Diagn Interv Radiol 2023; DOI: 10.4274/dir.2023.232404

CHEST IMAGING

ORIGINAL ARTICLE

Nomogram based on clinical characteristics and radiological features for the preoperative prediction of spread through air spaces in patients with clinical stage IA non-small cell lung cancer: a multicenter study

Yun Wang*
Deng Lyu*
Di Zhang*
Lei Hu
Junhong Wu
Wenting Tu
Yi Xiao
Li Fan
Shiyuan Liu

***Authors contributed equally to this work.**

From the Department of Radiology (Y.W., D.L., D.Z., W.T., Y.X., L.F. ✉ fanli0930@163.com, S.L. ✉ radiology_cz@163.com), Second Affiliated Hospital of Navy Medical University, Shanghai, China; Department of Radiology Medicine (L.H.), The People's Hospital of Chizhou, Anhui, China; Department of Radiology Medicine (J.W.), The People's Hospital of Guigang, Guangxi, China.

Received 08 July 2023; revision requested 06 August 2023; accepted 30 August 2023.

Epub: 19.09.2023

Publication date: 07.11.2023

DOI: 10.4274/dir.2023.232404

PURPOSE

To investigate the value of clinical characteristics and radiological features for predicting spread through air spaces (STAS) in patients with clinical stage IA non-small cell lung cancer (NSCLC).

METHODS

A total of 336 patients with NSCLC from our hospital were randomly divided into two groups, i.e., the training cohort (n = 236) and the internal validation cohort (n = 100) (7:3 ratio). Furthermore, 69 patients from two other hospitals were collected as the external validation cohort. Eight clinical patient characteristics were recorded, and 20 tumor radiological features were quantitatively measured and qualitatively analyzed. In the training cohort, the differences in clinical characteristics and radiological features were compared using univariate and multivariate analysis. A nomogram was created, and the predictive efficacy of the model was evaluated in the validation cohorts. The receiver operating characteristic curve and area under the curve (AUC) value were used to evaluate the discriminative ability of the model. In addition, the Hosmer–Lemeshow test and calibration curve were used to evaluate the goodness-of-fit of the model, and the decision curve was used to analyze the model's clinical application value.

RESULTS

The best predictors included gender, the carcinoembryonic antigen (CEA), consolidation-to-tumor ratio (CTR), density type, and distal ribbon sign. Among these, the tumor density type [odds ratio (OR): 6.738] and distal ribbon sign (OR: 5.141) were independent risk factors for predicting the STAS status. Moreover, three different STAS prediction models were constructed, i.e., a clinical, radiological, and combined model. The clinical model comprised gender and the CEA, the radiological model included the CTR, density type, and distal ribbon sign, and the combined model comprised the above two models. A DeLong test results revealed that the combined model was superior to the clinical model in all three cohorts and superior to the radiological model in the external validation cohort; the cohort AUC values were 0.874, 0.822, and 0.810, respectively. The results also showed that the combined model had the highest diagnostic efficacy among the models. The Hosmer–Lemeshow test showed that the combined model showed a good fit in all three cohorts, and the calibration curve showed that the predicted probability value of the combined model was in good agreement with the actual STAS status. Finally, the decision curve showed that the combined model had a better clinical application value than the clinical and radiological models.

CONCLUSION

The nomogram created in this study, based on clinical characteristics and radiological features, has a high diagnostic efficiency for predicting the STAS status in patients with clinical stage IA NSCLC and may support the creation of personalized treatment strategies before surgery.

KEYWORDS

Spread through air spaces, nomogram, radiological, prediction, non-small cell lung cancer

You may cite this article as: Wang Y, Lyu D, Zhang D, et al. Nomogram based on clinical characteristics and radiological features for the preoperative prediction of spread through air spaces in patients with clinical stage IA non-small cell lung cancer: a multicenter study. *Diagn Interv Radiol.* 2023;29(6):771-785.

The global cancer incidence statistics for 2020 reported lung cancer as one of the leading causes of cancer-induced death, with non-small cell lung cancer (NSCLC) accounting for approximately 80–85% of lung cancers.[1] According to the latest National Comprehensive Cancer Network guidelines,[2] lobectomy has been the standard surgical procedure for clinical-stage IA NSCLC; however, sublobar resection, including segmentectomy and wedge resection, is considered acceptable for lower-risk patients in early stages of the disease, such as those with a tumor diameter ≤2 cm, a ground-glass opacity (GGO) component and other standards.[3-5]

In 2015, the World Health Organization (WHO) formally defined "spread through air spaces (STAS)" as tumor cells appearing in the form of micropapillary cell clusters, solid cancer nests, or single tumor cells in the lung tissue surrounding the main tumor. It further identified STAS as the fourth type of invasion mode in lung adenocarcinoma.[6] Recent studies showed that STAS was a prognostic factor of poor outcomes for sublobar resection in patients with clinical or pathological stage IA lung cancer.[7,8] The presence of STAS was an independent risk factor for recurrence in patients with stage IA part-solid adenocarcinoma after sublobar resection.[9] Therefore, the use of relatively radical surgery, such as lobectomy rather than sublobar resection, as well as a wider surgical resection margin may be appropriate in the presence of STAS.[7-10] Effective preoperative evaluation of the STAS status may help to improve patient outcomes.

At present, postoperative pathological section,[6] which does not facilitate preoperative surgical planning, remains the gold standard for STAS diagnosis. However, preoperative puncture and intraoperative frozen pathological examinations are limited by their low sensitivity, small tissue sample size, and short diagnosis time.[11,12] Therefore, accurate assessment of the STAS status using preoperative imaging methods and other valuable clinical information could significantly impact surgical plans.

Scholar-led studies on the STAS status of lung cancer based on preoperative computed tomography (CT) results revealed a correlation between certain radiological features of lung cancer and the STAS status.[13-19] Through the use of different multiple regression models, these studies showed an association between STAS and larger tumor size,[13,14] larger solid component size,[13-16] a larger ratio of solid component size to total tumor size [consolidation-to-tumor ratio (CTR)],[13,14,16-18] air bronchogram sign,[13] vacuole sign,[15] spiculation sign,[15] lobulation sign,[18] and notch sign.[19] Additionally, Kim et al.[17] found that STAS was absent in pure GG lesions. Most existing studies included lesions in any stage of lung adenocarcinoma;[15-17,19] only two studies included stage IA lung adenocarcinoma,[13,18] and one study included only part-solid nodules.[14] Furthermore, the above-described studies only included histologic adenocarcinoma tumors. However, STAS was reportedly associated with the poor prognosis of other types of lung cancer, such as lung squamous cell carcinoma,[20] lung pleomorphic carcinoma,[21] and lung neuroendocrine tumors.[22] Jia et al.[23] also showed that clinical characteristics, such as age, gender, and the carcinoembryonic antigen (CEA), were predictors of STAS in lung cancer. To the best of the present authors' knowledge, studies focusing on clinical stage IA NSCLC [excluding pure GG nodules (pGGNs)] and incorporating the relevant CT radiological features and valuable clinical information for predicting the STAS status are rare. Therefore, the current study explored the risk factors of STAS in clinical stage IA NSCLC with a focus on clinical characteristics and radiological features. A prediction model was constructed, and a nomogram was produced. The model underwent internal and external validation to provide a basis for accurately assessing the presence or absence of STAS in patients with preoperative clinical stage IA NSCLC.

Main points

- The density type and distal ribbon sign were independent risk factors for predicting the spread through air spaces (STAS) status in patients with clinical stage IA non-small cell lung cancer.
- The designed nomogram based on clinical characteristics and radiological features has a high diagnostic efficiency in predicting the STAS status.
- Compared with the clinical and radiological models, the nomogram showed a better discriminative ability and clinical application value.

Methods

Patients

The present study was approved by the Ethics Committee of Changzheng Hospital, Naval Medical University (decision number: CZ-20220712-03). Due to the retrospective nature of this study, the need for informed consent was waived.

A retrospective patient analysis was performed; the included patients (1) had a clinical stage IA NSCLC status assessed by postoperative pathology and (2) had undergone a preoperative chest CT examination one week before surgical resection at our hospital and two other hospitals between September 2019 to September 2022. The authors collected 290 patients presenting as pGGNs, none of which were positive for STAS. As reported in the references,[17] these pGGNs were excluded.

The inclusion criteria were as follows: patients with (i) a thin-slice (≤1.5 mm) chest CT with no artifacts within one week before surgery; (ii) complete clinical and pathological data; (iii) clinical stage IA NSCLC (cT1N0M0, with a maximum tumor diameter of ≤3 cm); and (iv) solid or mixed GG nodules (mGGNs).

The exclusion criteria were as follows: patients with (i) poor image quality; (ii) incomplete clinical and pathological data; (iii) a maximum tumor diameter of >3 cm; (iv) tumors with lymph node or distant metastasis; (v) a pathological type other than NSCLC; (vi) a history of preoperative neoadjuvant treatment and chemotherapy; and (vii) pGGNs.

