Journal of Cardiothoracic Surgery

RESEARCH

Open Access

Predicting spread through air space of lung adenocarcinoma based on deep learning and machine learning models

Zengming Wang[1], Lingxin Kong[1], Bin Li[1], Qingtao Zhao[1], Xiaopeng Zhang[1], Huanfen Zhao[2], Wenfei Xue[1], Wei Li[3], Shun Xu[4*] and Guochen Duan[1*]

Abstract

Objective The aim of this study was to develop a machine learning model that can predict spread through air space (STAS) of lung adenocarcinoma preoperatively. STAS is associated with poor prognosis in invasive lung adenocarcinoma. Therefore non-invasive and accurate pre-surgical prediction of STAS in patients with lung adenocarcinoma is essential for individualised patient management.

Methods We included 138 patients with invasive lung adenocarcinoma who underwent lobectomy, collected their preoperative imaging data and clinical features, built a model for predicting STAS using machine learning and deep learning methods, and validated the efficacy of the model. Finally a nomogram was created based on logistic regression (LR).

Results Imaging histology features showed good model efficacy in both the training set (LR AUC = 0.764) and the test set (LR AUC = 0.776), and we combined the imaging histology and clinical features to jointly build a nomogram graph (AUC = 0.878), extracted the deep learning features, and built a machine learning model based on the ResNET50 algorithm, where the LR AUC = 0.918.

Conclusions This presented radiomics model can be served as a non-invasive for predicting STAS in Infiltrating lung adenocarcinoma.

Introduction

Lung adenocarcinoma (LUAD) is the most common type of pathology and has a high incidence [1]. In general, tumour spread is an important factor in determining the prognosis of almost all malignancies. The concept of STAS was formally incorporated into the modes of invasion of lung adenocarcinoma in by the World Health Organization (WHO) in 2015, and STAS is defined as the spread of single or clustered cancer cells in the air spaces outside the main body of the tumour [2]. STAS is an independent influence on the poor prognosis of lung cancer patients. It is associated with both recurrence and survival of patients after surgery [2, 3].

*Correspondence:
Shun Xu
xushuncmu1h@sina.com
Guochen Duan
duanguoc@126.com
[1]Department of Thoracic Surgery, Hebei General Hospital, NO. 348 Heping West Road, Xinhua District, Shijiazhuang, Hebei 050000, P.R. China
[2]Department of Pathology, Hebei General Hospital, Shijiazhuang, Hebei 050000, China
[3]Department of Geriatric Cardiovascular, Hebei General Hospital, Shijiazhuang, Hebei 050000, China
[4]Department of Thoracic Surgery, The First Affiliated Hospital of China Medical University, No. 155 Nanjing North Street, Shenyang, Liaoning 110000, China

Previous researchers described 3 morphologic patterns of STAS: micropapillary or ring-like clusters; solid nests or tumor islands; single cells [4]. STAS for microscopic diagnosis of lung cancer has been receiving attention from pathologists. Both preoperative fibreoptic bronchoscopy and percutaneous lung aspiration biopsy have shown low sensitivity for diagnosing the presence of STAS. Low sensitivity means high false positives, which can lead to misjudgment of the patient's condition by the clinician and delays in treatment. Intraoperative frozen pathology is also often difficult for pathologists to determine accurately due to time constraints [5–7]. Previous research has shown that the size of the lesion and the size and proportion of the solid component correlate with STAS [8]. The observation of CT features is highly subjective and relies on the experience of the radiologist. Imagingomics transforms regions of interest (ROI) in medical images into feature space data with high resolution and can be mined by using automated data characterisation algorithms, and ultimately transformed into quantitative data that can be used to characterise the ROI. Therefore, in this study, features of radiomic of CT and 2D deep learning (DL) features are extracted for modelling.

In clinical practice, more and more lung cancer patients with smaller tumours (< 2 cm) are opting for sublobar resection as a surgical procedure. However, in patients with STAS-positive stage I lung ADC, the recurrence risk of sublobar resection is higher than that of lobectomy [9]. If we can predict STAS before the patient's surgery, we can specify the surgical procedure more precisely, which will benefit the patient. In this study, we collected preoperative imaging and clinical data of postoperative pathologically confirmed invasive adenocarcinoma, constructed the STAS model and validated it.

Methods

Patient screening and collection of characteristics

138 patients who attended the Department of Thoracic Surgery of Hebei Provincial People's Hospital between July 2022 and May 2023 were retrospectively included in this study. Inclusion criteria: 1). All patients underwent lobectomy and standardised lymph node dissection, and postoperative pathology confirmed stage I-IIIA invasive lung adenocarcinoma; 2). Patients had complete preoperative imaging and clinical data and were ≥ 18 years old at the time of surgery. 3). The patient did not receive adjuvant therapy. Exclusion criteria: 1). Poor quality of CT images and artefacts affecting the experimental process can result from inadequate breath holding or excessive respiratory activity during the patient's CT examination. patients with incomplete or unclear preoperative chest CT images; 2). patients that did not undergo surgery due to their own or various reasons.

This retrospective study was approved by the Ethics Committee of the Hebei Provincial General Hospital, which waived informed consent (NO.2023128). The study was conducted in accordance with the Declaration of Helsinki to ensure compliance with ethical standards and patient confidentiality by removing all identifying information from the data.

Image processing and feature extraction

All chest plain CT images in this study were done by Siemens dual source thin layer CT scanning machine with fixed interlayer spacing. The CT images were used to outline the region of interest (ROI) and cross-checked against each other by two Chinese Academy of Management Sciences (CMC) certified imaging technology engineers using ITK-SNAP version 3.8 software. A rigorous standardised methodology was used for all processes including identification of pathology sections.

In the field of image processing, the ROI is an image region selected from an image that is the focus of your image analysis. The region is circled for further processing. Using the ROI to circle the target you want to read reduces processing time and increases accuracy. The ROI region for this study is the lesion area in the patient's chest CT image [10]. Using Pyradiomics, image paths and mask paths are specified to perform feature extraction. Geometric, intensity and texture features together make up the manual features. The geometric features are used to describe the 3D shape of the tumor and the intensity features are used to describe the distribution of voxel intensities within the tumor. Texture features are extracted based on intensity distributions, including second-order and higher-order distributions. The texture features include grey level co-occurrence matrix (GLCM), grey level run length matrix (GLRLM), grey level size zone matrix (GLSZM) and neighbourhood grey-tone difference matrix (NGTDM).

Baseline characteristics

Statistical methods such as Mann-Whitney U test, t-test, and Chi-square test were applied for data analysis of

patients' clinical baseline characteristics. Table 1 summarises the baseline characteristics of the included population, which included gender, age, physical baseline and preoperative pulmonary function. The aim was to assess the correlation between the clinical characteristics and the study results to ensure that the outcomes were valid and reliable. We divided the inclusion population into training and validation sets as we went along in a 4:1 ratio. The grouping scheme was completely randomised.

