A Subsolid Nodules Imaging Reporting System (SSN-IRS) for Classifying 3 Subtypes of Pulmonary Adenocarcinoma

Cui, Xiaonan; Heuvelmans, Marjolein A; Fan, Shuxuan; Han, Daiwei; Zheng, Sunyi; Du, Yihui; Zhao, Yingru; Sidorenkov, Grigory; Groen, Harry J M; Dorrius, Monique D

Published in:
Clinical lung cancer

DOI:
10.1016/j.cllc.2020.01.014

Publication date:
2020

Original Study

Check for updates

A Subsolid Nodules Imaging Reporting System (SSN-IRS) for Classifying 3 Subtypes of Pulmonary Adenocarcinoma

Xiaonan Cui,[1,2] Marjolein A. Heuvelmans,[3,4] Shuxuan Fan,[1] Daiwei Han,[2] Sunyi Zheng,[5] Yihui Du,[3] Yingru Zhao,[1] Grigory Sidorenkov,[3] Harry J.M. Groen,[6] Monique D. Dorrius,[2] Matthijs Oudkerk,[7,8] Geertruida H. de Bock,[3] Rozemarijn Vliegenthart,[2] Zhaoxiang Ye[1]

Abstract

It is essential to identify the subsolid nodules subtype preoperatively to select the optimal treatment algorithm. We developed and validated an imaging reporting system using a classification and regression tree model that based on computed tomography imaging characteristics (291 cases in training group, 146 cases in testing group). The model showed high sensitivity and accuracy of classification. Our model can help clinicians to make follow-up recommendations or decisions for surgery for clinical patients with a subsolid nodule.

Objectives: To develop an imaging reporting system for the classification of 3 adenocarcinoma subtypes of computed tomography (CT)-detected subsolid pulmonary nodules (SSNs) in clinical patients. **Methods:** Between November 2011 and October 2017, 437 pathologically confirmed SSNs were retrospectively identified. SSNs were randomly divided 2:1 into a training group (291 cases) and a testing group (146 cases). CT-imaging characteristics were analyzed using multinomial univariable and multivariable logistic regression analysis to identify discriminating factors for the 3 adenocarcinoma subtypes (pre-invasive lesions, minimally invasive adenocarcinoma, and invasive adenocarcinoma). These factors were used to develop a classification and regression tree model. Finally, an SSN Imaging Reporting System (SSN-IRS) was constructed based on the optimized classification model. For validation, the classification performance was evaluated in the testing group. **Results:** Of the CT-derived characteristics of SSNs, qualitative density (nonsolid or part-solid), core (non-core or core), semantic features (pleural indentation, vacuole sign, vascular invasion), and diameter of solid component ($\leq$6 mm or $>$6 mm), were the most important factors for the SSN-IRS. The total sensitivity, specificity, and diagnostic accuracy of the SSN-IRS was 89.0% (95% confidence interval [CI], 84.8%-92.4%), 74.6% (95% CI, 70.8%-78.1%), and 79.4% (95% CI, 76.5%-82.0%) in the training group and 84.9% (95% CI, 78.1%-90.3%), 68.5% (95% CI, 62.8%-73.8%), and 74.0% (95% CI, 69.6%-78.0%) in the testing group, respectively. **Conclusions:** The SSN-IRS can classify 3 adenocarcinoma subtypes using CT-based characteristics of subsolid pulmonary nodules. This classification tool can help clinicians to make follow-up recommendations or decisions for surgery in clinical patients with SSNs.

Clinical Lung Cancer, Vol. 21, No. 4, 314-25

Keywords: Decision trees, Diagnosis, Lung, Solitary pulmonary nodule, X-ray computed tomography

[1]Department of Radiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Centre of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin, The People's Republic of China
[2]Department of Radiology
[3]Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[4]Department of Pulmonology, Medisch Spectrum Twente, Enschede, The Netherlands
[5]Department of Radiotherapy
[6]Department of Pulmonary Diseases, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[7]i-DNA, Groningen, The Netherlands
[8]Faculty of Medical Sciences, University of Groningen, Groningen, The Netherlands

Submitted: Sep 3, 2019; Revised: Dec 24, 2019; Accepted: Jan 20, 2020; Epub: Feb 6, 2020

Address for correspondence: Zhaoxiang Ye, MD, Tianjin Medical University Cancer Institute and Hospital, Huan-Hu-Xi Road, Ti-Yuan-Bei, He Xi District, Tianjin 300060, China
E-mail contact: yezhaoxiang@163.com

https://doi.org/10.1016/j.cllc.2020.01.014

Introduction

The increased utilization of chest computed tomography (CT) examinations, especially in low-dose CT lung cancer screening programs, has improved our awareness of subsolid nodules (SSNs). SSNs include nonsolid nodules (synonymous with pure ground-glass nodules) and part-solid nodules.[1] Persistent SSNs, in particular, part-solid nodules are likely to be malignant and are usually considered to be early stages of primary lung adenocarcinoma. Lung adenocarcinoma is divided into pre-invasive lesions (PLs) (consisting of atypical adenomatous hyperplasia [AAH] and adenocarcinoma-in-situ [AIS]), minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (IA).[2] In 2018, the Eighth Edition Lung Cancer Stage Classification (8th TNM) classified AIS as Tis and MIA as T1a(mi).[3] This is the first time that MIA has been treated as a separate category in the TNM system. Patients with SSNs can benefit from the early classification of nodule invasiveness because patients with MIA are good candidates for sublobar resection with nearly 100% overall survival,[3-5] whereas IA usually requires lobectomy and lymph node dissection.[6,7] PLs, on the other hand, often grow slowly and can be managed by CT follow-up instead of resection until signs of invasiveness occur.[8-10] Therefore, it is essential to identify the adenocarcinoma subtype preoperatively to select the optimal treatment algorithm. Several studies have shown the discriminatory power of radiological characteristics to assess the probability of invasiveness in pulmonary SSNs. CT features like nodule size, solid proportion, CT attenuation, vascular convergence sign, pleural indentation sign, and air-bronchogram sign are considered important factors.[11-22] However, the previous studies only used a binary classifier (PLs [AAH, AIS, MIA] vs. IA; or [AAH, AIS] vs. [MIA, IA]), without recognizing the importance of separate MIA classification.

In this study, we developed a classification model based on CT features of SSNs with histologic information, using a classification and regression tree (CART) to distinguish the 3 adenocarcinoma subtypes.

Materials and Methods

Patients

The institutional review board waived the need for informed consent because of its retrospective design and the use of anonymized data. From November 2011 to October 2017, in 7785 patients at Tianjin Medical University Cancer Institute and Hospital, a pulmonary nodule was surgically resected and confirmed by postoperative histopathological specimens to be AAH, AIS, MIA, or IA. Inclusion criteria for this study were as follows: (1) pathology report with diagnosis of lung adenocarcinoma; (2) preoperative thin-section CT images within 1 month before surgery in the Picture Archiving and Communication System (PACS); and (3) SSNs on CT; lesions were viewed in CT lung window setting, with evaluation of a potential solid component in mediastinal window setting. We excluded 6918 solid nodule cases and 430 patients without available preoperative CT images.

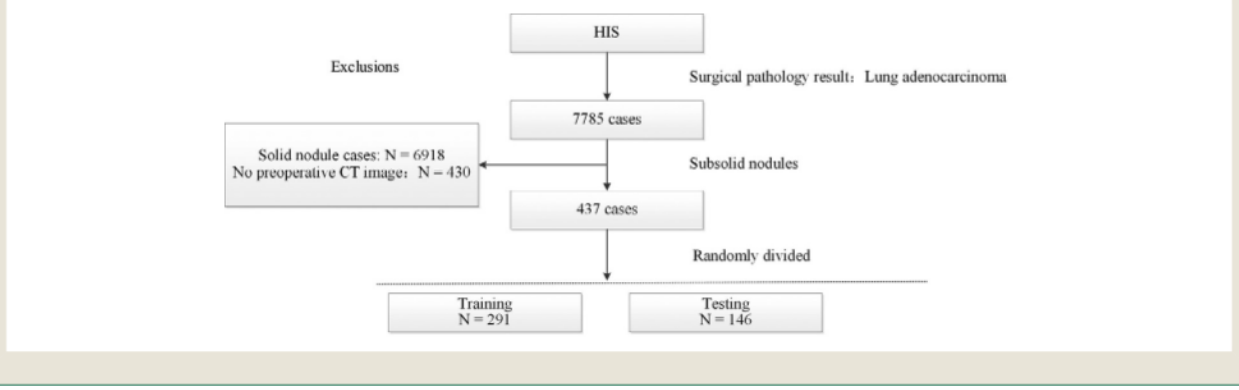
CT Examination

Noncontrast chest CT examinations were conducted using a Discovery CT 750 HD (GE Medical Systems, Milwaukee, WI), Lightspeed 16 (GE Medical Systems), or Somatom Sensation 64 (Siemens Medical Solutions, Forchheim, Germany) CT system. The scan range was from pulmonary apex level to below the diaphragm. The scanning protocol was as follows: tube voltage 120 kVp with automatic tube current modulation. For GE CT systems, reconstructed slice thickness was 1.25 mm, pitch 0.984. For the Siemens CT system, reconstructed slice thickness was 1.5 mm, pitch 0.95.