A total of 336 patients from our hospital (hospital 1) were included in the study and randomly divided into two groups, i.e., the training cohort (n = 236) and the internal validation cohort (n = 100) (7:3 ratio). Furthermore, 69 cases from two other hospitals (hospital 2 and hospital 3) were used as the external validation cohort, including 30 patients from hospital 2 and 39 patients from hospital 3. The detailed patient inclusion procedure is shown in Figure 1.

If multiple lesions in the same patient were surgically removed and met the inclusion criteria, the research conducted by Dercle et al.[24] was referred to for the selection of a representative lesion with the largest tumor size for analysis.

Clinical and pathological data collection

Patient data were collected, including gender, age, clinical symptoms, smoking status, family history of lung cancer, history of malignant tumors, history of multiple primary lung cancer, surgical type, pathological type, and CEA levels.

Equipment and parameters

Patients from hospital 1 underwent preoperative chest CT examinations with four types of CT machines, including the Toshiba Aquilion16 row, GE Light Speed VCT64 row,

Philips Ingenuity 64 row, and Brilliance iCT 128 row CT machines, from three vendors. In the external cohort, patients from hospital 2 were assessed with the American Light Speed 16, Light Speed VCT64 row, and Dutch Philips iCT 256-row CT machines. Hospital 3 utilized the German SOMATOM Definition Flash and SOMATOM Drive 64-row CT machines. The patients were instructed to lie down during the scan. The scanning range was set from the thoracic inlet to the middle portion of the kidneys; the tube voltage was set to 120 kVp, with a tube current of 150–250 mAs or automatic tube current regulation; the scanning slice thickness and slice increment were 5 mm; the reconstruction slice thickness and slice increment were (1) 1 mm in hospital 1 and hospital 2 and (2) 1.3 mm in hospital 3. The lung or standard algorithm reconstruction was selected, and non-contrast enhanced images were used for analysis.

Image evaluation

The CT images were imported into software (RadiAnt DICOM Viewer 4.2.1, Medixant, Poland) and analyzed by two independent radiologists with 2 and 10 years of experience, respectively, who were blinded to the pathological information. The lung window [width: 1500 Hounsfield scale (HU), level: −500 HU], mediastinal window (width: 300 HU, level: 50 HU), multiplanar reformation (MPR), and maximal intensity projection were used to analyze the lesion. For quantitative measures, the average measurements of two independent radiologists were used as the final data. For qualitative indicators, disagreements were discussed until a consensus was reached.

First, the longest diameter of the entire tumor and the consolidation part were measured at the lung window on the MPR images, and the proportion of the consolidation part (CTR) was calculated.[13] Clinical T-staging was performed according to the maximum diameter of the solid components of the tumor.[25]

Second, the following qualitative CT radiological features were assessed: the tumor location, density type (solid and mGGN), shape (round and irregular), tumor–lung interface (well-defined and ill-defined), margin (lobulation and spiculation), internal features (vacuole sign and cavity/cystic airspace), and external features (vascular convergence, bronchial change, pleural tags, pleural indentation, halo sign, satellite lesions, distal ribbon sign, and combination with emphysema).

The definitions of CT radiological features are described in Supplementary Table 1, and the CT radiological features are shown in Supplementary Figures 1-4. Most of the definitions concerning the pulmonary nodules' radiological features have previously been reported.[17,26-29]

Pathological diagnosis

The pathological diagnosis of each patient included in the present study was established, respectively, by two pathologists, a junior pathologist and a senior pathologist with more than 10 years of work experience collectively, based on the 2015 WHO definition of STAS.[6] The classification of lung cancer was based on the WHO's classification of lung cancer (2015 edition),[6] and the clinical and pathological staging was based on the TNM staging standard of lung cancer (8th edition),[25] as the pathological diagnosis was determined as part of routine clinical practice, and the specimens were not reviewed specifically for this study.

Statistical analysis

The SPSS (v.20.0) and R (v.4.2.2) statistical software programs were used for analyzing all statistical tests. The classified data were represented by the number of cases, and the comparison between the two groups was performed using Pearson's chi-square test, Yate's correction for continuity, or Fisher's exact test. A P value of <0.05 was considered statistically significant in the univariate analysis. Variables with $P < 0.1$ in the univariate analysis were included in the multivariate logistic regression analysis, backward stepwise selection was applied using the likelihood ratio test, with Akaike's information criterion as the stopping rule to select the best combination of variables for building the prediction model in the training cohort, and a corresponding nomogram was constructed. The interobserver agreement of numeric and categorical variables was assessed using intraclass correlation coefficients (ICCs) and κ-statistics, respectively. The receiver operating characteristic (ROC) curve with the

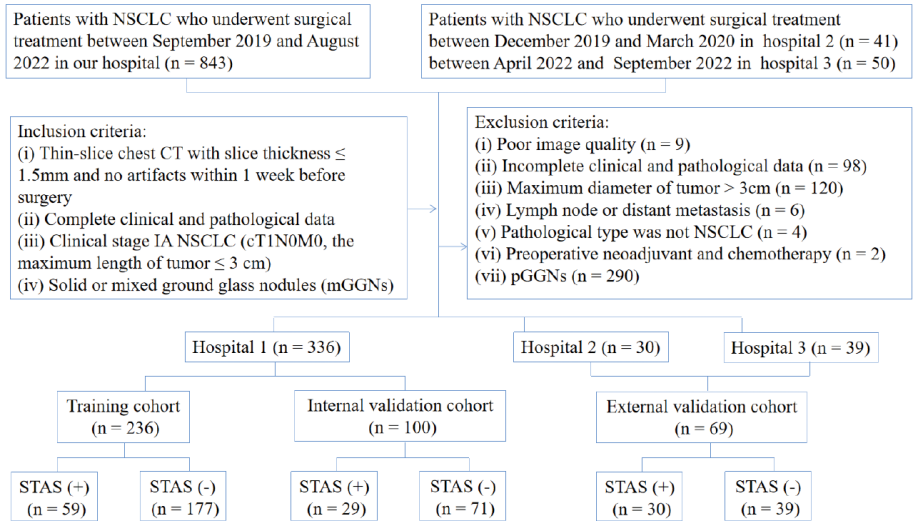


Figure 1. The flow chart for patient selection. NSCLC, non-small cell lung cancer; CT, computed tomography; STAS, spread through air spaces.

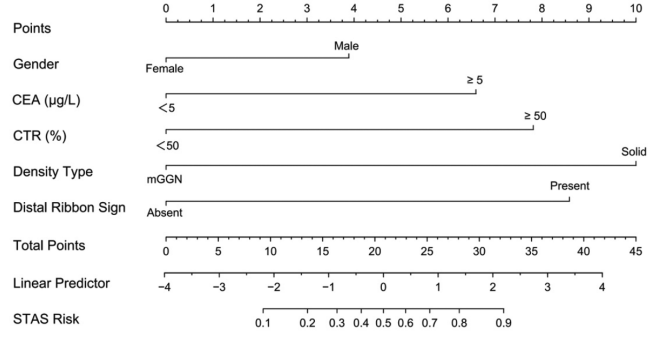


Figure 2. The nomogram for the preoperative prediction of the spread through air spaces status based on clinical characteristics and radiological features in clinical stage IA non-small cell lung cancer. CEA, carcinoembryonic antigen; CTR, consolidation-to-tumor ratio; STAS, spread through air spaces, mGGN, mixed ground-glass nodule.

corresponding area under the curve (AUC) value was used to evaluate the discriminative ability of the prediction models for predicting STAS in the training and validation cohorts. The DeLong test was used to assess AUC differences between models. Calibration curves and the Hosmer–Lemeshow test were used to evaluate the goodness-of-fit of the prediction model; a *P* value of >0.05 indicated a high goodness-of-fit. The decision curve analysis (DCA) was used to evaluate the clinical utility of the nomogram. Multivariate binary logistic regression, nomograms, and calibration plots were performed with the "rms" package of the R software. The ROC was performed using the "pROC" package. Validation was performed by the "rms" package. DCA was performed using the "gg.DCA" function.

Results

Clinicopathological characteristics of the study cohorts

Of the 405 patients with NSCLC included in the present study, 118 were STAS-positive and 287 were STAS-negative. Statistically significant gender differences were observed in the training and external validation cohorts ($P < 0.05$). Furthermore, the surgical method showed statistically significant differences in the training and internal validation cohorts ($P < 0.05$), whereas smoking status, CEA, and pathological type demonstrated statistically significant differences in the training cohort ($P < 0.05$). Additionally, there were statistically significant differences concerning clinical symptoms in the external validation cohort ($P < 0.05$) (Table 1).

Univariate and multivariate logistic regression analysis

For radiological features, good consistency was observed in terms of quantitative parameters between two observers (ICC: 0.934–0.935), with a strong consistency in qualitative indicators (Kappa value: 0.852–1.000). The interobserver agreement assessment results of each index are shown in Supplementary Table 2.