Table 1 Laboratory and clinical features

	Training Cohort				Testing Cohort			
Feature	**All**	**Non-STAS**	**STAS**	***P*-value**	**All**	**Non-STAS**	**STAS**	***P*-value**
Age (years)	62.20 ± 8.85	62.22 ± 8.64	62.15 ± 9.61	0.972	60.43 ± 8.23	60.00 ± 8.83	61.71 ± 6.47	0.69
BMI (kg/m^2)	26.07 ± 3.36	25.74 ± 3.45	27.08 ± 2.93	0.061	25.28 ± 3.74	25.91 ± 3.02	23.37 ± 5.18	0.06
CEA (ng/ml)	8.24 ± 22.72	4.72 ± 6.13	19.08 ± 43.40	0.361	11.10 ± 26.19	2.47 ± 1.06	36.96 ± 45.19	< 0.001
NSE (ng/ml)	13.45 ± 4.92	13.65 ± 5.36	12.84 ± 3.21	0.583	14.78 ± 3.45	14.62 ± 3.32	15.25 ± 4.07	0.533
CK19 (ng/ml)	2.63 ± 1.51	2.48 ± 1.26	3.10 ± 2.08	0.239	2.85 ± 1.39	3.04 ± 1.49	2.28 ± 0.90	0.339
Size (cm)	2.07 ± 1.04	2.11 ± 1.10	1.95 ± 0.85	0.581	1.96 ± 1.06	2.00 ± 1.08	1.84 ± 1.07	0.75
FVC (ml)	104.79 ± 13.97	104.06 ± 13.01	107.04 ± 16.66	0.629	110.36 ± 12.43	112.29 ± 12.67	104.57 ± 10.37	0.123
FEV1	101.93 ± 16.61	100.14 ± 15.75	107.41 ± 18.23	0.097	104.14 ± 14.46	106.86 ± 13.76	96.00 ± 14.39	0.093
FEV1/FVC (%)	97.04 ± 8.59	95.71 ± 7.91	101.11 ± 9.45	0.014	94.25 ± 5.76	94.81 ± 5.71	92.57 ± 6.05	0.409
MVV	86.54 ± 15.20	85.85 ± 14.63	88.67 ± 16.96	0.581	84.57 ± 23.00	84.52 ± 23.57	84.71 ± 23.00	1.0
DLCO	90.77 ± 17.71	89.73 ± 18.32	93.96 ± 15.57	0.23	86.75 ± 14.99	88.00 ± 12.44	83.00 ± 21.76	0.873
VA	94.89 ± 11.63	95.52 ± 11.68	92.96 ± 11.48	0.124	98.43 ± 9.70	99.19 ± 8.93	96.14 ± 12.23	0.41
DL/VA (%)	101.77 ± 19.62	100.35 ± 20.06	106.15 ± 17.83	0.154	93.07 ± 18.05	93.86 ± 13.51	90.71 ± 29.14	0.915
Gender				0.15				0.776
Female	56(50.91)	46(55.42)	10(37.04)		23(82.14)	18(85.71)	5(71.43)	
Male	54(49.09)	37(44.58)	17(62.96)		5(17.86)	3(14.29)	2(28.57)	
ECOGPS				0.905				0.298
0	8(7.27)	6(7.23)	2(7.41)		5(17.86)	5(23.81)	null	
1	65(59.09)	50(60.24)	15(55.56)		12(42.86)	9(42.86)	3(42.86)	
2	37(33.64)	27(32.53)	10(37.04)		11(39.29)	7(33.33)	4(57.14)	
Hypertension				0.185				0.281
No	55(50.00)	45(54.22)	10(37.04)		16(57.14)	13(61.90)	3(42.86)	
Grade 1	12(10.91)	7(8.43)	5(18.52)		3(10.71)	1(4.76)	2(28.57)	
Grade 2	19(17.27)	12(14.46)	7(25.93)		2(7.14)	2(9.52)	null	
Grade 3	24(21.82)	19(22.89)	5(18.52)		7(25.00)	5(23.81)	2(28.57)	
Diabetes				1.0				0.533
No	95(86.36)	72(86.75)	23(85.19)		24(85.71)	17(80.95)	7(100.00)	
Yes	15(13.64)	11(13.25)	4(14.81)		4(14.29)	4(19.05)	null	
CHD				0.231				0.533
No	81(73.64)	64(77.11)	17(62.96)		24(85.71)	17(80.95)	7(100.00)	
Yes	29(26.36)	19(22.89)	10(37.04)		4(14.29)	4(19.05)	null	
Smoking				0.299				0.29
No	79(71.82)	57(68.67)	22(81.48)		25(89.29)	20(95.24)	5(71.43)	
Yes	31(28.18)	26(31.33)	5(18.52)		3(10.71)	1(4.76)	2(28.57)	
c-Stage				0.481				0.191
1	81(73.64)	61(73.49)	20(74.07)		20(71.43)	16(76.19)	4(57.14)	
2	25(22.73)	18(21.69)	7(25.93)		7(25.00)	5(23.81)	2(28.57)	
3	4(3.64)	4(4.82)	null		1(3.57)	null	1(14.29)	
Position				0.106				0.469
LUL	32(29.09)	21(25.30)	11(40.74)		8(28.57)	6(28.57)	2(28.57)	
LLL	16(14.55)	11(13.25)	5(18.52)		6(21.43)	6(28.57)	null	
RUL	27(24.55)	25(30.12)	2(7.41)		6(21.43)	4(19.05)	2(28.57)	
RML	14(12.73)	9(10.84)	5(18.52)		4(14.29)	2(9.52)	2(28.57)	
RLL	21(19.09)	17(20.48)	4(14.81)		4(14.29)	3(14.29)	1(14.29)	

Abbreviations: CEA, Carcinoembryonic Antigen; NSE, neuron-specific enolase; CK19,Cytokeratin 19; Size, tumour size; BMI, body mass index; ECOG PS, Eastern Cooperative Oncology Group Performance status; CHD Coronary Heart Disease; LUL, Left Upper Lobe; LLL, Left Lower Lobe; RUL, Right Upper Lobe; RML, Right Middle Lobe; RLL, Right Lower Lobe

Four-fifths of them are the training set and the rest are the test set.

Feature selection

All extracted radiological features were screened by Mann-Whitney U test. Features with p-value less than 0.05 were finally included. For processing details, we calculated the correlation coefficient between all features two by two by Spearman's rank correlation coefficient. If the correlation coefficient was greater than 0.9, this one feature was retained. In order to determine the ability of good transformation between features and data, feature screening was done using a greedy recursive deletion strategy [11]. The features with the highest redundancy within the set in each iteration were eliminated. A total of 107 features were finally included. Recursive Feature Elimination (RFE) is a feature selection method based on the predictive performance of a model. It evaluates the importance of each feature by recursively removing features and retraining the model. In each iteration, it removes the least important features and retrains the model.

The feature data were modelled by LASSO regression. The regression coefficients were reduced to zero based on the moderating weights λ, and the coefficients of features with no correlation were set to 0. Fifth-fold cross-validation was used to determine the optimal λ, a minimum baseline was set, and the final value of λ was determined based on the minimum cross-validation error. The final non-zero coefficients that were filtered out were used to fit the regression model that comprised the imaging characteristics of this study. Plotting feature weights to visually analyse the relevance of different features for the occurrence of STAS. The LASSO regression model was implemented using the Python scikit-learn package.

Radiomics models

A total of 8 machine learning models, LR, SVM, KNN, Random Forest, ExtraTrees, XGBoost, LightGBM and MLP, were built using radiomics features after dimensionality reduction by Lasso regression, and their corresponding ROC curves were plotted. A five-fold cross-validation method was applied to draw the box plots of the effectiveness of each model. The key radiomics features of each model were analysed to produce bar charts to determine the importance of different features in the models.

Clinical models

The construction of clinical models was similar to the process used for radiomics characterisation models. Univariate and multivariate regression models were applied to this study to analyse the clinical features associated with the occurrence of STAS. In addition, the same machine learning algorithms as used to construct the radiomics models were applied, where features with a p-value threshold of less than 0.05 were included in the construction of the models. 5-fold cross-validation and test cohorts were kept constant to eliminate bias.

Nomogram

A nomogram for predicting STAS was created by combining the screened radiomics features and clinical features. The ROC curve of the nomogram was plotted, AUC was calculated and compared with a single radiomics model and a clinical features model. The calibration curve and Hosmer-Lemeshow analysis were used to assess the calibration efficiency of the nomogram. Calibration curves and Hosmer-Lemeshow analysis were used to assess the calibration efficiency of the nomogram. Decision Curve Analysis (DCA) was used to further assess model efficacy. All of these evaluation tools were performed on the training set species.

Radiomics deep learning features

We intercepted the maximum diameter plane after outlining from all CT image samples. 2D deep learning features were extracted from this plane based on the ResNet50 model algorithm in convolutional neural network. The application of downsampling completes the construction of the feature extraction network, which potentially serves to enhance the transform invariance of the features and to reduce the feature parameters to prevent overfitting, as evidenced by the gradual reduction of the feature image size and the gradual increase in the number of channels. 49 convolutional layers and one fully connected layer were included in the ResNet50 network. t-SNE was used to perform visualisation of the features. Lasso and machine learning models are applied to construct the radiomics deep learning model. Each node of the fully connected layer is connected to all the nodes of the previous layer, thus combining the previously extracted features, where Layer4 is the output layer.

Results

Baseline population characteristics

Table 1 describes the clinical characteristics of our included sample. A total of 138 patients were included in the study, out of which a total of 110 patients were listed in the training group and the rest were classified to the test group. A total of 56 females (50.91%) were in the training group and 23 (82.14%) in the test group. 34 patients had postoperative pathological findings suggestive of STAS. The pulmonary function profile of the patients FEV1/FVC showed some correlation in the training group ($p < 0.05$). CEA showed correlation in the test group ($p < 0.001$).

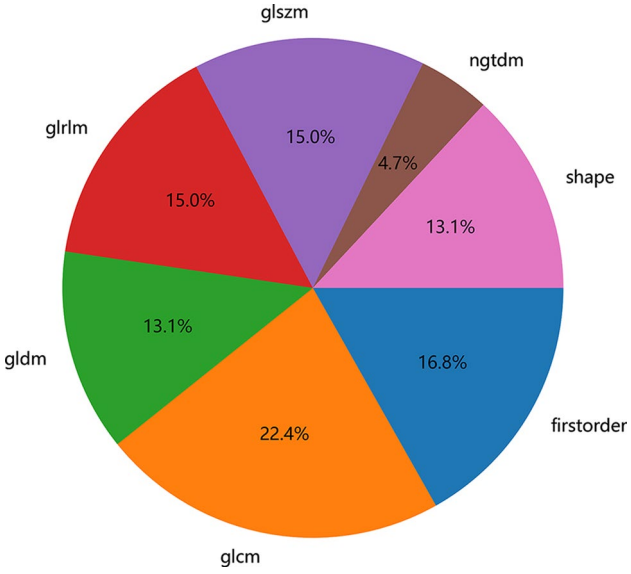

Fig. 1 Percentage of handcrafted features extracted based on CT images. GLCM, gray-level co-occur- rence matrix; GLDM, gray level dependence matrix; GLRLM, gray-level run length matrix; GLSZM, gray-level size zone matrix; NGTDM, neighborhood gray-tone difference matrix

