Original Article

Deep learning model for predicting spread through air spaces of lung adenocarcinoma based on transfer learning mechanism

Jia-Ning Zhang[1,2#], Zhuo-Fu Li[1,2#], Sun-Yi Zheng[1,2], Jiao-Jiao Li[3], Lei-Na Sun[2,4], Zhao-Xiang Ye[1,2]

[1]Department of Radiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin, China; [2]State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Key Laboratory of Cancer Prevention and Therapy, Tianjin, China; [3]Department of Radiology, First Affiliated Hospital of Hebei North University, Zhangjiakou, China; [4]Department of Pathology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin, China

Contributions: (I) Conception and design: JN Zhang, ZX Ye; (II) Administrative support: ZX Ye; (III) Provision of study materials or patients: LN Sun, ZX Ye; (IV) Collection and assembly of data: JN Zhang, ZF Li, JJ Li; (V) Data analysis and interpretation: JN Zhang, ZF Li, SY Zheng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Zhao-Xiang Ye, MD. Department of Radiology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Huanhuxi Road, Ti Yuan Bei, Hexi District, Tianjin 300060, China; State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Key Laboratory of Cancer Prevention and Therapy, Tianjin, China. Email: yezhaoxiang@163.com.

Background: Spread through air space (STAS) is a novel invasive pattern of lung adenocarcinoma (LUAD) associated with poor prognosis. Preoperative predicting of STAS helps choose an appropriate surgical and therapeutic strategy. This study aimed to develop and validate an STAS prediction model in LUAD based on deep learning algorithms.

Methods: A dataset of 290 patients with preoperative chest computed tomography (CT) images and confirmed STAS status was retrospectively selected. Optimal semantic features were selected by logistic regression. Image features were learned from cubic patches containing lung tumors and the area around the tumor within 5/10/15 mm extracted from CT scans. ResNet50 architecture was used to train deep learning models based on the transfer learning mechanism. The optimal semantic features are combined with the deep learning model to construct a hybrid model. Receiver operating characteristic (ROC) curves were used to evaluate the performance.

Results: Patients were randomly partitioned into a training set (70%, n=203) and a test set (30%, n=87). The International Association for the Study of Lung Cancer (IASLC) grade, maximum tumor diameter, tumor density, spiculated sign, vacuole sign, and peritumor obstructive inflammation were incorporated into the hybrid model as independent predictors. The STAS-HYBRID$_{t10}$ proved to be the optimal STAS prediction model with an area under the curve (AUC) value of 0.874 in the training set and 0.801 in the test set. The sensitivity, specificity, and accuracy of STAS-HYBRID$_{t10}$ were 0.659/0.526, 0.904/0.837, and 0.798/0.701 in the training set and test set, respectively.

Conclusions: The STAS-HYBRID$_{t10}$ has great potential for the preoperative prediction of STAS and may support decision-making for surgical and therapeutic planning in LUAD.

Keywords: Spread through air spaces (STAS); deep learning (DL); ResNet50; transfer learning; lung adenocarcinoma (LUAD)

Submitted Oct 22, 2024. Accepted for publication Feb 19, 2025. Published online Apr 18, 2025.
doi: 10.21037/tlcr-24-985
View this article at: https://dx.doi.org/10.21037/tlcr-24-985

Introduction

Lung cancer remained the leading cause of cancer morbidity and mortality, responsible for 12.4% of all cancers diagnosed worldwide (1-3). Lung adenocarcinoma (LUAD) accounts for approximately 60% of newly diagnosed cases of lung cancer, making it the most prevalent subtype (1,4). For a long time, lobectomy has remained the standard surgical procedure for patients with resectable stage I/II diseases and functional operability (C-level evidence) (5-7). However, with the advancement of screening programs, more relatively early-stage nodules are being detected (8-10), and as a result, an increasing number of studies work on the feasibility of using sub-lobectomy such as wedge resection and segmental resection for early non-small cell lung cancer (NSCLC) (11-14).

Existing studies have demonstrated that LUAD has a variety of invasive pathological features, resulting in significant differences in the true prognosis for individuals at the same stage. Tumor spread through air spaces (STAS) is one of the invasive pathological features, which may indicate an unfavorable prognosis of NSCLC (15,16). STAS is a unique invasion pattern of lung cancer compared with malignancies originating in other organs. In 2015, STAS was defined as "the spread of micropapillary clusters, solid nests, or single cancer cells into air spaces in surrounding lung parenchyma beyond the edge of the main tumor" by the World Health Organization (WHO) (17). Distinct from the traditional invasion in lung cancer (nonlepidic histologic pattern of growth, myofibroblastic proliferation with desmoplasia, and vascular or pleural invasion), STAS represents a unique airspace invasion of the lung (18). Many studies have confirmed the association of STAS and its equivalent with shorter recurrence-free survival (RFS) and/or overall survival (OS) (15,19,20). Accurate preoperative STAS status prediction is of great significance for patient prognosis stratification and therapeutic plan optimization. However, the accuracy of intraoperative frozen section for STAS detection is still limited (21-23), and STAS proved by postoperative pathology has a certain lag.

The value of preoperative computed tomography (CT) to evaluate STAS status in treatment decision-making has been the subject of comprehensive research. However, due to the limitation of spatial resolution of CT scanners, high-resolution CT images cannot show STAS directly. Deep learning (DL) algorithms can autonomously identify intrinsic correlations within specific medical imaging datasets and learn these features, then generate models with exceptional performance (24-26). This study developed and validated a novel preoperative CT-based ResNet50 model to predict the STAS status in LUAD. We present this article in accordance with the TRIPOD reporting checklist (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-985/rc).

Highlight box

Key findings

- The transfer learning-based ResNet50 model, combining semantic independent predictors, could extract multi-scale features to improve the accuracy of tumor spread through air spaces (STAS) status prediction. Pre-surgical stratification for STAS status could guide therapeutic decisions.

What is known and what is new?

- STAS is a unique invasion pattern of lung cancer that has been confirmed to be associated with shorter recurrence-free survival (RFS) and/or overall survival (OS). Accurate preoperative STAS status prediction is of great significance for patient prognosis stratification and therapeutic plan optimization.
- The hybrid deep learning model showed remarkable performance in the identification of STAS status in lung adenocarcinoma (LUAD) patients.

What is the implication, and what should change now?

- Such a hybrid model can effectively predict STAS status preoperatively, thereby guiding clinical surgical decisions and optimizing the clinical translation of precision medicine. Larger multicentre studies are necessary to address this hypothesis. In addition, subsequent work will include collecting patient prognostic information to elucidate the role of deep learning models in STAS patient prognosis.

Methods

Study population

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This retrospective study was approved by the Ethics Committee of Tianjin Medical University Cancer Institute and Hospital (No. Ek2023044) and informed consent was waived from all the patients. None of the participants was previously reported.