Evaluation of the CT Characteristics

CTs were retrospectively reviewed. All images were blindly and independently evaluated by 2 experienced radiologists (X.C., S.F., with 5 and 9 years of experience in chest imaging, respectively). The final radiological characteristics were determined by the consensus of the 2 radiologists in case of disagreement. They

Figure 1 Flowchart of the Study



Abbreviations: CT = computed tomography; HIS = hospital information system.

Table 1 Characteristics of Patients in the Training and Testing Group

Characteristics	Total (n = 437)	Training (n = 291)	Testing (n = 146)	*P*
Age, mean (SD), y	57.0 ± 8.9	57.3 ± 8.8	56.4 ± 9.1	.325[a]
Sex				
Man	109	70	39	.559[b]
Woman	328	221	107	
Histologic spectrum				
PLs	88	52	36	.246[b]
MIA	233	159	74	
IA	116	80	36	
Morphological				
Qualitative density				
Nonsolid nodule	233	150	83	.295[b]
Part-solid nodule	204	141	63	
Lobe				
Left upper lobe	115	67	48	.230[b]
Left lower lobe	45	32	13	
Right upper lobe	169	114	55	
Right middle lobe	35	25	10	
Right lower lobe	73	53	20	
Core				
Non-core	76	48	28	.505[b]
Core	361	243	118	
Shape				
Round/oval	308	204	104	.721[b]
Irregular	129	87	42	
Margin				
Smooth	243	159	27	.839[b]
Lobulated	83	56	35	
Spiculated	111	76	52	
Semantic features score, median (IQR)	2 (1-4)	2 (1-4)	2 (0-3)	.341[a]
Size, median (IQR), mm				
Diameter of nodule	13.3 (9.6-18.4)	13.5 (9.9-19.0)	12.7 (8.5-16.8)	.077[a]
Diameter of solid component	0.0 (0.0-5.1)	0.0 (0.0-5.7)	0.0 (0.0-4.2)	.365[a]
CT attenuation (HU), mean (SD)	-369 ± 194	-359 ± 193	-388 ± 195	.147[a]

Abbreviations: HU = Hounsfield units; IA = invasive adenocarcinoma; IQR = interquartile range; MIA = minimally invasive adenocarcinoma; PLs = pre-invasive lesions.
[a]Student *t* test (normally distributed) and Kolmogorov-Smirnov test (non-normally distributed).
[b]Pearson χ^2 test and Fisher exact test.

measured the whole size of SSNs on lung window setting (width, 1450 HU; level, −500 HU), the solid size on mediastinal window setting (width, 350 HU; level, 40 HU), and if the solid component absent, then use the "core" window setting (width, 1250 HU; level, 40 HU) to determine whether the SSN has a minimally invasive lesion component.[18] The following radiological characteristics were included: (1) qualitative density; (2) shape; (3) location; (4) margin; (5) core; (6) semantic feature score (pleural indentation, vacuole sign and vascular invasion); (7) CT attenuation; (8) nodule diameter; and (9) diameter of solid component. The details of the assessment method are described in Supplemental Figures 1 to 8 in the online version.

Histopathological Evaluation

The diagnosis and categorization of the pathological specimens were evaluated by 2 lung pathologists (one junior, one senior). The specimens were fixed in formalin and embedded in paraffin. The adenocarcinoma classification was followed by the IASLC/ATS/ERS adenocarcinoma classification.[2] Immunohistochemistry was performed in case of equivocal histopathological classification under light microscopy.

Statistical Analysis

All nodules (n = 437) were randomly divided (2:1) into a training group (n = 291) and testing group (n = 146) by using a

Table 2 Characteristics of Patients in Different Pathology Classification in the Training Group

Characteristics	PLs		MIA (n = 159)	IA (n = 80)	*P*
	AAH (n = 10)	AIS (n = 42)			
Age, mean (SD), y	58.5 ± 6.0	57.3 ± 9.3	56.5 ± 8.8	58.9 ± 8.9	.143[a]
Sex					
Man	0	15	37	18	.666[b]
Woman	10	27	122	62	
Morphological					
Density					
Nonsolid	10	36	100	4	<.001[b]
Part-solid	0	6	59	76	
Size, median (IQR), mm					
Diameter of nodule	9.2 (8.7-11.2)	9.6 (7.0-10.8)	12.6 (10.2-16.8)	20.4 (17.1-25.4)	<.001[a]
Diameter of solid component	-	0 (0-0)	0 (0-2.5)	10.9 (5.6-15.6)	<.001[a]
CT attenuation (HU), mean (SD)					
Location	-614.2 ± 129.4	-489.5 ± 139.1	-401.3 ± 162.5	-177.6 ± 134.2	<.001[a]
Left upper lobe	2	6	41	18	.405[b]
Left lower lobe	0	3	20	9	
Right upper lobe	6	23	55	30	
Right middle lobe	0	2	17	6	
Right lower lobe	2	8	26	17	
Core					
No	8	22	18	0	<.001[b]
Yes	2	20	141	80	
Shape					
Round/oval	9	37	124	34	<.001[b]
Irregular	1	5	35	46	
Margin					
Smooth	9	35	98	17	<.001[b]
Lobulate	0	1	29	26	
Spiculate	1	6	32	37	

Table 2 Continued

Characteristics	PLs							MIA (n = 159)			IA (n = 80)			*P*
	AAH (n = 10)			AIS (n = 42)										
Semantic features	0	1	2	0	1	2		0	1	2	0	1	2	
Pleural indentation sign	9	1	0	33	8	1		77	50	32	13	5	62	<.001[b]
Vacuole sign	8	2	0	35	6	1		102	38	19	28	31	21	<.001[b]
Vascular invasion sign	9	1	0	35	6	1		62	72	25	5	23	52	<.001[b]
Score, median (IQR)	0 (0-1)			0 (0-1)				2 (1-3)			4 (3-5)			<.001[a]

Abbreviations: AAH = atypical adenomatous hyperplasia; AIS = adenocarcinoma-in-situ; HU, Hounsfield units; IA = invasive adenocarcinoma; IQR = interquartile range; MIA = minimally invasive adenocarcinoma; PLs = pre-invasive lesions.
[a]One-way analysis of variance and Kruskal-Wallis.
[b]Pearson χ^2 test.

random number generator in SPSS. Student *t* test, Kolmogorov-Smirnov test, and χ^2 test were used to assess differences in variables between the training and testing group. Next, 1-way analysis of variance, Kruskal-Wallis test, Pearson's χ^2 test, and Fisher exact test were used to evaluate differences in categorical and continuous variables among the 3 groups (PLs, MIA, IA) in the training group. Third, multinomial multivariable logistic regression was performed to assess the association between the variables that might potentially classify ($P < .05$ in univariable analysis) adenocarcinoma subtype. Variables were entered into the model one by one step using the forward stepwise method. At each step, the most significant variable $P < .05$ is added to the model until none of the stepwise variable left out of the model would have a statistically significant contribution if added to the model, thus resulting in the best performing model for classification of adenocarcinoma subtype. Finally, to develop a discriminatory tool, a CART method was constructed.[23,24] Because in the tree growing algorithm only univariable splits are considered,[25] the variables for the classification tree were included only if previously preselected for the final multinomial multivariable logistic regression model. The depth of the classification tree was set to be no more than 4, because in our study we assumed at least 80% accuracy as the cutoff value for each subtype. After testing classification depth of 3 to 5 nodes, we found that 4 nodes had a relatively good value with high classification accuracy and still resulting in a relatively simple structure.[26] The node will stop splitting when the size of the parent node is less than 15 or a child node is less than 5. Because the classification tree can only give the result of the classification but cannot show the accuracy of the classification, we used concise tables to simplify the classification tree diagram and made 5 thresholds to describe the percentage of adenocarcinoma subtype in each node. We assumed ≥80% accuracy as the cutoff value for each subtype. Five threshold grades (SSN1 to SSN5) were made to describe the accuracy of adenocarcinoma pathology classification in each child node (SSN1 = IF PLs%≥80%; SSN2 = IF PLs%<80% AND MIA%<80% AND IA%<80% AND PLs%> IA%; SSN3 = IF MIA%≥80%; SSN4 = IF PLs%<80% AND MIA%<80% AND IA%<80% AND PLs%≤IA%; SSN5 = IF IA%≥80%). Finally, the Subsolid Nodules Imaging Reporting System (SSN-IRS) was developed based on the optimized classification tree model. The classification performance of the SSN-IRS was examined in the testing group. Statistical analyses were performed using SPSS software version 20.0 (IBM Corp, New York, NY). $P < .05$ was considered to be statistically significant and all tests were 2-tailed, unless otherwise indicated.