Radiomics and clinical characteristics

A total of 10 categories radiomics features were extracted. The percentage of different radiomics features is shown in Fig. 1. Figure 2 shows the distribution of radiomics features. Among them, GLCM (Gray - level co - occurrence matrix) has the highest percentage of 22.4%. NGTDM (Neighbourhood grey - level difference matrix) has the lowest percentage of 4.7%. Figure 3 shows the result of the image after visualisation of all the features.Lasso regression result is shown in Fig. 4 with $\lambda = 0.0036$. Figure 5 shows the mean standard error (MSE) of the regression. Figure 6 reacts to the weights of different features. original_gldm_LargeDependenceLowGrayLevelEmphasis and original_firstorder_Skewness have Coefficients greater than 0.1. Figure 7a and b show the results of regression after applying the machine learning algorithm on the case of the models built in the training and test sets. Table 2 shows the details of the models. The models in the training set, such as LR (AUC = 0.764), SVM (AUC = 0.903) and KNN (AUC = 0.827), show good model performance, while RandomForest, ExtraTrees and XGBoost show overfitting. The

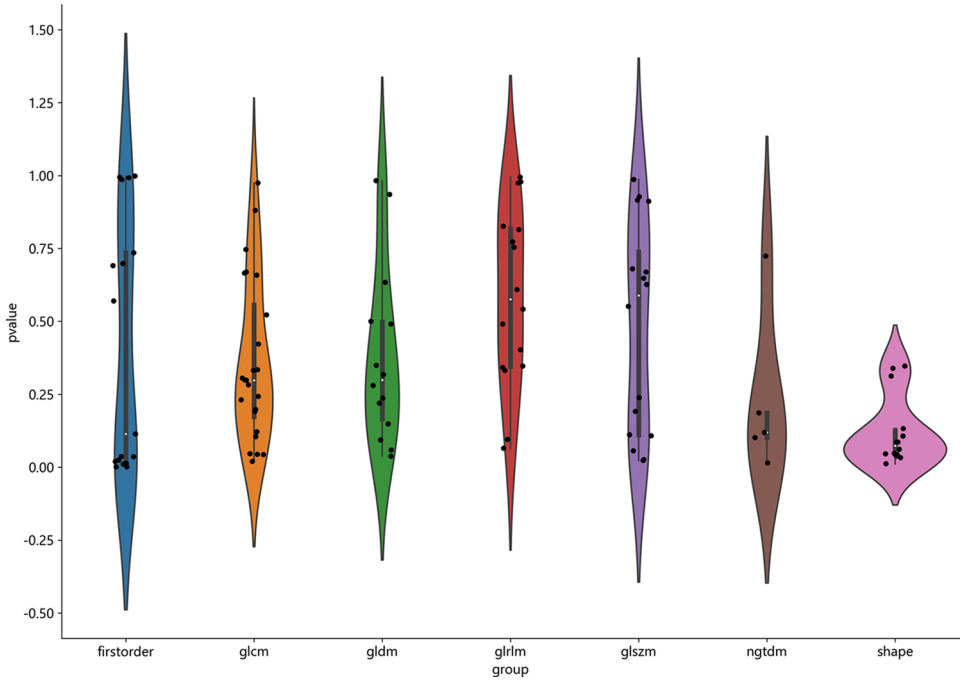

Fig. 2 Distribution diagram of manual features. GLCM, gray-level co-occur- rence Matrix; GLDM, gray level dependence matrix; GLRLM, gray-level run length matrix; GLSZM, gray-level size zone matrix; NGTDM, neighborhood gray-tone difference matrix

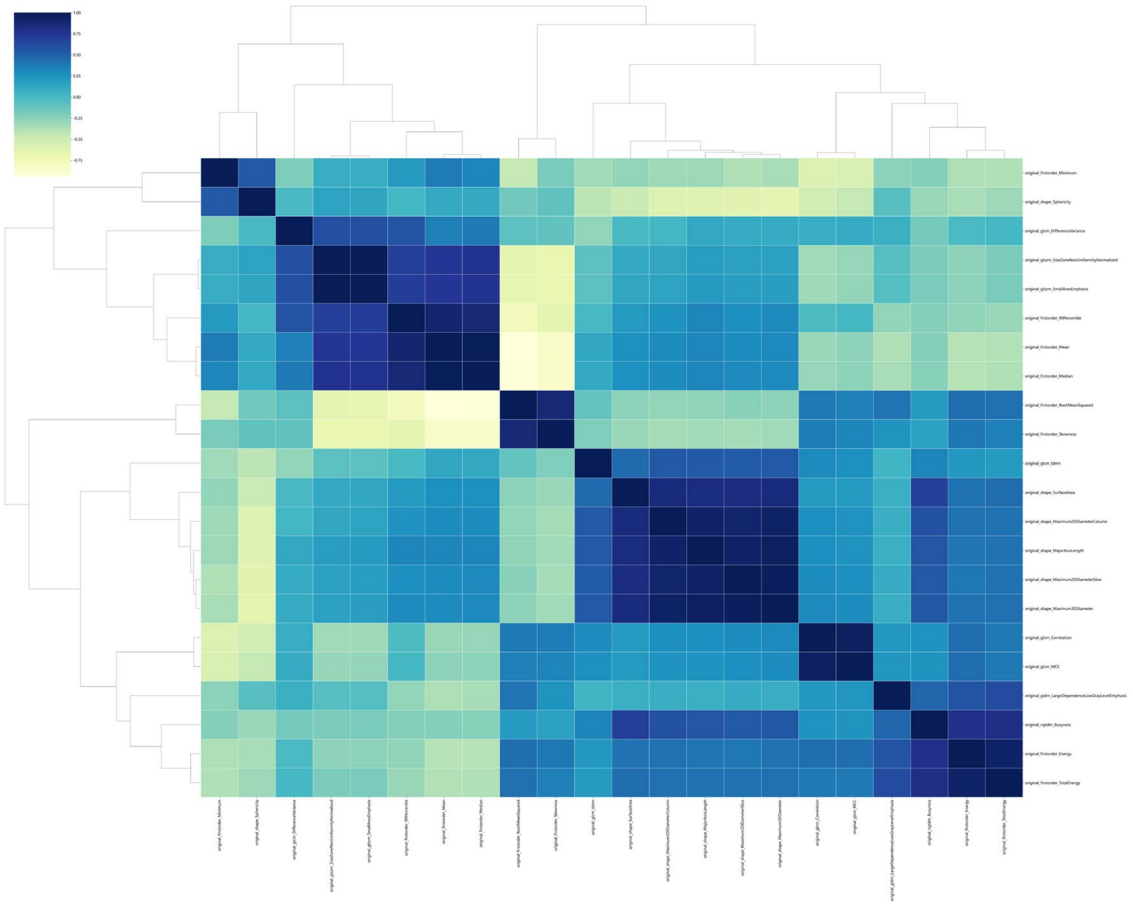

Fig. 3 Visualization matrix for different features

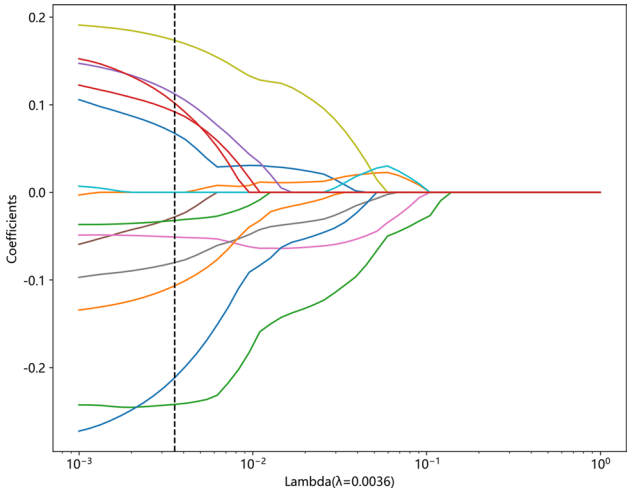

Fig. 4 Lasso regression curves for handcrafted features. Different colors represent different image features before dimensionality reduction

AUC of the ROC curves for the RF, ET and XGB models is 1. All models in the test set showed good model efficacy, with the LR model having an AUC = 0.764. Figure 8 illustrates the five-fold stochastic cross-validation of the models. Figure 9 shows the sample prediction histogram for the LR model. The calculation formula is listed as below.

$$
\begin{aligned}
&label \\
&= 0.23918625380490918 + 0.088333 \\
&* original\_firstorder\_90Percentile - 0.261142 \\
&* original\_firstorder\_RootMeanSquared \\
&+ 0.148982 * original\_firstorder\_Skewness \\
&+ 0.074416 * original\_firstorder\_TotalEnergy \\
&- 0.039804 * original\_glcm\_Difference \\
&Variance - 0.034953 * original\_glcm\_Idmn \\
&- 0.075475 * original\_glcm\_MCC + 0.174171 \\
&* original\_gldm\_LargeDependenceLowGray \\
&LevelEmphasis - 0.171767 \\
&* original\_ngtdm\_Busyness - 0.077408 \\
&* original\_shape\_Maximum3DDiameter \\
&- 0.006219 * original\_shape\_Sphericity \\
&+ 0.084126 * original\_shape\_SurfaceArea
\end{aligned}
$$

