshow test showed a good fit between the nomogram and the actual trend (*Figure 5*). The DeLong test showed that the AUC of the nomogram was significantly different to that of the clinical

model (P<0.01), but there was no statistical difference in the AUC between the radiomics signature and the nomogram (P=0.77).

Assessment of clinical value using DCA

DCA was used to assess the clinical value of the 3 models (*Figure 6*). The DCA showed that for patients with threshold probabilities of 10% to 78%, the net benefit

Yan et al. Predicting Ki-67 expression in pulmonary subsolid nodules



Figure 4 The nomogram incorporating the radiomics signature and the diameter to predict the risk of high Ki-67 PI expression. The value of the diameter and the rad-score can be converted into a risk score according to the "Points". After adding up the individual risk scores of these predictors, the corresponding prediction probability of Ki-67 is shown as "Risk". PI, proliferation index.

Table 2 The predictive values in the different models in the training cohort, internal validation cohort, and external validation cohort

Models	AUC	95% CI	Accuracy	Sensitivity	Specificity	PPV	NPV
Training cohort							
Nomogram	0.859	0.772-0.947	0.773	0.675	0.885	0.871	0.705
Radiomics signature	0.856	0.765–0.948	0.862	0.875	0.771	0.814	0.844
Clinical features	0.701	0.583–0.819	0.667	0.750	0.571	0.667	0.667
Internal validation cohort							
Nomogram	0.804	0.636-0.972	0.813	0.923	0.781	0.706	0.933
Radiomics signature	0.812	0.642-0.981	0.813	0.823	0.800	0.824	0.824
Clinical features	0.669	0.473-0.864	0.625	0.588	0.667	0.667	0.588
External validation cohort							
Nomogram	0.794	0.651–0.936	0.761	0.852	0.632	0.767	0.750
Radiomics signature	0.767	0.618–0.915	0.717	0.733	0.688	0.815	0.579
Clinical features	0.708	0.542-0.874	0.674	0.667	0.688	0.800	0.524

AUC, the area under the curve; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.

was maximized by the nomogram and radiomics signature approach.

Discussion

The aim of this study was to develop a nomogram to predict the Ki-67 PI in pulmonary adenocarcinoma patients presenting with SSNs on a conventional CT scan. The investigation demonstrated that the predictive performance of radiomics could act as a noninvasive predictor of Ki67 expression for early LUAD patients. The nomogram showed a superior performance in predicting the Ki-67 PI. This concurs with a previous study by Gu *et al.*, which showed that the radiomic classifier could predict the Ki-67 PI in patients with non-small cell lung cancer (NSCLC) (19).

In the current study, there was no significant difference in gender, age, or nodule type between patients with high Ki-67 PI and patients with low Ki-67 PI. In contrast, previous studies have shown that a high Ki-67 PI was



Figure 5 The goodness-of-fit curve of the nomogram in the training cohort, internal validation cohort, and external validation cohort. The 45° dotted line represents the ideal prediction, the dotted curve indicates the performance of the nomogram in Ki-67 PI prediction, and the solid curve indicates the bias-corrected nomogram. PI, proliferation index.



Figure 6 Decision curves of the clinical value assessment of the 3 models in the training cohort, internal validation cohort, and external validation cohort. The X-axis represents the high-risk threshold and the Y-axis represents the net benefit. The black solid line represents high Ki-67 PI, and the yellow line represents low Ki-67 PI. The red, blue, and green lines represent the nomogram, radiomics signature, and clinical model, respectively. PI, proliferation index.

commonly observed in older male subjects (19). The differences in results may be attributed to the fact that the study population in the latter were early LUAD patients with SSN as an inert growth. However, in agreement with McWilliams *et al.* (26), our current study found that larger diameter nodules with a high cTNM stage were associated

with elevated Ki-67 expression.

In accordance with a previous study (22), a cutoff value of 5% (the median of our data) was used for Ki-67 expression. Due to the heterogeneity of tumors, a tumor biopsy does not represent the whole tumor. Thus, to ensure that Ki-67 expression represented the entire SSN, our study only

selected LUADs with SSNs that were completely surgically removed. Meanwhile, Warth *et al.* found that different types of lung cancer have different Ki-67 expression and different Ki-67 cutoff values (17). This may explain the lack of an agreed upon standard for the Ki-67 PI cutoff value.