Results

Data Randomization and the General Population

In total, 437 patients were included, 75.1% of whom were women (328/437), with a mean age of 57.0 years. The most common adenocarcinoma subtype was MIA (n = 233, 53.3%), then IA (n = 116, 26.5%), AIS (n = 71, 16.2%) and AAH (n = 17, 3.9%) (Figure 1, Table 1). No significant differences were found between the training and testing group in terms of patient and nodule characteristics.

Table 3 Multinomial Logistic Regression Analyses

Adenocarcinoma Subtype	Factors	*P*	OR	95% CI Low	95% CI up
PLs	Core	<.01	8.52	3.58	20.27
	Score	<.01	0.35	0.22	0.55
	Diameter of solid component	.37	1.13	0.87	1.46
IA	Core	-	-	-	-
	Score	.16	1.26	0.92	1.72
	Diameter of solid component	<.01	1.46	1.28	1.67

The reference category is: MIA.
Abbreviations: CI = confidence interval; IA = invasive adenocarcinoma; MIA = minimally invasive adenocarcinoma; OR = odds ratio; PLs = pre-invasive lesions.

Development of the Classification Model

In the training group, 8 variables (density, shape, margin, core, score, nodule diameter, diameter of solid component, CT attenuation) were potentially different among the 3 subtype groups with $P < .05$ (Table 2). Of these 8 variables, 3 were selected by the forward stepwise method for entering to the multinomial multivariable regression models shown in Table 3. The core and semantic feature score were significant factors to differentiate between the PLs and MIA groups (odds ratio [OR], 8.5; 95% confidence interval [CI], 3.6-20.3, and OR, 0.3; 95% CI, 0.2-0.5); diameter of solid component was a distinguishing factor for the IA and MIA group (OR, 1.5; 95% CI, 1.3-1.7; $P < .001$). These 3 factors (core, semantic feature score, and diameter of solid component) were entered into the CART. Four categorical variables (diameter of solid component [≤6.5 mm, >6.5 mm]; diameter of solid component [≤0.5 mm, >0.5 mm]; core [non-core, core]; semantic feature score [0,1-2,3-6]) were automatically generated by the CART (Figure 2).

Development of the SSN-IRS

In our data, the diameter was an integer in millimeters. To facilitate the measurement, we used 0 mm and 6 mm instead of 0.5 mm and 6.5 mm as the cutoff of the diameter of solid component, which did not change the results. We considered the diameter of solid component ≤0 as nonsolid nodule and ≥0 as part-solid nodules. Five threshold grades SSN 1-5 were used to make a concise table to describe the percentage of adenocarcinoma subtype in each node. The SSN-IRS based on the optimized classification diagrams is shown in Figure 3. We considered SSN1 as PLs, SSN2 as PLs or MIA, SSN3 as MIA, SSN4 as MIA or IA, SSN5 as IA. The sensitivity, specificity, and classification accuracy of the MIA group was 91.2% (85.7%-95.1%), 59.9% (51.0%-68.3%), and 77.0% (71.7%-81.7%). The total sensitivity, specificity, and accuracy of the SSN-IRS was 89.0% (95% CI, 84.8%-92.4%), 74.6% (95% CI, 70.8%-78.1%), and 79.4% (95% CI, 76.5%-82.0%), respectively (Table 4).

Validation of the SSN-IRS

The classification performance of SSN-IRS was evaluated using the testing group. The sensitivity, specificity, and accuracy of the MIA group was 91.9% (83.2%-97.0%), 43.1% (31.4%-55.3%), and 67.8% (59.6%-75.3%). The total sensitivity, specificity, and accuracy was 84.9% (95% CI, 78.1%-90.3%), 68.5% (95% CI, 62.8%-73.8%), and 74.0% (95% CI 69.6%-78.0%), respectively (Table 5).

Discussion

In this retrospective study, 4 radiological characteristics of SSNs were identified as distinguishing factors of 3 adenocarcinoma subtypes preoperatively (PLs, MIA, IA). These characteristics were qualitative density, core, semantic feature score, and diameter of the solid. The SSN-IRS resulted in a high sensitivity, specificity, and classification accuracy 84.9% (95% CI, 78.1%-90.3%), 68.5% (95% CI, 62.8%-73.8%), and 74.0% (95% CI, 69.6%-78.0%), respectively. Thus, the SSN-IRS can support clinicians in making treatment recommendations based on expected pathology.

Several recent studies focused on discriminating the invasiveness of adenocarcinoma. Table 6 shows an overview of these studies with characteristics and classification performance. The diameter of the solid component measured in mediastinal window setting has been found to be one of the most important factors in discriminating between IA and PLs (MIA, AIS, AAH) in most studies. Zhang et al,[17] Wu et al,[19] Liu et al,[20] Yue et al,[22] and our study have shown cutoff values for the diameter of the solid component of 6.7 mm, 3.0 mm, 8.1 mm, 7.0 mm, and 6.0 mm, respectively. It is interesting to note that the Lung CT screening reporting and data system (Lung-RADS 1.1)[27] categorizes part-solid nodule with solid diameter larger than 6 mm as 4A, which is considered suspicious for malignancy. Small solid or semisolid components that represent early signs of IA may be rendered invisible in mediastinal window setting.[8] This is the reason that some groups have studied the optimization of the window setting for detection of a solid core in correlation with pathological invasiveness.[28-30] In our study, we have not devised a new window setting, but have used the optimized window setting as derived in the study by Mao et al.[18] This study showed that a window width of 1250 HU and window level of 40 HU was useful for viewing the invasive components of SSNs. Based on the previous studies and our clinical practice, we prefer to use mediastinal window setting to measure the solid size, which considered as invasive lesion, and if absent, used the "core" window setting (WW 1250 HU; WL 40 HU) to determine whether the SSN had minimally invasive lesions. Our results showed the presence of a core was an important feature in classifying MIA. The semantic features (pleural indentation, vacuole sign, and vascular invasion) are also considered as important factors that closely relate

Figure 2 A Classification Tree Model for the Classification of Adenocarcinoma Types by Classification and Regression Tree (CART) Analysis in the Training Group