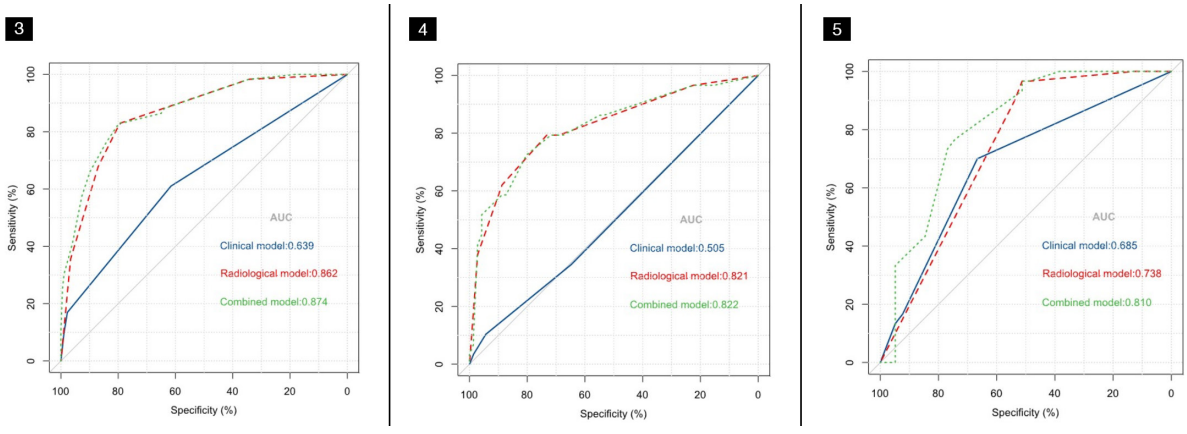
In the training cohort, univariate analysis showed statistically significant differences in gender, smoking status, CEA, clinical T-stage, CTR, density type, spiculation, bronchial change, vascular convergence, halo sign, distal ribbon sign, pleural indentation, and pleural tags between the STAS-positive and the STAS-negative groups ($P < 0.05$) (Table 2). No multicollinearity was observed (see Supplementary Table 3). The best combination of variables selected by multivariate logistic regression analysis included gender, the CEA, CTR, density type, and distal ribbon sign, among which the tumor density type [odds ratio (OR: 6.738] and distal ribbon sign (OR: 5.141) were independent risk factors for predicting the STAS status (Table 3).

Model development and evaluation

Based on the best combination of variables selected by multivariate logistic regression analysis, three different models were constructed: (1) a clinical model comprising gender and the CEA; (2) a radiological model comprising the CTR, density type, and distal ribbon sign; and (3) a combined model including of all the above variables. Then, predictive diagnostic efficacy was compared among the different models. The results indicated that the AUC values in the three cohorts were 0.874, 0.822, and 0.810 in the combined model, 0.862, 0.821, and 0.738 in the radiological model, and 0.639, 0.505, and 0.685 in the clinical model. The predictive efficacy is shown in Table 4, and the ROC curves are shown in Figures 3-5. The DeLong test demonstrated that the combined model was superior to the clinical model in the three cohorts (Z = 6.315, 4.969, and 2.085; $P < 0.05$) and that the combined model was superior to the radiological model in the external validation cohort (Z = 2.529; $P < 0.05$). The radiological model was superior to the clinical model in the training cohort and the internal validation cohort (Z = 5.065 and 4.306; $P < 0.05$). Based on the regression coefficients of the variables selected by multivariate logistic regression analysis, a nomogram was constructed to evaluate the STAS risk intuitively, as shown in Figure 2. The regression equation was ln (P/1−P) = −3.97 + 0.74 × gender + 1.25 × CEA + 1.49 × CTR + 1.90 × density type + 1.64 × distal ribbon sign, where P represents the probability of predicting the occurrence of STAS. The Hosmer–Lemeshow test showed that the combined model was a good fit in all three cohorts ($P = 0.22$, 0.94, and 0.51, respectively). Moreover, the calibration curve showed that the predicted probabilities of the combined model were in acceptable agreement with the actual probabilities in the three cohorts (Figures 6-8). The DCA showed that the combined model had a better clinical application value than the clinical and radiological models (Figures 9-11).

Discussion

With the extensive application of low-dose CT screening for lung cancer, the detection and surgical rates of the disease have been continuously improving.[30] STAS is a risk factor for postoperative tumor recurrence and metastasis in patients with early lung cancer, and lobectomy can achieve a better clinical prognosis than sublobectomy.[7,8,31] Therefore, an accurate prediction of the STAS status before surgery has important guiding significance for the selection of the surgical procedure. This study revealed that the tumor density type and distal



Figures 3-5. The receiver operation characteristic curve analysis of the clinical, radiological, and combined models in the three cohorts. **(Figure 3)** The training cohort, **(Figure 4)** the internal validation cohort, and **(Figure 5)** the external validation cohort. AUC, area under the curve.

Table 1. The clinical and pathological characteristics of patients in the training cohort and the two validation cohorts

Characteristics	Training cohort (n = 236)			Internal validation cohort (n = 100)			External validation cohort (n = 69)		
	STAS (–) (n = 177)	STAS (+) (n = 59)	*P* value	STAS (–) (n = 71)	STAS (+) (n = 29)	*P* value	STAS (–) (n = 39)	STAS (+) (n = 30)	*P* value
Gender									
Male	66 (37.3%)	32 (54.2%)	0.022[a]	49 (69.0%)	21 (72.4%)	0.736[a]	12 (30.8%)	20 (66.7%)	0.003[a]
Female	111 (62.7%)	27 (45.8%)		22 (31.0%)	8 (27.6%)		27 (69.2%)	10 (33.3%)	
Age (year)									
<65	125 (70.6%)	35 (59.3%)	0.108[a]	50 (70.4%)	17 (58.6%)	0.255[a]	30 (76.9%)	19 (63.3%)	0.217[a]
≥65	52 (29.4%)	24 (40.7%)		21 (29.6%)	12 (41.4%)		9 (23.1%)	11 (36.7%)	
Clinical symptoms									
Absent	124 (70.1%)	47 (79.7%)	0.153[a]	49 (69.0%)	16 (55.2%)	0.188[a]	31 (79.5%)	17 (56.7%)	0.041[a]
Present	53 (29.9%)	12 (20.3%)		22 (31.0%)	13 (44.8%)		8 (20.5%)	13 (43.3%)	
Smoking status									
Non-smoker	151 (85.3%)	42 (71.2%)	0.015[a]	62 (87.3%)	25 (86.2%)	1.000[b]	34 (87.2%)	21 (70.0%)	0.079[a]
Smoker	26 (14.7%)	17 (28.8%)		9 (12.7%)	4 (13.8%)		5 (12.8%)	9 (30.0%)	
Family history of lung cancer									
Absent	164 (92.7%)	57 (96.6%)	0.441[b]	67 (94.4%)	28 (96.6%)	1.000[b]	39 (100.0%)	30 (100.0%)	N/A
Present	13 (7.3%)	2 (3.4%)		4 (5.6%)	1 (3.4%)		0 (0.0%)	0 (0.0%)	
History of malignancy									
Absent	151 (85.3%)	51 (86.4%)	0.831[a]	60 (84.5%)	25 (86.2%)	1.000[b]	35 (89.7%)	29 (96.7%)	0.528[b]
Present	26 (14.7%)	8 (13.6%)		11 (15.5%)	4 (13.8%)		4 (10.3%)	1 (3.3%)	
History of multiple primary lung cancer									
Absent	145 (81.9%)	51 (86.4%)	0.423[a]	55 (77.5%)	25 (86.2%)	0.321[a]	32 (82.1%)	26 (86.7%)	0.851[b]
Present	32 (18.1%)	8 (13.6%)		16 (22.5%)	4 (13.8%)		7 (17.9%)	4 (13.3%)	
CEA (μg/L)									
<5	173 (97.7%)	49 (83.1%)	<0.001[b]	67 (94.4%)	26 (89.7%)	0.685[b]	36 (92.3%)	25 (83.3%)	0.438[b]
≥5	4 (2.3%)	10 (16.9%)		4 (5.6%)	3 (10.3%)		3 (7.7%)	5 (16.7%)	
Surgery type									
Sublobectomy	80 (45.2%)	14 (23.7%)	0.004[a]	25 (35.2%)	3 (10.3%)	0.023[b]	12 (30.8%)	3 (10.0%)	0.075[b]
Lobectomy	97 (54.8%)	45 (76.3%)		46 (64.8%)	26 (89.7%)		27 (69.2%)	27 (90.0%)	
Pathological type									
MIA	14 (7.9%)	0 (0.0%)		7 (9.9%)	0 (0.0%)		1 (2.6%)	0 (0.0%)	
IA	154 (87.0%)	53 (89.8%)		64 (90.1%)	28 (96.6%)		37 (94.8%)	28 (93.3%)	
IMA	5 (2.8%)	4 (6.8%)	0.021[c]	0 (0.0%)	0 (0.0%)	0.069[c]	1 (2.6%)	0 (0.0%)	0.588[c]
SCC	3 (1.7%)	0 (0.0%)		0 (0.0%)	0 (0.0%)		0 (0.0%)	0 (0.0%)	
ASC	1 (0.6%)	2 (3.4%)		0 (0.0%)	0 (0.0%)		0 (0.0%)	1 (3.4%)	
NSCLC-NOS	0 (0.0%)	0 (0.0%)		0 (0.0%)	1 (3.4%)		0 (0.0%)	1 (3.4%)	

The *P* value represents the univariate analysis; data are presented as n (%). [a], Pearson's chi-square; [b], Yates' correction for continuity; [c], Fisher's exact test; STAS, spread through air spaces; STAS (–), STAS-negative; STAS (+), STAS-positive; CEA, carcinoembryonic antigen; MIA, minimally invasive adenocarcinoma; IMA, invasive mucinous adenocarcinoma; SCC, squamous cell carcinoma; ASC, adenosquamous carcinoma; NSCLC-NOS: non-small cell lung cancer, not otherwise specified.