The study initially collected preoperative chest CT images from consecutive patients with confirmed LUAD between March 2016 and January 2020 at Tianjin Medical University Cancer Institute and Hospital. The flowchart of patient selection is presented in *Figure 1*. The inclusion

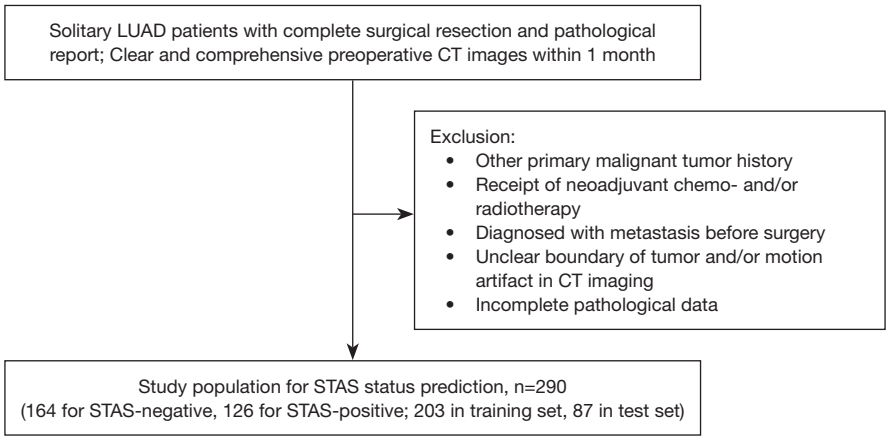

Figure 1 Flowchart for the study cohort. CT, computed tomography; LUAD, lung adenocarcinoma; STAS, spread through air spaces.

criteria were as follows: (I) solitary lesions with complete surgical resection; (II) preoperative CT scans performed within one month, without prior treatment such as neoadjuvant radiotherapy/chemotherapy, etc.; (III) clear and comprehensive CT images devoid of artifacts; and (IV) STAS status confirmed by postoperative pathology. The exclusion criteria were as follows: (I) patient with a history of other primary malignant tumors; (II) patient diagnosed with metastasis to other organs and tissues before surgery; and (III) incomplete pathological data. Finally, 290 cases were included in this study and randomly split into the training set and the testing set in a 7:3 ratio.

Morphological features assessment

All CT images were acquired in axial orientation using scanning parameters of 120 kVp, 150–200 mA, and 1.25 mm thickness. One senior radiologist and two junior radiologists independently carried out a visual evaluation of morphological CT features using continuous value or a binomial classification of 0 (absent) *vs.* 1 (present). All radiologists were blinded to the final pathologic results of all patients. In case of controversy, a consensus was reached after discussion. The evaluated features include (I) maximum tumor diameter; (II) tumor location (peripheral/central type); (III) lung lobe of tumor; (IV) tumor density; (V) lobulated sign; (VI) spiculated sign; (VII) air bronchogram; (VIII) vacuole sign; (IX) pleural indentation; (X) vessel convergence; (XI) peritumoral obstructive inflammation.

Clinical and pathological assessment

Clinical data collection includes gender, age, smoking status, tumor stage, and surgical methods. LUAD tumor staging was based on the International Association for the Study of Lung Cancer (IASLC) Lung Cancer Staging Project, Ninth Edition (27). Pathological data collection includes STAS status and the proportion of each histologic pattern contained in the tumor. The assessment of STAS status was conducted by two senior thoracic pathologists, according to the WHO-defined criteria. The pathological grading referred to the triple grading system for invasive nonmucinous adenocarcinoma proposed by the IASLC Pathology Committee. This grading scheme was based on a combination of the predominant histologic pattern plus the worst pattern: (I) grade I (well differentiated), lepidic predominant with no or <20% high-grade pattern (micropapillary, solid, cribriform, and complex glandular patterns); (II) grade II (moderately differentiated), acinar or papillary predominant with no or <20% high-grade pattern; (III) grade III (poorly differentiated), any tumor with ≥20% high-grade pattern (4,28).

Machine learning on morphological features and clinicopathologic factors

The univariable and multivariable logistic regression were used to select semantic independent predictors and establish a semantic model for STAS status prediction. The Bootstrap method was used to resample the training set

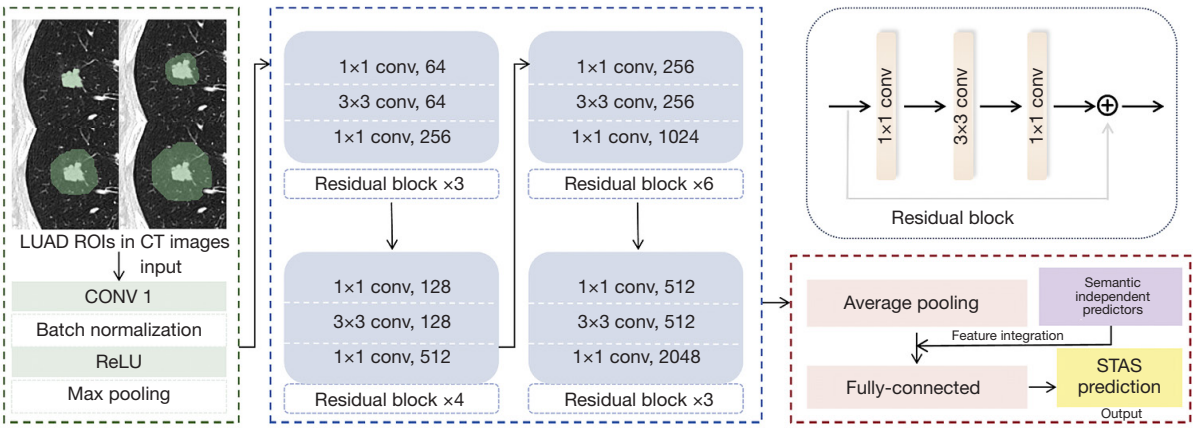

Figure 2 Schematic diagram of ResNet50. ResNet50 started with conventional convolutional layers, batch normalization, ReLU functions, and max-pooling layers for initial feature extraction. The model then iterated through residual blocks to capture more higher-level features. Each residual block contained 3 convolutional layers. The grey arrow indicated residual connection. ResNet50 output the predicted probability of STAS-DL after the fully connected layer. In contrast, the semantic independent predictors were integrated into a feature combination process after the global average pooling layer, which then generated the predicted probability of STAS-HYBRID. CT, computed tomography; CONV, convolution; DL, deep learning; LUAD, lung adenocarcinoma; ROIs, regions of interest; ReLU, rectified linear unit; STAS, spread through air space.

data 1,000 times to obtain the average area under the curve (AUC) value, which can be used to evaluate the stability of the semantic model and thus correct possible overfitting in the training set.

