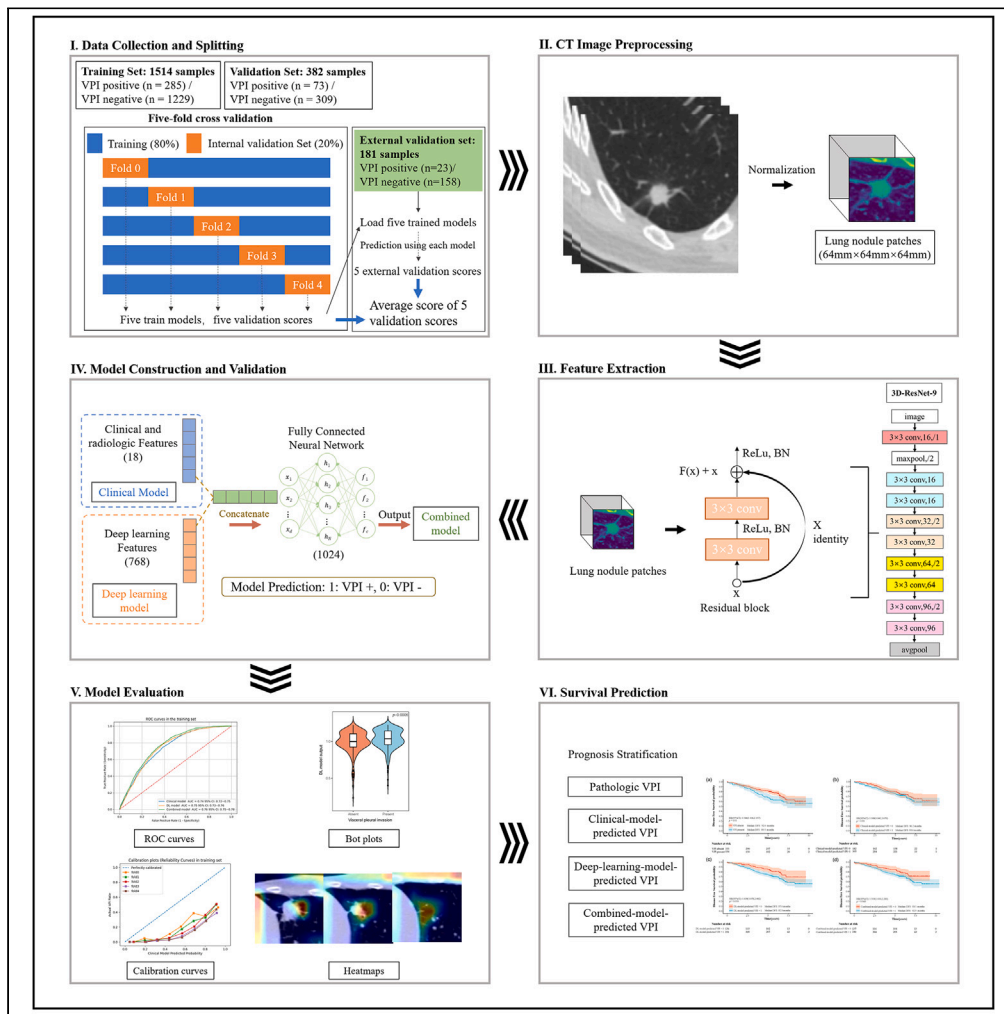**Article**

A CT-based deep learning model: visceral pleural invasion and survival prediction in clinical stage IA lung adenocarcinoma

Xiaofeng Lin,
Kunfeng Liu,
Kunwei Li, ...,
Sheng Li, Huai
Chen, Li Li

lisheng@sysucc.org.cn (S.L.)
chenhuai1977@163.com (H.C.)
lil@sysucc.org.cn (L.L.)

Highlights

A CT-based deep learning
signature was extracted
with a 3D residual neural
network

The deep learning model
showed good predictive
efficacy for visceral pleural
invasion

Visceral pleural invasion
prediction outputs are
predictors of the disease-
free survival

The deep learning
signature had a wide scope
of clinical application

Article

A CT-based deep learning model: visceral pleural invasion and survival prediction in clinical stage IA lung adenocarcinoma

Xiaofeng Lin,[1,5] Kunfeng Liu,[1,5] Kunwei Li,[2,5] Xiaojuan Chen,[3] Biyun Chen,[1] Sheng Li,[1,*] Huai Chen,[4,*] and Li Li[1,6,*]

SUMMARY

Pathologic visceral pleural invasion (VPI) in patients with early-stage lung cancer can result in the upstaging of T1 to T2, in addition to having implications for surgical resection and prognostic outcomes. This study was designed with the goal of establishing and validating a CT-based deep learning (DL) model capable of predicting VPI status and stratifying patients based on their prognostic outcomes. In total, 2077 patients from three centers with pathologically confirmed clinical stage IA lung adenocarcinoma were enrolled. DL signatures were extracted with a 3D residual neural network. DL model was able to effectively predict VPI status. VPI predicted by the DL models, as well as pathologic VPI, was associated with shorter disease-free survival. The established deep learning signature provides a tool capable of aiding the accurate prediction of VPI in patients with clinical stage IA lung adenocarcinoma, thus enabling prognostic stratification.

INTRODUCTION

Lung cancer is among the most prevalent forms of cancer in the world, contributing to high morbidity among patients.[1] Of these cases, 85–90% are of the non-small cell lung cancer (NSCLC) subtype, and the majority of patients with NSCLC exhibit lung adenocarcinoma as their diagnosed histological subtype.[1] In patients with stage IA lung cancer, the 5-year overall survival rate is 82%, as compared to 6% in individuals with stage IV disease.[2] Diagnosing lung cancer at an early stage can thus contribute to improved prognostic outcomes. Surgery is the most common means of treating individuals with stage IA lung cancer.[3] Visceral pleural invasion (VPI) includes the PL1 and PL2 levels, which respectively correspond to the invasion of the tumor beyond the elastic layer or to the pleural surface.[4] In patients with NSCLC with tumors ≤3 cm in size, VPI is associated with a poorer prognosis and can aid in predicting lymph node metastasis and postoperative recurrence.[5] The eighth edition of the TNM classification system recommends upstaging NSCLC tumors ≤3.0 cm in size with VPI from stage T1 to stage T2 and from stage IA to stage IB owing to these differences in patient prognosis.[6] Even in those patients exhibiting clinical stage IA disease, lobectomy with mediastinal lymph node dissection is recommended as an alternative to solely undergoing segmentectomy.[7]

At present, diagnosing VPI prior to surgery based solely on computed tomography (CT) findings remains difficult. A range of CT features have been reported to be indicative of VPI, including pleural retraction, pleural tags, pleural indentation, and contact between the tumor and the pleura.[8–12] The exact diagnostic accuracy of these features, however, reportedly ranges from 62.7 to 72.3%.[9–12]

Recent advances in deep learning (DL) technologies have enabled the establishment of algorithms capable of characterizing data in an automated manner such that imaging data associated with particular regions of interest can be transformed into high-resolution spatial data from which features can be excavated. The resultant algorithms are conducive to the analysis of imaging data in a high-throughput manner, enabling comprehensive evaluations of a range of tumor phenotypes in a high-fidelity manner.[13–15] Owing to their ability to readily learn features, DL strategies have exhibited great progress as approaches to classifying pulmonary nodules, predicting lymph node metastasis, and evaluating patient prognostic outcomes.[16–19] At present, however, few studies have sought to leverage these DL techniques as a means of predicting VPI status in patients with lung cancer.

