Diagnosis of Invasive Lung Adenocarcinoma Based on Chest CT Radiomic Features of Part-Solid Pulmonary Nodules

Radiology

ORIGINAL RESEARCH • THORACIC IMAGING

Diagnosis of Invasive Lung Adenocarcinoma Based on Chest CT Radiomic Features of Part-Solid Pulmonary Nodules: A Multicenter Study

Guangyao Wu, MD • Henry C. Woodruff, PhD • Jing Shen, MD • Turkey Refaee, MS • Sebastian Sanduleanu, MD • Abdalla Ibrahim, MD • Ralph T. H. Leijenaar, PhD • Rui Wang, MD • Jingtong Xiong, MD • Jie Bian, MD, PhD • Jianlin Wu, MD, PhD • Philippe Lambin, MD, PhD

From the Departments of Precision Medicine (G.W., H.C.W., T.R., S.S., A.I., R.T.H.L., P.L.) and Radiology and Nuclear Medicine (H.C.W., A.I., P.L.), GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center+, Maastricht, the Netherlands; Department of Radiology, Affiliated Zhongshan Hospital of Dalian University, 6 Jiefang Street, Dalian 116001, People's Republic of China (G.W., J.S., J.W.); Department of Radiology, The Fifth Hospital of Dalian, Dalian, People's Republic of China (R.W.); and Department of Radiology, The Second Affiliated Hospital of Dalian Medical University, Dalian, People's Republic of China (J.X., J.B.). Received November 1, 2019; revision requested January 6, 2020; revision received April 30; accepted May 21. **Address correspondence to** J.W. (e-mail: *cjr.wujianlin@vip.163.com*).

Supported by the China Scholarships Council (grant 201808210318), ERC Advanced Grant Project HYPOXIMMUNO (grant 694812), Technology Foundation STW (grant P14-19 Radiomics Strategy), Technology Programme of the Ministry of Economic Affairs, SME Phase 2 (RAIL grant 673780), Eurostars (DART, DECIDE, COMPACT12053), European Program H2020-2015-17 (ImmunoSABR grant 733008, PREDICT-ITN grant 766276), TRANSCAN Joint Transnational Call 2016 (JTC2016 "CLEARLY" grant UM 2017-8295), and Interreg V-A Euregio Meuse-Rhine ("Euradiomics").

Conflicts of interest are listed at the end of this article.

See also the editorial by Nishino in this issue.

Radiology 2020; 297:451–458 • https://doi.org/10.1148/radiol.2020192431 • Content code: CH

Background: Solid components of part-solid nodules (PSNs) at CT are reflective of invasive adenocarcinoma, but studies describing radiomic features of PSNs and the perinodular region are lacking.

Purpose: To develop and to validate radiomic signatures diagnosing invasive lung adenocarcinoma in PSNs compared with the Brock, clinical-semantic features, and volumetric models.

Materials and Methods: This retrospective multicenter study (*https://ClinicalTrials.gov*, NCT03872362) included 291 patients (median age, 60 years; interquartile range, 55–65 years; 191 women) from January 2013 to October 2017 with 297 PSN lung adenocarcinomas split into training (n = 229) and test (n = 68) data sets. Radiomic features were extracted from the different regions (gross tumor volume [GTV], solid, ground-glass, and perinodular). Random-forest models were trained using clinical-semantic, volumetric, and radiomic features, and an online nodule calculator was used to compute the Brock model. Performances of models were evaluated using standard metrics such as area under the curve (AUC), accuracy, and calibration. The integrated discrimination improvement was applied to assess model performance changes after the addition of perinodular features.

Results: The radiomics model based on ground-glass and solid features yielded an AUC of 0.98 (95% confidence interval [CI]: 0.96, 1.00) on the test data set, which was significantly higher than the Brock (AUC, 0.83 [95% CI: 0.72, 0.94]; P = .007), clinical-semantic (AUC, 0.90 [95% CI: 0.83, 0.98]; P = .03), volumetric GTV (AUC, 0.87 [95% CI: 0.78, 0.96]; P = .008), and radiomics GTV (AUC, 0.88 [95% CI: 0.80, 0.96]; P = .01) models. It also achieved the best accuracy (93% [95% CI: 84%, 98%]). Both this model and the model with added perinodular features showed good calibration, whereas adding perinodular features did not improve the performance (integrated discrimination improvement, −0.02; P = .56).

Conclusion: Separating ground-glass and solid CT radiomic features of part-solid nodules was useful in diagnosing the invasiveness of lung adenocarcinoma, yielding a better predictive performance than the Brock, clinical-semantic, volumetric, and radiomics gross tumor volume models.

Online supplemental material is available for this article.

The reduction of lung cancer mortality by almost 20% in the National Lung Screening Trial can be partially attributed to the extensive use of low-dose CT for lung cancer screening in high-risk populations, which led to the improved detection of pulmonary nodules and early stage lung cancers (1,2). Pulmonary nodules are classified as solid, pure ground-glass, and part-solid nodules (PSNs) based on CT phenotyping, with PSNs being an important cancer predictor in the Brock model that is widely used to assess the malignant risk of pulmonary nodules (3).

Moreover, adenocarcinomas manifesting as PSNs have been suggested to be a distinct subtype, most of which are confirmed as adenocarcinoma in situ (AIS), minimally invasive adenocarcinoma (MIA), or invasive adenocarcinoma (IA) by abnormality, requiring a different management strategy due to different clinical-pathologic characteristics (4). Furthermore, evidence from histological specimens suggests that the solid components of lung nodules have a close-knit association with the invasive component of adenocarcinomas (5–7). Among the different subtypes of lung adenocarcinoma, IA has the worst prognosis, with the others having an almost 100% survival probability (8). Therefore, lobectomy is often recommended for patients with IA, whereas limited resections are suggested for patients with AIS or MIA (9).

Abbreviations

AIS = adenocarcinoma in situ, AUC = area under the curve, CI = confidence interval, GTV = gross tumor volume, IA = invasive adenocarcinoma, MIA = minimally invasive adenocarcinoma, PSN = part-solid nodule

Summary

CT radiomic ground-glass and solid features from part-solid nodules predicted invasiveness of lung adenocarcinoma better than conventional Brock, clinical-semantic, volumetric, and radiomics gross tumor volume models.