Transfer learning-based DL model and hybrid model

Tumor segmentation was performed on each patient's image by 3D Slicer software (https://download.slicer.org), which was subsequently automatically extended to the peri-tumor area within 5, 10, and 15 mm based on the tumor itself, respectively. Four 3D tumor regions of interest (ROIs) were finally formed for each lesion. This study used a pre-trained convolutional neural network (CNN), Residual Network with 50 parameter layers (ResNet50), to train models based on the transfer learning mechanism. Transfer learning can improve model accuracy by utilizing learned features from pre-trained models. The architecture of ResNet50 is illustrated in *Figure 2*. Linear interpolation was used to unify the resolution of CT images, and the cube patches containing ROIs were extracted from the interpolated images as input. The ResNet50 architecture contained 16 residual blocks and ended with a global average pooling layer and a 1000-way fully-connected layer with softmax. Connected with the fully-connected layer, the output generated probabilities to enable STAS status prediction. A hybrid model was then developed. The feature integration of semantic independent predictors was performed after the global average pooling layer, ultimately obtaining the prediction probability of the hybrid model.

Statistical analysis

Statistical analysis was performed using SPSS Statistics 25.0 and Python (version 3.7). P values smaller than 0.05 were considered statistically significant. The disparities of variables between the STAS-positive and STAS-negative groups concerning both morphological and clinicopathologic features were evaluated. The independent samples *t*-test or Mann-Whitney *U* test were used for continuous variables and the Chi-square test and Fisher's exact test were used for categorical variables. Continuous variables are presented as mean ± standard deviation. Categorical variables are expressed as numerical values and percentages. The performance of the prediction model is evaluated using the receiver operating characteristic (ROC) curve, sensitivity, specificity, and accuracy. The predictive performance between different models is compared by the

Table 1 Baseline demographics and clinicopathological features

Variables	STAS– (n=164)	STAS+ (n=126)	P	Training set (n=203)	Test set (n=87)	P
Gender			0.15			0.35
Female	101 (61.6)	67 (53.2)		114 (56.2)	54 (62.1)	
Male	63 (38.4)	59 (46.8)		89 (43.8)	33 (37.9)	
Age, years	58.84±8.58	59.16±8.67	0.82	59.30±8.97	58.23±7.68	0.23
Smoking status			0.04*			0.85
No	111 (67.7)	70 (55.6)		126 (62.1)	55 (63.2)	
Yes	53 (32.3)	56 (44.4)		77 (37.9)	32 (36.8)	
Tumor stage			0.052			0.45
IA	157 (95.7)	112 (88.9)		185 (91.1)	82 (94.3)	
IB	6 (3.7)	9 (7.1)		15 (7.4)	5 (5.7)	
IIA	1 (0.6)	5 (4.0)		3 (1.5)	0 (0.0)	
IASLC grade			<0.001*			0.65
I	103 (62.8)	28 (22.2)		94 (46.3)	38 (43.7)	
II	40 (24.4)	16 (12.7)		43 (21.2)	16 (18.4)	
III	21 (12.8)	82 (65.1)		66 (32.5)	33 (37.9)	
Surgical method			0.22			0.49
Lobectomy	136 (82.9)	111 (88.1)		171 (84.2)	76 (87.4)	
Sub-lobectomy	28 (17.1)	15 (11.9)		32 (15.8)	11 (12.6)	

Continuous variables are presented as mean ± standard deviation. Categorical variables are expressed as numerical values and percentages. *, statistically significant. STAS, spread through air space; IASLC, International Association for the Study of Lung Cancer.

DeLong test.

Results

Patient demographics and clinicopathologic features

In this study, 290 LUAD patients (STAS-positive: n=126; STAS-negative: n=164) were enrolled including 122 males (42.1%) and 168 females (57.9%). The proportion of smokers in the STAS-positive group (n=56, 44.4%) was higher than that in the STAS-negative group (n=53, 32.3%) with a P value of 0.04. The predominant IASLC grade of the STAS-negative group was grade I (n=103, 62.8%), while the predominant IASLC grade of the STAS-positive group was grade III (n=82, 65.1%). The difference in IASLC grade distribution between the two groups was statistically significant (P<0.001). After dividing patients into the training set (n=203) and the test set (n=87), no significant difference was found in any clinicopathologic features between the two sets. Within the training set, 88 patients were STAS-positive, and 115 patients were STAS-negative. Within the test set, 38 and 49 patients were STAS-positive and STAS-negative respectively. Detailed information on the total study cohort is summarized in *Table 1*.

Morphological features evaluation

The morphological features in *Table 2* showed that the maximum diameter of STAS-positive tumors (2.53±0.74 cm) was larger than that of STAS-negative tumors (1.49±0.69 cm, P<0.001). Significant differences were also shown in tumor density (P<0.001), spiculated sign (P<0.001), air bronchogram (P=0.02), vacuole sign (P<0.001), vessel convergence (P=0.003), and obstructive inflammation (P<0.001) between STAS-positive tumors and STAS-negative tumors. There were no significant statistical differences in tumor location, lobulated sign, or pleural indentation between two different STAS status groups (P>0.05).

Table 2 Morphological features of patients

Variables	STAS– (n=164)	STAS+ (n=126)	P	Training set (n=203)	Test set (n=87)	P
Size, cm	1.49±0.69	2.53±0.74	<0.001*	1.84±0.76	2.11±0.82	0.39
Location			0.30			0.44
Peripheral type	161 (98.2)	121 (96.0)		196 (96.6)	86 (98.9)	
Central type	3 (1.8)	5 (4.0)		7 (3.4)	1 (1.1)	
Lobular location			0.59			0.39
RUL	66 (40.2)	48 (38.1)		82 (40.4)	32 (36.8)	
RML	6 (3.7)	7 (5.6)		7 (3.4)	6 (6.9)	
RLL	22 (13.4)	23 (18.3)		28 (13.8)	17 (19.5)	
LUL	53 (32.3)	33 (26.2)		61 (30.0)	25 (28.7)	
LLL	17 (10.4)	15 (11.9)		25 (12.3)	7 (8.0)	
Density			<0.001*			0.27
GGN	91 (55.5)	11 (8.7)		79 (38.9)	33 (37.9)	
Part-solid nodule	41 (25.0)	30 (23.8)		32 (15.8)	8 (9.2)	
Solid nodule	32 (19.5)	85 (67.5)		92 (45.3)	46 (52.9)	
Lobulated sign			0.051			0.55
No	80 (48.8)	76 (60.3)		47 (23.2)	23 (26.4)	
Yes	84 (51.2)	50 (39.7)		156 (76.8)	64 (73.6)	
Spiculated sign			<0.001*			0.35
No	113 (68.9)	21 (16.7)		86 (42.4)	42 (48.3)	
Yes	51 (31.1)	105 (83.3)		117 (57.6)	45 (51.7)	
Air bronchogram			0.02*			0.73
No	130 (79.3)	84 (66.7)		151 (74.4)	63 (72.4)	
Yes	34 (20.7)	42 (33.3)		52 (25.6)	24 (27.6)	
Vacuole sign			<0.001*			0.68
No	156 (95.1)	86 (68.3)		183 (90.1)	77 (88.5)	
Yes	8 (4.9)	40 (31.7)		20 (9.9)	10 (11.5)	
Pleural indentation			0.15			0.88
No	89 (54.3)	79 (62.7)		117 (57.6)	51 (58.6)	
Yes	75 (45.7)	47 (37.3)		86 (42.4)	36 (41.4)	
Vessel convergence			0.003*			0.64
No	134 (81.7)	84 (66.7)		151 (74.4)	67 (77.0)	
Yes	30 (18.3)	42 (33.3)		52 (25.6)	20 (23.0)	
Obstructive inflammation			<0.001*			0.47
No	155 (94.5)	78 (61.9)		168 (82.8)	75 (86.2)	
Yes	9 (5.5)	48 (38.1)		35 (17.2)	12 (13.8)	