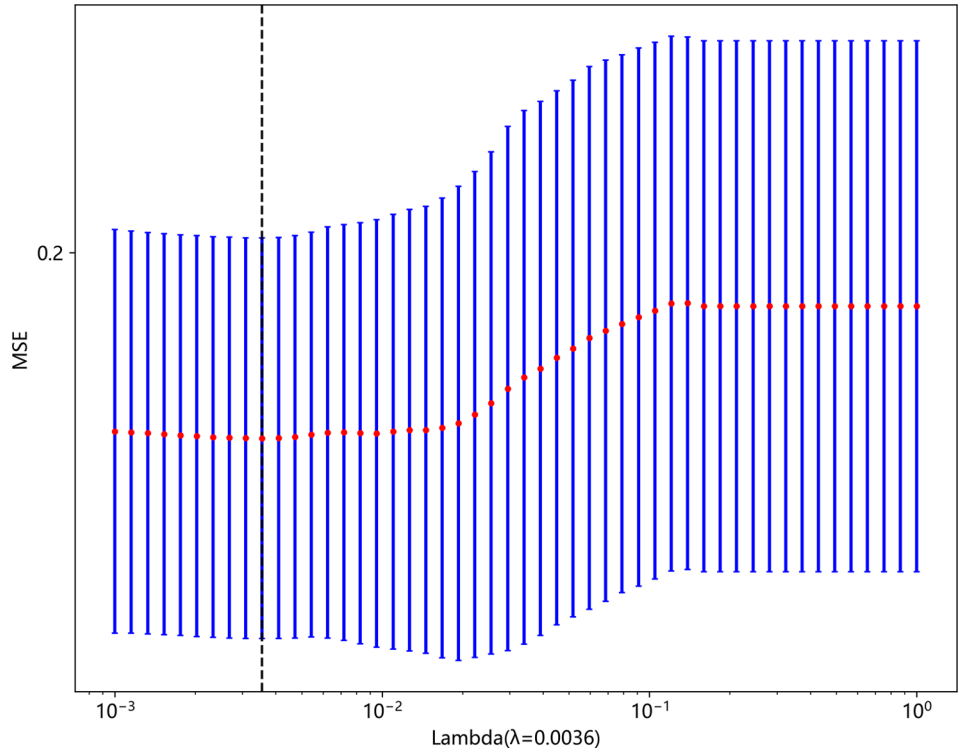

Fig. 5 MSE of handcrafted features. MSE, mean standard error

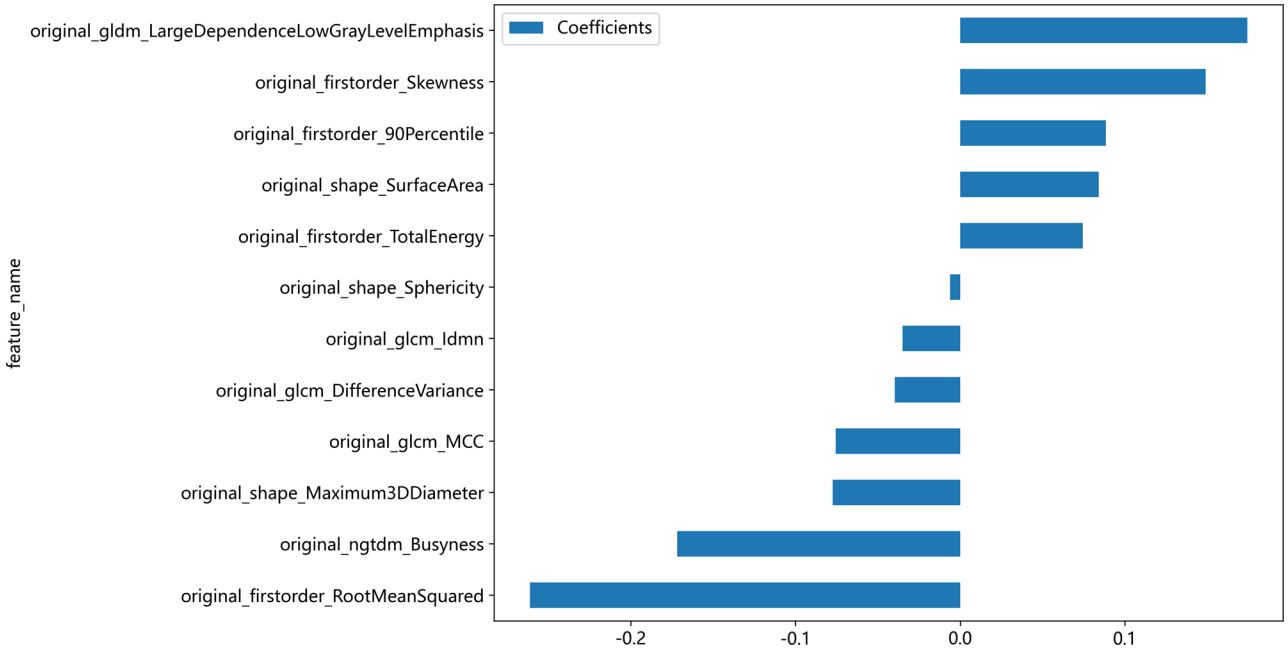

Fig. 6 Feature weight graph for filtered features

Table 3 shows the logistic regression univariate and multivariate analyses of clinical characteristics and whether STAS occurred. Both CEA and FEV1/FVC were significantly associated with STAS ($p \leq 0.001$). We also visualize this (Fig. 9a and b).

Nomogram

We performed front-end fusion with clinical data and radiomics data to build 1 nomogram in the training set, as shown in Fig. 10. It can effectively guide surgeons to evaluate STAS in patients preoperatively. Figure 11 shows the comparison of the efficacy of the nomograms,

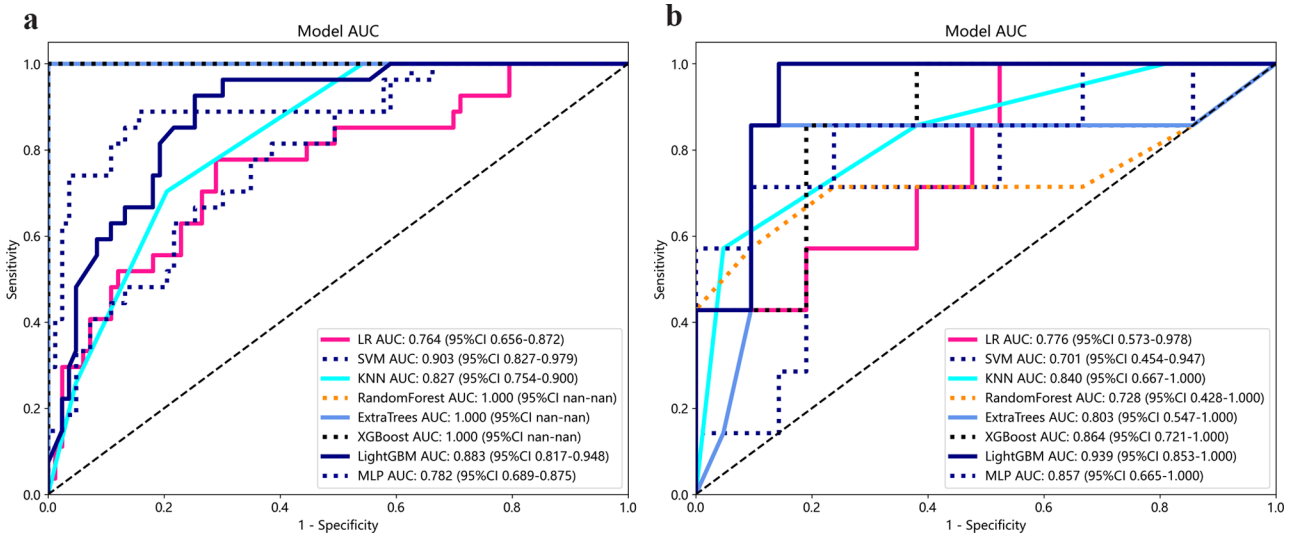

Fig. 7 (**a**) ROC curves for different machine learning models in the training set. (**b**) ROC curves for different machine learning models in the testing set

in which the AUC of clinical combined with imaging = 0.878. ROC suggests that the AUC of a combined nomogram is greater than that of a single nomogram. Detailed data in Table 4. To ensure the accuracy of the model, we built Decision Curve Analysis (DCA) and calibrated curves as shown in Figs. 12 and 13. The net benefit is the value of the benefit that balances the predicted return and risk through a threshold probability. Threshold Probability is in the range of about 0.2 to 0.4, and the composite evaluation models all have higher net benefit ratios than the simple model. Points is the score reference for each variable. clinical_Sig is the clinical score. rad_Sig is the radiological score. The final total score is derived, and a plumb line downwards gives the risk of STAS in patients with lung adenocarcinoma.