Key Results

- The radiomics model combining ground-glass and solid features had a higher area under the curve (AUC, 0.98) than the Brock (AUC, 0.83; $P = .007$), clinical-semantic (AUC, 0.90; $P = .03$), volumetric (AUC, 0.87; $P = .008$), and radiomics gross tumor volume (AUC, 0.88; $P = .01$) models.
- Adding ground-glass radiomic features to the solid features improved the AUC by 0.14 ($P = .03$).
- Perinodular features did not help predict invasiveness of part-solid nodules ($P = .11$).

Although radiologists are able to assess the invasive probability of PSNs using semantic or morphologic CT features (eg, diameter, solid proportion) and medical history in clinical practice, this method is limited by subjectivity and creates a work burden (10–12). Automated volumetric analysis facilitates the quantification of the nodule volume and mean density with higher repeatability than manual measurement (13). Quantitative image analysis, or radiomics, the automated extraction of hundreds or thousands of imaging features from medical images and their correlation with the underlying biologic and clinical outcomes, has shown promise in further automating image analysis (14). Recent studies (15,16) show the potential of radiomics to distinguish AIS or MIA from IA in subsolid pulmonary nodules. Additionally, radiomic features extracted from the perinodular region have been examined for their diagnostic power (17,18).

Our purpose was to develop and to validate radiomic signatures based on different regions of the nodule (gross, solid, ground-glass, and perinodular) and compare the predictive performance with clinical-semantic features, volumetric analysis, and the Brock model for the differentiation of AIS or MIA and IA with PSNs.

Materials and Methods

Patient Selection

The institutional review boards approved this retrospective study (*http://ClinicalTrials.gov* identifier: NCT03872362), and the requirement for written informed consent was waived. A total of 1428 patients (1614 nodules) from three centers (center 1: The Affiliated Zhongshan Hospital of Dalian University; center 2: The Second Affiliated Hospital of Dalian Medical University; and center 3: The Fifth Hospital of Dalian) who underwent lung surgery from January 2013 to October 2017 were considered consecutively. The inclusion criteria were presence of PSNs and pathologically confirmed lung adenocarcinoma. The exclusion criteria were *(a)* previous history of radiation therapy, chemotherapy, or biopsy before baseline CT; *(b)* time interval greater than 2 weeks between the CT examination and surgery; *(c)* insufficient CT quality (severe motion artifacts and outlying scan modes or reconstruction algorithms); *(d)* atypical adenomatous hyperplasia; and *(e)* contrast material–enhanced CT. For multiple PSNs, only the nodules with conclusive pathologic results were included. All patients from center 1 and center 2 were allocated to the training data set, and patients from center 3 were used as the test data set (Fig 1).

CT Scanning and Semantic Features

Nonenhanced CT was performed on a 64- or 128-detector row CT system in the axial plane with 1- to 1.5-mm slice thickness and without reconstruction interval. Detailed acquisition and reconstruction parameters are presented in Table E1 (online).

The 18 clinical and semantic features obtained from the clinical record and CT images are listed in Appendix E1 (online). Two radiologists (G.W. and J.S., each with 8 years of experience in lung imaging) blinded to the clinical and pathologic results evaluated semantic features in the lung window setting (level, −600 HU; width, 1200 HU) and the mediastinal window setting (level, 40 HU; width, 350 HU). Any disagreement in describing semantic features was resolved by a consensus read. An online Brock University nodule calculator was also used to calculate the invasive probability of each PSN in this study (3).

Segmentation and Volumetry

The regions of interest were manually contoured slicewise by one radiologist (G.W.) to arrive at a three-dimensional segmentation using MIM (version 6.9.4; *https://www.mimsoftware.com*). The gross tumor volume (GTV), ground-glass, solid (19), and perinodular (20) regions were contoured following the process depicted in Figure 2. To assess the robustness of radiomic features in segmentation, 50 randomly selected nodules were segmented twice by the same radiologist (G.W.) in a 1-week period, as well as by another physician (S.S., with 5 years of contouring experience in lung imaging). Only the location of the nodules was provided and other information such as age, sex, and pathologic outcome was blinded from the readers during segmentation. Volume and attenuation for the GTV, ground-glass, and solid regions were calculated automatically using MIM.

Radiomic Features

All images were resampled to $1 \times 1 \times 1$-mm voxel size using linear interpolation to partially counter the heterogeneous reconstruction settings (21). For nonfiltered features excluding first-order statistics features, voxels values were aggregated into bins 25 HU wide to reduce noise and interscanner variability (22). Filtered features used a fixed number of bins equal to the amount calculated for nonfiltered features. Feature extraction was performed using the RadiomiX Discovery Toolbox (version, October 2019; *https://www.oncoradiomics.com*). Descriptions and mathematical definitions of the features have been described in detail previously (14). A posterior feature harmonization method (ComBat) was used to correct the radiomic features for the batch effect introduced by different scanners (23).