The present study was developed with the goal of designing a DL algorithm capable of predicting VPI in patients with stage IA lung adenocarcinoma in a noninvasive manner before surgical resection, with the overall goal of providing clinicians with information that can guide decision-making related to patient care.

[1]Department of Medical Imaging, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, P.R. China
[2]Department of Radiology, The Fifth Affiliated Hospital of Sun Yat-sen University, Zhuhai 519000, P.R. China
[3]Department of Radiology, First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510260, P.R. China
[4]Department of Radiology, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou 510260, P.R. China
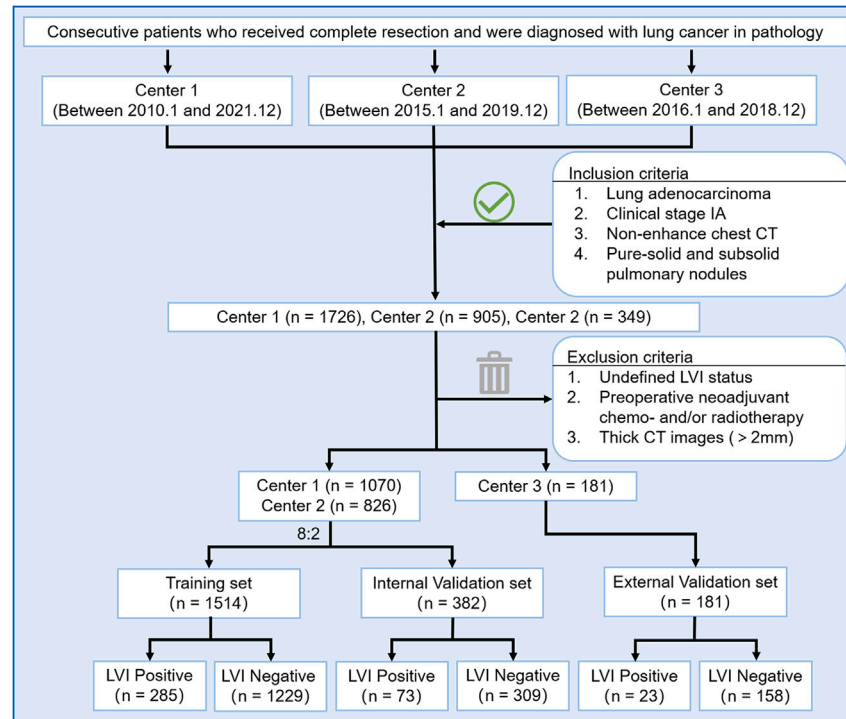[5]These authors contributed equally
[6]Lead contact
*Correspondence: lisheng@sysucc.org.cn (S.L.), chenhuai1977@163.com (H.C.), lil@sysucc.org.cn (L.L.)
https://doi.org/10.1016/j.isci.2023.108712

Figure 1. A flowchart detailing the inclusion and exclusion of patients in this study
VPI, visceral pleural invasion.

RESULTS

Patient characteristics

The patient selection process is detailed in Figure 1, and the characteristics of included subjects at baseline are provided in Tables 1 and 2. In total, this retrospective analysis incorporated 2077 patients, of whom 1514, 382 and 181 were respectively assigned to the training, internal validation, and external validation cohorts. In the overall patient cohort, 59years ± 9.6 was the mean age, 47% (n = 977) of the participants were males, and 18.3% (n = 381) had received a pathologic VPI diagnosis. The three chest radiologists responsible for evaluating radiologic characteristics exhibited a clear consensus (p < 0.05) (see Tables S1 and S2), and no differences in baseline characteristics were noted when comparing the training and internal validation cohorts (p > 0.05). When comparing patients that were positive and negative for VPI, significant associations were noted with family history of cancer, clinical T stage, histologic subtype, pathologic stage, LVI, lymph node metastasis, whole maximum diameter, consolidation maximum diameter, CTR, density, pleural CT subtype, pleural indentation, consolidation contact with the pleura, minimum distance from the lesion to the pleura, the maximum diameter of contact with pleura, the maximum diameter of consolidation in contact with the pleura, and CTR at the pleural end (p < 0.05) in the training and internal validation cohorts. While in the external validation cohorts, clinical T stage, histologic subtype, pathologic stage, distant metastasis, consolidation maximum diameter, spiculated sign, pleural CT subtype, pleural indentation and consolidation contact with the pleura were significant correlated with VPI status. Prior to the construction of a clinical model capable of preoperatively predicting VPI status, univariate logistic regression analyses were employed to identify relevant clinical or radiologic features. Screened variables significantly associated with VPI status in these analyses (p < 0.05) included family history of cancer, clinical T stage, overall tumor maximum diameter, consolidation maximum diameter, CTR, density, spiculated sign, lobulated sign, vessel convergence, pleural CT subtype, pleural indentation, consolidation contact with the pleura, minimum distance between the lesion and the pleura, maximum diameter of pleural contact, maximum diameter of consolidation contact with the pleura, and CTR at the pleural end. Multivariate logistic regression analyses ultimately revealed that lobulated sign, pleural CT subtype (Type 2, Type 3, and Type 5), and pleural indentation were all independently predictive of VPI status (p < 0.05). For further details regarding the univariate and multivariate analyses performed in the training cohort, see Table 3.

Analyses of model efficiency

The DL model exhibited AUC values of 0.75 (95% confidence interval [CI]: 0.73–0.76), 0.73 (95%CI: 0.69–0.76), and 0.62 (95%CI: 0.54–0.68) in the training, internal validation, and external validation cohorts, respectively, comparable to the clinical model [AUC = 0.74 (95%CI: 0.72–0.75), 0.73 (95%CI: 0.70–0.77), 0.66 (95%CI: 0.59–0.72)]. However, after the incorporation of DL signature and clinical-radiological factors, combined model