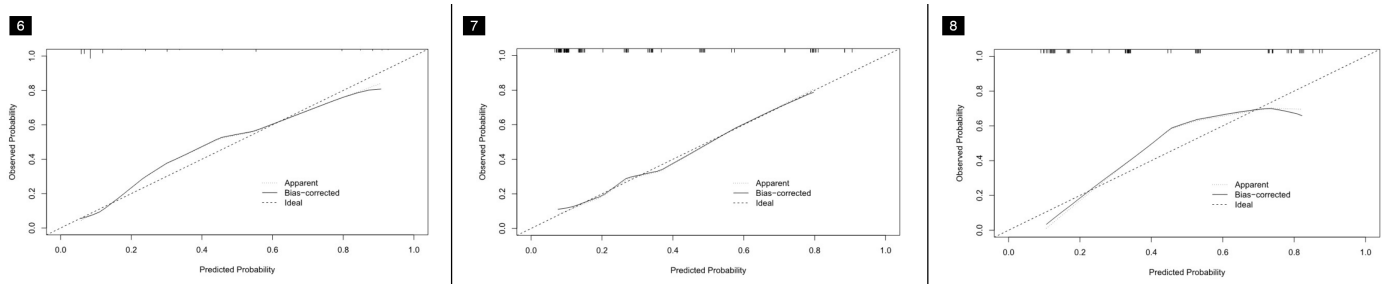
Table 2. Radiological features of the tumor in the training cohort and the two validation cohorts

Features	Training cohort (n = 236)			Internal validation cohort (n = 100)			External validation cohort (n = 69)		
	STAS (–) (n = 177)	STAS (+) (n = 59)	*P* value	STAS (–) (n = 71)	STAS (+) (n = 29)	*P* value	STAS (–) (n = 39)	STAS (+) (n = 30)	*P* value
Clinical T-stage									
cT1mi	17 (9.6%)	0 (0.0%)		5 (7.0%)	0 (0.0%)		1 (2.6%)	0 (0.0%)	
cT1a	75 (42.4%)	11 (18.6%)	<0.001[c]	28 (39.4%)	1 (3.4%)	<0.001[c]	11 (28.2%)	1 (3.3%)	<0.001[c]
cT1b	69 (39.0%)	22 (37.3%)		32 (45.1%)	9 (31.0%)		25 (64.1%)	18 (60.0%)	
cT1c	16 (9.0%)	26 (44.1%)		6 (8.5%)	19 (65.6%)		2 (5.1%)	11 (36.7%)	
CTR (%)									
<50	63 (35.6%)	2 (3.4%)	<0.001[a]	19 (26.8%)	1 (3.4%)	0.008[a]	6 (15.4%)	0 (0.0%)	0.032[c]
≥50	114 (64.4%)	57 (96.6%)		52 (73.2%)	28 (96.6%)		33 (84.6%)	30 (100.0%)	
Density type									
mGGN	154 (87.0%)	19 (32.2%)	<0.001[a]	63 (88.7%)	11 (37.9%)	<0.001[a]	21 (53.8%)	3 (10.0%)	<0.001[b]
Solid	23 (13.0%)	40 (67.8%)		8 (11.3%)	18 (62.1%)		18 (46.2%)	27 (90.0%)	
Location									
RUL	58 (32.8%)	12 (20.3%)		19 (26.8%)	11 (37.9%)		11 (28.2%)	8 (26.7%)	
RML	16 (9.0%)	5 (8.6%)		7 (9.9%)	3 (10.5%)		4 (10.3%)	2 (6.7%)	
RLL	27 (15.3%)	16 (27.1%)	0.174[a]	17 (23.9%)	5 (17.2%)	0.783[a]	7 (17.9%)	6 (20.0%)	0.737[a]
LUL	53 (29.9%)	16 (27.1%)		17 (23.9%)	5 (17.2%)		10 (25.7%)	5 (16.6%)	
LLL	23 (13.0%)	10 (16.9%)		11 (15.5%)	5 (17.2%)		7 (17.9%)	9 (30.0%)	
Shape									
Irregular	26 (14.7%)	12 (20.3%)	0.307[a]	12 (16.9%)	5 (17.2%)	1.000[b]	20 (51.3%)	14 (46.7%)	0.704[a]
Round/oval	151 (85.3%)	47 (79.7%)		59 (83.1%)	24 (82.8%)		19 (48.7%)	16 (53.3%)	
Tumor-lung interface									
Well-defined	167 (94.4%)	51 (86.4%)	0.089[b]	68 (95.8%)	23 (79.3%)	0.026[b]	34 (87.2%)	25 (83.3%)	0.653[a]
Ill-defined	10 (5.6%)	8 (13.6%)		3 (4.2%)	6 (20.7%)		5 (12.8%)	5 (16.7%)	
Lobulation									
Absent	32 (18.1%)	6 (10.2%)	0.152[a]	18 (25.4%)	3 (10.3%)	0.095[a]	14 (35.9%)	11 (36.7%)	0.947[a]
Present	145 (81.9%)	53 (89.8%)		53 (74.6%)	26 (89.7%)		25 (64.1%)	19 (63.3%)	
Spiculation									
Absent	156 (88.1%)	34 (57.6%)	<0.001[a]	64 (90.1%)	18 (62.1%)	0.001[a]	32 (82.1%)	18 (60.0%)	0.042[a]
Present	21 (11.9%)	25 (42.4%)		7 (9.9%)	11 (37.9%)		7 (17.9%)	12 (40.0%)	
Vacuole									
Absent	127 (71.8%)	38 (64.4%)	0.287[a]	49 (69.0%)	20 (69.0%)	0.996[a]	31 (79.5%)	25 (83.3%)	0.685[a]
Present	50 (28.2%)	21 (35.6%)		22 (31.0%)	9 (31.0%)		8 (20.5%)	5 (16.7%)	
Cavity/cystic airspace									
Absent	161 (91.0%)	52 (88.1%)	0.526[a]	66 (93.0%)	26 (89.7%)	0.884[b]	37 (94.9%)	30 (100.0%)	0.501[c]
Present	16 (9.0%)	7 (11.9%)		5 (7.0%)	3 (10.3%)		2 (5.1%)	0 (0.0%)	
Bronchial change									
Absent	84 (47.5%)	19 (32.2%)	0.041[a]	29 (40.8%)	9 (31.0%)	0.359[a]	19 (48.7%)	16 (53.3%)	0.704[a]
Present	93 (52.5%)	40 (67.8%)		42 (59.2%)	20 (69.0%)		20 (51.3%)	14 (46.7%)	

Table 2. Continued

Features	Training cohort (n = 236)			Internal validation cohort (n = 100)			External validation cohort (n = 69)		
	STAS (–) (n = 177)	STAS (+) (n = 59)	*P* value	STAS (–) (n = 71)	STAS (+) (n = 29)	*P* value	STAS (–) (n = 39)	STAS (+) (n = 30)	*P* value
Vascular convergence									
Absent	150 (84.7%)	43 (72.9%)	0.041[a]	53 (74.6%)	19 (65.5%)	0.356[a]	35 (89.7%)	18 (60.0%)	0.009[b]
Present	27 (15.3%)	16 (27.1%)		18 (25.4%)	10 (34.5%)		4 (10.3%)	12 (40.0%)	
Pleural tags									
Absent	97 (54.8%)	18 (30.5%)	0.001[a]	31 (43.7%)	8 (27.6%)	0.135[a]	14 (35.9%)	11 (36.7%)	0.947[a]
Present	80 (45.2%)	41 (69.5%)		40 (56.3%)	21 (72.4%)		25 (64.1%)	19 (63.3%)	
Pleural indentation									
Absent	115 (65.0%)	23 (39.0%)	<0.001[a]	38 (53.5%)	12 (41.4%)	0.271[a]	17 (43.6%)	13 (43.3%)	0.983[a]
Present	62 (35.0%)	36 (61.0%)		33 (46.5%)	17 (58.6%)		22 (56.4%)	17 (56.7%)	
Halo sign									
Absent	175 (98.9%)	54 (91.5%)	0.015[b]	70 (98.6%)	24 (82.8%)	0.010[b]	37 (94.9%)	29 (96.7%)	1.000[b]
Present	2 (1.1%)	5 (8.5%)		1 (1.4%)	5 (17.2%)		2 (5.1%)	1 (3.3%)	
Satellite lesions									
Absent	165 (93.2%)	56 (94.9%)	0.878[b]	63 (88.7%)	23 (79.3%)	0.360[b]	34 (87.2%)	25 (83.3%)	0.653[a]
Present	12 (6.8%)	3 (5.1%)		8 (11.3%)	6 (20.7%)		5 (12.8%)	5 (16.7%)	
Distal ribbon sign									
Absent	157 (88.7%)	29 (49.2%)	<0.001[a]	55 (77.5%)	13 (44.8%)	0.001[a]	31 (79.5%)	19 (63.3%)	0.136[a]
Present	20 (11.3%)	30 (50.8%)		16 (22.5%)	16 (55.2%)		8 (20.5%)	11 (36.7%)	
ELLC									
Absent	163 (92.1%)	54 (91.5%)	1.000[b]	69 (97.2%)	27 (93.1%)	0.702[b]	38 (97.4%)	25 (83.3%)	0.103[b]
Present	14 (7.9%)	5 (8.5%)		2 (2.8%)	2 (6.9%)		1 (2.6%)	5 (16.7%)	
ERL									
Absent	163 (92.1%)	54 (91.5%)	1.000[b]	70 (98.6%)	27 (93.1%)	0.416[b]	38 (97.4%)	25 (83.3%)	0.103[b]
Present	14 (7.9%)	5 (8.5%)		1 (1.4%)	2 (6.9%)		1 (2.6%)	5 (16.7%)	