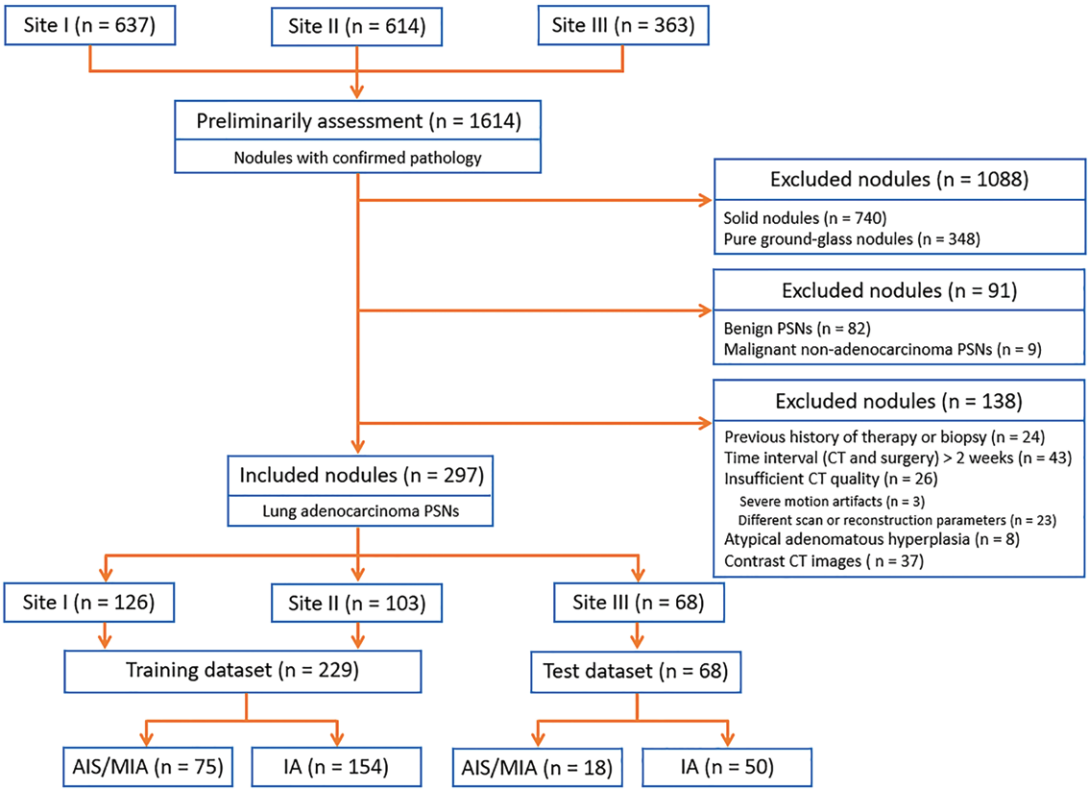

Figure 1: Flowchart shows the patient selection process. AIS = adenocarcinoma in situ, IA = invasive adenocarcinoma, MIA = minimally invasive adenocarcinoma, PSN = part-solid nodule.

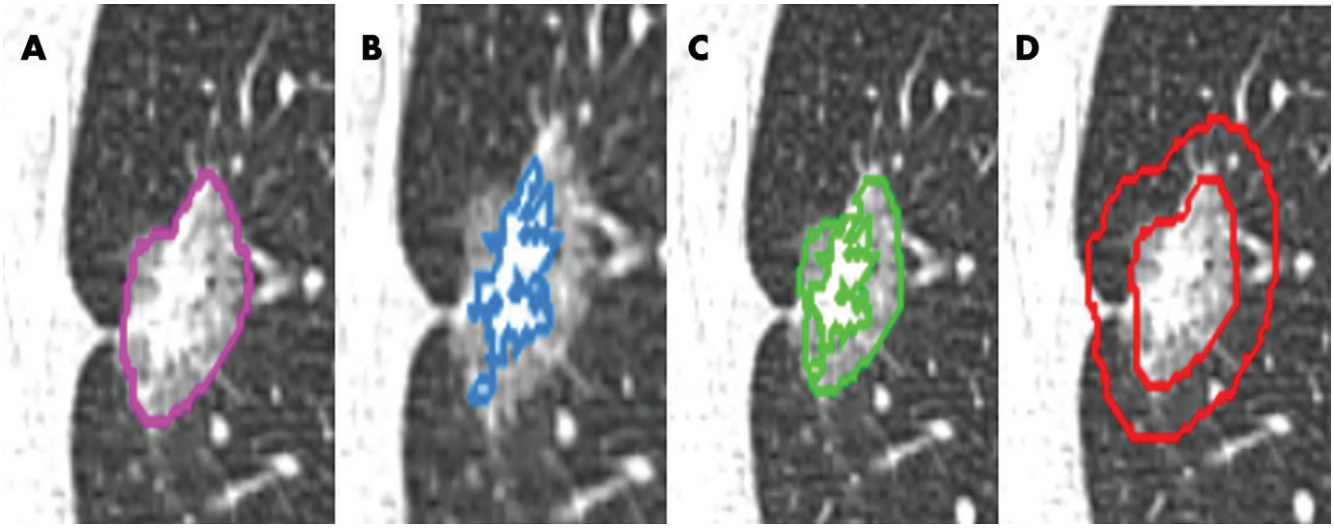

Figure 2: Images show delineation of the various region of interest habitats. *A*, The gross tumor volume (GTV) is contoured along the border of the nodules in lung window setting; the large vessels and bronchus were excluded. *B*, The solid region is identified using a thresholding method within the GTV (>−50 HU), and only the largest solid portion was kept. *C*, The ground-glass region is obtained by subtracting the solid component from the GTV. *D*, The perinodular region is defined by extending the GTV by 5 mm from the nodule boundary in three dimensions and subtracting the GTV.

Feature Selection and Modeling

The feature selection process, such as Boruta (24), for clinical-semantic, volumetric, and radiomic features is described in Appendix E1 (online) and shown in Figure E1 (online). Random-forest binary classification models based on selected features were trained on the training data, including clinical-semantic models, volumetric models, and radiomics models (Fig 3).

Statistical Analysis

Comparisons between data sets are summarized using a Wilcoxon rank sum test for the continuous variables and χ^2 or Fisher exact test for the categorical variables. Correlations between features were performed by means of the Spearman correlation. Model performance was evaluated by means of the area under the curve (AUC), and the 95% confidence interval