Table 1. Clinical, pathological and radiological characteristics

	Training set (n = 1514)			Internal validation set (n = 382)			External validation set (n = 181)			
	VPI positive (n = 285)	VPI negative (n = 1229)	P1 value	VPI positive (n = 73)	VPI negative (n = 309)	P1 value	VPI positive (n = 23)	VPI negative (n = 158)	P1 value	P2 value
Clinical characteristics										
Age (years)[a]	58.9 ± 9.8	58.3 ± 9.9	0.422	59.6 ± 10.2	58.0 ± 10.4	0.235	63.5 ± 7.8	58.4 ± 10.6	0.087	0.812
Sex			0.630			0.265			0.712	0.856
Male	138 (48.4%)	573 (46.6%)		30 (41.1%)	152 (49.2%)		12 (52.2%)	72 (45.6%)		
Female	147 (51.6%)	656 (53.4%)		43 (58.9%)	157 (50.8%)		11 (47.8%)	86 (54.4%)		
Symptom			0.985			0.200			0.084	0.855
Yes	178 (62.5%)	771 (62.7%)		41 (56.2%)	201 (65.0%)		6 (26.1%)	17 (10.8%)		
No	107 (37.5%)	458 (37.3%)		32 (43.8%)	108 (35.0%)		17 (73.9%)	141 (89.2%)		
Smoking history			1.000			0.963			0.479	0.564
Yes	76 (26.7%)	329 (26.8%)		19 (26.0%)	77 (24.9%)		8 (34.8%)	40 (25.3%)		
No	209 (73.3%)	900 (73.2%)		54 (74.0%)	232 (75.1%)		15 (65.2%)	118 (74.7%)		
Family history of cancer			0.009			0.008			0.714	0.077
Yes	56 (19.6%)	164 (13.3%)		11 (15.1%)	24 (7.8%)		2 (8.7%)	7 (4.4%)		
No	229 (80.4%)	1065 (86.7%)		62 (84.9%)	285 (92.2%)		21 (91.3%)	151 (95.6%)		
Clinical T stage			<0.001			<0.001			0.035	0.444
cT1a	16 (5.6%)	341 (27.7%)		5 (6.8%)	97 (31.4%)		2 (8.7%)	55 (34.8%)		
cT1b	125 (43.9%)	571 (46.5%)		32 (43.8%)	136 (44.0%)		12 (52.2%)	65 (41.1%)		
cT1c	144 (50.5%)	317 (25.8%)		36 (49.3%)	76 (24.6%)		9 (39.1%)	38 (24.1%)		
Pathological characteristics										
Histologic subtype			<0.001			<0.001			0.001	0.116
Lepidic predominant	10 (3.5%)	162 (13.2%)		2 (2.7%)	42 (13.6%)		0 (0%)	6 (3.8%)		
Acinar predominant	168 (58.9%)	746 (60.7%)		36 (49.3%)	173 (56.0%)		5 (21.7%)	83 (52.5%)		
Papillary predominant	44 (15.4%)	147 (12.0%)		19 (26.0%)	41 (13.3%)		8 (34.8%)	43 (27.2%)		
Micropapillary predominant	10 (3.5%)	27 (2.2%)		4 (5.5%)	10 (3.2%)		0 (0%)	3 (1.9%)		
Solid predominant	21 (7.4%)	62 (5.0%)		7 (9.6%)	17 (5.5%)		4 (17.4%)	7 (4.4%)		
Mucous predominant	13 (4.6%)	37 (3.0%)		1 (1.4%)	8 (2.6%)		0 (0%)	5 (3.2%)		
Special type	1 (0.4%)	6 (0.5%)		0 (0%)	0 (0%)		0 (0%)	6 (3.8%)		
MIA	1 (0.4%)	8 (0.7%)		0 (0%)	3 (1.0%)		0 (0%)	1 (0.6%)		
AIS	0 (0%)	0 (0%)		0 (0%)	1 (0.3%)		0 (0%)	3 (1.9%)		
Non-classified	17 (6.0%)	34 (2.8%)		4 (5.5%)	14 (4.5%)		6 (26.1%)	7 (4.4%)		
Pathologic stage			<0.001			<0.001			<0.001	0.432
IA	0 (0%)	1123 (91.4%)		0 (0%)	288 (93.2%)		2 (8.7%)	143 (90.5%)		
IB	219 (76.8%)	0 (0%)		52 (71.2%)	0 (0%)		15 (65.2%)	2 (1.3%)		
IIB	25 (8.8%)	39 (3.2%)		9 (12.3%)	14 (4.5%)		0 (0%)	5 (3.2%)		
IIIA	38 (13.3%)	66 (5.4%)		11 (15.1%)	7 (2.3%)		1 (4.3%)	7 (4.4%)		
IIIB	1 (0.4%)	0 (0%)		0 (0%)	0 (0%)		1 (4.3%)	0 (0%)		

(Continued on next page)

Table 1. *Continued*

	Training set (n = 1514)			Internal validation set (n = 382)			External validation set (n = 181)			
	VPI positive (n = 285)	VPI negative (n = 1229)	P1 value	VPI positive (n = 73)	VPI negative (n = 309)	P1 value	VPI positive (n = 23)	VPI negative (n = 158)	P1 value	P2 value
IVA	2 (0.7%)	1 (0.1%)		1 (1.4%)	0 (0%)		4 (17.4%)	1 (0.6%)		
LVI			<0.001			<0.001			0.823	0.451
Yes	87 (30.5%)	149 (12.1%)		25 (34.2%)	28 (9.1%)		2 (8.7%)	8 (5.1%)		
No	198 (69.5%)	1080 (87.9%)		48 (65.8%)	281 (90.9%)		21 (91.3%)	150 (94.9%)		
Lymph node metastasis			<0.001			<0.001			1.000	0.854
Yes	65 (22.8%)	105 (8.5%)		20 (27.4%)	21 (6.8%)		2 (8.7%)	12 (7.6%)		
No	220 (77.2%)	1124 (91.5%)		53 (72.6%)	288 (93.2%)		21 (91.3%)	146 (92.4%)		
Distant metastasis			0.052			0.431			<0.001	1.000
Yes	3 (1.1%)	1 (0.1%)		1 (1.4%)	0 (0%)		4 (17.4%)	1 (0.6%)		
No	282 (98.9%)	1228 (99.9%)		72 (98.6%)	309 (100%)		19 (82.6%)	157 (99.4%)		
Radiological characteristics										
Whole maximum diameter (mm)[a]	21.1 ± 5.47	18.7 ± 5.85	<0.001	20.9 ± 5.42	18.1 ± 5.85	<0.001	21.9 ± 6.08	19.8 ± 7.67	0.140	0.139
Consolidation maximum diameter (mm)[a]	20.2 ± 5.87	15.4 ± 7.04	<0.001	19.7 ± 6.46	15.1 ± 7.06	<0.001	18.5 ± 7.72	14.8 ± 7.68	0.043	0.408
CTR[a]	0.958 ± 0.131	0.821 ± 0.254	<0.001	0.933 ± 0.175	0.822 ± 0.251	<0.001	0.853 ± 0.278	0.775 ± 0.301	0.223	0.809
Location			0.051			0.057			0.719	0.276
RUL	93 (32.6%)	426 (34.7%)		23 (31.5%)	100 (32.4%)		8 (34.8%)	65 (41.1%)		
RML	32 (11.2%)	81 (6.6%)		3 (4.1%)	34 (11.0%)		2 (8.7%)	11 (7.0%)		
RLL	61 (21.4%)	239 (19.4%)		8 (11.0%)	61 (19.7%)		2 (8.7%)	26 (16.5%)		
LUL	69 (24.2%)	299 (24.3%)		17 (23.3%)	70 (22.7%)		7 (30.4%)	39 (24.7%)		
LLL	30 (10.5%)	184 (15.0%)		22 (30.1%)	44 (14.2%)		4 (17.4%)	17 (10.8%)		
Density			<0.001			0.002			0.089	0.766
Subsolid	39 (13.7%)	495 (40.3%)		13 (17.8%)	118 (38.2%)		6 (26.1%)	75 (47.5%)		
Pure-solid	246 (86.3%)	734 (59.7%)		60 (82.2%)	191 (61.8%)		17 (73.9%)	83 (52.5%)		
Shape			0.405			0.774			1.000	0.154
Regular	5 (1.8%)	35 (2.8%)		4 (5.5%)	12 (3.9%)		9 (39.1%)	61 (38.6%)		
Irregular	280 (98.2%)	1194 (97.2%)		69 (94.5%)	297 (96.1%)		14 (60.9%)	97 (61.4%)		
Boundary			0.001			0.934			0.352	1.000
Clear	48 (16.8%)	323 (26.3%)		17 (23.3%)	76 (24.6%)		6 (26.1%)	61 (38.6%)		
Vague	237 (83.2%)	906 (73.7%)		56 (76.7%)	233 (75.4%)		17 (73.9%)	97 (61.4%)		
Vacuole sign			0.280			0.956			1.000	0.324
Yes	59 (20.7%)	218 (17.7%)		11 (15.1%)	50 (16.2%)		4 (17.4%)	28 (17.7%)		
No	226 (79.3%)	1011 (82.3%)		62 (84.9%)	259 (83.8%)		19 (82.6%)	130 (82.3%)		
Cavity			1.000			0.877			1.000	0.996
Yes	8 (2.8%)	37 (3.0%)		3 (4.1%)	9 (2.9%)		0 (0%)	2 (1.3%)		
No	277 (97.2%)	1192 (97.0%)		70 (95.9%)	300 (97.1%)		23 (100%)	156 (98.7%)		
Calcification			0.060			1.000			0.262	0.934
Yes	5 (1.8%)	6 (0.5%)		0 (0%)	2 (0.6%)		1 (4.3%)	0 (0%)		
No	280 (98.2%)	1223 (99.5%)		73 (100%)	307 (99.4%)		22 (95.7%)	158 (100%)		