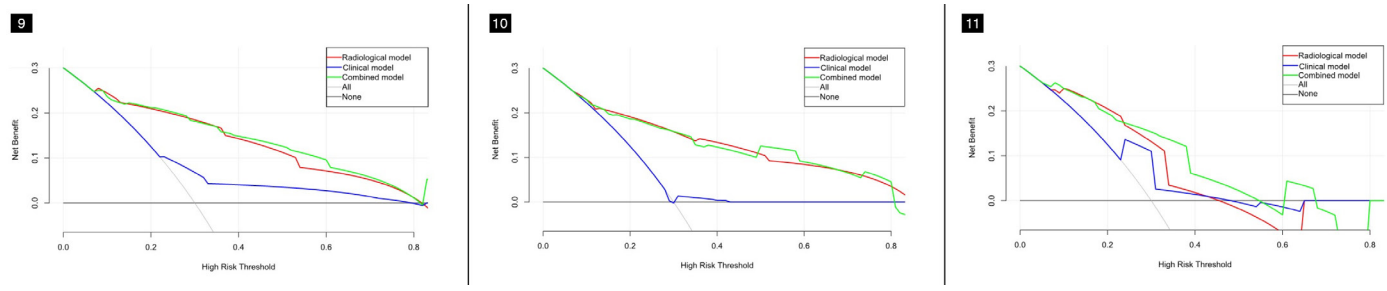
The *P* value represents the univariate analysis; data are presented as n (%). [a], Pearson's chi-square; [b], Yates' correction for continuity; [c], Fisher's exact test; STAS, spread through air spaces; STAS (–), STAS-negative; STAS (+), STAS-positive; cT1mi, tumor with a solid component size of <0.5 cm and an entire tumor size of <3.0 cm; cT1a, tumor with a solid component size of 0.6–1.0 cm and entire tumor size of 0.6–3.0 cm; cT1b, tumor with a solid component size of 1.1–2.0 cm and entire tumor size of 1.1–3.0 cm; cT1b, tumor with a solid component size of 2.1–3.0 cm and an entire tumor size of 2.1–3.0 cm; CTR, consolidation-to-tumor ratio; mGGN, mixed ground-glass nodule; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe; LUL, left upper lobe; LLL, left lower lobe; ELLC, emphysema in the lobe of lung cancer; ERL, emphysema in the remaining lobes.



Figures 6-8. The calibration curves of the combined model in the three cohorts. **(Figure 6)** The training cohort, **(Figure 7)** the internal validation cohort, and (**Figure 8)** the external validation cohort.

ribbon sign were independent risk factors for predicting the STAS status. The combined model was constructed based on the best combination of variables selected by multivariate logistic regression analysis, including gender, the CEA, CTR, density type, and distal ribbon sign, and had the highest diagnostic efficacy of the models. The AUC values in the training cohort, internal validation cohort, and external validation cohorts were 0.874, 0.822, and 0.810, respectively. The accuracy of the model was 80.51%, 78.00%, and 75.36%, respectively, the sensitivity was 83.05%, 72.41%, and 76.67%, respectively, and the specificity was 79.66%, 80.28%, and 74.36%, respectively.

In the present study, the spiculation sign, pleural indentation sign, and vascular convergence sign were more common in the STAS-positive than in the STAS-negative group; this is consistent with previous lit-



Figures 9-11. The decision curve shows that the combined model has better clinical application value than the clinical and radiological models in the three cohorts. **(Figure 9)** The training cohort, **(Figure 10)** the internal validation cohort, and **(Figure 11)** the external validation cohort.

Table 3. Univariate and multivariate logistic regression analysis of factors in the training cohort

Factors	Univariate logistic regression analysis		Multivariate logistic regression analysis	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Gender	1.993 (1.100–3.638)	0.023	2.100 (0.977–4.585)	0.058
Smoking status	2.351 (1.154–4.718)	0.017		
CEA	8.827 (2.821–33.31)	<0.001	3.519 (0.826–17.32)	0.101
Clinical T-stage	3.576 (2.344–5.684)	<0.001		
CTR	15.75 (4.690–98.11)	<0.001	4.441 (1.159–29.24)	0.057
Density type	14.10 (7.125–29.03)	<0.001	6.738 (3.107–15.18)	<0.001
Tumor-lung interface	2.620 (0.954–6.993)	0.054		
Spiculation	5.462 (2.757–10.99)	<0.001		
Bronchial change	1.902 (1.034–3.596)	0.042		
Pleural tags	2.762 (1.493–5.275)	0.002		
Pleural indentation	2.903 (1.592–5.388)	0.001		
Vascular convergence	2.067 (1.007–4.158)	0.044		
Halo sign	8.102 (1.694–57.70)	0.014		
Distal ribbon sign	8.121 (4.114–16.46)	<0.001	5.141 (2.272–12.00)	<0.001

OR, odds ratio; CI, confidence interval; CEA, carcinoembryonic antigen; CTR, consolidation-to-tumor ratio.

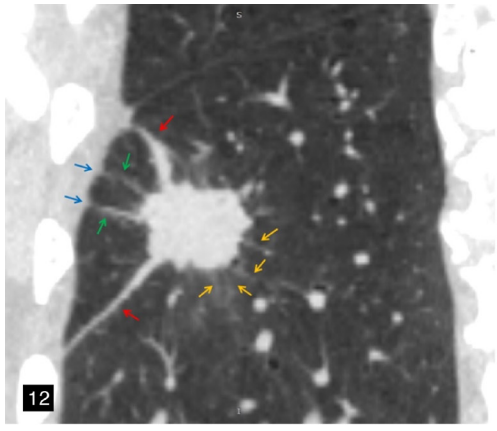
Table 4. The predictive efficacy of the clinical, radiological, and combined models in the three cohorts

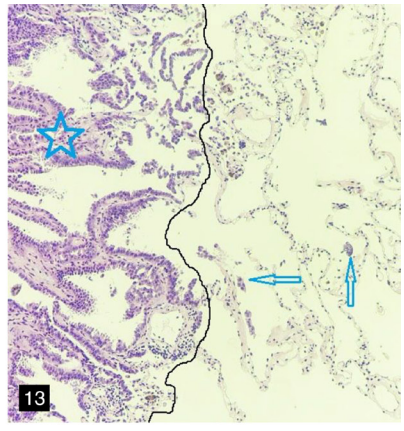
Model	Cohort	Cut-off	AUC (95% CI)	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Clinical	Training	0.231	0.639 (0.574–0.700)	61.44%	61.02%	61.58%	34.62%	82.58%
	Internal validation		0.505 (0.403–0.607)	70.00%	10.34%	94.37%	42.86%	72.04%
	External validation		0.685 (0.562–0.792)	68.12%	70.00%	66.67%	61.76%	74.29%
Radiological	Training	0.113	0.862 (0.811–0.903)	80.08%	83.05%	79.10%	56.98%	93.33%
	Internal validation		0.821 (0.731–0.890)	75.00%	79.31%	73.24%	54.76%	89.66%
	External validation		0.738 (0.618–0.836)	71.01%	96.67%	51.28%	60.42%	95.24%
Combined	Training	0.264	0.874 (0.825–0.914)	80.51%	83.05%	79.66%	57.65%	93.37%
	Internal validation		0.822 (0.733–0.891)	78.00%	72.41%	80.28%	60.00%	87.69%
	External validation		0.810 (0.697–0.894)	75.36%	76.67%	74.36%	70.00%	80.56%

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

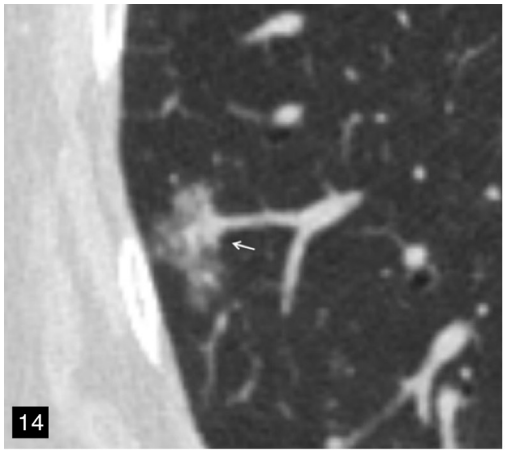
erature reports[19,32] and indicates a possible relation of STAS-positivity to the disease's pathological mechanism. Spiculation is associated with tumor cell infiltration into adjacent blood and lymphatic vessels, suggesting that lung cancer is relatively aggressive;[33] meanwhile, pleural indentation arises from intratumor reactive fibrous hyperplasia, pulling the adjacent pleura and causing deviation from its original position.[34] The vascular convergence sign is also caused by the reactive fibrous hyperplasia of the tumor, which pulls the adjacent pulmonary vessels to converge with the tumor.[35] A relatively high degree of tumor infiltration is associated with serious internal reactive fibroplasia as well as with a higher probability of pleural indentation and vascular convergence, indicating a greater possibility of STAS.