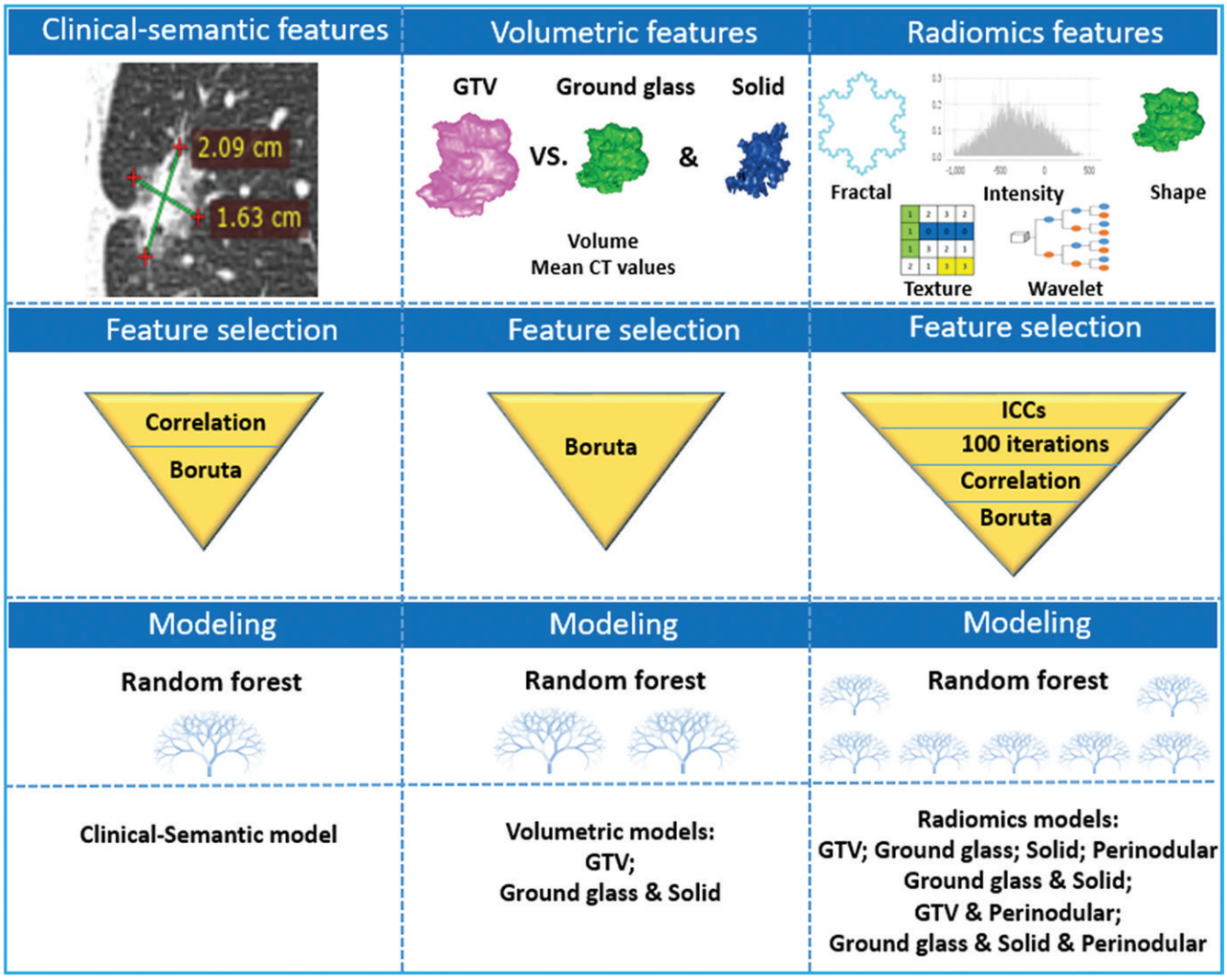

Figure 3: Graph shows the process of semantic, volumetric, and radiomics models. GTV = gross tumor volume, ICC = intraclass correlation coefficient.

(CI) was reported. Accuracy, sensitivity, and specificity were calculated from the confusion matrix.

The significance of the difference between the radiomics model for the ground-glass and solid regions and other models was evaluated using the DeLong method (25). The improvement of the mixed radiomics models with added features compared with models without additional features was evaluated by using the integrated discrimination improvement (26). The Hosmer-Lemeshow test was used to estimate the goodness-of-fit of models, and the calibration plots were generated to visualize the consistency of models (27). The feature selection, modeling, and statistical analysis were performed in R (version 3.5.3; *http://www.r-project.org*). Comprehensive package descriptions are found in Table E2 (online). R codes are available at GitHub *(https://github.com/wu576587441/g.wu)*. $P < .05$ indicated a statistically significant difference. This study followed the Standards for Reporting of Diagnostic Accuracy Studies statement (28) and was assessed using the radiomics quality score (14), a checklist system that evaluates the validity and completeness of radiomics workflows and reporting.

Results

Demographic Characteristics

Of the 1614 pulmonary nodules included in this study, 526 manifested as PSNs, 425 of which were confirmed as adenocarcinomas by their abnormalities. A total of 291 patients (100 men, 191 women; median age of 61 years [interquartile range, 55–65 years]) with 297 PSNs (AIS or MIA, 93; IA, 204) were included after application of exclusion criteria (Fig 1). The demographic characteristics for subgroups grouped by adenocarcinoma type are summarized in Table 1.

Feature Selection

Seven clinical-semantic features were retained, while no volumetric feature was removed. From the 106 radiomic features removed because of low segmentation–based reproducibility, only three features belonged to the nonwavelet group (Figs E2, E3 [online]). The seven (seven, six, three) most robust radiomic features were kept from the GTV region (ground-glass, solid, perinodular regions) of which three (four, three, three) features were selected to build single region of interest–based radiomics models. Between