(Continued on next page)

Table 1. *Continued*

	Training set (n = 1514)			Internal validation set (n = 382)			External validation set (n = 181)			
	VPI positive (n = 285)	VPI negative (n = 1229)	P1 value	VPI positive (n = 73)	VPI negative (n = 309)	P1 value	VPI positive (n = 23)	VPI negative (n = 158)	P1 value	P2 value
Spiculated sign			<0.001			0.086			0.005	0801
Yes	230 (80.7%)	736 (59.9%)		54 (74.0%)	193 (62.5%)		18 (78.3%)	71 (44.9%)		
No	55 (19.3%)	493 (40.1%)		19 (26.0%)	116 (37.5%)		5 (21.7%)	87 (55.1%)		
Lobulated sign			<0.001			0.104			0.077	0.609
Yes	281 (98.6%)	1048 (85.3%)		68 (93.2%)	263 (85.1%)		21 (91.3%)	113 (71.5%)		
No	4 (1.4%)	181 (14.7%)		5 (6.8%)	46 (14.9%)		2 (8.7%)	45 (28.5%)		
Air bronchogram			0.005			0.056			0.253	0.250
Yes	46 (16.1%)	295 (24.0%)		8 (11.0%)	67 (21.7%)		7 (30.4%)	72 (45.6%)		
No	239 (83.9%)	934 (76.0%)		65 (89.0%)	242 (78.3%)		16 (69.6%)	86 (54.4%)		
Vessel convergence			<0.001			0.113			0.812	1.000
Yes	262 (91.9%)	957 (77.9%)		64 (87.7%)	243 (78.6%)		6 (26.1%)	49 (31.0%)		
No	23 (8.1%)	272 (22.1%)		9 (12.3%)	66 (21.4%)		17 (73.9%)	109 (69.0%)		

— Unless otherwise noted, values are numbers of patients, with percentages in parentheses. AIS, adenocarcinoma *in situ*; CTR, consolidation tumor ratio; LLL, left lower lobe; LUL, left upper lobe; LVI, lymphovascular invasion; MIA, minimally invasive adenocarcinoma; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe; VPI, visceral pleural invasion.

[a]Data are means ± standard deviations. *P*1 value was derived from the univariable association analyses between VPI positive and VPI negative. *P*2 value was derived from the univariable association analyses between training set and internal validation set.

[AUC = 0.76 (95% CI: 0.75–0.78), 0.75 (95%CI: 0.71–0.78), 0.69 (95%CI: 0.62–0.75)] achieved little improvement. The prediction performance of the three models in three cohorts is exhibited in Table 4. More details about the additional standard deviation of 5-fold cross-validation are shown in supplementary data (see Table S3). ROC curves for these three models when used to predict the VPI status of patients with clinical stage IA lung adenocarcinoma in the three cohorts are shown in Figure 2. No significant differences in AUC values for these three models were detected when using DeLong's test in either the training or the two validation cohorts (see Table S4). Significant associations between the prediction scores of these three models and VPI status were noted (p < 0.05) (Figure 3), emphasizing the reliability of these models as a means of predicting VPI. When assessing model accuracy, sensitivity, specificity, PPV, NPV, and F1 score values, the DL model outperformed the clinical models, while the combined model was similar to the DL model in terms of performance. Calibration curves for these models are provided in Figure S1. Hosmer-Lemeshow tests showed that there were no significant difference in the goodness of fit of the three models ($P > 0.05$).

Clinical significance of model

Clinicopathological details, CT images, and corresponding heatmaps for VPI-positive and VPI-negative patients which similar clinicopathological characteristics are shown in Figure 4. The DFS of the VPI-positive patients exhibiting bone metastases were shorter than those of VPI-negative patients free of recurrent or metastatic disease. Heatmaps highlighted clear differences between tumors that were positive and negative for VPI. The regions most closely associated with predicting VPI status were the regions in contact with the adjacent pleura and the peritumoral zone, which was distant from the pleura in some cases. The DL features extracted from these images were of value when seeking to predict VPI status and to stratify patients into low- and high-risk groups.

LVI and lymph node metastasis had significant statistical differences between the VPI-negative and -positive predicted by the three models, whether in the training or the internal validation sets. However, there was no difference in the external validation set. This is consistent with the above three pathological factors presented in Table1. When comparing distant metastasis three-model-predicted VPI status was not statistically different in any of the three datasets. Assessment of the above three pathological characteristics for the three models is shown in supplementary data (see Table S5).

Survival prediction

For survival comparison analysis, 487 patients with clinical stage IA lung adenocarcinoma were evaluated. These individuals exhibited a median DFS of 67.4 months, and 29.3% (143/487) of these individuals experienced recurrent disease, metastasis, or death following complete surgical resection. These patients exhibited a 67.4-month median follow-up duration. Univariate Cox regression analyses results are summarized, see Table S6, revealing that DFS was associated with pathologic VPI and with VPI status as predicted by both the DL and combined models (p < 0.05 for all), while VPI status as predicted with the clinical model was not related to DFS (p = 0.327). After multivariate analyses, CTR, clinical T stage (cT1b and cT1c), pathologic VPI, and VPI status predicted by DL and combined models remained significant predictors