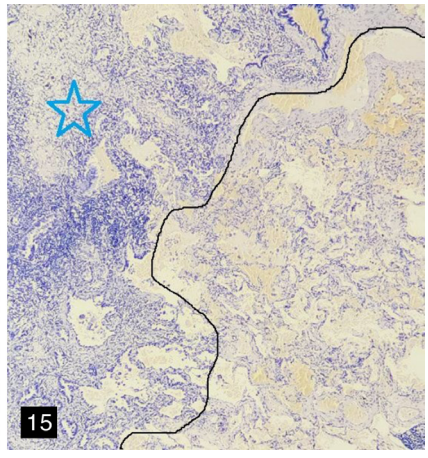
Kim et al.[17] reported that STAS did not exist in the pGGNs of lung adenocarcinoma. The authors of the present study collected 290 patients presenting as pGGNs, none of which were positive for STAS. Therefore, this study excluded patients with pGGNs but included those with mGGNs and solid nodules, which are at potential risk of STAS. The strict inclusion and exclusion criteria promoted an objective prediction efficiency of the model constructed in this study. Yin et al.[36] showed that STAS was more common in lung adenocarcinoma presenting as solid nodules than mGGNs or pGGNs. The results of the present study revealed that the tumor density type was an independent risk factor for predicting STAS. The incidence of STAS in patients with solid nodules showing on CT was 63.4% (85/134), while that of mGGNs was 12.1% (33/271); this is consistent with the results of previous studies.[19,36]





Figures 12, 13. A 58-year-old male patient with lung adenocarcinoma and a positive spread through air spaces status. The sagittal non-contrast computed tomography image shows a solid nodule in the right lower lobe of the lung **(Figure 12)**, consolidation-to-tumor ratio ≥50%, with distal ribbon sign (red arrow), lobulation sign, spiculation sign (yellow arrow), and multiple pleural tags (green arrow). **(Figure 13)** The photomicrograph (hematoxylin and eosin stained, magnification x100) shows detached micropapillary clusters of tumor cells (arrows) in the alveolar beyond the edge (dark line) of the main tumor (star).





Figures 14, 15. A 67-year-old female patient with lung adenocarcinoma and a negative spread through air spaces (STAS) status. **(Figure 14)** The axial non-contrast computed tomography image shows a mixed ground-glass nodule in the right lower lobe of the lung, with the longest interface length of the entire tumor and solid component being 1.80 cm and 5.57 mm, respectively; consolidation-to-tumor ratio <50%, with an irregular shape and a well-defined interface. **(Figure 15)** The pathological section indicated a negative STAS status; that is, there were no free tumor cell clusters in the alveolar cavity outside the edge of the main lesion. The photomicrograph (hematoxylin and eosin stained, magnification x40) shows clean alveolar spaces adjacent to the boundary (dashed line) of the tumor (star).

Solid components typically represent the more aggressive part of the tumor.[25] Previous studies have shown a positive correlation between CTR and STAS,[13,14,16,18] and higher CTR values were associated with more aggressive tumors, a higher likelihood of a STAS-positive status, and a worse patient prognosis.[27] In the present study, the STAS-positive rate in the CTR ≥50% group was 36.6% (115/314), and the STAS-positive rate in the CTR <50% group was 3.3% (3/91), indicating a higher STAS incidence in cases with higher solid components showing on CT, which is in agreement with existing research. This study also found that the distal ribbon sign was another independent risk factor for predicting STAS. This can possibly be attributed to tumor cells escaping from the primary lesion, redistributing through the airway, and growing along the surrounding alveolar wall, thus resulting in parenchymal obstruction of the surrounding lung or obstruction of the terminal bronchioles and reducing the gas content in the alveoli.

Qi et al.[32] proposed for the first time that GG ribbons were an independent risk factor for predicting STAS; a GG ribbon was defined as a band-shaped GGO with a blurred edge emitting from the edge of the nodule and extending into the adjacent lung. This is similar to the distal ribbon sign described in the present study (Figures 12-15). In general, a higher degree of tumor invasion is associated with a higher incidence of STAS, which is indicated by a relatively high proportion of tumor solid components on CT and a relatively high number of malignant radiological features. Lobectomy is recommended for this type of early lung cancer.

The current study found STAS-positivity correlated with gender and the CEA level. Among the patients included in the study, 37.5% (60/160) of male patients and 23.7% (58/245) of female patients were STAS-positive; furthermore, 26.6% (100/376) of patients with CEA <5 μg/L and 62.1% (18/29) of patients with CEA ≥5 μg/L were STAS-positive. The results showed that male patients with lung cancer and CEA ≥5 μg/L had a heightened likelihood of being associated with STAS-positivity; this is similar to the conclusions presented by Jia et al.[23]

Many previous studies have investigated STAS based on preoperative CT examination;[13-19] however, these studies only included lung adenocarcinoma, and the variables selected to construct the model were mostly based on CT features. The diagnostic efficiency of these models ranged from 0.726 to 0.803, and most of them lacked external validation. Focusing on clinical stage IA NSCLC, this study included a wider spectrum of lung cancer, representing relatively high clinical significance. The best combination of variables selected by multivariate logistic regression analysis included preoperative clinical characteristics and radiological features, and a corresponding clinical model, radiological model, and combined model were constructed, respectively. Moreover, this study adopted a multicenter external dataset, which improved the generalizability of the model. Subsequently, the diagnostic efficacy of the different models was assessed in the three cohorts, indicating that some clinical characteristics and radiological features of patients could be used to predict the STAS status before surgery. The combined model yielded the highest diagnostic efficacy and showed the best clinical application value among the models. Furthermore, a nomogram was drawn to illustrate the complex regression equation of the model in a visual graph, providing an intuitive, easy-to-understand, and convenient method to evaluate patients. The nomogram also guides the accurate judgment of the STAS status before surgery.

The limitations of this study should be acknowledged. First, this research represents a retrospective study, and a degree of selection bias was inevitable. Second, this study only analyzed the clinical characteristics and radiological features of patients, and further studies related to radiomics and artificial intelligence can be performed to explore whether the diagnostic efficiency could be further improved. Third, due to the lack of detailed follow-up data, whether STAS is an independent prognostic factor, as well as its effect on patients undergoing sublobectomy, should be further confirmed; the authors aim to conduct another study in the future when more follow-up data has been collected.

In conclusion, the nomogram presented in this study incorporating clinical characteristics and radiological features allows for the preoperative prediction of the STAS status in patients with clinical stage IA NSCLC and can assist thoracic surgeons in rationally selecting surgical methods.

Conflict of interest disclosure

The authors declared no conflicts of interest.

Funding

This research was funded by the National Key R&D Program of China (2022YFC2010000, 2022YFC2010002), the Key Program of National Natural Science Foundation of China (81930049), the National Natural Science Foundation of China (82171926, 82202140), the Shanghai Sailing Program (20YF1449000), the Shanghai Science and Technology Innovation Action Plan Program (19411951300), the Clinical Innovative Project of Shanghai Changzheng Hospital (2020YLCYJ-Y24), and the Program of Science and Technology Commission of Shanghai Municipality (21DZ2202600).