Continuous variables are presented as mean ± standard deviation. Categorical variables are expressed as numerical values and percentages. *, statistically significant. GGN, ground glass nodule; LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe; STAS, spread through air space.

Table 3 Logistic regression analysis based on the training set

Variables	Univariable		Multivariable	
	OR (95% CI)	P	OR (95% CI)	P
IASLC grade				
I	Reference		Reference	
II	1.30 (0.54–3.12)	0.56	0.70 (0.21–2.32)	0.56
III	20.89 (9.22–47.32)	<0.001*	3.58 (1.07–11.96)	0.04*
Size, cm	8.92 (4.91–16.19)	<0.001*	2.46 (1.07–5.66)	0.03*
Density				
GGN	Reference		Reference	
Part-solid nodule	6.62 (2.42–18.13)	<0.001*	2.13 (0.58–7.77)	0.25
Solid nodule	55.73 (19.98–155.46)	<0.001*	4.52 (1.02–19.98)	0.046*
Spiculated sign				
No	Reference		Reference	
Yes	16.68 (7.89–35.23)	<0.001*	3.29 (1.16–9.31)	0.03*
Air bronchogram				
No	Reference		Reference	
Yes	1.97 (1.04–3.72)	0.04*	0.49 (0.15–1.57)	0.23
Vacuole sign				
No	Reference		Reference	
Yes	9.40 (3.70–23.88)	<0.001*	4.81 (1.19–19.44)	0.03*
Vessel convergence				
No	Reference		Reference	
Yes	2.43 (1.28–4.64)	0.007*	1.16 (0.39–3.43)	0.79
Obstructive inflammation				
No	Reference		Reference	
Yes	11.73 (4.90–28.07)	<0.001*	4.06 (1.20–13.73)	0.02*

*, statistically significant. CI, confidence interval; GGN, ground glass nodule; IASLC, International Association for the Study of Lung Cancer; OR, odds ratio.

Performance of semantic signature

Logistic regression analysis based on the training set showed that the IASLC grade, tumor maximum diameter, density, spiculated sign, vacuole sign, and obstructive inflammation were independent predictors of STAS status (*Table 3*). The constructed morphological and clinicopathological signature (MCS) model predicted STAS status with AUCs of 0.842 (95% CI: 0.789–0.895) in the training set and 0.714 (95% CI: 0.604–0.823) in the test set, respectively (*Figure 3*, *Table 4*).

Performance of DL models and hybrid models

The ROC curves for the DL and hybrid models are presented in *Table 5* and *Figure 4*. The hybrid models achieved better predictive efficacy than the DL models, with the STAS-HYBRID$_{t10}$ obtaining the largest AUC value (0.801, 95% CI: 0.704–0.884) in the test set. The DeLong

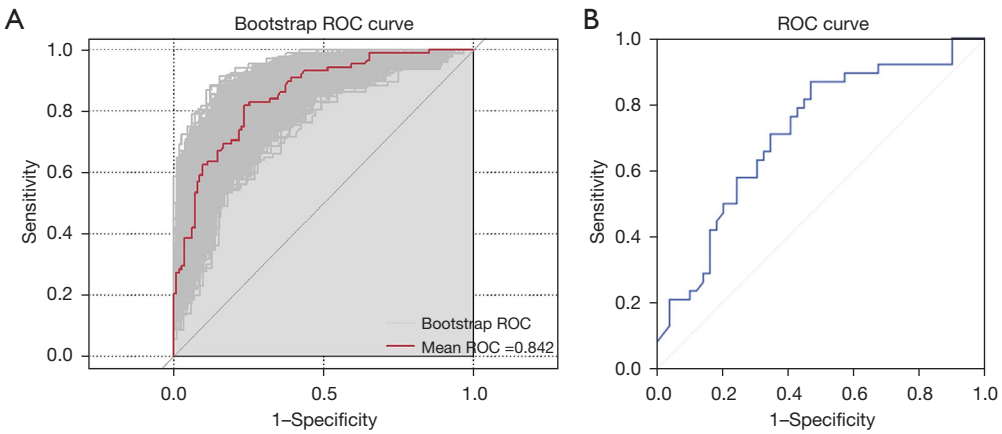

Figure 3 ROC curves of STAS-MCS for STAS status prediction. (A) ROC curves for the training set of the STAS-MCS. The grey line indicates AUC values for resampling 1,000 times. (B) ROC curves for the test set of the STAS-MCS. AUC, area under the curve; MCS, morphological and clinicopathological signature; ROC, receiver operating characteristic; STAS, spread through air space.

Table 4 Performance of semantic features and STAS-MCS model

Variables	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV
IASLC grade	0.751 (0.681–0.821)	0.744	0.534	0.904	0.810	0.717
Size, cm	0.710 (0.640–0.780)	0.626	0.909	0.409	0.541	0.855
Density	0.673 (0.599–0.747)	0.645	0.568	0.704	0.595	0.681
Spiculated sign	0.708 (0.636–0.780)	0.694	0.807	0.609	0.612	0.805
Vacuole sign	0.595 (0.514–0.675)	0.640	0.250	0.939	0.759	0.621
Obstructive inflammation	0.600 (0.520–0.680)	0.645	0.261	0.939	0.767	0.624
STAS-MCS train	0.842 (0.789–0.895)	0.773	0.670	0.852	0.776	0.772
STAS-MCS test	0.714 (0.604–0.823)	0.291	0.868	0.531	0.589	0.838

AUC, area under the curve; CI, confidence interval; IASLC, International Association for the Study of Lung Cancer; MCS, morphological and clinicopathological signature; NPV, negative predictive value; PPV, positive predictive value; STAS, spread through air space.

test suggested that the predictive efficacy of the STAS-HYBRID$_{t10}$ is significantly higher than that of the STAS-MCS in the test set (P=0.03).