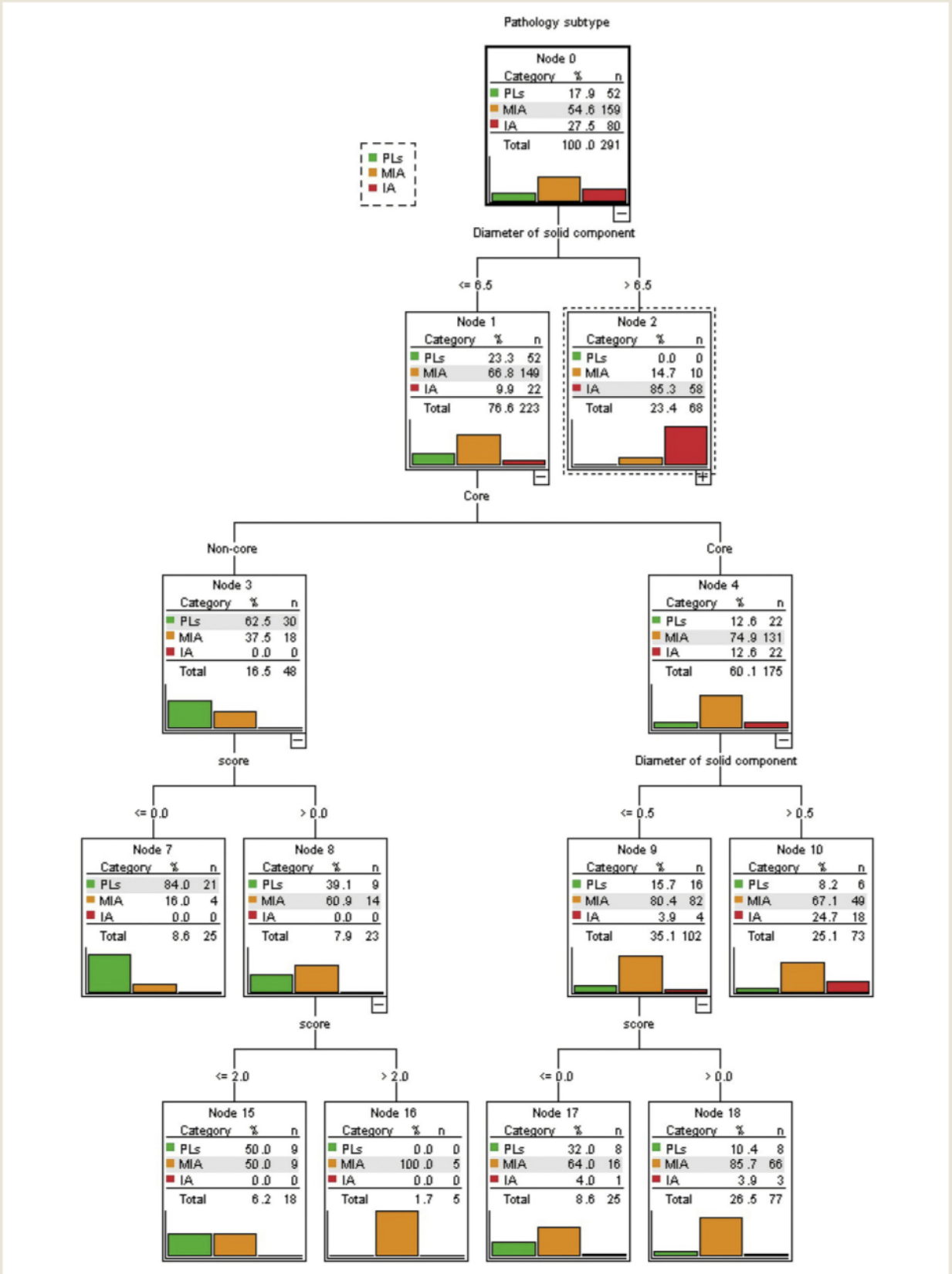
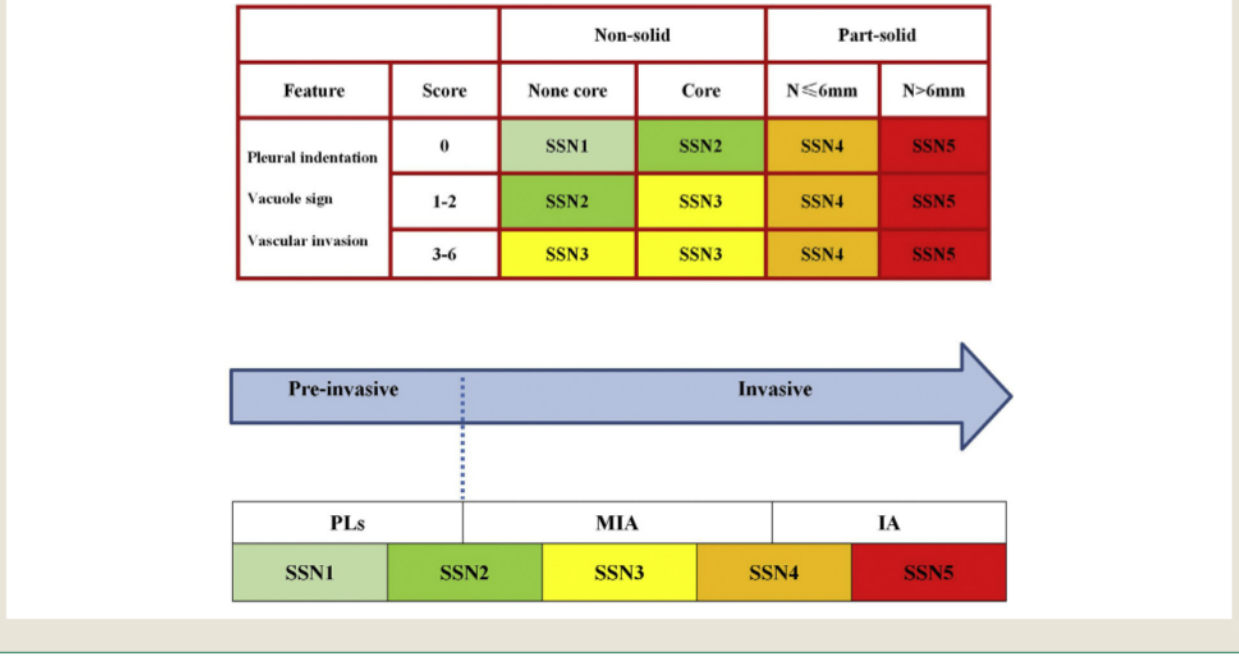


Figure 3 Subsolid Nodules Imaging Reporting System (SSN-IRS)



Abbreviations: IA = invasive adenocarcinoma; MIA = minimally invasive adenocarcinoma; PL = pre-invasive lesion.

to the invasive diagnosis of SSNs.[11-22] All these studies used a binary classifier (yes/no), but in our clinical practice, we found that the invasiveness of SSNs seems also related to the apparent degree of feature presence. In our study, we therefore used 3 scores for each feature, score 0 = no, and score 1 = yes, and score 2 = obvious. As shown in Table 2, nodules with the highest score (score 2) were most likely to represent an invasive lesion. The sum of 3 semantic feature scores instead of 3 independent factors in our SSN-IRS also showed a significant difference between all subtypes in our data. With regard to the classification performance, prior studies have shown sensitivity from 53.8% to 86.7%, and specificity from 44.4% to 84.6%. Unfortunately, we cannot directly compare the performance of the entire SSN-IRS with previous studies because our data used MIA as an independent category, whereas others included MIA into the PL category or invasive category. In addition, most prior studies did not use a separate training and testing group, as we did. The separate sensitivity and specificity of the SSN-IRS for discriminating IA was 95.0% and 69.2% in the training group, and 86.1% and 70.9% in the testing group; thus, our SSN-IRS showed relatively high classification performance, in particular for sensitivity. This may be because of the relatively large sample size and the application of the CART model.

Table 4 The Crossover Table of Prediction of SSN-IRS Threshold Grade in the Training Group

Adenocarcinoma Subtype	SSN-IRS Threshold Grade					Total	Sensitivity 95% CI	Specificity 95% CI	Accuracy 95% CI
	SSN1	SSN2	SSN3	SSN4	SSN5				
PLs	21 40.4	17 32.7	8 15.4	6 11.5	0 0.0	52 100.0	73.1 59.0-84.4	87.5 82.6-91.4	84.9 80.2-88.8
MIA	4 2.5	25 15.7	71 44.7	49 30.8	10 6.3	159 100.0	91.2 85.7-95.1	59.9 51.0-68.3	77.0 71.7-81.7
IA	0 0.0	1 1.2	3 3.8	18 22.5	58 72.5	80 100.0	95.0 87.7-98.6	69.2 62.5-75.4	76.3 71.0-81.1
Overall	25 8.6	43 14.8	82 28.2	73 25.1	68 23.4	291 100.0	89.0 84.8-92.4	74.6 70.8-78.1	79.4 76.5-82.0

All values are %.
Abbreviations: CI = confidence interval; IA = invasive adenocarcinoma; MIA = minimally invasive adenocarcinoma; PLs = pre-invasive lesions; SSN = subsolid nodule; SSN-IRS = Subsolid Nodules Imaging Reporting System.

Table 5 The Crossover Table of Prediction of SSN-IRS Threshold Grade and Pathology Classification in the Testing Group

Adenocarcinoma Subtype	SSN-IRS Threshold Grade					Total	Sensitivity 95% CI	Specificity 95% CI	Accuracy 95% CI
	SSN1	SSN2	SSN3	SSN4	SSN5				
PLs	13 36.1	12 33.3	5 13.9	6 16.7	0 0.0	36 100.0	69.4 51.9-83.7	82.7 74.4-89.3	79.5 72.0-85.7
MIA	1 1.4	16 21.6	31 41.9	21 27.4	5 6.8	74 100.0	91.9 83.2-97.0	43.1 31.4-55.3	67.8 59.6-75.3
IA	1 2.8	1 2.8	3 8.3	14 38.9	17 47.2	36 100.0	86.1 70.5-95.3	70.9 61.5-79.2	74.7 66.8-81.5
Overall	15 10.3	29 19.9	39 26.7	41 28.1	22 15.1	146 100.0	84.9 78.1-90.3	68.5 62.8-73.8	74.0 69.6-78.0

All values are %.
Abbreviations: CI = confidence interval; IA = invasive adenocarcinoma; MIA = minimally invasive adenocarcinoma; PLs = pre-invasive lesions; SSN = subsolid nodule; SSN-IRS = Subsolid Nodules Imaging Reporting System.