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249. [CrossRef]
2. Ettinger DS, Wood DE, Aisner DL, et al. NCCN Guidelines® Insights: non-small cell lung cancer, version 2.2023. *J Natl Compr Canc Netw*. 2023;21(4):340-350. [CrossRef]
3. Saji H, Okada M, Tsuboi M, et al. Segmentectomy versus lobectomy in small-sized peripheral non-small-cell lung cancer (JCOG0802/WJOG4607L): a multicentre, open-label, phase 3, randomised, controlled, non-inferiority trial. *Lancet*. 2022;399(10335):1607-1617. [CrossRef]
4. Suzuki K, Saji H, Aokage K, et al. Comparison of pulmonary segmentectomy and lobectomy: safety results of a randomized trial. *J Thorac Cardiovasc Surg*. 2019;158(3):895-907. [CrossRef]
5. Suzuki K, Watanabe SI, Wakabayashi M, et al. A single-arm study of sublobar resection for ground-glass opacity dominant peripheral lung cancer. *J Thorac Cardiovasc Surg*. 2022;163(1):289-301. [CrossRef]
6. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol*. 2015;10(9):1243-1260. [CrossRef]
7. Shiono S, Endo M, Suzuki K, Yarimizu K, Hayasaka K, Yanagawa N. Spread through air spaces is a prognostic factor in sublobar resection of non-small cell lung cancer. *Ann Thorac Surg*. 2018;106(2):354-360. [CrossRef]
8. Ren Y, Xie H, Dai C, et al. Prognostic impact of tumor spread through air spaces in sublobar resection for 1A lung adenocarcinoma patients. *Ann Surg Oncol*. 2019;26(6):1901-1908. [CrossRef]
9. Chae M, Jeon JH, Chung JH, et al. Prognostic significance of tumor spread through air spaces in patients with stage IA part-solid lung adenocarcinoma after sublobar resection. *Lung Cancer*. 2021;152:21-26. [CrossRef]
10. Kadota K, Nitadori JI, Sima CS, et al. Tumor spread through air spaces is an important pattern of invasion and impacts the frequency and location of recurrences after limited resection for small stage I lung adenocarcinomas. *J Thorac Oncol*. 2015;10(5):806-814. [CrossRef]
11. Cao L, Jia M, Sun PL, Gao H. Histopathologic features from preoperative biopsies to predict spread through air spaces in early-stage lung adenocarcinoma: a retrospective study. *BMC Cancer*. 2021;21(1):913. [CrossRef]
12. Villalba JA, Shih AR, Sayo TMS, et al. Accuracy and reproducibility of intraoperative assessment on tumor spread through air spaces in stage 1 lung adenocarcinomas. *J Thorac Oncol*. 2021;16(4):619-629. [CrossRef]
13. Zhang Z, Liu Z, Feng H, et al. Predictive value of radiological features on spread through air space in stage cIA lung adenocarcinoma. *J Thorac Dis*. 2020;12(11):6494-6504. [CrossRef]
14. de Margerie-Mellon C, Onken A, Heidinger BH, VanderLaan PA, Bankier AA. CT manifestations of tumor spread through airspaces in pulmonary adenocarcinomas presenting as subsolid nodules. *J Thorac Imaging*. 2018;33(6):402-408. [CrossRef]
15. Qin L, Sun Y, Zhu R, Hu B, Wu J. Clinicopathological and CT features of tumor spread through air space in invasive lung adenocarcinoma. *Front Oncol*. 2022;12:959113. [CrossRef]
16. Li C, Jiang C, Gong J, Wu X, Luo Y, Sun G. A CT-based logistic regression model to predict spread through air space in lung adenocarcinoma. *Quant Imaging Med Surg*. 2020;10(10):1984-1993. [CrossRef]
17. Kim SK, Kim TJ, Chung MJ, et al. Lung adenocarcinoma: CT features associated with spread through air spaces. *Radiology*. 2018;289(3):831-840. [CrossRef]
18. Chen Y, Jiang C, Kang W, et al. Development and validation of a CT-based nomogram to predict spread through air space (STAS) in peripheral stage IA lung adenocarcinoma. *Jpn J Radiol*. 2022;40(6):586-594. [CrossRef]
19. Toyokawa G, Yamada Y, Tagawa T, et al. Computed tomography features of resected lung adenocarcinomas with spread through air spaces. *J Thorac Cardiovasc Surg*. 2018;156(4):1670-1676. [CrossRef]
20. Yanagawa N, Shiono S, Endo M, Ogata SY. Tumor spread through air spaces is a useful predictor of recurrence and prognosis in stage I lung squamous cell carcinoma, but not

in stage II and III. *Lung Cancer*. 2018;120:14-21. [CrossRef]

21. Yokoyama S, Murakami T, Tao H, et al. Tumor spread through air spaces identifies a distinct subgroup with poor prognosis in surgically resected lung pleomorphic carcinoma. *Chest*. 2018;154(4):838-847. [CrossRef]
22. Aly RG, Rekhtman N, Li X, et al. Spread through air spaces (STAS) is prognostic in atypical carcinoid, large cell neuroendocrine carcinoma, and small cell carcinoma of the lung. *J Thorac Oncol*. 2019;14(9):1583-1593. [CrossRef]
23. Jia M, Yu S, Gao H, Sun PL. Spread through air spaces (STAS) in lung cancer: a multiple-perspective and update review. *Cancer Manag Res*. 2020;12:2743-2752. [CrossRef]
24. Dercle L, Fronheiser M, Lu L, et al. Identification of non-small cell lung cancer sensitive to systemic cancer therapies using radiomics. *Clin Cancer Res*. 2020;26(9):2151-2162. [CrossRef]
25. Travis WD, Asamura H, Bankier AA, et al. The IASLC lung cancer staging project: proposals for coding T categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the TNM classification of lung cancer. *J Thorac Oncol*. 2016;11(8):1204-1223. [CrossRef]
26. Fan L, Liu SY, Li QC, Yu H, Xiao XS. Multidetector CT features of pulmonary focal ground-glass opacity: differences between benign and malignant. *Br J Radiol*. 2012;85(1015):897-904. [CrossRef]
27. Tu W, Sun G, Fan L, et al. Radiomics signature: a potential and incremental predictor for EGFR mutation status in NSCLC patients, comparison with CT morphology. *Lung Cancer*. 2019;132:28-35. [CrossRef]
28. Hsu JS, Han IT, Tsai TH, et al. Pleural Tags on CT scans to predict visceral pleural invasion of non-small cell lung cancer that does not abut the pleura. *Radiology*. 2016;279(2):590-596. [CrossRef]
29. Kim HJ, Cho JY, Lee YJ, et al. Clinical significance of pleural attachment and indentation of subsolid nodule lung cancer. *Cancer Res Treat*. 2019;51(4):1540-1548. [CrossRef]
30. 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. [CrossRef]
31. Kadota K, Kushida Y, Kagawa S, et al. Limited resection is associated with a higher risk of locoregional recurrence than lobectomy in stage I lung adenocarcinoma with tumor spread through air spaces. *Am J Surg Pathol*. 2019;43(8):1033-1041. [CrossRef]
32. Qi L, Xue K, Cai Y, Lu J, Li X, Li M. Predictors of CT morphologic features to identify spread through air spaces preoperatively in small-sized lung adenocarcinoma. *Front Oncol*. 2021;10:548430. [CrossRef]
33. 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(2):462-469. [CrossRef]
34. Gruden JF. What is the significance of pleural tags? *AJR Am J Roentgenol*. 1995;164(2):503-504. [CrossRef]
35. Zhang Y, Qiang JW, Ye JD, Ye XD, Zhang J. High resolution CT in differentiating minimally invasive component in early lung adenocarcinoma. *Lung Cancer*. 2014;84(3):236-241. [CrossRef]
36. Yin Q, Wang H, Cui H, et al. Meta-analysis of association between CT-based features and tumor spread through air spaces in lung adenocarcinoma. *J Cardiothorac Surg*. 2020;15(1):243. [CrossRef]

Supplementary Table 1. The definitions of radiological features

Feature	Definition
Tumor size	The longest diameter of the whole tumor at the lung window on the MPR image
Solid component size	The longest diameter of the solid component of the tumor at the lung window on the MPR image
CTR	The proportion of the solid component part (consolidation-to-tumor ratio, CTR)
Location	Lobe of tumor originated from
Density	mGGN, presence of ground-glass opacity and solid density component; solid, absence of ground-glass opacity, contains solid density component only
Shape	Shape of tumor on the MPR images, including round/oval, irregular shape
Tumor-lung interface	Interface of the tumor-lung, including well-defined or ill-defined interface
Marginal characteristics	
Lobulation sign	Petaloid or wavy appearance at the tumor's margins
Spiculation sign	Short, thin linear strands radiating around the surface of the tumor without reaching the pleural surface
Internal characteristics	
Vacuole sign	A small air containing space (≤5 mm) in the tumor, referring to lung tissue not invaded by the tumor
Cavity or cystic airspace	A larger air containing space (>5 mm) in the tumor because of intratumoral necrosis or represents spared parenchyma, normal or ectatic bronchi, or focal emphysema
External characteristics	
Bronchial change	Air-filled bronchus manifesting as natural, dilated/distorted or cut-off within the lesions, or cut-off at the edge of the lesions
Vascular convergence sign	The convergence of pulmonary vessels around the tumor towards the lesion
Pleural tags sigh	One or multiple high-density linear strands connecting the tumor margin and the pleura
Pleural indentation sigh	The deviation of the pleura from its original position due to tumor traction at the lung window
Halo sigh	Ill-defined peripheral ground-glass opacity or consolidation around the tumor, which should be distinguished from the well-defined ground-glass opacity of a part-solid lesion
Satellite lesion sigh	Smaller nodules located within 2 cm of the primary tumor
Distal ribbon sign	A long, thick ribbon strand extending from the distal part of the tumor into the surrounding lung tissue
ELLC	Presence of emphysema in the lobe of lung cancer with visual observation
ERL	Presence of emphysema in the remaining lobes with visual observation

CTR, consolidation-to-tumor ratio; ELLC, emphysema in the lobe of lung cancer; ERL, emphysema in the remaining lobes; MPR, multiplanar reformation.