Table 1: Demographic and Clinical Characteristics of All Patients

Variable	AIS or MIA (n = 93)	IA (n = 204)	P Value
Age (y)*	58 (53–64)	62 (57–66)	.003
No. of men	29 (31)	73 (36)	.44
No. of smokers	9 (10)	26 (13)	.45
Pack-years (only for smokers)*	26.0 (25.5–31.0)	30.5 (28.0–35.8)	.18
Patients with family history	6 (6)	5 (2)	.17
Patients with emphysema	22 (24)	47 (23)	.91
Maximal diameter*	10.6 (8.1–13.6)	17.9 (14.1–24.3)	<.001
Minimal diameter*	7.5 (6.1–9.1)	12.1 (9.3–16.0)	<.001
Solid component size*	4.6 (3.2–6.0)	10.1 (7.0–15.0)	<.001
Solid proportion (%)*	43 (33–55)	61 (45–75)	<.001
Lesion location			.57
Right upper lobe	37 (40)	74 (36)	
Right middle lobe	9 (10)	12 (6)	
Right lower lobe	17 (18)	34 (17)	
Left upper lobe	17 (18)	47 (23)	
Left lower lobe	13 (14)	37 (18)	
Two-dimensional attenuation*	−448.9 (−526.3 to −348.5)	−293.2 (−447.0 to −191.6)	<.001
Shape			.01
Round	85 (91)	162 (79)	
Irregular	8 (9)	42 (21)	
Margin			<.001
Smooth	87 (94)	135 (66)	
Lobulated	5 (5)	55 (27)	
Spiculated	1 (1)	14 (7)	
No. of present bubble lucency	14 (15)	46 (23)	.14
No. of present pleural indentation	32 (34)	132 (65)	<.001
No. of present air bronchogram	16 (17)	67 (33)	.005
No. of present vessel dilatation	1 (1)	19 (9)	.009
Nodule count*	1 (1–2)	1 (1–2)	.18
GTV volume*	0.40 (0.21–0.86)	1.5 (0.8–3.8)	<.001
GTV attenuation*	−503.1 (−575.3 to −438.2)	−401.9 (−554.3 to −303.6)	<.001
Ground-glass volume*	0.36 (0.18–0.77)	1.2 (0.64–2.6)	<.001
Ground-glass attenuation*	−530.5 (−601.7 to −487.3)	−484.2 (−554.3 to −437.7)	<.001
Solid volume*	0.02 (0.01–0.05)	0.25 (0.08–0.69)	<.001
Solid attenuation*	−38.7 (−77.2 to −2.6)	−6.4 (−30.2 to 63.2)	<.001

Note.—Unless otherwise indicated, data are numbers of patients and data in parentheses are percentages. AIS = adenocarcinoma in situ, GTV = gross tumor volume, IA = invasive adenocarcinoma, MIA = minimally invasive adenocarcinoma. P values are a comparison between participants with AIS or MIA and IA groups.

*Data are the median, and data in parentheses are the interquartile range.

six and nine features were retained for the radiomics models based on multiple regions of interest using the aforementioned methods (Appendix E1 [online]). The chosen features are shown in Table E3 (online). The training sets were resampled in the balancing step, except for the Brock model (Table E4 [online]).

Performance of Brock, Clinical-Semantic, and Volumetric Models

All results are reported on the test data set. The online Brock model yielded an AUC of 0.83 (95% CI: 0.72, 0.94) with an accuracy of 76% (95% CI: 65%, 86%), whereas the clinical-semantic signature yielded an AUC and accuracy of 0.90 (95% CI: 0.83, 0.98) and 74% (95% CI: 61%, 84%), respectively. The volumetric model based on both the GTV and ground-glass with solid regions yielded AUC values of 0.87 (95% CI: 0.78, 0.96) and 0.90 (95% CI: 0.83, 0.97), and an accuracy of 74% (95% CI: 61%, 84%) for both models.

Performance of Radiomics Models

The radiomics models trained on single regions of interest (GTV, ground-glass, solid, and perinodular regions) achieved AUC values of 0.88 (95% CI: 0.80, 0.96), 0.96 (95% CI: 0.91, 1.00), 0.82 (95% CI: 0.67, 0.98), and 0.66 (95% CI: 0.49, 0.82) as well as accuracies of 72% (95% CI: 60%, 82%), 90% (95% CI: 80%, 96%), 75% (95% CI: 63%, 85%), and 54% (95% CI: 42%, 67%), respectively.

For radiomics models based on multiple regions of interest (ground-glass with solid region, GTV, and perinodular region

as well as ground-glass with solid and perinodular regions), the AUCs were 0.98 (95% CI: 0.96, 1.00), 0.90 (95% CI: 0.83, 0.98), and 0.95 (95% CI: 0.89, 1.00) (Fig 4), and the accuracies were 93% (95% CI: 84%, 98%), 72% (95% CI: 60%, 82%), and 88% (95% CI: 78%, 95%), respectively. The detailed sensitivities and specificities of the models are summarized in Table 2 (test data set) and Table E4 (online; training data set).

Performance Comparison

There were no significant differences in AUC between the radiomics model trained on features from ground-glass and solid regions and the radiomics model with additional perinodular features (P = .22), as well as the radiomics model based on ground-glass region alone (P = .08). The AUCs of all other models were significantly lower than the radiomics model informed by the combined features extracted from the ground-glass and solid regions (Table 2).

For radiomics models, the integrated discrimination improvement value of the GTV model with additional features extracted from the perinodular region had a nonsignificantly positive integrated discrimination improvement value (0.05; P = .11), and these features improved the AUC value slightly (0.02) compared with the GTV model alone. The model trained with features extracted from both the ground-glass and solid regions adding perinodular features did not improve predictive performance (integrated discrimination improvement: −0.02, P = .56, and AUC value decreased by 0.03). The integrated discrimination improvement values and the AUC improvements are summarized in Table 3.

Calibration and Radiomics Quality Score

Among all models, only two models (radiomics: ground-glass with solid region; radiomics: ground-glass, solid, and perinodular regions) showed proper calibrations with P = .10 and P = .12, respectively (calibration plots depicted in Fig 5). The radiomics quality score of this study is 50% (18 of 36).

Discussion

This multicenter study analyzed individually and in combination the perinodular, ground-glass, and solid components from part-solid nodules (PSNs) to develop and to validate several radiomic signatures to help classify invasive adenocarcinomas in PSNs. We found that a radiomics model combining ground-glass and solid features had a higher area under the curve (AUC, 0.98) and achieved the best accuracy (93%)

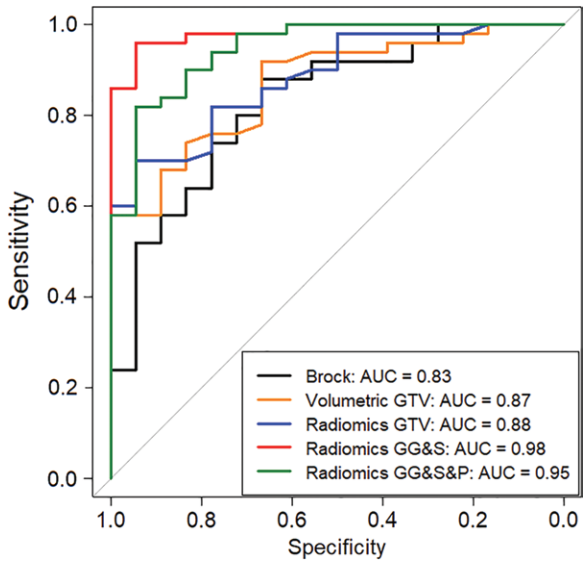

Figure 4: Graph shows the area under the receiver operating characteristic (AUC) curve of different models. GG&S = ground-glass and solid regions, GG&S&P = ground-glass and solid regions with perinodular regions, GTV = gross tumor volume.