Table 2. CT findings for visceral pleural

	Training set (n = 1514)			Internal validation set (n = 382)			External validation set (n = 181)			
	VPI positive (n = 285)	VPI negative (n = 1229)	P1 value	VPI positive (n = 73)	VPI negative (n = 309)	P1 value	VPI positive (n = 23)	VPI negative (n = 158)	P1 value	P2 value
CT subtype of pleura			<0.001			<0.001			0.025	0.663
Type 1	63 (22.1%)	154 (12.5%)		22 (30.1%)	30 (9.7%)		7 (30.4%)	29 (18.4%)		
Type 2	83 (29.1%)	224 (18.2%)		20 (27.4%)	57 (18.4%)		3 (13.0%)	27 (17.1%)		
Type 3	96 (33.7%)	407 (33.1%)		24 (32.9%)	105 (34.0%)		6 (26.1%)	24 (15.2%)		
Type 4	23 (8.1%)	34 (2.8%)		3 (4.1%)	8 (2.6%)		0 (0%)	5 (3.2%)		
Type 5	10 (3.5%)	85 (6.9%)		3 (4.1%)	18 (5.8%)		6 (26.1%)	25 (15.8%)		
Type 6	10 (3.5%)	325 (26.4%)		1 (1.4%)	91 (29.4%)		1 (4.3%)	48 (30.4%)		
Pleural indentation			<0.001			<0.001			<0.001	0.695
Yes	265 (93.0%)	819 (66.6%)		69 (94.5%)	200 (64.7%)		14 (60.9%)	56 (35.4%)		
No	20 (7.0%)	410 (33.4%)		4 (5.5%)	109 (35.3%)		9 (39.1%)	102 (64.6%)		
Consolidation contact with the pleura			<0.001			<0.001			<0.001	0.828
Yes	135 (47.4%)	290 (23.6%)		40 (54.8%)	70 (22.7%)		14 (60.9%)	44 (27.8%)		
No	150 (52.6%)	939 (76.4%)		33 (45.2%)	239 (77.3%)		9 (39.1%)	114 (72.2%)		
Minimum distance from the lesion to the pleura (mm)[a]	3.31 ± 4.82	6.98 ± 7.62	<0.001	2.00 ± 2.92	8.19 ± 9.39	<0.001	4.00 ± 5.57	6.39 ± 7.78	0.077	0.144
Maximum diameter of contact with pleura (mm)[a]	6.35 ± 7.69	3.39 ± 6.63	<0.001	8.63 ± 9.67	2.93 ± 5.83	<0.001	7.09 ± 10.3	4.61 ± 7.85	0.280	0.858
Maximum diameter of consolidation contact with the pleura (mm)[a]	6.12 ± 7.54	2.87 ± 6.11	<0.001	8.27 ± 9.58	2.40 ± 5.19	<0.001	6.52 ± 9.93	3.16 ± 6.27	0.128	0.913
CTR et the pleural end[a]	0.461 ± 0.493	0.214 ± 0.399	<0.001	0.521 ± 0.489	0.203 ± 0.391	<0.001	0.363 ± 0.481	0.233 ± 0.403	0.229	0.897

Note. — Unless otherwise noted, values are numbers of patients, with percentages in parentheses. CTR, consolidation tumor ratio; VPI, visceral pleural invasion.
[a]Data are means ± standard deviations. P1 value was derived from the univariable association analyses between VPI positive and VPI negative. P2 value was derived from the univariable association analyses between training set and internal validation set.

even following adjustment for clinical variables (Table 5). Survival curves for these patients are shown in Figure 5. The median DFS of patients with and without pathological VPI was 89.5 and 92.0 months, respectively (p = 0.01) (Figure 5A), with corresponding predicted 5-year survival rates of 68.2% and 81.5%. Patients were additionally assigned to two groups using the clinical, DL, and combined models. The median DFS of patients predicted to be VPI present and absent using the clinical model was 93.6 and 96.2 months, respectively (p = 0.33) (Figure 5B), with corresponding predicted 5-year survival rates of 75.9% and 79.5%. The median DFS of patients predicted to be VPI present and absent using the DL model was 92.3 and 97.0 months, respectively (p = 0.019) (Figure 5C), with corresponding predicted 5-year survival rates of 74.7% and 84.5%. The median DFS of patients predicted to be VPI present and absent using the combined model, the median DFS was 92.5 and 93.1 months, respectively (p = 0.043) (Figure 5D), with corresponding predicted 5-year survival rates of 74.5% and 84.9%.

There were 355 patients with pathologic stage I lung adenocarcinoma were evaluated for subgroup survival analyses. The median DFS was 67.4 months, and 24.8% (88/355) exhibited unfavorable prognosis. These patients exhibited a 68.8-month median follow-up duration. Univariate Cox regression analyses results are showed in Table S7. After multivariate analyses, smoking history, CTR, clinical T stage (cT1b and cT1c), pathologic VPI, and VPI status predicted by DL and combined models remained significant predictors even following adjustment for clinical

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

Factors	Univariate logistic regression		multivariate logistic regression	
	OR (95%CI)	p value	OR (95%CI)	p value
Age	1.005 (0.992–1.018)	0.423		
Sex	0.930 (0.718–1.205)	0.583		
Symptom	0.988 (0.758–1.292)	0.930		
Smoking history	0.995 (0.740–1.326)	0.972		
Family history of cancer	1.588 (1.128–2.209)	0.006	1.390 (0.952–2.012)	0.083
Clinical T stage				
cT1b	4.666 (2.807–8.282)	<0.001	1.162 (0.525–2.689)	0.717
cT1c	9.681 (5.819–17.21)	<0.001	0.977 (0.324–3.122)	0.996
Whole maximum diameter	1.074 (1.050–1.100)	<0.001	1.041 (0.861–1.265)	0.674
Consolidation maximum diameter	1.108 (1.086–1.131)	<0.001	1.004 (0.818–1.230)	0.968
Density	4.254 (3.013–6.157)	<0.001	1.407 (0.832–2.420)	0.208
CTR	46.07 (18.29–130.8)	<0.001	2.975 (0.037–341.0)	0.638
Location				
RML	1.810 (1.124–2.866)	0.053		
RLL	1.169 (0.813–1.672)	0.394		
LUL	1.057 (0.747–1.490)	0.752		
LLL	0.746 (0.472–1.155)	0.110		
Shape	1.642 (0.697–4.817)	0.304		
Boundary	1.760 (1.269–2.485)	<0.001	1.194 (0.813–1.778)	0.372
Vacuole sign	1.211 (0.872–1.661)	0.244		
Cavity	0.930 (0.399–1.919)	0.855		
Calcification	3.640 (1.042–12.17)	0.054		
Spiculated sign	2.801 (2.058–3.872)	<0.001	1.342 (0.913–1.993)	0.139
Lobulated sign	12.13 (5.088–39.63)	<0.001	4.919 (1.880–16.97)	0.004
Air bronchogram	0.609 (0.429–0.850)	0.004	0.853 (0.567–1.267)	0.437
Vessel convergence	3.238 (2.115–5.191)	<0.001	0.778 (0.444–1.395)	0.389
CT subtype of pleura				
Type 1	13.20 (6.937–28.20)	<0.001	0.401 (0.145–1.063)	0.071
Type 2	12.04 (6.405–25.20)	<0.001	0.354 (0.138–0.866)	0.026
Type 3	7.666 (4.127–15.90)	<0.001	0.439 (0.238–0.819)	0.009
Type 4	21.99 (9.919–52.06)	<0.001	NA	
Type 5	3.824 (1.522–9.610)	0.004	2.644 (0.995–7.051)	0.049
Pleural indentation	6.633 (4.253–10.94)	<0.001	9.162 (3.443–25.93)	<0.001
Consolidation contact with the pleura	2.914 (2.230–3.808)	<0.001	2.657 (0.196–57.38)	0.484
Minimum distance from the lesion to the pleura	0.903 (0.878–0.927)	<0.001	0.986 (0.943–1.028)	0.523
Maximum diameter of contact with pleura	1.054 (1.037–1.072)	<0.001	0.990 (0.813–1.127)	0.896
Maximum diameter of consolidation contact with the pleura	1.066 (1.048–1.085)	<0.001	0.980 (0.855–1.199)	0.811
CTR at the pleural end	3.323 (2.520–4.382)	<0.001	1.286 (0.054–20.35)	0.864