Deep learning models

Figure 14 shows the visualised image of the deep learning model implemented by the t-SNE method. Label 1 represents the presence of airway dissemination in lung adenocarcinoma. label 0 represents the absence. We performed Lasso regression and MSE on these deep learning features with $\lambda = 0.0295$, as shown in Figs. 15 and 16. The downscaled feature data were modelled for correlation with STAS, as shown in Figs. 17 and 18 where each model showed excellent model efficacy.LR AUC = 0.918, which is significantly better than the efficacy of traditional imaging histology feature modelling.

Discussion

In this study, we developed and validated a tumour-derived radiomics model based on chest CT images for pre-surgical prediction of STAS profiles in lung adenocarcinoma. We incorporated a deep learning approach while fusing clinical data with radiomics data. The predictive performance of the model is excellent compared with general regression models, which can help thoracic surgeons to select the most suitable treatment for patients before surgery.

Spread through air spaces (STAS) is an aggressive behavior of tumors. This aggressive behavior was first identified in lung adenocarcinoma. In subsequent studies, STAS was also found in squamous cell carcinoma and endocrine-related tumors [12–14]. The occurrence of STAS is thought to be associated with advanced age, male patients, serum carcinoembryonic antigen levels and larger tumor size [15]. Several studies have confirmed that the occurrence of STAS is closely associated with poor prognosis in lung cancer patients. This observation has been made in samples of lung adenocarcinoma with different stages [16, 17]. Although STAS is now considered an independent prognostic factor, there are no clinical data that consent the surgeon and the oncologist to classify the tumour with a higher stage. STAS is an independent prognostic factor for both recurrence-free survival (RFS) and overall survival (OS). If we can predict in advance whether STAS will occur or not, this benefits the patient's long-term prognosis and survival.

Table 2 Radiomics model features

Model	Task	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Precision	Recall	F1	Threshold
LR	Training	0.727	0.764	0.6563–0.8716	0.778	0.711	0.467	0.908	0.467	0.778	0.583	0.256
LR	Testing	0.607	0.776	0.5735–0.9776	1.000	0.476	0.389	1.000	0.389	1.000	0.560	0.190
SVM	Training	0.855	0.903	0.8274–0.9789	0.889	0.843	0.649	0.959	0.649	0.889	0.750	0.217
SVM	Testing	0.786	0.701	0.4539–0.9475	0.714	0.810	0.556	0.895	0.556	0.714	0.625	0.251
KNN	Training	0.773	0.827	0.7535–0.9002	0.704	0.835	0.528	0.892	0.528	0.704	0.603	0.400
KNN	Testing	0.857	0.840	0.6668–1.0000	0.571	1.000	0.800	0.870	0.800	0.571	0.667	0.600
RandomForest	Training	1.000	1.000	1.0000–1.0000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.500
RandomForest	Testing	0.750	0.728	0.4284–1.0000	0.714	0.762	0.500	0.889	0.500	0.714	0.588	0.400
ExtraTrees	Training	1.000	1.000	1.0000–1.0000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ExtraTrees	Testing	0.893	0.803	0.5465–1.0000	0.857	0.950	0.750	0.950	0.750	0.857	0.800	0.400
XGBoost	Training	1.000	1.000	1.0000–1.0000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.495
XGBoost	Testing	0.821	0.864	0.7207–1.0000	0.857	0.810	0.600	0.944	0.600	0.857	0.706	0.286
LightGBM	Training	0.791	0.883	0.8174–0.9479	0.926	0.747	0.543	0.969	0.543	0.926	0.685	0.263
LightGBM	Testing	0.893	0.939	0.8528–1.0000	1.000	0.857	0.700	1.000	0.700	1.000	0.824	0.304
MLP	Training	0.664	0.782	0.6886–0.8750	0.815	0.614	0.407	0.911	0.407	0.815	0.543	0.256
MLP	Testing	0.857	0.857	0.6649–1.0000	0.714	0.905	0.714	0.905	0.714	0.714	0.714	0.308

Abbreviations: AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; F1 = 2 x (precision x recall)/(precision + recall); LR, logistic regression; SVM, support vector machine; KNN, K-nearest neighbors; XG Boost, extreme gradient boosting; Light GBM, light gradient boosting machine; MLP, multi-layer perceptron

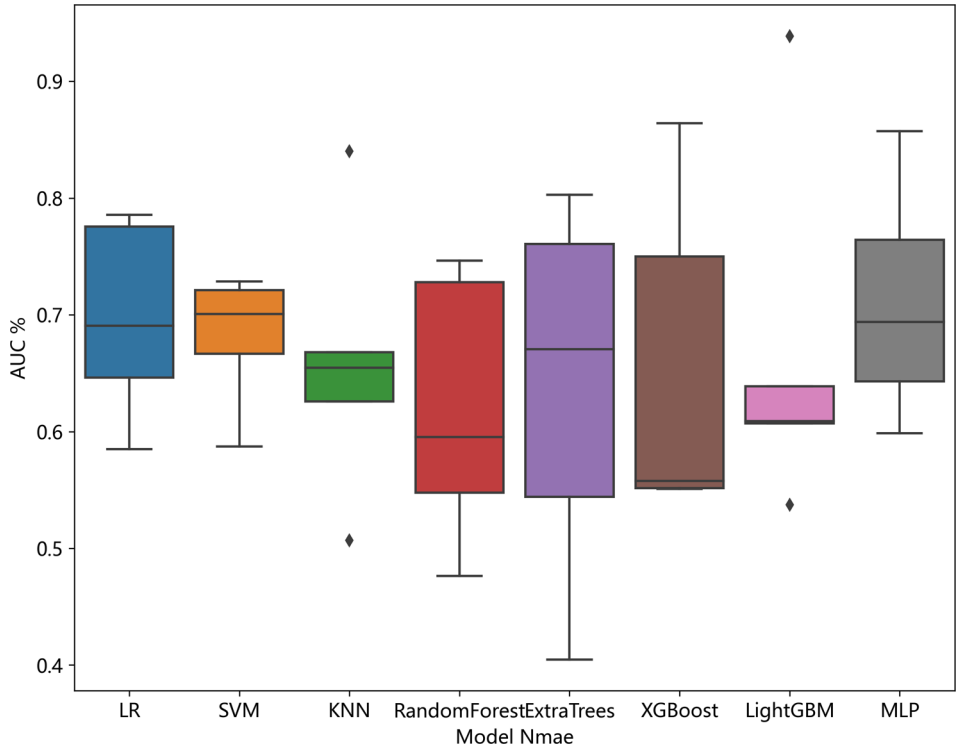

Fig. 8 Fifty-fold randomized cross-validation box plots for different machine learning models

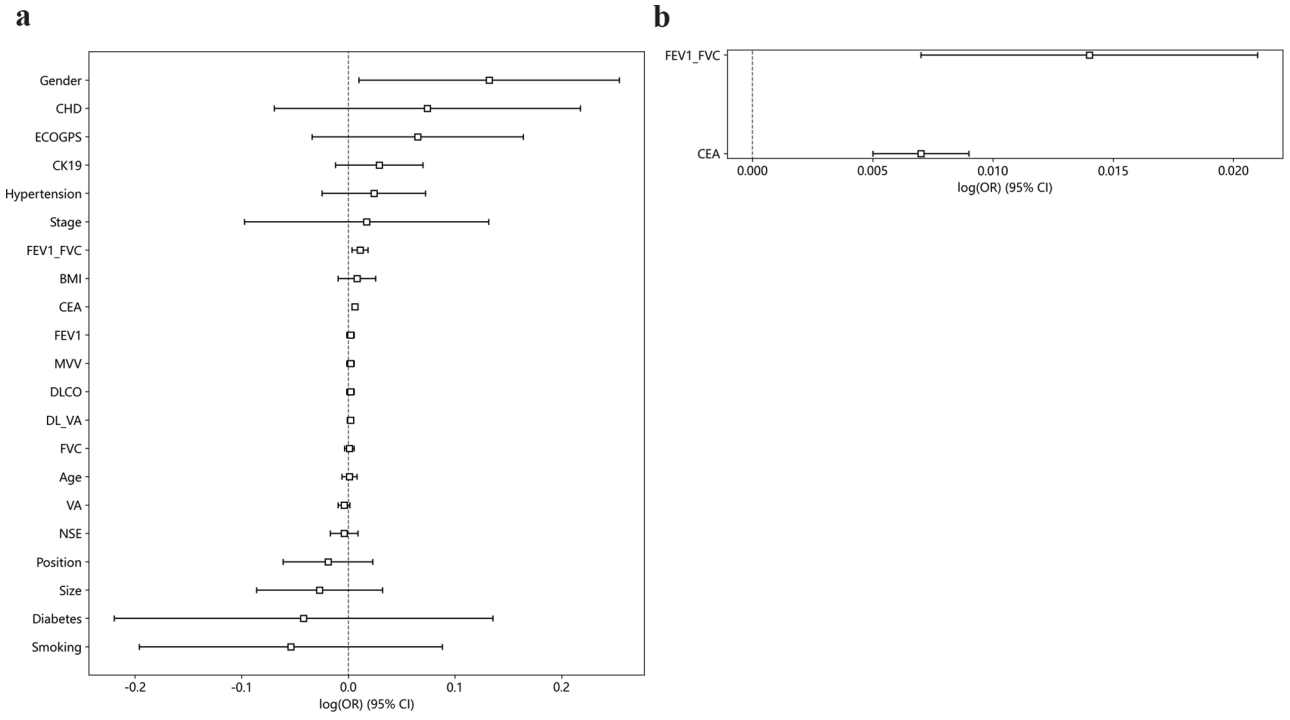

Fig. 9 (**a**) Univariate analysis of clinical and laboratory data. (**b**) Multifactorial analysis of clinical and laboratory data