Recently, texture analysis has become an emergent field in radiology. Radiomic features can provide a plethora of information, especially in oncology images (14,27,28). Radiomic features have been shown to reflect numerous clinically relevant characteristics of lung cancer (29) and have been shown to be associated with biomarkers which may act as potential clinical prognostic indicators (16).

The Ki-67 PI is clinically the most commonly used biomarker for proliferative activity in tumors. Several studies have shown that a high Ki-67 PI indicates tumor cell proliferation, metastasis, poor prognosis, and high grade in NSCLC patients, and all this information may be beneficial during the clinical decision-making process. Furthermore, Ishibashi *et al.* reported that NSCLC patients with a high Ki-67 PI were more sensitive to radiation (30).

This study hypothesized that the Ki-67 PI in tumors may be related to radiomics signatures and a predictive nomogram was constructed. By comparing patients with a high Ki-67 PI (\geq 5%) and patients with a low Ki-67 PI (<5%), the report demonstrated that both the radiomics signature and the nomogram showed a better AUC compared to the predicted values. The nomogram was able to provide a risk nomo-score for each patient. This method may represent a noninvasive strategy to dynamically monitor the biological behavior of SSNs and provide a guideline for individualized clinical treatment.

There were several limitations to this study. First, this retrospective study involving 2 centers had a relatively a small sample size. Future large-scale studies should be conducted with data from multiple centers. Second, the lesions in this study were from early-stage LUAD patients with SSNs, and future studies should also examine LUAD with solid nodules. In addition, the data gathered herein were obtained from different scanners with different slice thicknesses. However, the AUC analysis of both the training and validation cohorts indirectly indicated that the radiomics signature and the nomogram had better generalization ability than did the clinical model.

Conclusions

This report demonstrated that the radiomics signature extracted from SNNs of the lung may represent a potential

method for predicting the Ki-67 PI. Furthermore, a nomogram was developed to predict the individual patient risk score and this may contribute to the personalized management of patients with LUAD.

Acknowledgments

We thank Mr. Michael C. Lin for language editing and Mr. Liang Li for technical support in the software code. *Funding:* This work was supported by the Zhejiang Provincial Natural Science Foundation of China (No. LSY19H180003) and the Scientific Research Project of the Zhejiang Education Department (No. Y202045271).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi. org/10.21037/qims-20-1385). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was approved by the institutional ethics board of the First Affiliated Hospital of Zhejiang Chinese Medical University and the First Affiliated Hospital, Zhejiang University School of Medicine. Individual patient consent was waived due to the retrospective nature of this study.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

 Yankelevitz DF, Yip R, Smith JP, Liang M, Liu Y, Xu DM, Salvatore MM, Wolf AS, Flores RM, Henschke CI; International Early Lung Cancer Action Program Investigators Group. CT Screening for Lung Cancer: Nonsolid Nodules in Baseline and Annual Repeat Rounds. Radiology 2015;277:555-64.