CI, confidence interval; CTR, consolidation tumor ratio; OR, odds ratio.

variables (Table 6). Survival curves for these patients are shown in Figure 6. The median DFS of patients with and without pathological VPI was 90.6 and 91.8 months, respectively (p = 0.029) (Figure 5A), with corresponding predicted 5-year survival rates of 69.9% and 81.3%. Patients were additionally assigned to VPI present and absent groups using the clinical, DL, and combined models. The median DFS of patients predicted to be VPI present and absent using the clinical model was 93.6 and 96.9 months, respectively (p = 0.28) (Figure 5B), with corresponding

Table 4. Prediction performance of the three models

Models	Dataset	AUC (95% CI)	ACC	SEN	SPE	PPV	NPV	F1 score
Clinical model	Training set	0.74 (0.72–0.75)	0.60	0.82	0.55	0.30	0.93	0.43
	Internal validation set	0.73 (0.70–0.77)	0.63	0.78	0.60	0.31	0.92	0.44
	External validation set	0.66 (0.59–0.72)	0.55	0.78	0.51	0.22	0.94	0.33
DL model	Training set	0.75 (0.73–0.76)	0.65	0.79	0.79	0.32	0.93	0.46
	Internal validation set	0.73 (0.69–0.76)	0.60	0.80	0.56	0.30	0.93	0.43
	External validation set	0.62 (0.54–0.68)	0.54	0.75	0.51	0.18	0.94	0.29
Combined model	Training set	0.76 (0.75–0.78)	0.66	0.78	0.63	0.33	0.93	0.46
	Internal validation set	0.75 (0.71–0.78)	0.63	0.83	0.58	0.32	0.94	0.46
	External validation set	0.69 (0.62–0.75)	0.69	0.65	0.70	0.23	0.93	0.35

ACC, accuracy; AUC, area under the receiver operating characteristic curve; CI, confidence interval; DL, deep learning; NPV, negative predictive value; PPV, positive predictive value; SEN, sensitivity; SPE, specificity.

predicted 5-year survival rates of 76.6% and 80.3%. The median DFS of patients predicted to be VPI present and absent using the DL model was 92.6 and 98.7 months, respectively (p = 0.017) (Figure 5C), with corresponding 5-year survival rates of 75.4% and 86.7%. The median DFS of patients predicted to be VPI present and absent using the combined model, the median DFS was 92.7 and 95 months, respectively (p = 0.026) (Figure 5D), with corresponding 5-year survival rates of 74.9% and 88%.

DISCUSSION

In the present study, DL features were extracted from CT images and used to develop a 3D-ResNet-9 DL model that was subsequently validated for its ability to predict the VPI status of patients with clinical stage IA lung adenocarcinoma in a noninvasive manner. These data highlight the value of DL features, as the resultant model was the most effective predictor of VPI with respective AUC values of 0.75, 0.73, and 0.62 in the training, internal validation, and external validation cohorts, respectively. While the combined model did not markedly outperform the DL or clinical models, the DL model nonetheless yielded predictive accuracy similar to that of a model based solely on clinical-radiologic characteristics. High-risk status as predicted with the DL model (HR for clinical stage IA DFS, 1.639 [p = 0.021], HR for pathologic stage I DFS, 10.67 [p = 0.008]) or the combined model (HR for clinical stage IA DFS, 1.519 [p = 0.037], HR for clinical stage IA DFS, 10.66 [p = 0.014]) and the presence of VPI (HR for clinical stage IA DFS, 1.544 [p = 0.011], HR for pathologic stage I DFS, 2.344 [p = 0.013]) were all found to be related to poor prognostic outcomes. In line with these results, VPI has repeatedly been identified as a predictor of poor prognostic outcomes.[4,20,21] Moreover, VPI status as predicted by both the DL and combined models was independently associated with patient DFS in multivariate analyses. DL model has the advantages of being simpler and assisting in prognostic assessment compared with the clinical model due to it could captures useful attributes, which are independent from human-defined semantic features from imaging data without any feature engineering.