Our design has several innovative points. First, the most important novelty of our study is classifying MIA as a separate adenocarcinoma subgroup in our model. We used concise tables to simplify classification diagrams and used 5 thresholds, SSN 1-5, to describe the percentage of adenocarcinoma subtypes (PLs, MIA, and IA), instead of a binary classifier (PLs and IA) in previous studies. This is relevant because recommended treatment for PLs, MIA, and IA differ in the most recent guidelines.[3,31] In our SSN-IRS, the category SSN1 is considered as pre-invasive (AAH or AIS); the SSN2 is considered as suspicious minimally invasive (AIS or MIA); the SSN3 is considered as minimally invasive (MIA); the SSN4 is considered as suspicious invasive (MIA or IA), and the SSN5 is considered as high-risk invasive (IA). The SSN-IRS could be useful for clinicians to make accurate personalized

Table 6 Overview of Previous Studies on the CT-based Classification of SSN Invasiveness

Study	Year	N(SSNs)	Classification	Feature (Cutoff Value)	Research Cohort	Sensitivity, %	Specificity, %	AUC
Zhang et al[17]	2016	237	(AIS, MIA) vs. IA	1. Diameter of nodule (12.2 mm)	Single cohort	85	62	0.75
				2. Diameter of solid component (6.7 mm)		79	62	0.79
				3. CT value of solid component (-192 HU)		77	62	0.74
				4. Air-bronchogram		-	-	0.64
Mao et al[18]	2016	209	(AAH, AIS) vs. (MIA, IA)	1. Diameter of nodule (8.9 mm)	Single cohort	60.7	71.6	0.683
				2. Window width (1250 HU)		57.8	85.0	0.749
Wu et al[19]	2017	141	(AAH, AIS, MIA) vs. IA	1. Diameter of nodule (12 mm)	Single cohort	84.6	76.3	0.891
				2. Diameter of solid component (3 mm)		76.9	94.7	0.881
				3. Density (part-solid)		81.5	88.2	0.886
				4. Air-bronchogram		53.8	89.5	0.717
Liu et al[20]	2017	334	(AAH, AIS, MIA) vs. IA	1. Tumor volume (1125 mm^3)	Single cohort	71.43	84.54	0.809
				2. Tumor mass (386)		82.14	78.35	0.829
			(AAH, AIS, MIA) vs. IA in part-solid	1. Diameter of solid component (8.1 mm)	Single cohort	76.80	90.48	0.904
				2. CT attenuation (-222 HU)		82.84	85.71	0.867
			AIS vs. AAH in nonsolid	1. Tumor mass (70)	Single cohort	86.67	48.48	0.664
				2. Standard deviation of CT attenuation (68)		95.00	44.44	0.656
Jin et al[21]	2017	273	(Benign, AAH, AIS, MIA) vs. IA	CT features model (diameter of nodule, solid proportion, pleural indentation, vacuole sign and vascular invasion, CT attenuation, margin, age, family history of lung cancer)	Single cohort (Threshold = 65%)	78.6	82.5	-
Yue et al[22]	2018	260	MIA vs. IA	1. Diameter of nodule (14.7 mm)	Single cohort	90	81	0.91
				2. Diameter of solid component (7 mm)		95	82	0.95

Table 6 Continued

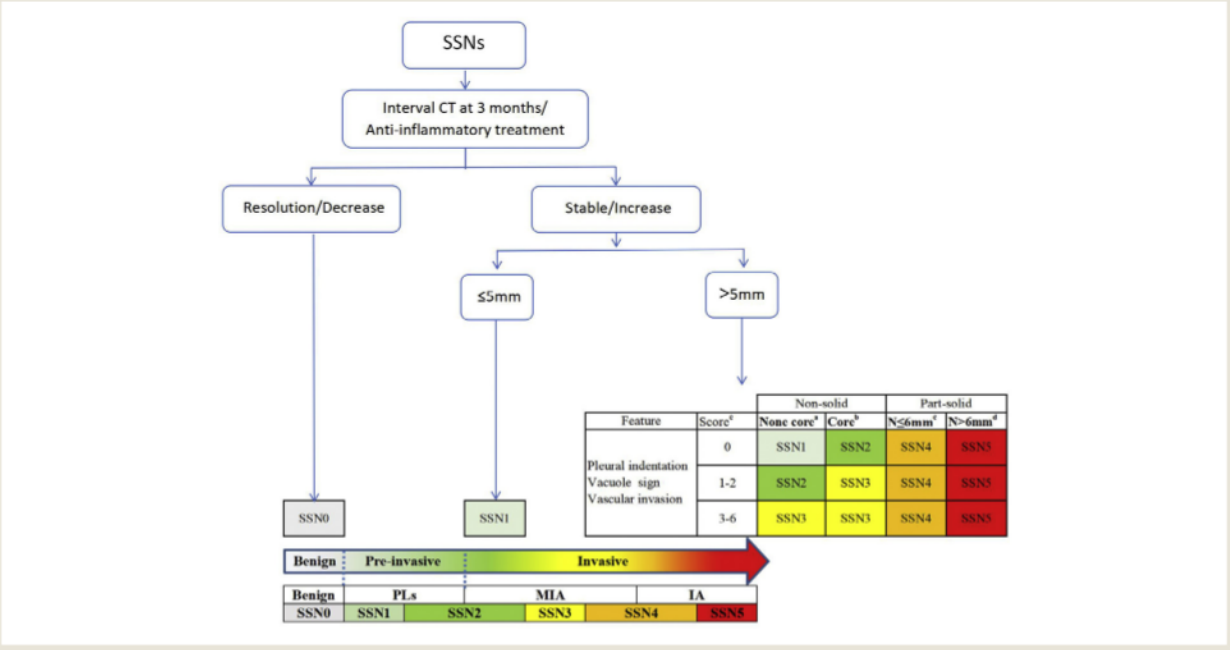
Study	Year	N(SSNs)	Classification	Feature (Cutoff Value)	Research Cohort	Sensitivity, %	Specificity, %	AUC
				3. Proportion of the solid component (<50%)		-	-	0.81
				4. CT value of solid component (-107 HU)		92	77	0.91
				5. Air-bronchogram		-	-	0.60
				6. Vascular invasion		-	-	0.63
Fan et al[33]	2018	395	(AAH, AIS, MIA) vs. IA	1. Radiomics features model	Primary cohort	83.1	89.6	0.917
					Intra-cross validation cohort	87.5	94.4	0.971
					External validation cohort 1	82.2	86.7	0.849
					External validation cohort 2	85.7	89.8	0.869
				2. CT features model (lobulation, spiculation, spine-like, pleural indentation, air-bronchogram, CT attenuation)	Primary cohort	-	-	0.857
Mei et al[37]	2018	1062	(AAH, AIS, MIA) vs. IA	CT features random forest model (lobulation, pleural indentation, vacuole sign and vascular invasion, air-bronchogram, calcification, spiculation, nodule density, CT attenuation, diameter of nodule, diameter of solid component)	Single cohort	80.7	84.6	-
Oikonomou et al[38]	2019	109	(AAH, AIS, MIA) vs. IA	CT features logistic model (mean CT attenuation; volume; and diameter ratio)	Single cohort	80.0	90.9	0.89
Zhan et al[39]	2019	313	(AIS, MIA) vs. IA	CT features logistic model (nodule density, vacuole sign, diameter of nodule, CT attenuation, tumor-lung interface)	Single cohort	80.3	81.0	0.847
Our study SSN-IRS		437	(AAH, AIS) vs. MIA vs. IA	CT features logistic and CART model [Nodule density, core, pleural indentation, vacuole sign, vascular invasion, and diameter of solid component (6 mm)]	Training cohort	89.0	74.6	-
					Testing cohort	84.9	68.5	-

Abbreviations: AAH = atypical adenomatous hyperplasia; AIS = adenocarcinoma-in-situ; AUC = area under the curve; CT = computed tomography; HU, Hounsfield units; IA = invasive adenocarcinoma; MIA = minimally invasive adenocarcinoma; PLs = pre-invasive lesions; SSN = subsolid nodule; SSN-IRS = Subsolid Nodules Imaging Reporting System.

treatment recommendations, because the current treatment strategy differs by adenocarcinoma subtype, especially for MIA.[3,6,7,31] Second, we used a classification model (CART) to visualize the diagnostic process; different from previous models, our SSN-IRS is easier to apply in clinical practice by using 4 radiological factors visualized by radiologists, instead of computerized radiomics features.[32-34] We recently proposed an initial SSN management (see Figure 4) based on the Chinese expert consensus on the management of ground-glass nodules suspected as lung adenocarcinoma,[31] the Eighth Edition Lung Cancer Stage Classification (8th TNM) in 2018,[3] and the Fleischner Society 2017 guidelines for the management of incidentally detected small pulmonary nodules.[8]