- Chang B, Hwang JH, Choi YH, Chung MP, Kim H, Kwon OJ, Lee HY, Lee KS, Shim YM, Han J, Um SW. Natural history of pure ground-glass opacity lung nodules detected by low-dose CT scan. Chest 2013;143:172-8.
- Hasegawa M, Sone S, Takashima S, Li F, Yang ZG, Maruyama Y, Watanabe T. Growth rate of small lung cancers detected on mass CT screening. Br J Radiol 2000;73:1252-9.
- Hammer MM, Palazzo LL, Kong CY, Hunsaker AR. Cancer Risk in Subsolid Nodules in the National Lung Screening Trial. Radiology 2019;293:441-8.
- 5. Sherr CJ. Cancer cell cycles. Science 1996;274:1672-7.
- Hu XX, Yang ZX, Liang HY, Ding Y, Grimm R, Fu CX, Liu H, Yan X, Ji Y, Zeng MS, Rao SX. Whole-tumor MRI histogram analyses of hepatocellular carcinoma: Correlations with Ki-67 labeling index. J Magn Reson Imaging 2017;46:383-92.
- Xu M, Tang Q, Li M, Liu Y, Li F. An analysis of Ki-67 expression in stage 1 invasive ductal breast carcinoma using apparent diffusion coefficient histograms. Quant Imaging Med Surg 2021;11:1518-31.
- Li Y, Qian Z, Xu K, Wang K, Fan X, Li S, Liu X, Wang Y, Jiang T. Radiomic features predict Ki-67 expression level and survival in lower grade gliomas. J Neurooncol 2017;135:317-24.
- Marchevsky AM, Hendifar A, Walts AE. The use of Ki-67 labeling index to grade pulmonary well-differentiated neuroendocrine neoplasms: current best evidence. Mod Pathol 2018;31:1523-31.
- Wiener RS, Schwartz LM, Woloshin S, Welch HG. Population-based risk for complications after transthoracic needle lung biopsy of a pulmonary nodule: an analysis of discharge records. Ann Intern Med 2011;155:137-44.
- Liu F, Wang M, Li H. Role of perfusion parameters on DCE-MRI and ADC values on DWMRI for invasive ductal carcinoma at 3.0 Tesla. World J Surg Oncol 2018;16:239.
- Zhao J, Zhang C, Li D, Niu J. Combining multi-scale feature fusion with multi-attribute grading, a CNN model for benign and malignant classification of pulmonary nodules. J Digit Imaging 2020;33:869-78.
- Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441-6.
- 14. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ,

Mehrtash A, Allison T, Arnaout O, Abbosh C, Dunn IF, Mak RH, Tamimi RM, Tempany CM, Swanton C, Hoffmann U, Schwartz LH, Gillies RJ, Huang RY, Aerts HJWL. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA Cancer J Clin 2019;69:127-57.

- Gates EDH, Lin JS, Weinberg JS, Hamilton J, Prabhu SS, Hazle JD, Fuller GN, Baladandayuthapani V, Fuentes D, Schellingerhout D. Guiding the first biopsy in glioma patients using estimated Ki-67 maps derived from MRI: conventional versus advanced imaging. Neuro Oncol 2019;21:527-36.
- 16. Woo T, Okudela K, Yazawa T, Wada N, Ogawa N, Ishiwa N, Tajiri M, Rino Y, Kitamura H, Masuda M. Prognostic value of KRAS mutations and Ki-67 expression in stage I lung adenocarcinomas. Lung Cancer 2009;65:355-62.
- Warth A, Cortis J, Soltermann A, Meister M, Budczies J, Stenzinger A, Goeppert B, Thomas M, Herth FJ, Schirmacher P, Schnabel PA, Hoffmann H, Dienemann H, Muley T, Weichert W. Tumour cell proliferation (Ki-67) in non-small cell lung cancer: a critical reappraisal of its prognostic role. Br J Cancer 2014;111:1222-9.
- Chen C, Zhu WD, Zhang XH, Zhu YH, Huang JA. Value of Ki-67 and computed tomography in the assessment of peripheral lung adenocarcinoma. Br J Biomed Sci 2016;73:32-7.
- Gu Q, Feng Z, Liang Q, Li M, Deng J, Ma M, Wang W, Liu J, Liu P, Rong P. Machine learning-based radiomics strategy for prediction of cell proliferation in non-small cell lung cancer. Eur J Radiol 2019;118:32-7.
- 20. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, Powell CA, Beer D, Riely G, Garg K, Austin JH, Rusch VW, Hirsch FR, Jett J, Yang PC, Gould M; American Thoracic Society. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. Proc Am Thorac Soc 2011;8:381-5.
- 21. Ahn HK, Jung M, Ha SY, Lee JI, Park I, Kim YS, Hong J, Sym SJ, Park J, Shin DB, Lee JH, Cho EK. Clinical significance of Ki-67 and p53 expression in curatively resected non-small cell lung cancer. Tumour Biol 2014;35:5735-40.
- 22. He LY, Zhang H, Wang ZK, Zhang HZ. Diagnostic and prognostic significance of E-cadherin and Ki-67 expression in non-small cell lung cancer patients. Eur Rev Med Pharmacol Sci 2016;20:3812-7.