Table 2: Detailed Predictive and Diagnostic Values among the Various Models Studied Using the Test Data Set (*n* = 68)

Model	AUC*	Accuracy (%)	Sensitivity (%)	Specificity (%)	*P* Value
Brock	0.83 (0.72, 0.94)	52/68 (76 [65, 86])	39/50 (78 [64, 88])	13/18 (72 [46, 89])	.007
Clinical-semantic	0.90 (0.83, 0.98)	50/68 (74 [61, 84])	33/50 (66 [51, 78])	17/18 (94 [71, 100])	.03
Volumetric					
GTV	0.87 (0.78, 0.96)	50/68 (74 [61, 84])	34/50 (68 [53, 80])	16/18 (89 [64, 98])	.008
GG&S	0.90 (0.83, 0.97)	50/68 (74 [61, 84])	32/50 (64 [49, 77])	18/18 (100 [78, 100])	.02
Radiomics					
GTV	0.88 (0.80, 0.96)	49/68 (72[60, 82])	32/50 (64 [49, 77])	17/18 (94 [71, 100])	.01
Ground glass region	0.96 (0.91, 1.00)	61/68 (90 [80, 96])	44/50 (88 [75, 95])	17/18 (94 [71, 100])	.22
Solid region	0.82 (0.67, 0.98)	51/68 (75 [63, 85])	36/50 (72 [57, 83])	15/18 (83 [58, 96])	.03
Perinodular region	0.66 (0.49, 0.82)	37/68 (54 [42, 67])	27/50 (54 [39, 68])	10/18 (56 [31, 78])	<.001
GG&S	0.98 (0.96, 1.00)	63/68 (93 [84, 98])	49/50 (98 [88, 100])	14/18 (78 [52, 93])	Reference
GTV&P	0.90 (0.83, 0.98)	49/68 (72 [60, 82])	32/50 (64 [49, 77])	17/18 (94 [71, 100])	.02
GG&S&P	0.95 (0.89, 1.00)	60/68 (88 [78, 95])	45/50 (90 [77, 96])	15/18 (83 [58, 96])	.08

Note.—Unless otherwise specified, data are numerators and denominators, with percentages in parentheses and the 95% confidence intervals in brackets. The ground-glass and solid nodule model had the highest area under curve (AUC) of 0.98. *P* values are for comparison of the radiomics ground-glass and solid regions (GG&S) model to each other model using the Delong method. GG&S&P = ground-glass, solid, and perinodular regions; GTV = gross tumor volume; GTV&P = GTV and perinodular regions.

* Data in parentheses are the 95% confidence intervals.

when compared with the other methods: radiomics gross tumor volume (GTV) (AUC, 0.88; P = .01), conventional Brock University (AUC, 0.83; P = .007), clinical-semantic (AUC, 0.90; P = .03), and volumetric GTV (AUC, 0.87; P = .008) models.

A previous study demonstrated that the Brock model has the potential to differentiate IA from AIS or MIA, with a reported AUC of 0.79 (3,29). In our study, the Brock model achieved an AUC of 0.83 but still performed worse than our clinical-semantic model (AUC, 0.90). Several researchers have suggested that size and proportion measures of solid components could be prognostic indicators (30,31). In our study, semantic features containing information from the solid component (solid component size and solid proportion) were selected, with five other selected features, to achieve high classification performance with an AUC of 0.90. Similarly, Luo et al (15) selected three CT features (pleural indentation, solid component size, and solid component proportion) in their model for distinguishing IA from non-IA in PSNs (AUC = 0.85). Likewise, Weng et al (32) used two CT features (lesion shape and solid component size) to yield an AUC of 0.76.

Table 3: Evaluating the Improved Performance of Models on the Test Data Set by Adding Radiomic Features from the Perinodular Region.

Model	Integrated Discrimination Improvement	*P* Value	Improved AUC
GTV	Reference	Reference	Reference
GTV&P	0.05 (−0.01, 0.12)	.11	0.02
GG&S	Reference	Reference	Reference
GG&S&P	−0.02 (−0.09, 0.05)	.56	−0.03

Note.—Data in parentheses are 95% confidence intervals. *P* values are for integrated discrimination improvement. AUC = area under curve; GG&S = ground-glass and solid regions; GG&S&P = ground-glass, solid, and perinodular regions; GTV = gross tumor volume; GTV&P = GTV and perinodular regions.

Volumetric measures can achieve high correlations with clinical outcomes (33). Our volumetric model based on the entire GTV (AUC, 0.87) performed worse than the clinical-semantic model, whereas the volumetric model consisting of measures from both the solid and ground-glass components (volumetric ground-glass and solid regions, AUC of 0.90) had a better predictive performance than the volumetric GTV model and a similar AUC as the clinical-semantic model (with the clinical-semantic model having a slightly higher sensitivity [2%] and the volumetric ground-glass and solid regions having a higher specificity [6%]). This result may indicate that the volumetric ground-glass and solid regions have added value compared with only the GTV and have comparable overall performance with the clinical-semantic model.

In this study, a radiomics model informed by GTV features yielded an AUC value of 0.88 on the test data set. Welch et al (34) suggested that radiomic models should be measured against recognized clinical factors (eg, tumor volume) to demonstrate clinical usefulness. The performance of the radiomics GTV was not higher than the clinical-semantic and volumetric models, and the model did not calibrate well. Previous studies reported lower AUC values than our model, possibly due to smaller sample sizes, higher heterogeneity in the data analyzed, and differing feature selection and modeling workflows (15,32).