Table 3 Univariate and multivariate analyses of clinical characteristics

Feature	Log(OR)	lower 95%CI	upper 95%CI	OR	OR lower 95%CI	OR upper 95%CI	*P*-value
Univariate analyses							
Gender	0.132	0.010	0.254	1.141	1.010	1.289	0.076
Age	0.001	-0.006	0.008	1.001	0.994	1.008	0.867
BMI	0.008	-0.009	0.026	1.009	0.991	1.026	0.432
ECOGPS	0.065	-0.034	0.164	1.067	0.967	1.178	0.281
Hypertension	0.024	-0.025	0.072	1.024	0.975	1.075	0.428
Diabetes	-0.042	-0.219	0.136	0.959	0.803	1.146	0.699
CHD	0.074	-0.069	0.218	1.077	0.933	1.244	0.390
Smoking	-0.054	-0.196	0.088	0.948	0.822	1.092	0.531
CEA	0.006	0.004	0.009	1.006	1.004	1.009	0.000
NSE	-0.004	-0.017	0.009	0.996	0.983	1.009	0.583
CK19	0.029	-0.012	0.070	1.030	0.988	1.073	0.242
c-Stage	0.017	-0.098	0.131	1.017	0.907	1.140	0.811
Size	-0.027	-0.086	0.032	0.973	0.918	1.033	0.446
Position	-0.019	-0.061	0.023	0.981	0.941	1.023	0.448
FVC	0.001	-0.004	0.005	1.001	0.996	1.005	0.768
FEV1	0.002	-0.001	0.006	1.003	0.999	1.006	0.267
FEV1/FVC	0.011	0.003	0.018	1.011	1.003	1.018	0.017
MVV	0.002	-0.002	0.005	1.002	0.998	1.005	0.500
DLCO	0.002	-0.002	0.005	1.002	0.998	1.005	0.497
VA	-0.004	-0.009	0.002	0.996	0.991	1.002	0.239
DL/VA	0.002	-0.001	0.005	1.002	0.999	1.005	0.311
Multivariate analyses							
CEA	0.007	0.005	0.009	1.007	1.005	1.009	0.000
FEV1/FVC	0.014	0.007	0.021	1.014	1.007	1.021	0.001

Abbreviations: CEA, Carcinoembryonic Antigen; NSE, neuron-specific enolase; CK19,Cytokeratin 19; Size, tumour size; BMI, body mass index; ECOG PS, Eastern Cooperative Oncology Group Performance status; CHD Coronary Heart Disease

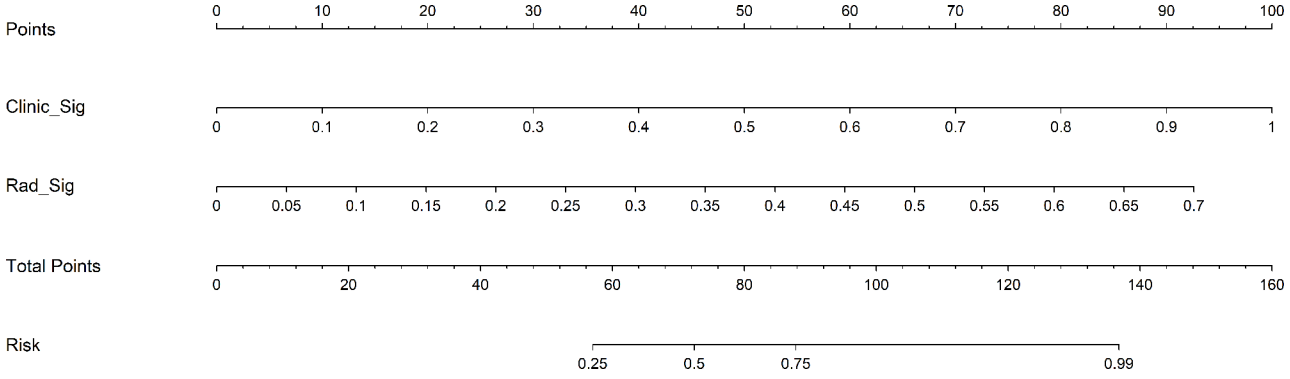

Fig. 10 A nomogram constructed on the basis of clinical features and imaging histologic features

Eguchi T et al. found that in patients with lung adenocarcinoma who developed STAS, undergoing lobectomy had a better outcome than sublobar resection [18]. Therefore, accurate preoperative prediction of STAS status is important for the selection of surgical options for lung cancer patients. Avoiding some STAS lung cancer patients with small tumor sizes will face higher recurrence and metastasis after undergoing local lung resection. With the preoperative prediction of the model, high-risk STAS patients should directly choose lobectomy to replace the currently common clinical practice of wedge resection of the lesion followed by waiting for intraoperative frozen pathology results. It has been shown that the sensitivity of frozen pathological sections to detect the occurrence of STAS is only 59%-86% [7]. This reduces the probability of intracavitary tumor dissemination on the one hand. On the other hand, it also shortens the operation time and reduces the injury of unnecessary anaesthetic drugs and prolonged tracheal intubation.

It has been shown that the diameter of the tumor and the diameter and percentage of the solid component

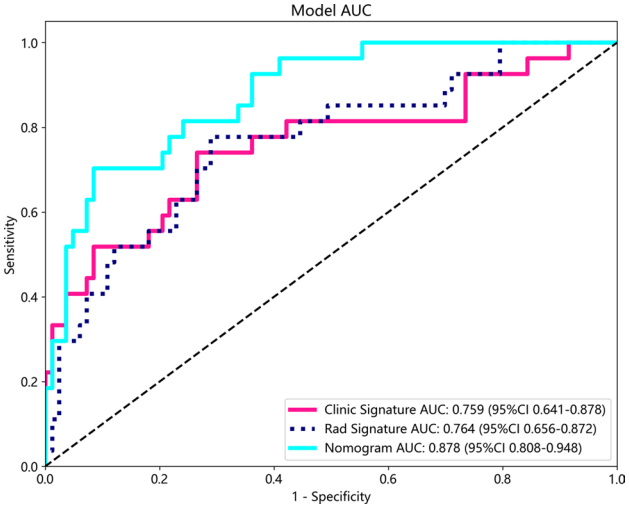

Fig. 11 ROC of the different components of the nomogram in the training set

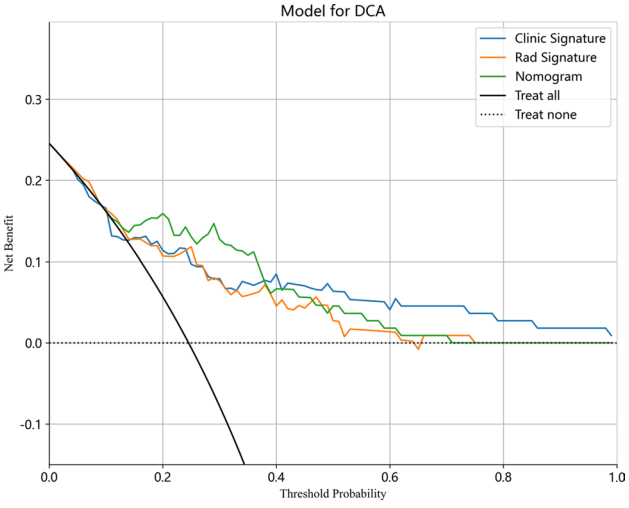

Fig. 12 DCA in the testing cohort. DCA, decision curve analysis

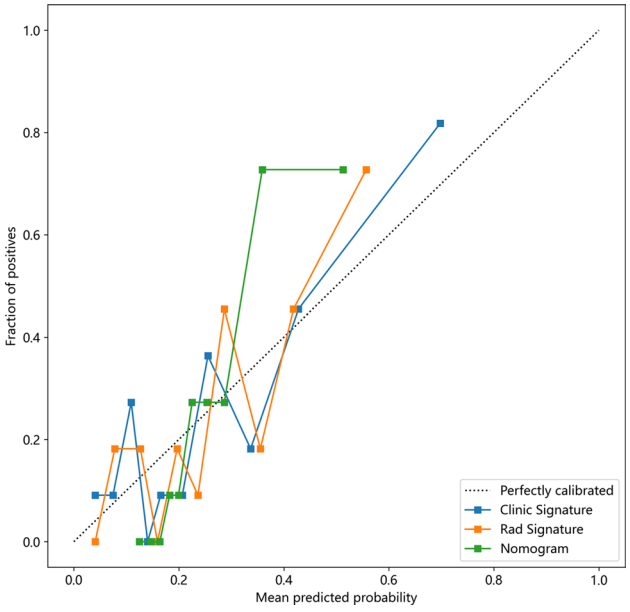

Fig. 13 Calibrated curve in the testing cohort

correlate with the development of STAS [19]. Tumor demarcation, pleural retroflexion and tumor contribution have also been used to assess VPI [20–22]. Qi L et al. concluded that the AUC = 0.76 in the model applying consolidation tumor ratio (CTR) to predict STAS [23]. This result is similar to the AUC of the imaging histology-only model in our study. There are previous studies that included preoperative 18 F-FDG PET/CT findings [24]. However, due to the high cost of PET/CT, most patients with CT manifestations of small pulmonary nodules do not choose to undergo PET/CT in the first instance. Radiomics has also been applied to predict STAS [25, 26]. Chen D applied imaging histological features for prediction in stage I lung adenocarcinoma and obtained good performance [27]. Currently there are fewer studies on deep learning applied to predict STAS. DL shows great promise in predicting invasiveness of lung cancer and shows good model efficacy in predicting infiltration-non-infiltration scenarios as well as lymph node metastasis. Junli T applies a 3D convolutional neural network model to predict STAS in non-small cell lung cancer, with efficacy better than the general base model [28]. However, the identification and screening of single imaging features is not an easy task even for highly qualified radiology staff.