Yan et al. Predicting Ki-67 expression in pulmonary subsolid nodules

- Gui J, Li H. Penalized Cox regression analysis in the highdimensional and low-sample size settings, with applications to microarray gene expression data. Bioinformatics 2005;21:3001-8.
- Kattan M, Vickers AJ. Decision Curve Analysis. Encyclopedia of Medical Decision Making 2012;313:409-10.
- Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565-74.
- 26. McWilliams A, Tammemagi MC, Mayo JR, Roberts H, Liu G, Soghrati K, et al. Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med 2013;369:910-9.
- 27. Liu H, Jiao Z, Han W, Jing B. Identifying the histologic subtypes of non-small cell lung cancer with computed

Cite this article as: Yan J, Xue X, Gao C, Guo Y, Wu L, Zhou C, Chen F, Xu M. Predicting the Ki-67 proliferation index in pulmonary adenocarcinoma patients presenting with subsolid nodules: construction of a nomogram based on CT images. Quant Imaging Med Surg 2022;12(1):642-652. doi: 10.21037/qims-20-1385 tomography imaging: a comparative study of capsule net, convolutional neural network, and radiomics. Quant Imaging Med Surg 2021;11:2756-65.

- Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.
- Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, Liang C, Tian J, Liang C. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. Radiology 2016;281:947-57.
- Ishibashi N, Maebayashi T, Aizawa T, Sakaguchi M, Nishimaki H, Masuda S. Correlation between the Ki-67 proliferation index and response to radiation therapy in small cell lung cancer. Radiat Oncol 2017;12:16.

652

Supplementary

The number of features extraction:

shape: 14 first of order: 252 GLCM: 336 GLRLM: 22 NGTDM: 70 GLDM:196 GLSZM: 224

Radscore

Radscore was calculated by summing the selected features weighted by their coefficients.

The final formula for the rad-score is as follows: "Radscore =

-1.117*wavelet-HHL_firstorder_Mean+0.3*log-sigma-3-0-mm-3D_glszm_GrayLevelNonUniformity+-0.609*wavelet-LHH_ gldm_SmallDependenceLowGrayLevelEmphasis+0.598*lbp-3D-k_glszm_ZoneEntropy+-.172*original_gldm_SmallDepen denceLowGrayLevelEmphasis+-0.346*log-sigma-3-0-mm-3D_glszm_SmallAreaLowGrayLevelEmphasis+-0.232*wavelet-HHL_firstorder_Median+0.024*lbp-3D-m2_glszm_SmallAreaEmphasis+0.324*lbp-3D-m2_firstorder_Range+0.714*original_ glcm_ClusterShade+0.545*lbp-3D-k_ngtdm_Coarseness+1.046*original_shape_Flatness+0.509*log-sigma-2-0-mm-3D_gldm_ LowGrayLevelEmphasis+0.284*log-sigma-3-0-mm-3D_glszm_ZoneVariance+0.065*wavelet-HHH_firstorder_Median+-1.236*wavelet-LLL_glszm_SmallAreaLowGrayLevelEmphasis+0.868*wavelet-HLH_ngtdm_Contrast+0.466*wavelet-LHH_ glszm_SmallAreaLowGrayLevelEmphasis+-0.051*log-sigma-3-0-mm-3D_gldm_DependenceNonUniformityNormalized + 0.304" [1]

Nomoscore

Nomo-score was generated by weighing the coefficients of rad-score and the diameter. The final formula for the nomo-score is as follows: "Nomoscore=-.741789249917205+Diameter*0.0510583869212873+rad score*0.695081609627141"

[2]

Table 51 The basic chinical characteristics of patients with hing auchocarcinoma							
Clinical characteristics	Low Ki-67 (n=66)	High Ki-67 (n=87)	P value				
Female/male	46/20	55/32	0.039				
Age (y)	57.0±12.7	55.6±9.5	0.872				
Nodule type			0.013				
pGGN	53 (80.3)	55 (63.2)					
Part-solid GGN	13 (19.7)	32 (36.8)					
Diameter(mm)	11.4±4.6	15.6±6.5	0.021				
cTNM			0.015				
IA1	10 (62.5)	6 (20.0)					
IA2	4 (25.0)	15 (50.0)					
IA3	2 (12.5)	9 (30.0)					

Table S1 The basic clinical characteristics of patients with lung adenocarcinoma

cTNM: clinical TNM stage according to The 8th Edition Lung Cancer Stage Classification; pGGN: pure ground-glass nodule.