Our study has some limitations. First, our study population may reflect a degree of selection bias because our data only included SSNs with known pathology, which may present more malignancy features than nonresected SSNs. The current study comprises incidentally detected cases in a clinical setting; whether the findings can be generalized to other settings needs further investigation. In future research, we plan to verify our SSN-IRS in a lung cancer screening program. Second, our study is a single-center retrospective study. The results may be limited to a local population of Chinese patients.[35,36] Part of the categorization factors, such as age cutoff above 40 years, could be specific for the Chinese setting.[31] The generalizability of our SSN-IRS needs to be verified in different populations. Finally, the volume radiological characteristics of nodules and the clinical information of patients were incomplete. We did not have information on

Figure 4 Subsolid Nodule Management Initial Proposal Based on the Subsolid Nodules Imaging Reporting System (SSN-IRS). a: Non-core: For Nonsolid Nodules, There Was No Solid Component Observed on Core Window Setting (Width, 1250 HU; Level, 40 HU). b: Core: For Nonsolid Nodules, There Was a Solid Component Observed on Core Window Setting (Width, 1250 HU; Level, 40 HU). c: N ≤ 6 mm: The Diameter of the Solid Component Observed on Mediastinal Window Setting (Width, 350 HU; Level, 40 HU) is Less than 7mm. d: N > 6 mm: The Diameter of the Solid Component Observed on Mediastinal Window Setting (Width, 350 HU; Level, 40 HU) Is Larger than 7mm. e: Score: The Sum of the 3 Feature Scores; Pleural Indentation (Score 0: There is No Pleural Indentation or the Nodule is Not Located Close to the Pleura; Score 1: Nodules Adhering to Pleura or Pleural Indentation With 1 Stripe; Score 2: Typical Pleural Indentation With ≥2 Stripes); Vacuole Sign (Score 0: There is no Vacuole; Score 1:air-Bronchograms in Nonsolid Nodule or Single Cystic Cavity; Score 2: Dilated Air-Bronchograms or Multiple Cystic Cavities); Vascular Invasion (Score 0: no Vascular in the Nodule or With Preservation of Normal Vessels; Score 1: Distorted or Dilated Vessels; Score 2: Coexistence of Irregular Vascular Dilation or Vascular Convergence From Multiple Supplying Vessels)



Abbreviations: IA = invasive adenocarcinoma; MIA = minimally invasive adenocarcinoma; PL = pre-invasive lesion.

smoking history, cancer history, and family history of cancer. Future studies should evaluate whether adding clinical factors can improve the model.

In conclusion, the SSN-IRS is a classification model for classifying 3 adenocarcinoma pathologic subtypes. It can assist radiologists in making a more accurate diagnosis and help clinicians to make management decisions for patients with an SSN.

Clinical Practice Points

- It is essential to identify the SSN subtype preoperatively to select the optimal treatment algorithm, because recommended treatment differs for patients with PLs, MIA, and IA.
- Several studies have shown the discriminatory power of radiological characteristics to classify PLs and IA (so binary classification) in pulmonary SSNs.
- CT features like nodule size, solid proportion, CT attenuation, vascular convergence sign, pleural indentation sign, and air-bronchogram sign are considered important factors; however, there are still many controversies about how these features should be applied in the classification of 3 subtypes of adenocarcinoma.
- In our study, the SSN-IRS based on CT-imaging characteristics has high classification performance for the 3 adenocarcinoma subtypes and could be useful for clinicians to make accurate personalized treatment recommendations.

Acknowledgments

This study has received funding by the Royal Netherlands Academy of Arts and Sciences (grant number. PSA_SA_BD_01) and Ministry of Science and Technology of the People's Republic of China, National Key R&D Program of China (grant number. 2016YFE0103000).

We thank Dr. Estelle Noach for her linguistic assistance during the preparation of this manuscript and thank Yisong Wang for her full support for my work.

Disclosure

M. Oudkerk has a financial interest in the company iDNA B.V. All other authors state that they have no conflicts of interest.

Supplemental Data

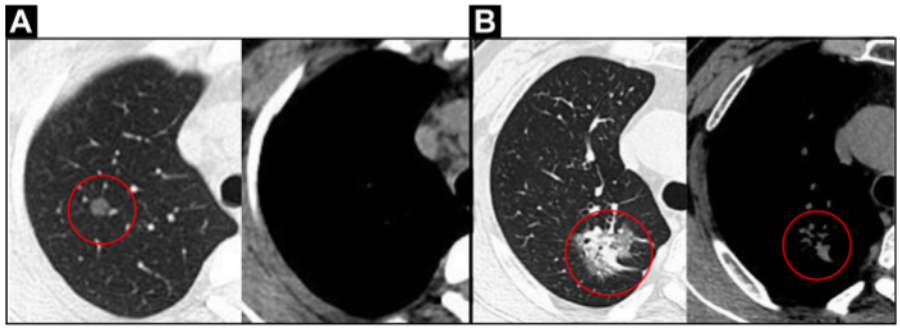
Supplemental figures accompanying this article can be found in the online version at https://doi.org/10.1016/j.cllc.2020.01.014.