Radiomic features may differ in value and clinical usefulness, depending on tumor habitat (35). Our radiomics model based on ground-glass and solid features yielded high AUC and accuracy values. The selected ground-glass features reflect the gray-level run and the distribution, and three solid features describe the heterogeneity of the solid component. The combination of ground-glass and solid features yielded the best AUC (0.98) and accuracy (93%) values while maintaining a proper calibration curve. Additionally, this model was significantly better than the clinical-semantic, volumetric, and radiomics GTV models, and there was an improvement compared with the radiomics models based on ground-glass or solid region features alone. This radiomic signature may be utilized for the

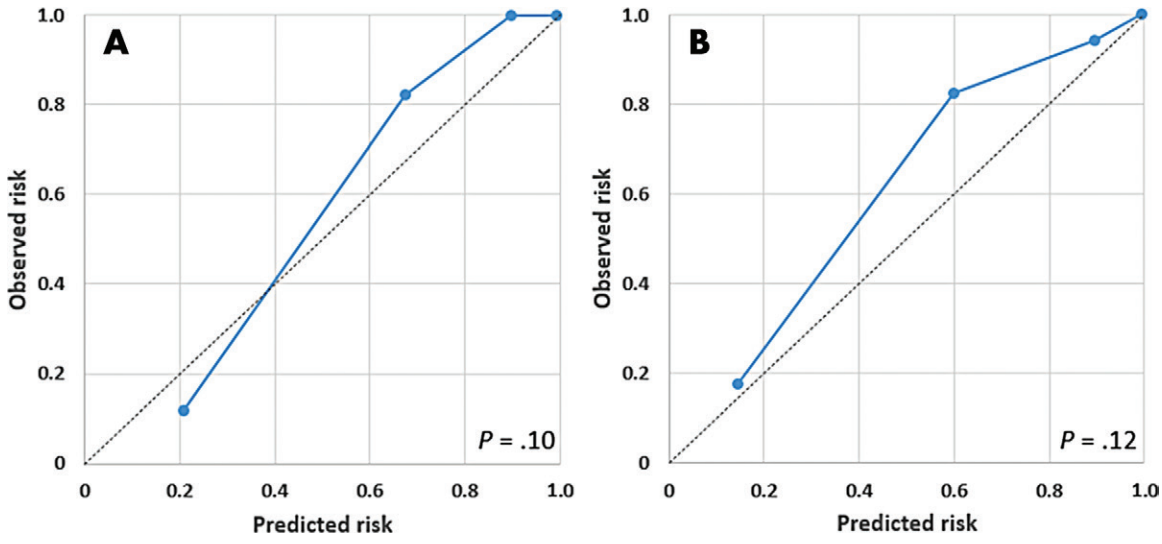

Figure 5: Images show the calibration plots of two radiomics models on the test data set. *A*, The radiomics model based on the features from ground-glass and solid regions. *B*, The radiomics model based on the features from ground-glass, solid, and perinodular regions.

diagnosis of IA, which can be managed with follow-up CT rather than radical surgical intervention.

Previous studies have reported that radiomic features extracted from the peritumoral region could provide additional information when predicting lymph node metastasis and abnormality type (17,18). Our results, however, indicate that the perinodular features do not contribute to radiomics model performance. This is further corroborated by the rarity of perinodular features (ie, vessel convergence) in PSNs.

This study had limitations. First, selection bias is unavoidable in this type of study, which is further confounded by a modest sample size. Second, patients from different countries may have diverse clinical and radiologic characteristics, possibly rendering the models less widely applicable. The same reasoning also applies to varying CT acquisitions and reconstructions that are known to influence the values of handcrafted radiomic feature. Third, although a semiautomatic method was used to contour the solid region, manual segmentation of the GTV is time-consuming and subjective and should ideally be repeated by separate radiologists. Finally, the clinical-semantic model should include more clinical information (eg, genetic information) for a truly detailed clinical model.

In conclusion, we show that a radiomics model combining the ground-glass and solid features from part-solid nodules is better than the five other models that were tested: Brock University, clinical-semantic, volumetric models, radiomic signature based solely on gross tumor volume features, and perinodular features. Further prospective multicenter studies are necessary to address how much margin around the nodule is optimal for diagnosis of the invasiveness of part-solid nodules and to assess the utility of deep learning models for this application.

Author contributions: Guarantors of integrity of entire study, H.C.W., J.W., P.L.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, G.W., H.C.W., J.S., J.X., J.W., J.B.; clinical studies, G.W., R.T.H.L., J.X., J.W., J.B.; statistical analysis, G.W., H.C.W., T.R., R.T.H.L., S.S., A.I., ; and manuscript editing, G.W., H.C.W., T.R., S.S., A.I.