The predictions for some case examples are shown in *Figure 5*. *Figure 5A* showed a solid lesion predicted to be STAS-positive by STAS-HYBRID$_{t10}$ and ultimately confirmed to be a true positive (TP) case with typical spiculated sign and IASLC III. The predicted probabilities of the STAS-MCS, STAS-DL$_{t10}$, and STAS-HYBRID$_{t10}$ were 0.64, 0.68 and 0.71, respectively. *Figure 5B* showed a ground glass predominant nodule with small vacuole. The STAS-MCS predicted negative, but the STAS-DL$_{t10}$ and STAS-HYBRID$_{t10}$ predicted positive. A case that was ultimately predicted to be negative by STAS-HYBRID$_{t10}$ was illustrated in *Figure 5C*. The lesion was surrounded by a few inflammatory infiltrates and was misclassified as STAS-positive by the STAS-MCS and STAS-DL$_{t10}$. *Figure 5D* showed a false negative (FN) case with clear vacuoles. *Figure 5E,5F* showed two false positive (FP) cases predicted by STAS-HYBRID$_{t10}$. Subtle imaging features may lead to misclassification by DL models.

Discussion

This study trained and tested a hybrid DL model that included clinicopathologic factors and morphological features from preoperative CT images to predict STAS status in LUAD. Our results showed that the predictive

Table 5 Performance of STAS-DL and STAS-HYBRID model

Models	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV
DL model						
STAS-DL$_{tumor}$ train	0.751 (0.679–0.819)	0.685	0.670	0.696	0.628	0.734
STAS-DL$_{tumor}$ test	0.682 (0.572–0.793)	0.632	0.605	0.653	0.575	0.681
STAS-DL$_{t5}$ train	0.809 (0.749–0.867)	0.685	0.955	0.478	0.583	0.932
STAS-DL$_{t5}$ test	0.723 (0.614–0.823)	0.621	0.789	0.490	0.545	0.750
STAS-DL$_{t10}$ train	0.790 (0.723–0.847)	0.695	0.739	0.661	0.625	0.768
STAS-DL$_{t10}$ test	0.742 (0.633–0.845)	0.713	0.632	0.776	0.686	0.731
STAS-DL$_{t15}$ train	0.787 (0.723–0.848)	0.744	0.795	0.704	0.673	0.818
STAS-DL$_{t15}$ test	0.717 (0.607–0.818)	0.655	0.632	0.673	0.600	0.702
HYBRID model						
STAS-HYBRID$_{tumor}$ train	0.879 (0.826–0.928)	0.808	0.716	0.878	0.818	0.802
STAS-HYBRID$_{tumor}$ test	0.785 (0.679–0.879)	0.678	0.605	0.735	0.639	0.706
STAS-HYBRID$_{t5}$ train	0.888 (0.840–0.929)	0.813	0.693	0.904	0.847	0.794
STAS-HYBRID$_{t5}$ test	0.774 (0.666–0.869)	0.713	0.579	0.816	0.710	0.714
STAS-HYBRID$_{t10}$ train	0.874 (0.816–0.919)	0.798	0.659	0.904	0.841	0.776
STAS-HYBRID$_{t10}$ test	0.801 (0.704–0.884)	0.701	0.526	0.837	0.714	0.695
STAS-HYBRID$_{t15}$ train	0.863 (0.805–0.913)	0.778	0.716	0.826	0.759	0.792
STAS-HYBRID$_{t15}$ test	0.786 (0.682–0.876)	0.713	0.632	0.776	0.686	0.731

AUC, area under the curve; CI, confidence interval; DL, deep learning; NPV, negative predictive value; PPV, positive predictive value; STAS, spread through air space.

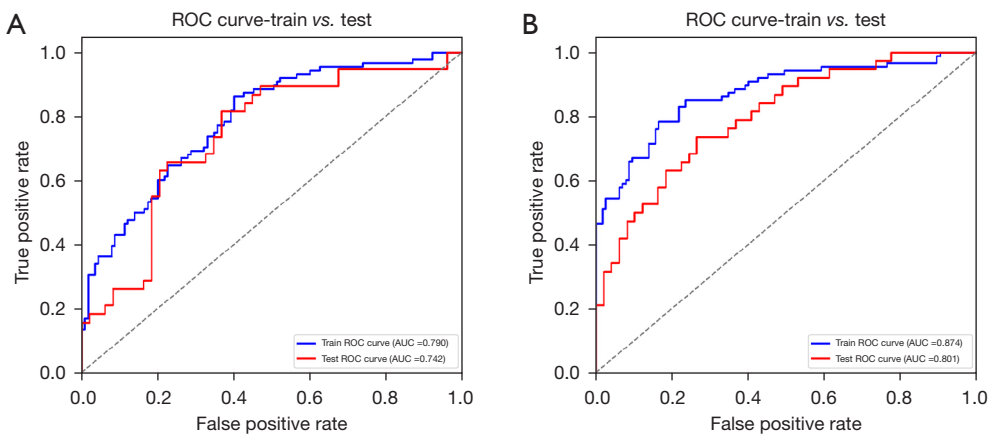

Figure 4 Comparison of ROC curves between STAS-DL$_{t10}$ and STAS-HYBRID$_{t10}$. (A) ROC curves for the training set and test set of the STAS-DL$_{t10}$. (B) ROC curves for the training set and test set of the STAS-HYBRID$_{t10}$. AUC, area under the curve; DL, deep learning; ROC, receiver operating characteristic; STAS, spread through air space.