However, there are still some limitations of this study: (1) The retrospective nature of this study may have some selection bias. (2) This is a single-centre study with a relatively small sample size, and further large studies are needed in the future. External validation was constructed to build a better model to predict STAS. 3.Imaging histology is usually reserved for specialised imaging physicians or clinical specialists. This results in a low reproducibility of the model and a high threshold.

Table 4 Training set nomogram results

Signature	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Precision	Recall	F1	Threshold
Clinic Signature	0.736	0.759	0.6412–0.8777	0.741	0.735	0.476	0.897	0.476	0.741	0.580	0.243
Rad Signature	0.727	0.764	0.6563–0.8716	0.778	0.711	0.467	0.908	0.467	0.778	0.583	0.256
Nomogram	0.864	0.878	0.8078–0.9477	0.704	0.916	0.731	0.905	0.731	0.704	0.717	0.292

Abbreviations: AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; F1 = 2 x (precision x recall)/(precision + recall)

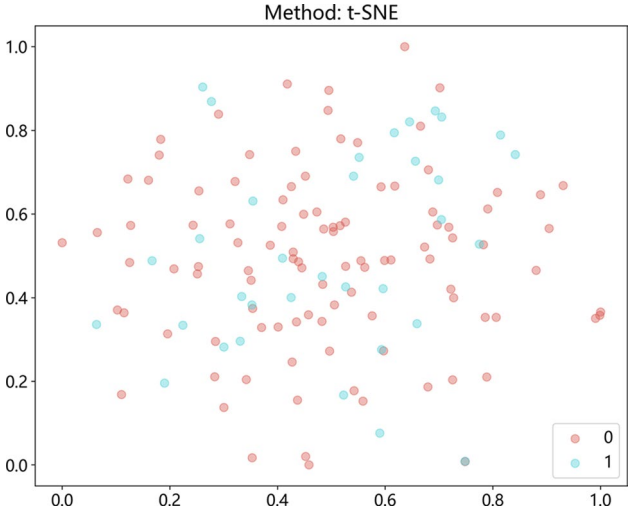

Fig. 14 2D Deep Learning Feature Visualization

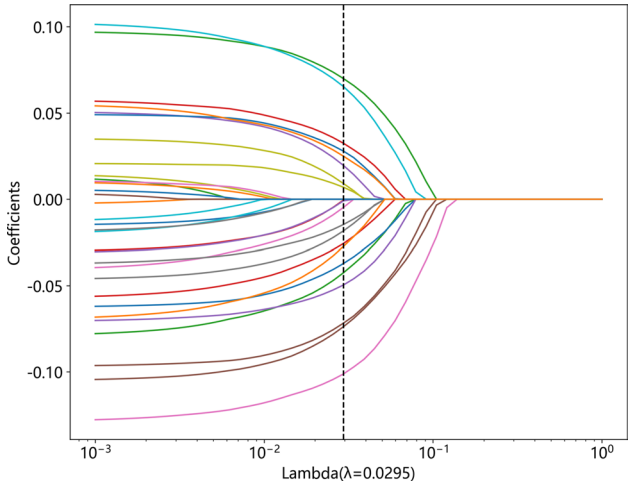

Fig. 15 Lasso regression curves for deep learning features

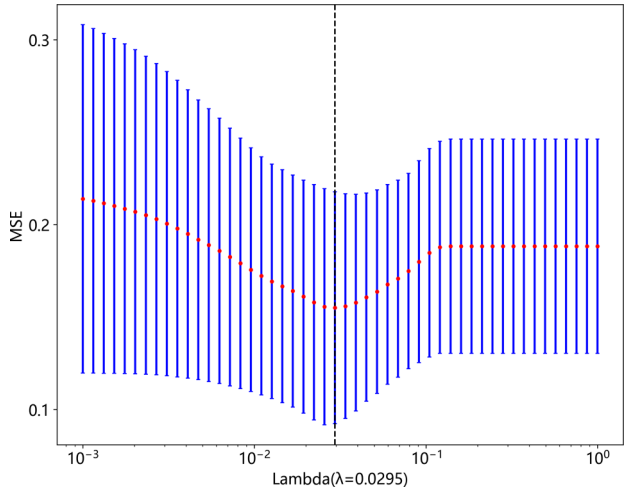

Fig. 16 MSE of deep learning features. MSE, mean standard error

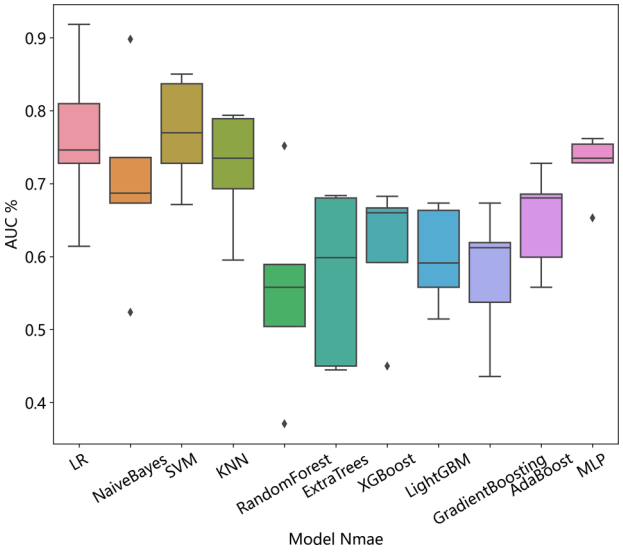

Fig. 17 Fifty-fold randomized cross-validation box plots for different DL machine learning models. DL, deep learning

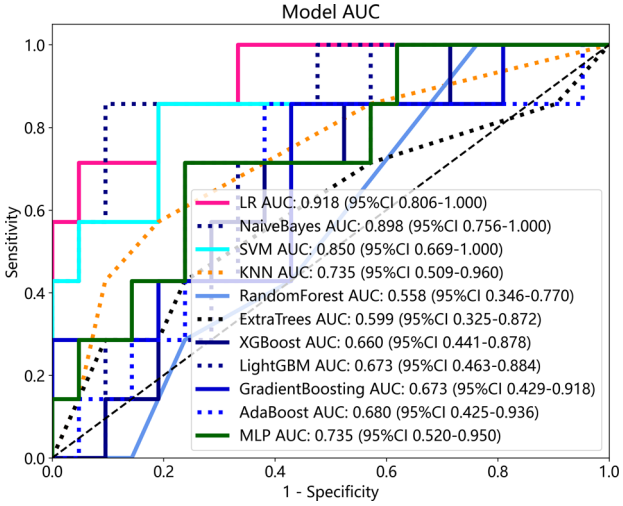

Fig. 18 ROC curves for different DL machine learning models

Conclusions

This study utilized a retrospective approach to investigate the correlation of imaging histology and clinical features with airway dissemination in patients with invasive lung adenocarcinoma. It incorporates machine learning and deep learning methods. A model with good efficacy was established and validated. A nomogram with clinical guidance was finally constructed.

Acknowledgements
Not applicable.

Author contributions
ZM Wang and LX Kong performed the data collection and statistics. GC Duan and Kong performed the ROI outlining of the imaging data. QT Zhao participated in the mapping. All other authors participated in writing and reviewing the article.

Funding
This research was funded by the Key Research and Development Program of Hebei Province (grant no. 22377790D).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
This study was approved by the Ethics Committee of Hebei Provincial General Hospital, approval number NO.2023128.

Patient consent for publication
This was a retrospective study, and patients were exempted from the informed consent application through the Ethics Committee of Hebei Provincial General Hospital.

Competing interests
The authors declare no competing interests.