VPI in patients with early-stage NSCLC is associated with an elevated risk of recurrence and mortality,[4,22] particularly in node-negative patients with tumors ≤3.0 cm in diameter.[23] VPI may be associated with tumor cell dissemination via the subpleural lymphatic pathway to the lymph nodes.[24] The eighth edition TNM classification system recommends upstaging NSCLC tumors ≤3.0 cm in size with VPI from stage T1 to stage T2 and from stage IA to stage IB.[6] In some prior reports, patients with stage IB disease have been suggested to benefit from adjuvant therapy.[25,26] The ability to preoperatively predict VPI status may thus be invaluable, and there is evidence that certain CT morphological features are correlated with VPI.[8–11,27–30] For example, analyses of pleural retraction, pleural attachment, and pleural tags have all been reported
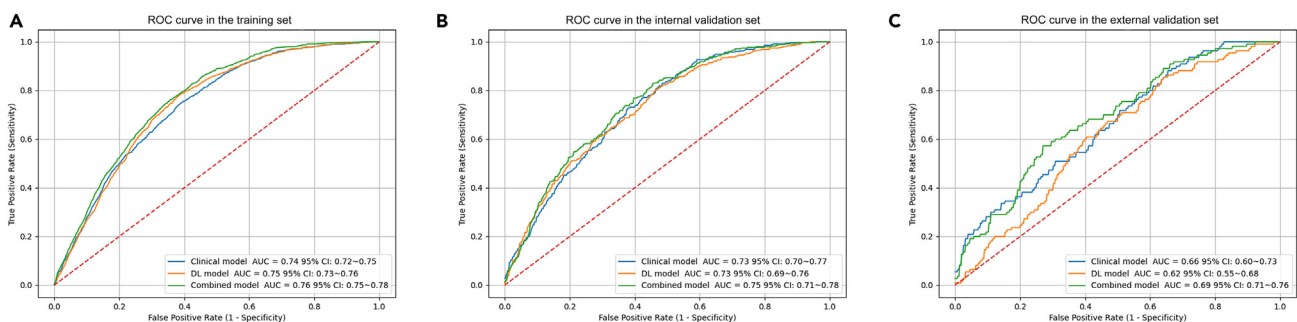
Figure 2. Comparison of the performance of three models for predicting VPI status in clinical stage IA lung adenocarcinoma
ROC curves for the models used to predict visceral pleural invasion status in the training (A), internal validation (B) and external validation (C) cohorts. AUC, area of curves; CI, confidence interval; DL, deep learning.
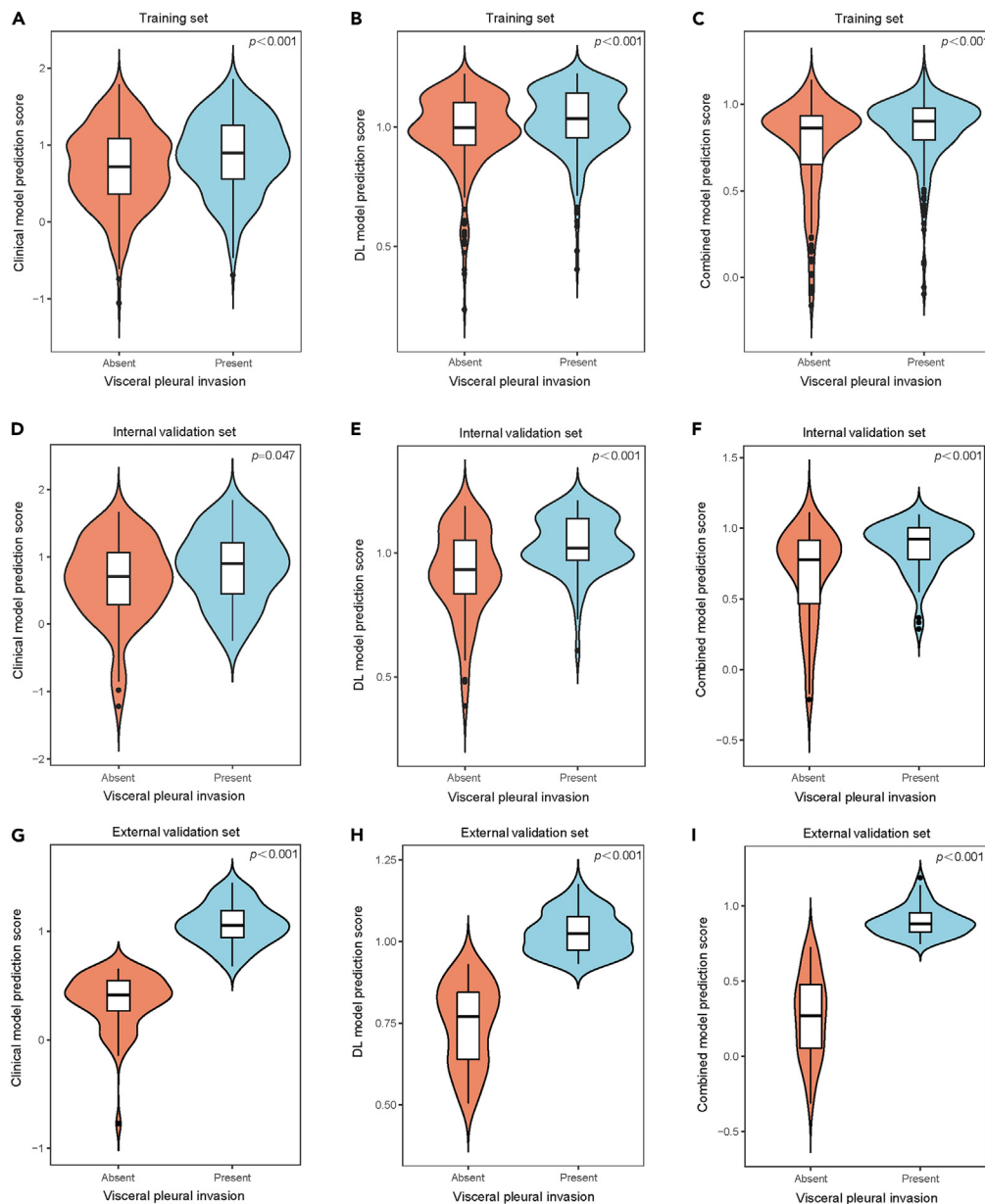
Figure 3. Difference in the three models prediction score between clinical stage IA lung adenocarcinoma with VPI and without VPI

(A–I) Boxplots corresponding to the clinical model (A, D, G), DL model (B, E, H), and combined model (C, F, I). Data are shown pertaining to visceral pleural invasion in the training cohort (A–C), the internal validation cohort (D–F), and external validation set (G–I). The prediction scores of these three models were significantly associated with visceral pleural invasion (all p < 0.05). DL, deep learning.

to be associated with VPI status, and Kim et al. found that CT features exhibit accuracy levels ranging from 62.7 to 72.3% when predicting VPI status, with corresponding PPV values from 44.1% to 56.4%, suggesting that roughly half of all CT-based predictions may yield false positive results.[25] When analyzing the association between VPI status and pleural tags in patients with NSCLC lacking CT evidence of contact with the pleural surface, Hsu et al. determined that the presence of one or more linear pleural tags with a soft tissue component at the pleural end was sufficient to diagnose VPI with an accuracy level of 71%,[9] similar to that for type 3 in this study. The etiological basis for type 3 may be related to tumor cell proliferation along lymphatic vessels together with peritumoral inflammatory exudation or fibrosis, ultimately resulting in peritumoral lymphatic vessel obstruction that contributes to VPI incidence.[9] Yang et al. evaluated the relationship between lesions and the pleura in two groups of patients defined according to the proximity of the lesion to the interlobar region, ultimately revealing a higher VPI positivity rate in type I patients in group B consistent with higher rates of pulling of the interlobar pleura.[29] This result aligns well with observations in type 2 cases in the present study. Imai et al. also observed positive correlations between the odds of pleural invasion and the length of focal contact
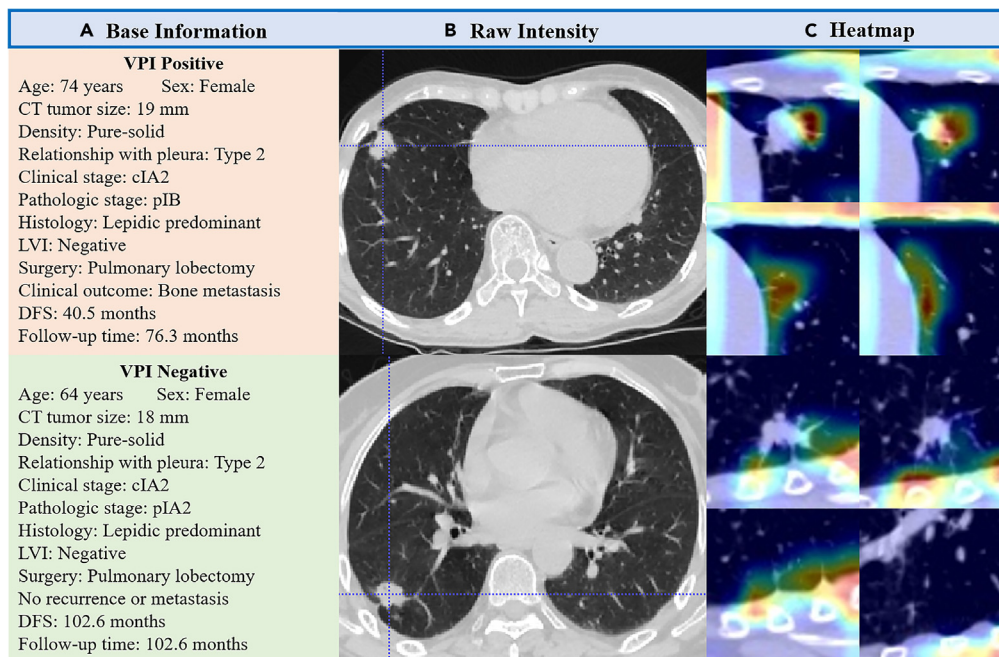
Figure 4. Gradient-weighted class activation maps for the interpretation of model output