References

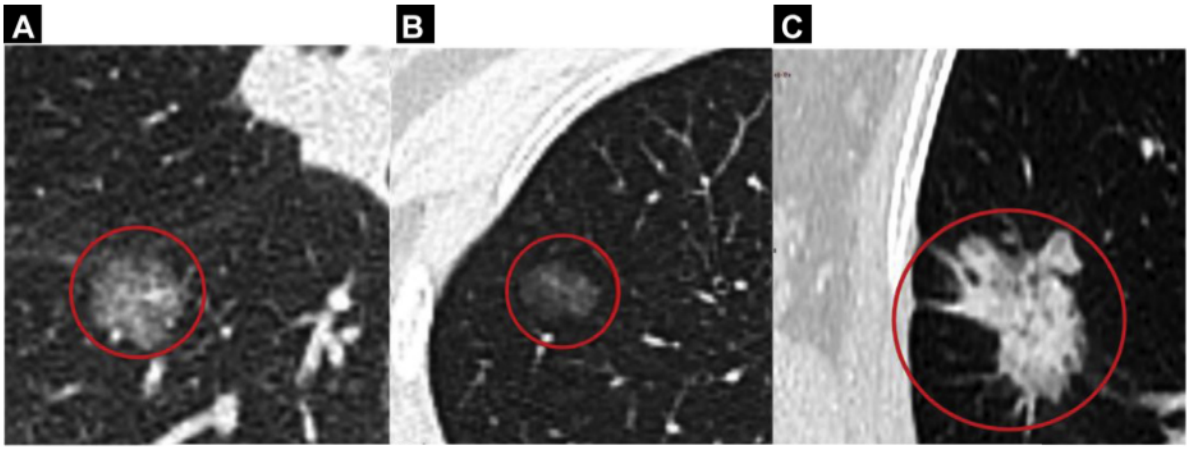
1. Callister ME, Baldwin DR, Akram AR, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. *Thorax* 2015; 70(Suppl 2):ii1-54.
2. Travis WD, Brambilla E, Noguchi M, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 2011; 6:244-85.
3. Detterbeck FC, Boffa DJ, Kim AW, et al. The eighth edition lung cancer stage classification. *Chest* 2017; 151:193-203.
4. Oka S, Hanagiri T, Uramoto H, et al. Surgical resection for patients with mucinous bronchioloalveolar carcinoma. *Asian J Surg* 2010; 33:89-93.
5. Kadota K, Villena-Vargas J, Yoshizawa A, 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.
6. Kodama K, Higashiyama M, Takami K, et al. Treatment strategy for patients with small peripheral lung lesion (s): intermediate-term results of prospective study. *Eur J Cardio-Thorac Surg* 2008; 34:1068-74.
7. 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; 34:307-13.
8. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology* 2017; 284:228-43.
9. Pedersen JH, Saghir Z, Wille MM, et al. Ground-glass opacity lung nodules in the era of lung cancer CT screening: radiology, pathology, and clinical management. *Oncology* 2016; 30:266-74.
10. Kobayashi Y, Sakao Y, Deshpande GA, et al. The association between baseline clinical-radiological characteristics and growth of pulmonary nodules with ground-glass opacity. *Lung Cancer* 2014; 83:61-6.
11. Ikeda K, Awai K, Mori T, et al. Differential diagnosis of ground-glass opacity nodules: CT number analysis by three-dimensional computerized quantification. *Chest* 2007; 132:984-90.
12. McWilliams A, Tammemagi MC, Mayo JR, et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med* 2013; 369:910-9.
13. Zhang Y, Qiang JW, Ye JD, et al. High resolution CT in differentiating minimally invasive component in early lung adenocarcinoma. *Lung Cancer* 2014; 84:236-41.
14. Xiang W, Xing Y, Jiang S, 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.
15. Xing Y, Li Z, Jiang S, et al. Analysis of pre-invasive lung adenocarcinoma lesions on thin-section computerized tomography. *Clin Respir J* 2015; 9:289-96.
16. Lee SM, Goo JM, Lee KH, et al. CT findings of minimally invasive adenocarcinoma (MIA) of the lung and comparison of solid portion measurement methods at CT in 52 patients. *Eur Radiol* 2015; 25:2318-25.
17. Zhang Y, Shen Y, Qiang JW, et al. HRCT features distinguishing pre-invasive from invasive pulmonary adenocarcinomas appearing as ground-glass nodules. *Eur Radiol* 2016; 26:2921-8.
18. Mao H, Labh K, Han F, et al. Diagnosis of the invasiveness of lung adenocarcinoma manifesting as ground glass opacities on high-resolution computed tomography. *Thorac Cancer* 2016; 7:129-35.
19. Wu FZ, Chen PA, Wu CC, et al. Semiquantative visual assessment of sub-solid pulmonary nodules ≦3 cm in differentiation of lung adenocarcinoma spectrum. *Sci Rep* 2017; 7:15790.
20. Liu Y, Sun H, Zhou F, et al. Imaging features of TSCT predict the classification of pulmonary preinvasive lesion, minimally and invasive adenocarcinoma presented as ground glass nodules. *Lung Cancer* 2017; 108:192-7.
21. Jin C, Cao J, Cai Y, et al. A nomogram for predicting the risk of invasive pulmonary adenocarcinoma for patients with solitary peripheral subsolid nodules. *J Thorac Cardiovasc Surg* 2017; 153:462-9.e461.
22. Yue X, Liu S, Liu S, et al. HRCT morphological characteristics distinguishing minimally invasive pulmonary adenocarcinoma from invasive pulmonary adenocarcinoma appearing as subsolid nodules with a diameter of <!=3 cm. *Clin Radiol* 2018; 73:411.e7-15.
23. Yang Z, Dong L, Zhang Y, et al. Prediction of severe acute pancreatitis using a decision tree model based on the revised atlanta classification of acute pancreatitis. *PLoS One* 2015; 10:e0143486.
24. Kim YH, Kim M-J, Shin HJ, et al. MRI-based decision tree model for diagnosis of biliary atresia. *Eur Radiol* 2018; 28:3422-31.
25. Steinberg D, Colla P. CART: classification and regression trees. *The Top Ten Algorithms Data Mining. Chapman and Hall/CRC* 2009; 9:179.
26. Tang T-I, Zheng G, Huang Y, Shu G, Wang P. A comparative study of medical data classification methods based on decision tree and system reconstruction analysis. *Ind Eng Manag Syst* 2005; 4:102-8.
27. Lung CT Screening Reporting & Data System (Lung-RADS). Available at: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads2019. Accessed: August 2, 2019.
28. Matsuguma H, Nakahara R, Anraku M, et al. Objective definition and measurement method of ground-glass opacity for planning limited resection in patients with clinical stage IA adenocarcinoma of the lung. *Eur J Cardiothorac Surg* 2004; 25:1102-6.
29. Matsuguma H, Oki I, Nakahara R, et al. Comparison of three measurements on computed tomography for the prediction of less invasiveness in patients with clinical stage I non–small cell lung cancer. *Ann Thorac Surg* 2013; 95:1878-84.
30. 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; 9:74-82.
31. Jiang G, Chen C, Zhu Y, et al. [Shanghai Pulmonary Hospital Experts Consensus on the Management of Ground-Glass Nodules Suspected as Lung Adenocarcinoma (Version 1)]. *Zhongguo Fei Ai Za Zhi* 2018; 21:147-59.
32. Xue X, Yang Y, Huang Q, et al. Use of a radiomics model to predict tumor invasiveness of pulmonary adenocarcinomas appearing as pulmonary ground-glass nodules. *BioMed Res Int* 2018; 2018:6803971.
33. 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; 29:889-97.
34. She Y, Zhang L, Zhu 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.
35. Kobayashi Y, Ambrogio C, Mitsudomi T. Ground-glass nodules of the lung in never-smokers and smokers: clinical and genetic insights. *Transl Lung Cancer Res* 2018; 7:487.
36. Zhou W, Christiani DC. East meets West: ethnic differences in epidemiology and clinical behaviors of lung cancer between East Asians and Caucasians. *Chin J Cancer* 2011; 30:287.
37. Mei X, Wang R, Yang W, et al. Predicting malignancy of pulmonary ground-glass nodules and their invasiveness by random forest. *J Thorac Dis* 2018; 10:458-63.
38. Oikonomou A, Salazar P, Zhang Y, et al. Histogram-based models on non-thin section chest CT predict invasiveness of primary lung adenocarcinoma subsolid nodules. *Sci Rep* 2019; 9:6009.
39. Zhan Y, Peng X, Shan F, et al. Attenuation and Morphologic Characteristics Distinguishing a Ground-Glass Nodule Measuring 5-10 mm in Diameter as Invasive Lung Adenocarcinoma on Thin-Slice CT. *Am J Roentgenol* 2019:1-9.

Supplemental Data

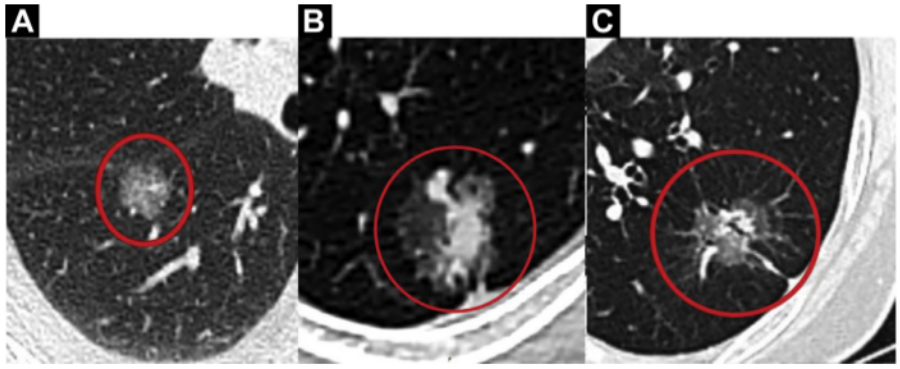
Supplemental Figure 1 Qualitative Density (Non-solid Nodule and Part-solid Nodule). (A) Nonsolid Nodule (NSN): Nodule as a Focal Area of Increased Lung Attenuation Viewed on Computed Tomography (CT) Lung Window Settings (Width, 1450 HU; level, −500 HU), Without a Solid Component Viewed on Mediastinal Window Setting (Width, 350 HU; Level, 40 HU). (B) Part-solid Nodule (PSN): Nodule With Lesions Viewed on CT Lung Window Settings, and With a Relatively Smaller Solid Component Viewed on Mediastinal Window Setting