Received: 4 July 2024 / Accepted: 1 August 2025
Published online: 14 August 2025

References

1. Sung Hyuna,Ferlay Jacques,Siegel, Rebecca L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 Countries.[J]. CA Cancer J Clin. 2021;71:209–49.
2. Warth A, Muley T, Kossakowski CA, Goeppert B, Schirmacher P, Dienemann H, Weichert W. Prognostic impact of intra-alveolar tumor spread in pulmonary adenocarcinoma. Am J Surg Pathol. 2015;39(6):793–801. https://doi.org/10.1097/PAS.0000000000000409. PMID: 25723114.
3. Okada M, Koike T, Higashiyama M, Yamato Y, Kodama K, Tsubota N. Radical sublobar resection for small-sized non-small cell lung cancer: a multicenter study. J Thorac Cardiovasc Surg. 2006;132(4):769– 75. https://doi.org/10.1016/j.jtcvs.2006.02.063. PMID: 17000286.
4. Kadota K, Nitadori JI, Sima CS, Ujiie H, Rizk NP, Jones DR, Adusumilli PS, Travis WD. 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):806–14. https://doi.org/10.1097/JTO.0000000000000486. PMID: 25629637; PMCID: PMC4500042.
5. Medina MA, Onken AM, de Margerie-Mellon C, Heidinger BH, Chen Y, Bankier AA, VanderLaan PA. Preoperative bronchial cytology for the assessment of tumor spread through air spaces in lung adenocarcinoma resection specimens. Cancer Cytopathol. 2020;128(4):278–286. https://doi.org/10.1002/cncy.22243. Epub 2020 Feb 3. PMID: 32012490.
6. 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. https://doi.org/10.1186/s12885-021-08648-0. PMID: 34380457; PMCID: PMC8359054.
7. Walts AE, Marchevsky AM. Current evidence does not warrant frozen section evaluation for the presence of tumor spread through alveolar spaces. Arch Pathol Lab Med. 2018;142(1):59–63. https://doi.org/10.5858/arpa.2016-0635-OA. Epub 2017 Oct 2. PMID: 28967802.
8. de Margerie-Mellon C, Onken A, Heidinger BH, et al. CT manifestations of tumor spread through airspaces in pulmonary adenocarcinomas presenting as subsolid nodules. J Thorac Imaging. 2018;33:402–8. https://doi.org/10.1097/RTI.0000000000000344.
9. Kadota K, Kushida Y, Kagawa S, Ishikawa R, Ibuki E, Inoue K, Go T, Yokomise H, Ishii T, Kadowaki N, Haba R. 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. https://doi.org/10.1097/PAS.0000000000001285. PMID: 31107717.
10. Gillies RJ, Kinahan PE, Hricak H, Radiomics. Radiology. 2016;278(2):563–77. https://doi.org/10.1148/radiol.2015151169. Epub 2015 Nov 18. PMID: 26579733; PMCID: PMC4734157. Images Are More than Pictures, They Are Data.
11. Jeffries CD, Ford JR, Tilson JL, Perkins DO, Bost DM, Filer DL, Wilhelmsen KC. A greedy regression algorithm with coarse weights offers novel advantages. Sci Rep. 2022;12(1):5440. https://doi.org/10.1038/s41598-022-09415-2. PMID: 35361850; PMCID: PMC8971398.
12. Kadota K, Kushida Y, Katsuki N, Ishikawa R, Ibuki E, Motoyama M, Nii K, Yokomise H, Bandoh S, Haba R. Tumor spread through air spaces is an independent predictor of recurrence-free survival in patients with resected lung squamous cell carcinoma. Am J Surg Pathol. 2017;41(8):1077–1086. https://doi.org/10.1097/PAS.0000000000000872. PMID: 28498282.
13. Lu S, Tan KS, Kadota K, Eguchi T, Bains S, Rekhtman N, Adusumilli PS, Travis WD. Spread through air spaces (STAS) is an independent predictor of recurrence and lung Cancer-Specific death in squamous cell carcinoma. J Thorac Oncol. 2017;12(2):223–34. https://doi.org/10.1016/j.jtho.2016.09.129. Epub 2016 Sep 28. PMID: 27693541; PMCID: PMC5639476.
14. Aly RG, Rekhtman N, Li X, Takahashi Y, Eguchi T, Tan KS, Rudin CM, Adusumilli PS, Travis WD. 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):1583–93. https://doi.org/10.1016/j.jtho.2019.05.009. Epub 2019 May 20. PMID: 31121325; PMCID: PMC8160527.
15. Lee JS, Kim EK, Kim M, Shim HS. Genetic and clinicopathologic characteristics of lung adenocarcinoma with tumor spread through air spaces. Lung Cancer. 2018;123:121–6. https://doi.org/10.1016/j.lungcan.2018.07.020. Epub 2018 Jul 17. PMID: 30089582.
16. Shiono S, Yanagawa N. Spread through air spaces is a predictive factor of recurrence and a prognostic factor in stage I lung adenocarcinoma. Interact Cardiovasc Thorac Surg. 2016;23(4):567–72. https://doi.org/10.1093/icvts/ivw211. Epub 2016 Jun 26. PMID: 27354463.
17. Terada Y, Takahashi T, Morita S, Kashiwabara K, Nagayama K, Nitadori JI, Anraku M, Sato M, Shinozaki-Ushiku A, Nakajima J. Spread through air spaces is an independent predictor of recurrence in stage III (N2) lung adenocarcinoma. Interact Cardiovasc Thorac Surg. 2019;29(3):442–448. https://doi.org/10.1093/icvts/ivz116. PMID: 31106332.
18. Eguchi T, Kameda K, Lu S, Bott MJ, Tan KS, Montecalvo J, Chang JC, Rekhtman N, Jones DR, Travis WD, Adusumilli PS. Lobectomy is associated with better outcomes than sublobar resection in spread through air spaces (STAS)-Positive T1 lung adenocarcinoma: A propensity Score-Matched analysis. J Thorac Oncol. 2019;14(1):87–98. https://doi.org/10.1016/j.jtho.2018.09.005. Epub 2018 Sep 19. PMID: 30244070; PMCID: PMC6309668.
19. 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. https://doi.org/10.1097/RTI.0000000000000344. PMID: 30067571.
20. Kim SK, Kim TJ, Chung MJ, Kim TS, Lee KS, Zo JI, Shim YM. Lung adenocarcinoma: CT features associated with spread through air spaces. Radiology. 2018;289(3):831–840. https://doi.org/10.1148/radiol.2018180431. Epub 2018 Sep 4. PMID: 30179108.
21. Toyokawa G, Yamada Y, Tagawa T, Kamitani T, Yamasaki Y, Shimokawa M, Oda Y, Maehara Y. Computed tomography features of resected lung adenocarcinomas with spread through air spaces. J Thorac Cardiovasc Surg. 2018;156(4):1670–e16764. https://doi.org/10.1016/j.jtcvs.2018.04.126. Epub 2018 Jun 6. PMID: 29961590.
22. Koezuka S, Mikami T, Tochigi N, Sano A, Azuma Y, Makino T, Otsuka H, Matsumoto K, Shiraga N, Iyoda A. Toward improving prognosis prediction in patients undergoing small lung adenocarcinoma resection: radiological and pathological assessment of diversity and intratumor heterogeneity. Lung Cancer. 2019;135:40–6. https://doi.org/10.1016/j.lungcan.2019.06.023. Epub 2019 Jun 25. PMID: 31447001.
23. 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. https://doi.org/10.3389/fonc.2020.548430. PMID: 33505903; PMCID: PMC7831277.
24. Wang XY, Zhao YF, Yang L, Liu Y, Yang YK, Wu N. Correlation analysis between metabolic tumor burden measured by positron emission tomography/computed tomography and the 2015 world health organization classification of lung adenocarcinoma, with a risk prediction model of tumor spread through air spaces. Transl Cancer Res. 2020;9(10):6412–22. https://doi.org/10.21037/tcr-20-1934. PMID: 35117249; PMCID: PMC8798457.
25. Zhuo Y, Feng M, Yang S, Zhou L, Ge D, Lu S, Liu L, Shan F, Zhang Z. Radiomics nomograms of tumors and peritumoral regions for the preoperative prediction of spread through air spaces in lung adenocarcinoma. Transl Oncol. 2020;13(10):100820. https://doi.org/10.1016/j.tranon.2020.100820. Epub 2020 Jul 1. PMID: 32622312; PMCID: PMC7334418.
26. Qi L, Li X, He L, Cheng G, Cai Y, Xue K, Li M. Comparison of diagnostic performance of spread through airspaces of lung adenocarcinoma based on

morphological analysis and perinodular and intranodular radiomic features on chest CT images. Front Oncol. 2021;11:654413. https://doi.org/10.3389/fonc.2021.654413. PMID: 34249691; PMCID: PMC8268002.
27. Chen D, She Y, Wang T, Xie H, Li J, Jiang G, Chen Y, Zhang L, Xie D, Chen C. Radiomics-based prediction for tumour spread through air spaces in stage I lung adenocarcinoma using machine learning. Eur J Cardiothorac Surg. 2020;58(1):51–58. https://doi.org/10.1093/ejcts/ezaa011. PMID: 32011674.
28. Tao J, Liang C, Yin K, Fang J, Chen B, Wang Z, Lan X, Zhang J. 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(11):535–44. https://doi.org/10.1016/j.diii.2022.06.002. Epub 2022 Jun 27. PMID: 35773100.

Publisher's note