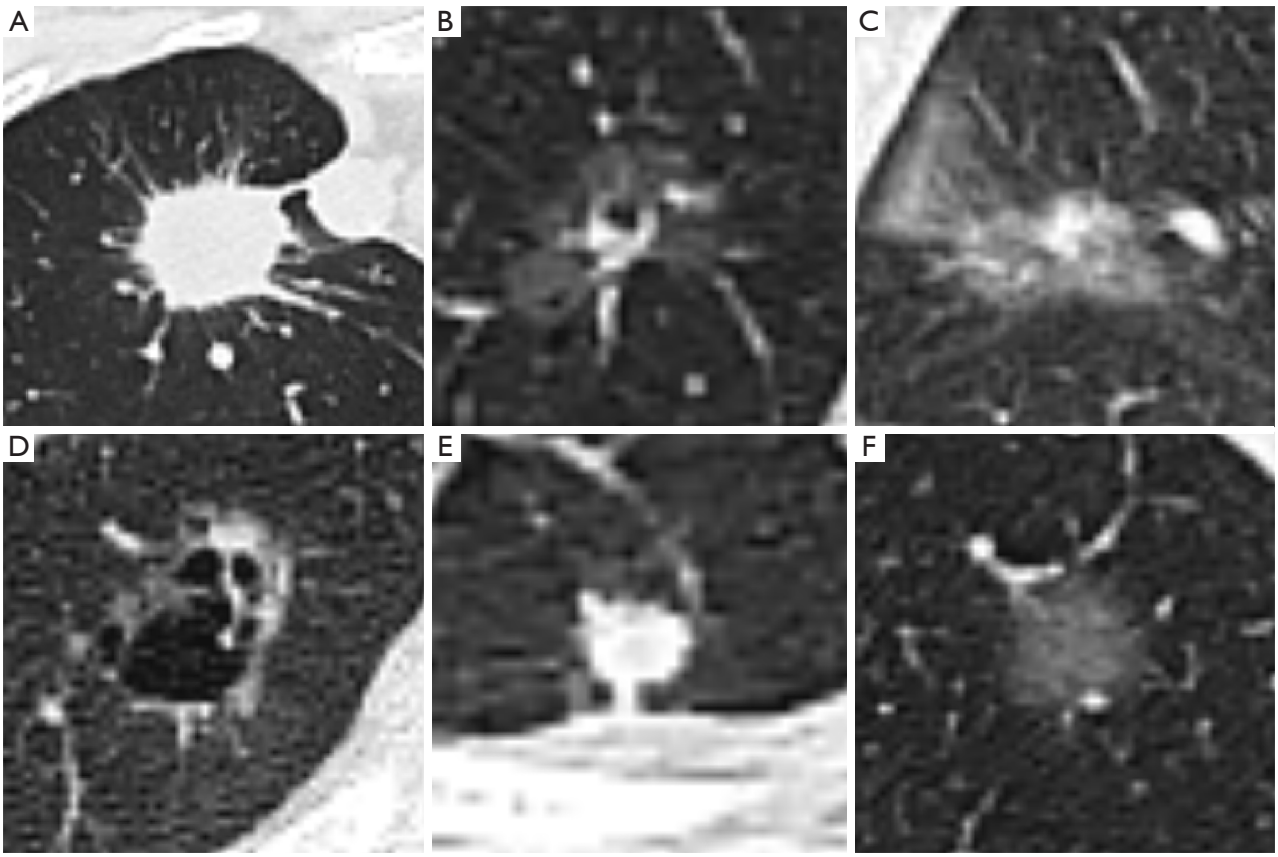

Figure 5 Representative case examples. (A) A solid lesion predicted to be STAS-positive by STAS-HYBRID$_{t10}$ and ultimately confirmed to be a true positive case. (B) A ground glass predominant nodule predicted as negative by STAS-MCS, but predicted as positive by STAS-DL$_{t10}$ and STAS-HYBRID$_{t10}$. (C) A case predicted to be negative by STAS-HYBRID$_{t10}$, but misclassified as STAS-positive by the STAS-MCS and STAS-DL$_{t10}$. (D) A false negative case. (E,F) Two false positive cases predicted by STAS-HYBRID$_{t10}$. DL, deep learning; MCS, morphological and clinicopathological signature; STAS, spread through air space.

efficacy of STAS-HYBRIDs outperformed the STAS-DLs and the STAS-MCS, and the proposed STAS-HYBRID$_{t10}$ model yielded the best performance with the AUCs of 0.874 (95% CI: 0.816–0.919) and 0.801 (95% CI: 0.704–0.884) in the training set and test set, respectively.

STAS is a lesion where tumor cells are microscopically identified as micropapillary clusters, solid nests, or solitary neoplastic cells beyond the edge of the main tumor in the adjacent alveolar parenchyma. The presence of STAS may lead to residual tumor cells at the resection margin after sub-lobectomy and the consequent poor prognosis, which emphasizes the importance of predicting STAS status before surgery to determine a reasonable surgical extension (29-31). As a pathological invasive behavior, STAS is a microscopic phenomenon that cannot be directly visualized with the resolution of CT. Hence, numerous studies have looked into indirect features reflecting aggressiveness for predicting STAS. Kim *et al.* (32) found that STAS was more likely to manifest in pure solid than subsolid lesions. The odds of STAS increased by 3.2 times for each 20% increase in the percentage of solid components, indicating that the proportion of solid components is a promising imaging biomarker for predicting STAS in LUAD, which may have a substantial impact on treatment decisions regarding the adequate extent of surgery and local ablative therapies. Consistent with the above conclusion, Tasnim *et al.* (31) used the variable importance for prediction (VIMP) measure to order the predicted response variables hierarchically. The results showed that the pure solid tumors had the highest STAS predicted probability of 40%, and part-solid tumors

had a predicted probability of 38%. Moreover, other studies have shown that the increasing tumor size was associated with an increasing probability of STAS (31,33,34). In line with these studies, our research showed that tumor density and the maximum tumor diameter were confirmed as independent predictors for STAS (AUC =0.673 and 0.710, respectively), typically indicating pathological infiltration.

Our results showed that tumor spiculation and peritumoral obstructive inflammation were independent predictors of STAS. Tumor spiculation may reflect the invasive growth of tumors at different speeds in various directions and the fibrous connective tissue proliferation caused by tumor stimulation of surrounding lung parenchyma. Tumor invasion into the lungs can lead to blockage of the bronchioles and alveolar spaces in the surrounding lung tissue, which may result in obstructive inflammation of the surrounding lung parenchyma. In addition, the vacuole sign was also an independent predictor in our study, consistent with Qin *et al.* (35). However, in Toyokawa's study, the vacuole sign was not an independent predictor (36). This may be due to different inclusion and exclusion criteria in different studies. Furthermore, the relationship between vacuole sign and STAS still need further investigation.

Previous studies have analyzed the relationship between STAS and the histological subtypes of LUAD. Kim *et al.* (32) and Qin *et al.* (35) noted that LUAD with STAS-positive were more likely to be non-lepidic or non-acinar subtypes. Nevertheless, research by Zhang *et al.* (37) showed that tumors predominantly acinar, solid, or micropapillary were more likely to exhibit STAS. The reason for the inconsistent results might be that these studies all classified tumors just according to the single predominant subtype by 2015 WHO criteria (17). A particular shortcoming of this classification scheme was that LUAD exhibited histological heterogeneity and may have multiple subtypes and combinations of different proportions. When classified based on a single predominant subtype, it may lead to the neglect of the "non-predominant" amounts of high-grade patterns, especially the micropapillary pattern. Our study used the triple grading system proposed by the IASLC pathology committee for pathological grading of invasive non-mucinous adenocarcinomas (28). This scheme took into account a combination of both major histological pattern and the worst pattern, which has proven practical and superior. The IASLC grade was confirmed as an independent predictor with an AUC of 0.751 in our study.