Supplementary Table 2. Consistency analysis of radiological features among observers

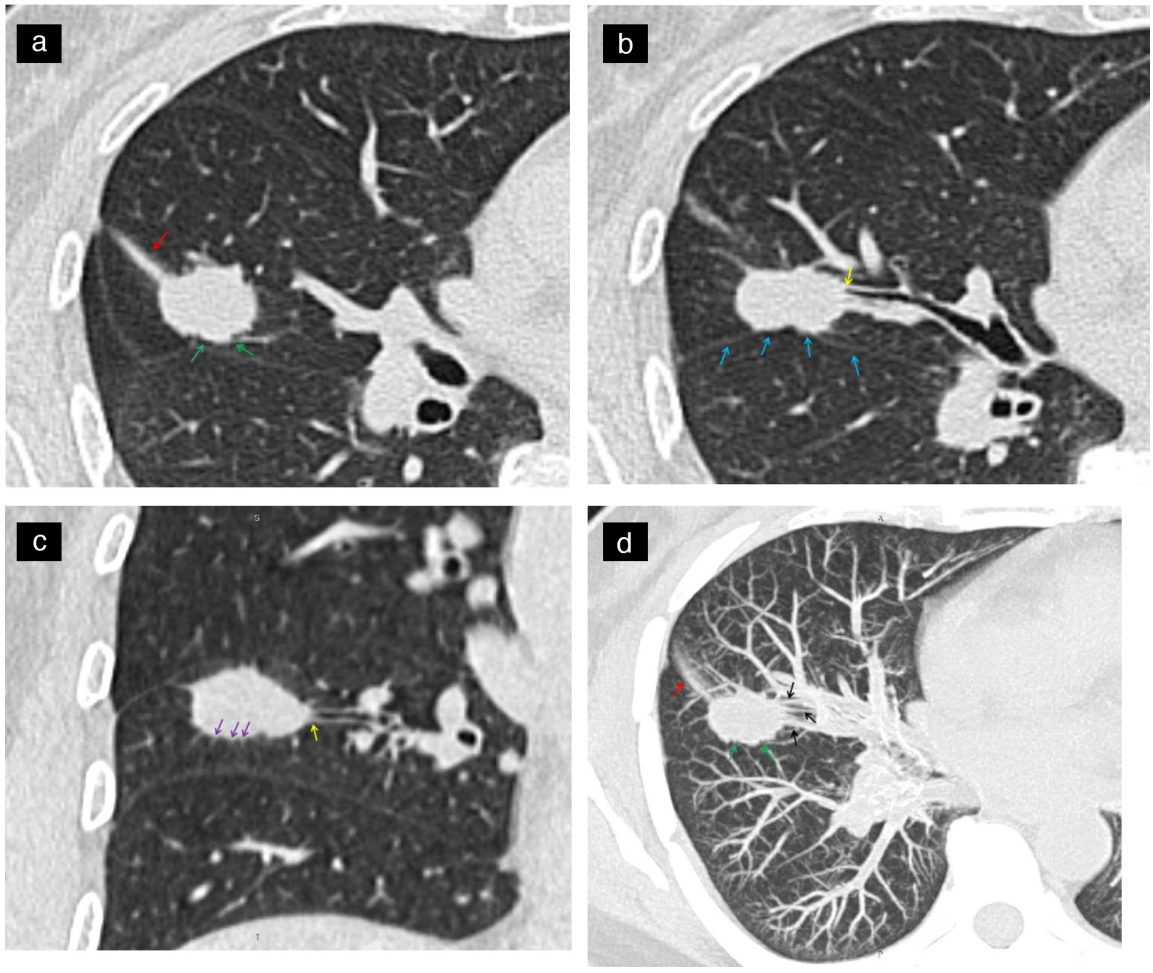
	Kappa (95% CI)		ICC (95% CI)
Density type	0.960 (0.931–0.989)	Tumor size (mm)	0.934 (0.911–0.950)
Distal ribbon sign	0.968 (0.939–0.996)	Solid component size (mm)	0.935 (0.919–0.947)
Shape	0.887 (0.834–0.941)		
Lobulation	0.949 (0.912–0.986)		
Spiculation	0.953 (0.916–0.990)		
Tumor-lung interface	0.886 (0.803–0.969)		
Bronchial change	0.925 (0.888–0.962)		
Vacuole sign	0.963 (0.934–0.992)		
Cavity or cystic airspace	0.881 (0.800–0.962)		
Vascular convergence	0.875 (0.817–0.933)		
Halo sign	0.852 (0.709–0.995)		
Pleural tags	0.960 (0.933–0.987)		
Pleural indentation	0.970 (0.947–0.994)		
Satellite lesions	0.925 (0.859–0.990)		
ELLC	1.000		
ERL	1.000		

CI, confidence interval; ICC, intraclass correlation coefficient; ELLC, emphysema in the lobe of lung cancer; ERL, emphysema in the remaining lobes.

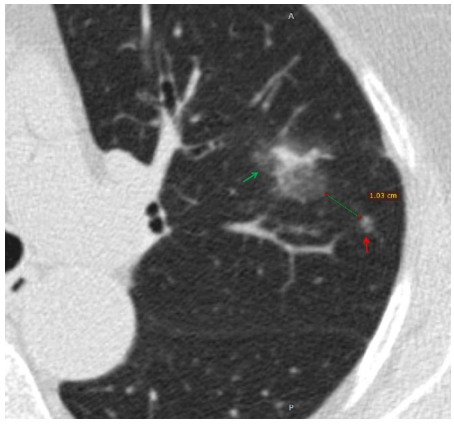
Supplementary Table 3. The training cohort included collinearity test of multivariate logistic regression analysis variables

	Collinearity test statistics	
	Tolerance	Variance inflation factor
Gender	0.669	1.494
Smoking status	0.664	1.505
CEA	0.873	1.146
Clinical T-stage	0.400	2.502
CTR	0.528	1.895
Density type	0.601	1.663
Tumor-lung interface	0.846	1.182
Spiculation	0.663	1.508
Bronchial change	0.844	1.185
Pleural tags	0.742	1.347
Pleural indentation	0.802	1.246
Vascular convergence	0.758	1.319
Halo sign	0.814	1.228
Distal ribbon sign	0.674	1.484

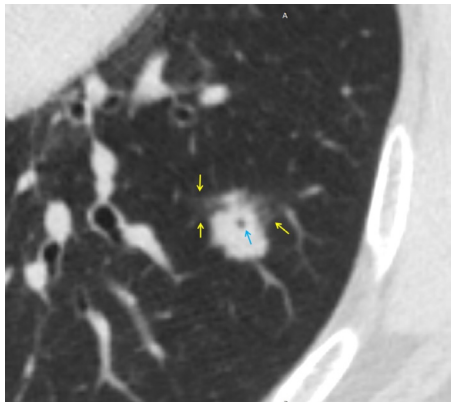
CEA, carcinoembryonic antigen; CTR, consolidation-to-tumor ratio.



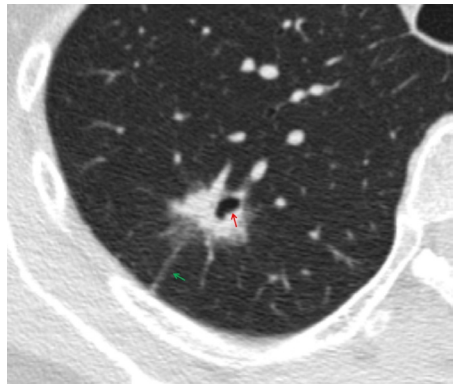
Supplementary Figure 1. Show a same patient, a 53-year-old female patient with lung adenocarcinoma and positive STAS status. The axial non-contrast computed tomography shows a solid nodule in the right middle lobe of the lung, CTR ≥50%, with distal ribbon sign (**a**, red arrow), lobulation sigh (**a**, green arrow), interlobar pleura indentation sigh (**b**, blue arrow), bronchial change (**b** and **c**, yellow arrow), spiculation sigh (**c**, purple arrow). The maximal intensity projection **(d)** shows a vascular convergence sign (black arrow), lobulation sigh (green arrow) and distal ribbon sigh (red arrow). STAS, spread through air spaces; CTR, consolidation-to-tumor ratio.



Supplementary Figure 2. Shows a 77-year-old female patient with lung adenocarcinoma and negative STAS status. The axial non-contrast computed tomography shows a mixed ground glass nodule in the left upper lobe of the lung (green arrow) with satellite lesion sign (red arrow), the distance between the nodule and the satellite lesion is 1.03 cm. STAS, spread through air spaces.



Supplementary Figure 3. Shows a 74-year-old male patient with invasive mucinous adenocarcinoma and negative STAS status. The axial non-contrast computed tomography shows a solid nodule in the left lower lobe of the lung with halo sign (yellow arrow) and vacuole sign (blue arrow). STAS, spread through air spaces.



Supplementary Figure 4. Shows a 62-year-old female patient with invasive adenocarcinoma and negative STAS status. The axial non-contrast computed tomography shows a mixed ground glass nodule in the right upper lobe of the lung with well-defined interface, pleural tags sigh (green arrow) and cavity or cystic airspace (red arrow). STAS, spread through air spaces.