Disclosures of Conflicts of Interest: **G.W.** disclosed no relevant relationships. **H.C.W.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: owns minority shares in Oncodadiomics. Other relationships: disclosed no relevant relationships. **J.S.** disclosed no relevant relationships. **T.R.** disclosed no relevant relationships. **S.S.** disclosed no relevant relationships. **A.I.** disclosed no relevant relationships. **R.T.H.L.** Activities related to the present article: received a grant from Interreg V-A Euregio Meuse-Rhine. Activities not related to the present article: is a shareholder and employee of Oncoradiomics. Other relationships: has been issued patent no. US10311571B2. **R.W.** disclosed no relevant relationships. **J.X.** disclosed no relevant relationships. **J.B.** disclosed no relevant relationships. **J.W.** disclosed no relevant relationships. **P.L.** disclosed no relevant relationships.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69(1):7–34.
2. National Lung Screening Trial Research Team; Church TR, Black WC, et al. Results of initial low-dose computed tomographic screening for lung cancer. N Engl J Med 2013;368(21):1980–1991.
3. McWilliams A, Tammemagi MC, Mayo JR, et al. Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med 2013;369(10):910–919.
4. Ye T, Deng L, Wang S, et al. Lung adenocarcinomas manifesting as radiological part-solid nodules define a special clinical subtype. J Thorac Oncol 2019;14(4):617–627.
5. Lee KH, Goo JM, Park SJ, et al. Correlation between the size of the solid component on thin-section CT and the invasive component on pathology in small lung adenocarcinomas manifesting as ground-glass nodules. J Thorac Oncol 2014;9(1):74–82.
6. Ko JP, Suh J, Ibidapo O, et al. Lung adenocarcinoma: correlation of quantitative CT findings with pathologic findings. Radiology 2016;280(3):931–939.
7. Matsunaga T, Suzuki K, Takamochi K, Oh S. What is the radiological definition of part-solid tumour in lung cancer?†. Eur J Cardiothorac Surg 2017;51(2):242–247.
8. Yoshizawa A, Motoi N, Riely GJ, et al. Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. Mod Pathol 2011;24(5):653–664.
9. Liu S, Wang R, Zhang Y, et al. Precise diagnosis of intraoperative frozen section is an effective method to guide resection strategy for peripheral small-sized lung adenocarcinoma. J Clin Oncol 2016;34(4):307–313.
10. Liang J, Xu XQ, Xu H, et al. Using the CT features to differentiate invasive pulmonary adenocarcinoma from pre-invasive lesion appearing as pure or mixed ground-glass nodules. Br J Radiol 2015;88(1053):20140811.
11. Nair A, Bartlett EC, Walsh SLF, et al. Variable radiological lung nodule evaluation leads to divergent management recommendations. Eur Respir J 2018;52(6):1801359.
12. Revel MP, Mannes I, Benzakoun J, et al. subsolid lung nodule classification: a CT criterion for improving interobserver agreement. Radiology 2018;286(1):316–325.
13. Kim H, Park CM, Woo S, et al. Pure and part-solid pulmonary ground-glass nodules: measurement variability of volume and mass in nodules with a solid portion less than or equal to 5 mm. Radiology 2013;269(2):585–593.
14. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14(12):749–762.
15. Luo T, Xu K, Zhang Z, Zhang L, Wu S. Radiomic features from computed tomography to differentiate invasive pulmonary adenocarcinomas from non-invasive pulmonary adenocarcinomas appearing as part-solid ground-glass nodules. Chin J Cancer Res 2019;31(2):329–338.
16. Fan L, Fang M, Li Z, et al. Radiomics signature: a biomarker for the preoperative discrimination of lung invasive adenocarcinoma manifesting as a ground-glass nodule. Eur Radiol 2019;29(2):889–897.
17. Wang X, Zhao X, Li Q, et al. Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? Eur Radiol 2019;29(11):6049–6058.
18. Dou TH, Coroller TP, van Griethuysen JJM, Mak RH, Aerts HJWL. Peritumoral radiomics features predict distant metastasis in locally advanced NSCLC. PLoS One 2018;13(11):e0206108.
19. Tu W, Li Z, Wang Y, et al. The "solid" component within subsolid nodules: imaging definition, display, and correlation with invasiveness of lung adenocarcinoma, a comparison of CT histograms and subjective evaluation. Eur Radiol 2019;29(4):1703–1713.
20. Beig N, Khorrami M, Alilou M, et al. Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas. Radiology 2019;290(3):783–792.
21. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. Med Phys 2017;44(3):1050–1062.
22. Mackin D, Fave X, Zhang L, et al. Measuring computed tomography scanner variability of radiomics features. Invest Radiol 2015;50(11):757–765.
23. Lucia F, Visvikis D, Vallières M, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2019;46(4):864–877.
24. Kursa MB, Rudnicki WR. Feature selection with the boruta package. J Stat Softw 2010;36(11):1–13.
25. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.
26. Pencina MJ, D'Agostino RB Sr, Demler OV. Novel metrics for evaluating improvement in discrimination: net reclassification and integrated discrimination improvement for normal variables and nested models. Stat Med 2012;31(2):101–113.
27. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. Crit Care Med 2007;35(9):2052–2056.
28. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Radiology 2015;277(3):826–832.
29. Kim H, Goo JM, Park CM. A simple prediction model using size measures for discrimination of invasive adenocarcinomas among incidental pulmonary subsolid nodules considered for resection. Eur Radiol 2019;29(4):1674–1683.
30. Kamiya S, Iwano S, Umakoshi H, et al. Computer-aided volumetry of part-solid lung cancers by using CT: solid component size predicts prognosis. Radiology 2018;287(3):1030–1040.
31. Hwang EJ, Park CM, Ryu Y, et al. Pulmonary adenocarcinomas appearing as part-solid ground-glass nodules: is measuring solid component size a better prognostic indicator? Eur Radiol 2015;25(2):558–567.
32. Weng Q, Zhou L, Wang H, et al. A radiomics model for determining the invasiveness of solitary pulmonary nodules that manifest as part-solid nodules. Clin Radiol 2019;74(12):933–943.
33. Mozley PD, Schwartz LH, Bendtsen C, Zhao B, Petrick N, Buckler AJ. Change in lung tumor volume as a biomarker of treatment response: a critical review of the evidence. Ann Oncol 2010;21(9):1751–1755.
34. Welch ML, McIntosh C, Haibe-Kains B, et al. Vulnerabilities of radiomic signature development: The need for safeguards. Radiother Oncol 2019;130(1):2–9.
35. Refaee T, Wu G, Ibrahim A, et al. The emerging role of radiomics in COPD and lung cancer. Respiration 2020;99(2):99–107.

Radiology

Erratum

Originally published online:
https://doi.org/10.1148/radiol.2020192431

Diagnosis of Invasive Lung Adenocarcinoma Based on Chest CT Radiomic Features of Part-Solid Pulmonary Nodules: A Multicenter Study

Guangyao Wu, Henry C. Woodruff, Jing Shen, Turkey Refaee, Sebastian Sanduleanu, Abdalla Ibrahim, Ralph T. H. Leijenaar, Rui Wang, Jingtong Xiong, Jie Bian, Jianlin Wu, Philippe Lambin

Erratum in:
DOI:10.1148/radiol.2020209019

An author name was published incorrectly. The corrected author name is **Abdalla Ibrahim**.