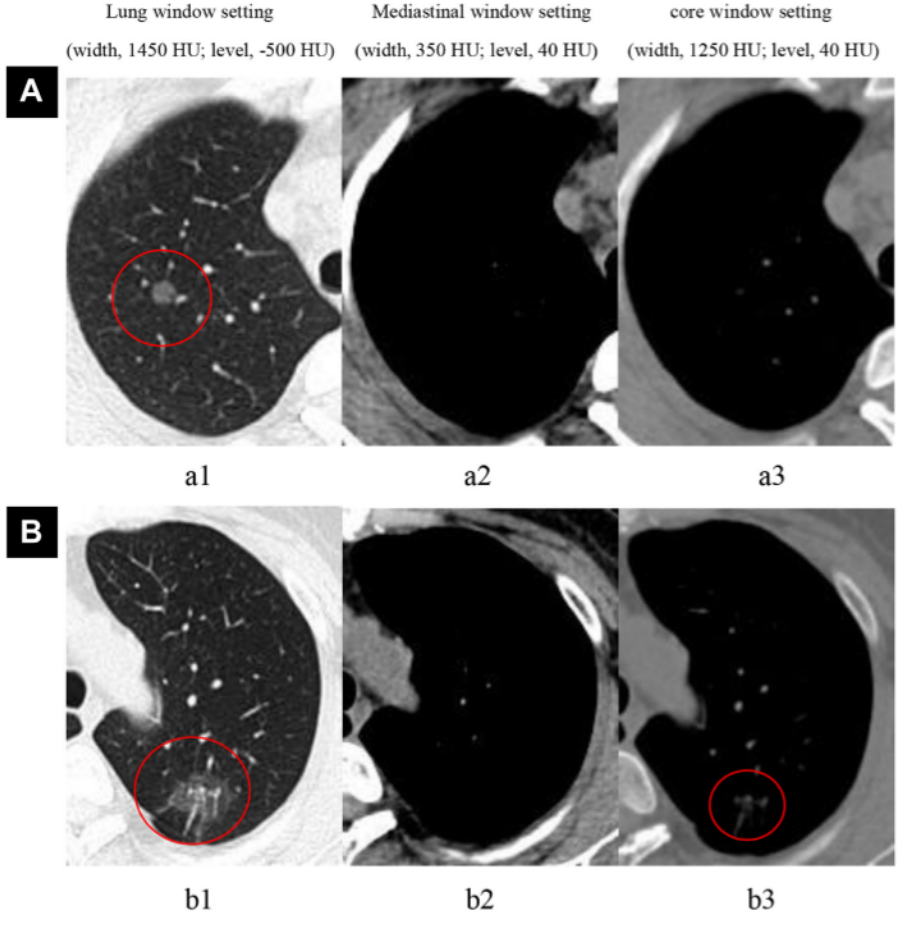
Supplemental Figure 2 Shapes. (A) Round. (B) Oval. (C) Irregular



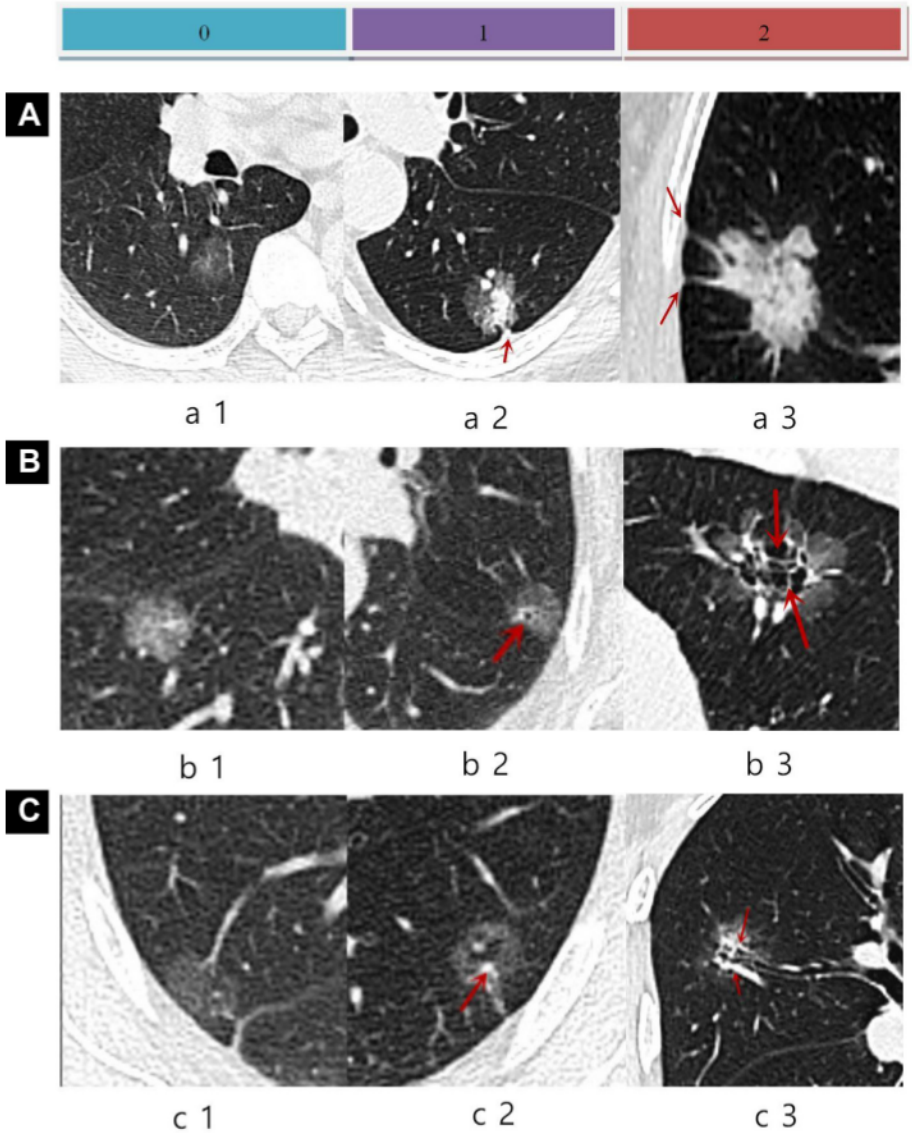
Supplemental Figure 3 Margins. (A) Smooth. (B) Lobulated. (C) Spiculated



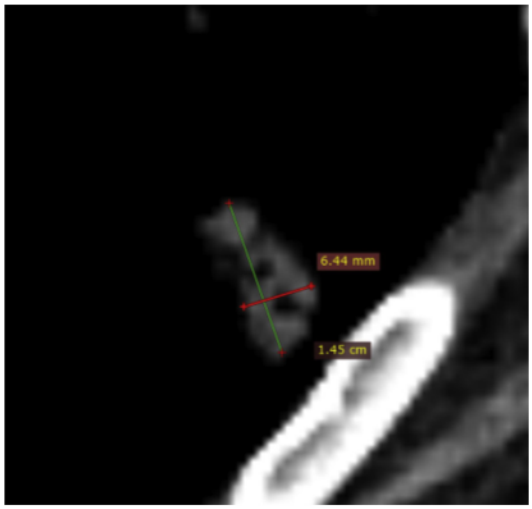
Supplemental Figure 4 Core (Noncore, Core). (A) Noncore: For Nonsolid Nodule, There was no Component Observed in Core Window Setting. (B) Core: For Nonsolid Nodule, There was Component Observed in Core Window Setting



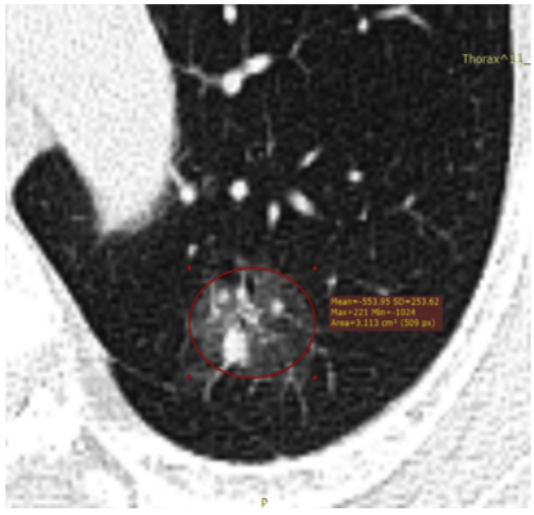
Supplemental Figure 5 Semantic Features Score. (A) Pleural Indentation: a1 (Score 0): There is no Pleural Indentation, or the Nodule is not Located Close to the Pleura; a2 (Score 1): Nodules Adhering to Pleura or Pleural Indentation with 1 Stripe; a3 (Score 2): Typical Pleural Indentation With ≥ 2 Stripes. (B) Vacuole Sign: b1 (Score 0): There is no Vacuole; b2 (Score 1): Air-bronchograms in Nonsolid Nodule or Single Cystic Cavity; b3 (Score 2): Dilated Air-bronchograms or Multiple Cystic Cavities. (C) Vascular Invasion: c1 (Score 0): No Vascular in the Nodule or With Preservation of Normal Vessels; c2 (Score1): Distorted or Dilated Vessels; c3 (Score 2): Coexistence of Irregular Vascular Dilation or Vascular Convergence From Multiple Supplying Vessels



Supplemental Figure 6 Computed Tomography (Hu): Attenuation Values of the Largest Region of Interest In the Largest Slice of Nodule



Supplemental Figure 8 Diameter of Solid Component. The Mean Value of Largest Long Diameter and Short Diameter That Measured in Mediastinal Window Setting. Diameter of Solid Component = (14.5 + 6.4)/2 = 10 mm



Supplemental Figure 7 Diameter of the Nodule. The Mean Value of Largest Long Diameter and Short Diameter, Which Measured in Lung Window Setting. Diameter = (15.5 + 15.4)/2 = 15 mm