Traditional semantic features evaluation of STAS relied on subjective judgment, which was more restricted in the practical application. Compared with conventional clinical-radiological models, some studies have constructed prediction models based on radiomics features (30,38,39), which have improved the diagnostic efficacy. However, the construction of radiomics prediction models required rigorous standard methods including feature extraction and selection. The time-consuming procedures need much manual work (29). DL, as an emerging technology, has achieved remarkable results in automatic feature extraction, image recognition, and differential diagnosis. Chest CT is a routine choice for diagnosing lung cancer, and chest CT-based DL techniques do not require any additional investment in evaluating tumor invasiveness and predicting prognosis, making it a new research hotspot of clinical interest. Studies have been conducted to predict STAS status based on DL models, but the number is still limited. A study based on ground-glass-predominant LUAD demonstrated that a DL model incorporating solid components gated was effective in predicting STAS in patients with ground-glass-predominant LUAD (40). Wang *et al.* (29) developed a super-resolution DL model with the attention mechanism. They demonstrated that the model has good predictive efficacy for estimating STAS status in patients with solid or partial-solid LUAD. Despite confirming the great predictive potential of DL models, both studies lacked comparisons with pure semantic models. Tao's study compared the capability of five models for predicting STAS in LUAD and squamous carcinoma (41). The results confirmed that the 3D CNN model achieved satisfactory discrimination performance with an AUC of 0.93/0.80 in the training and validation cohort compared to the other 4 models. Aligned with these studies, our results showed that the predictive performance of STAS-HYBRID models surpassed that of STAS-DL and STAS-MCS models. This suggested that STAS-HYBRID may provide a more accurate indication of STAS status, potentially guiding more effective clinical management strategies. Prediction of STAS status was made using ResNet50 architecture, which can automatically extract more objective and comprehensive multi-scale features. Compared to traditional CNN, ResNet50 introduced a residual module to solve the problem of gradient vanishing and degradation in deep network training, and its residual connection also made the training process more stable. In addition, small sample sizes tend to result in overfitting when training CNN models.

Transfer learning can solve the problem by transferring knowledge from source data to target data. Compared with other learning methods, transfer learning can reuse existing knowledge to accelerate model training speed and improve model accuracy. In our study, the accuracy of STAS-DL_{t10} in the test set was significantly improved compared to the STAS-MCS (0.713 *vs.* 0.291), and the accuracy of STAS-$HYBRID_{t10}$ combined with semantic independent predictive factors was higher than that of STAS-MCS in both training and test sets (0.798 *vs.* 0.773 in the training set and 0.701 *vs.* 0.291 in the test set). This may be because the STAS-DL model did not require explicit feature design and was able to capture more detailed features, such as the edges, corners, and texture information that human eyes cannot identify. The model including peritumoral region could supplement the heterogeneity information of the peritumoral microenvironment. It is noteworthy that the performance of DL models varies widely based on imaging quality (42). The finding calls for further optimization of imaging quality to obtain more features and details, enabling models to learn richer information, and thereby promoting the widespread application of DL techniques in clinical practice.

STAS is a phenomenon that occurs in the peritumoral lung tissue. In Dai's study (43), the distance between the tumor margin and furthest STAS was 1.35 cm in the training cohort and 0.87 cm in the validation cohort, respectively. Some studies have explored the efficacy of predicting STAS with peritumoral ROI. Takehana *et al.* (44) constructed a ring-shaped ROI 5 mm inward and 5 mm outward from the tumor surface. The predictive efficiency for STAS status was higher than that of the traditional C/T ratio model (P=0.045), confirming that the radiomics features of peritumoral ROI contained key information for predicting STAS. In the study by Wang *et al.* (45), the nomogram based on GPTV10 radscore and clinical-radiological features achieved the highest predictive efficiency. Our study constructed 4 ROIs based on tumor and different peritumoral ranges (5, 10, and 15 mm). Finally, STAS-$HYBRID_{t10}$ exhibited the highest AUC value, which is speculated to be related to the most accurate coverage of high-order features related to STAS.

This study had several limitations. First, it was a retrospective single-center study. The study was somewhat limited in terms of statistical power, potentially due to the subject sample size. Therefore, the hybrid model must be further explored and validated in multicenter studies in the future. Second, the DL model is not fully automated and requires semi-automatic segmentation of nodules. Finally, the impact of STAS on patient outcomes could not be assessed owing to the lack of follow-up data. Subsequent efforts will involve the collection of patient prognostic information to elucidate the role of DL models in STAS patient outcomes.

Conclusions

The hybrid DL model showed remarkable performance in the identification of STAS status in patients with LUAD. Such a hybrid model can be utilized to identify patients with STAS-positive, thereby guiding clinical surgical decisions and optimizing the clinical translation of precision medicine.

Acknowledgments

We thank the study participants and referring technicians for their participation in this study.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-985/rc

Data Sharing Statement: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-985/dss

Peer Review File: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-985/prf

Funding: This work was supported by The Chinese National Key Research and Development Project (No. 2021YFC2500400 and No. 2021YFC2500402), National Natural Science Foundation of China (No. 82171932 and No. 82302180), Tianjin Key Medical Discipline (Specialty) Construction Project (No. TJYXZDXK-012A and No. TJYXZDXK-009A), and Zhangjiakou City Self-financing Project of the 2019 Scientific Research Plan (No. 1921131H).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-24-985/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all

aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of Tianjin Medical University Cancer Institute and Hospital (No. Ek2023044) and informed consent was waived from all the patients.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74:229-63.
2. Lin L, Li Z, Yan L, et al. Global, regional, and national cancer incidence and death for 29 cancer groups in 2019 and trends analysis of the global cancer burden, 1990-2019. J Hematol Oncol 2021;14:197.
3. Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. Nat Rev Clin Oncol 2023;20:624-39.
4. Nicholson AG, Tsao MS, Beasley MB, et al. The 2021 WHO Classification of Lung Tumors: Impact of Advances Since 2015. J Thorac Oncol 2022;17:362-87.
5. Groome PA, Bolejack V, Crowley JJ, et al. The IASLC Lung Cancer Staging Project: validation of the proposals for revision of the T, N, and M descriptors and consequent stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumours. J Thorac Oncol 2007;2:694-705.
6. Vansteenkiste J, Crinò L, Dooms C, et al. 2nd ESMO Consensus Conference on Lung Cancer: early-stage non-small-cell lung cancer consensus on diagnosis, treatment and follow-up. Ann Oncol 2014;25:1462-74.
7. Sihoe AD, Van Schil P. Non-small cell lung cancer: when to offer sublobar resection. Lung Cancer 2014;86:115-20.
8. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. N Engl J Med 2020;382:503-13.
9. Rankin NM, McWilliams A, Marshall HM. Lung cancer screening implementation: Complexities and priorities. Respirology 2020;25 Suppl 2:5-23.
10. Wolf AMD, Oeffinger KC, Shih TY, et al. Screening for lung cancer: 2023 guideline update from the American Cancer Society. CA Cancer J Clin 2024;74:50-81.
11. 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;399:1607-17.
12. Aokage K, Suzuki K, Saji H, et al. Segmentectomy for ground-glass-dominant lung cancer with a tumour diameter of 3 cm or less including ground-glass opacity (JCOG1211): a multicentre, single-arm, confirmatory, phase 3 trial. Lancet Respir Med 2023;11:540-9.
13. Altorki N, Wang X, Kozono D, et al. Lobar or Sublobar Resection for Peripheral Stage IA Non-Small-Cell Lung Cancer. N Engl J Med 2023;388:489-98.
14. Stinchcombe TE, Wang X, Damman B, et al. Secondary Analysis of the Rate of Second Primary Lung Cancer From Cancer and Leukemia Group B 140503 (Alliance) Trial of Lobar Versus Sublobar Resection for T1aN0 Non-Small-Cell Lung Cancer. J Clin Oncol 2024;42:1110-3.
15. Laville D, Désage AL, Fournel P, et al. Spread Through Air Spaces in Stage I to III Resected Lung Adenocarcinomas: Should the Presence of Spread Through Air Spaces Lead to an Upstaging? Am J Surg Pathol 2024;48:596-604.
16. Takahara T, Satou A, Tsuyuki T, et al. Histology of Bronchiolar Tumor Spread Through Air Spaces. Am J Surg Pathol 2024;48:1052-9.
17. 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;10:1243-60.
18. Mino-Kenudson M. Significance of tumor spread through air spaces (STAS) in lung cancer from the pathologist perspective. Transl Lung Cancer Res 2020;9:847-59.
19. Gutierrez-Sainz L, López-Muñoz S, Cruz-Castellanos P, et al. Retrospective analysis of the prognostic implications of tumor spread through air spaces in lung adenocarcinoma patients treated with surgery. ESMO Open 2022;7:100568.
20. Chen S, Ye T, Yang S, et al. Prognostic implication

of tumor spread through air spaces in patients with pathologic N0 lung adenocarcinoma. Lung Cancer 2022;164:33-8.
21. 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;16:619-29.
22. Zhou F, Villalba JA, Sayo TMS, et al. Assessment of the feasibility of frozen sections for the detection of spread through air spaces (STAS) in pulmonary adenocarcinoma. Mod Pathol 2022;35:210-7.
23. Cao H, Zheng Q, Deng C, et al. Prediction of Spread Through Air Spaces (STAS) By Intraoperative Frozen Section for Patients with cT1N0M0 Invasive Lung Adenocarcinoma: A Multi-Center Observational Study (ECTOP-1016). Ann Surg 2024. [Epub ahead of print]. doi: 10.1097/SLA.0000000000006525.
24. Quanyang W, Yao H, Sicong W, et al. Artificial intelligence in lung cancer screening: Detection, classification, prediction, and prognosis. Cancer Med 2024;13:e7140.
25. Ye G, Wu G, Qi Y, et al. Non-invasive multimodal CT deep learning biomarker to predict pathological complete response of non-small cell lung cancer following neoadjuvant immunochemotherapy: a multicenter study. J Immunother Cancer 2024;12:e009348.
26. Chen M, Copley SJ, Viola P, et al. Radiomics and artificial intelligence for precision medicine in lung cancer treatment. Semin Cancer Biol 2023;93:97-113.
27. Rami-Porta R, Nishimura KK, Giroux DJ, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groups in the Forthcoming (Ninth) Edition of the TNM Classification for Lung Cancer. J Thorac Oncol 2024;19:1007-27.
28. Moreira AL, Ocampo PSS, Xia Y, et al. A Grading System for Invasive Pulmonary Adenocarcinoma: A Proposal From the International Association for the Study of Lung Cancer Pathology Committee. J Thorac Oncol 2020;15:1599-610.
29. Wang S, Liu X, Jiang C, et al. CT-Based Super-Resolution Deep Learning Models with Attention Mechanisms for Predicting Spread Through Air Spaces of Solid or Part-Solid Lung Adenocarcinoma. Acad Radiol 2024;31:2601-9.
30. Liu C, Meng A, Xue XQ, et al. Prediction of early lung adenocarcinoma spread through air spaces by machine learning radiomics: a cross-center cohort study. Transl Lung Cancer Res 2024;13:3443-59.
31. Tasnim S, Raja S, Mukhopadhyay S, et al. Preoperative predictors of spread through air spaces in lung cancer. J Thorac Cardiovasc Surg 2024;168:660-669.e4.
32. Kim SK, Kim TJ, Chung MJ, et al. Lung Adenocarcinoma: CT Features Associated with Spread through Air Spaces. Radiology 2018;289:831-40.
33. Li H, Li L, Liu Y, et al. Predictive value of CT and (18) F-FDG PET/CT features on spread through air space in lung adenocarcinoma. BMC Cancer 2024;24:434.
34. Liu BC, Ma HY, Huang J, et al. Does dual-layer spectral detector CT provide added value in predicting spread through air spaces in lung adenocarcinoma? A preliminary study. Eur Radiol 2024;34:4176-86.
35. Qin L, Sun Y, Zhu R, et al. Clinicopathological and CT features of tumor spread through air space in invasive lung adenocarcinoma. Front Oncol 2022;12:959113.
36. 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;156:1670-1676.e4.
37. 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;12:6494-504.
38. Jiang C, Luo Y, Yuan J, et al. CT-based radiomics and machine learning to predict spread through air space in lung adenocarcinoma. Eur Radiol 2020;30:4050-7.
39. Liao G, Huang L, Wu S, et al. Preoperative CT-based peritumoral and tumoral radiomic features prediction for tumor spread through air spaces in clinical stage I lung adenocarcinoma. Lung Cancer 2022;163:87-95.
40. Lin MW, Chen LW, Yang SM, et al. CT-Based Deep-Learning Model for Spread-Through-Air-Spaces Prediction in Ground Glass-Predominant Lung Adenocarcinoma. Ann Surg Oncol 2024;31:1536-45.
41. Tao J, Liang C, Yin K, et al. 3D convolutional neural network model from contrast-enhanced CT to predict spread through air spaces in non-small cell lung cancer. Diagn Interv Imaging 2022;103:535-44.
42. Wang TW, Hong JS, Chiu HY, et al. Standalone deep learning versus experts for diagnosis lung cancer on chest computed tomography: a systematic review. Eur Radiol 2024;34:7397-407.
43. Dai C, Xie H, Su H, et al. Tumor Spread through Air Spaces Affects the Recurrence and Overall Survival in Patients with Lung Adenocarcinoma >2 to 3 cm. J Thorac Oncol 2017;12:1052-60.
44. Takehana K, Sakamoto R, Fujimoto K, et al. Peritumoral radiomics features on preoperative thin-slice CT images can predict the spread through air spaces of lung adenocarcinoma. Sci Rep 2022;12:10323.

45. Wang Y, Lyu D, Hu L, et al. CT-Based Intratumoral and Peritumoral Radiomics Nomograms for the Preoperative Prediction of Spread Through Air Spaces in Clinical Stage IA Non-small Cell Lung Cancer. J Imaging Inform Med 2024;37:520-35.

Cite this article as: Zhang JN, Li ZF, Zheng SY, Li JJ, Sun LN, Ye ZX. Deep learning model for predicting spread through air spaces of lung adenocarcinoma based on transfer learning mechanism. Transl Lung Cancer Res 2025;14(4):1061-1075. doi: 10.21037/tlcr-24-985