(A–C) Clinicopathological details (A), CT images (B), and corresponding heatmaps (C) for two representative cases positive and negative for VPI. While the clinicopathological characteristics of these patients were similar other than their VPI status, VPI-positive patients with bone metastasis exhibited a significant drop in DFS relative to VPI-negative patients free of recurrent or metastatic disease. Heatmaps revealed the areas most relevant to the prediction of VPI status, which were partially in contact with the adjacent pleura, as well as in the peritumoral zone even if it was distant from the pleura.

with the pleura,[12] consistent with these results. Indeed, pleural tags and pleural indentation are radiological signs that have widely been recognized as indicators of pulmonary malignancy and possible VPI.[11] Here, pleural indentation was confirmed to be an independent predictor of VPI. Pleural indentation has been reported to arise as a result of intratumoral scar formation, which is related to VPI as a result of tumor-induced pulmonary atelectasis.[31] The present data suggest that lobulation can also predict VPI in patients diagnosed with clinical stage IA lung adenocarcinoma, consistent with evidence that lobulated contours are positively correlated with cell growth rates and are thus related to malignancy.[32] While most reports have focused on assessing VPI-related morphological characteristics in CT images, these studies have exhibited unsatisfactory accuracy, with feature identification being dependent on radiologist expertise.[9,27,33]

To date, few reports have sought to apply DL techniques as a means of predicting VPI status. Choi et al. proposed the development of a DL model for the prediction of VPI in patients with lung cancer, ultimately achieving an AUC of 0.75 that was similar to that achieved by radiologists.[34] This AUC was identical to that in the present report. However, in that study, radiologists made diagnoses based on a 5-point scoring system without any definitive criteria such that the reproducibility of these measurements was low. Choi et al. also included patients with a range of higher-stage lung cancers in addition to stage IA disease when conducting their analyses. In the present analysis, model outputs were useful as independent predictors of VPI status in individuals with clinical stage IA lung adenocarcinoma, providing an opportunity to discover latent characteristics underlying clinical assessments and semantic features. In addition, the VPI status predicted by our models

Table 5. Multivariate Cox regression analysis for disease-free survival in in clinical stage IA lung adenocarcinomas

Variable	Pathologic VPI		Clinical model		DL model		Combined model	
	HR (95%CI)	p value	HR (95%CI)	p value	HR (95%CI)	p value	HR (95%CI)	p value
Age	1.015 (0.997,1.033)	0.105	1.014 (0.996,1.032)	0.121	1.014 (0.996,1.032)	0.118	1.015 (0.997,1.033)	0.106
CTR	7.827 (1.011,60.61)	0.049	8.205 (1.056,63.74)	0.044	8.037 (1.032,62.59)	0.047	8.172 (1.057,63.18)	0.044
Clinical T stage		0.031		0.018		0.033		0.026
cT1b	1.896 (0.725,4.954)	0.192	1.933 (0.740,5.047)	0.179	1.849 (0.705,4.852)	0.212	1.860 (0.713,4.852)	0.205
cT1c	2.728 (1.033,7.205)	0.043	2.867 (1.089,7.551)	0.033	2.686 (1.008,7.157)	0.048	2.721 (1.033,7.169)	0.043
VPI	2.365 (1.004,6.912)	0.041	1.111 (0.786,1.571)	0.552	2.197 (1.778,7.843)	0.043	2.256 (1.032,6.896)	0.048

CI, confidence interval; DL, deep learning; HR, hazard ratio; VPI, visceral pleural invasion.
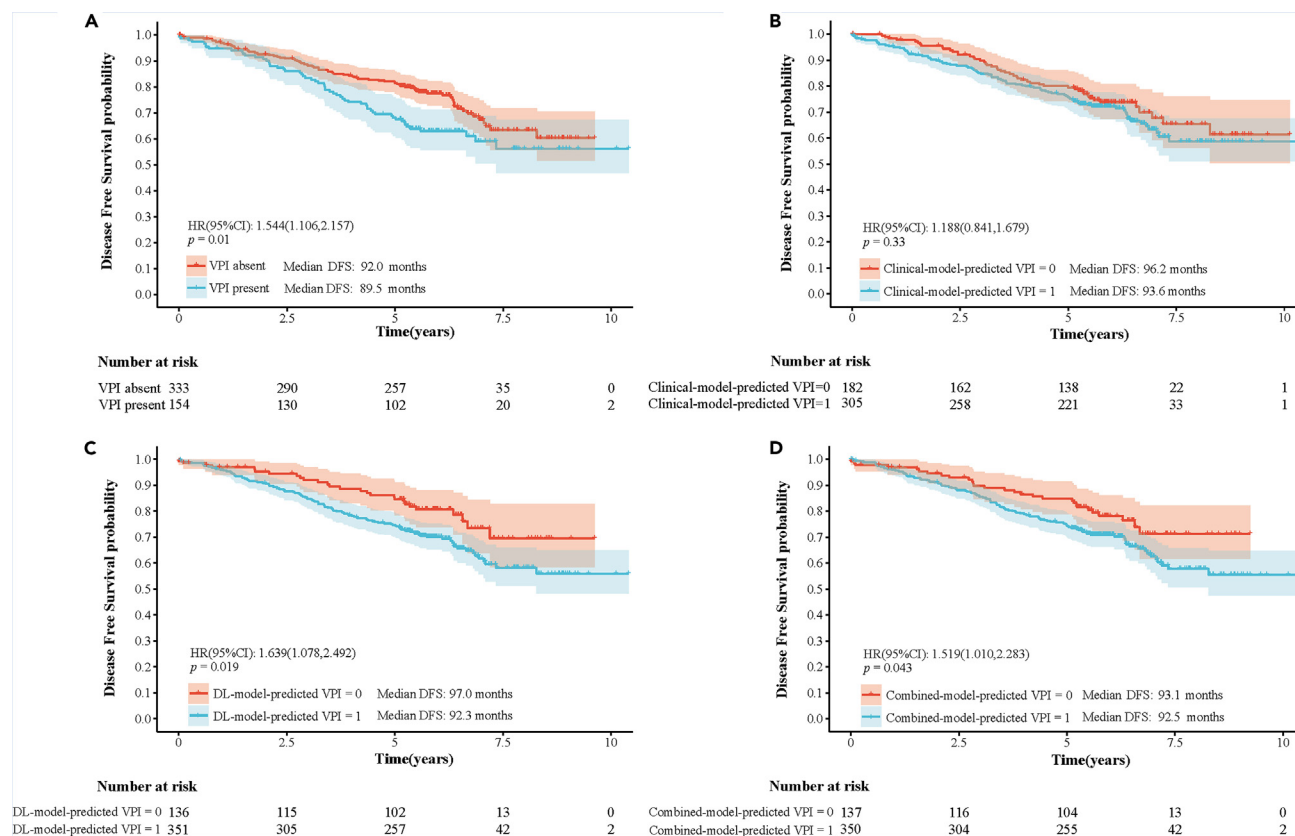
Figure 5. Prognostic value of pathological VPI status and VPI status predicted by the three models in clinical stage IA lung adenocarcinoma
(A–D) Kaplan-Meier survival curves evaluating the outcomes of clinical stage IA lung adenocarcinoma grouped according to their pathological VPI status (A), or their VPI status as predicted by the clinical model (B), DL model (C), and combined model (D). CI, confidence interval; DFS, disease-free survival; DL, deep learning; HR, hazard ratio; VPI, visceral pleural invasion.

was highly correlated with LVI and lymph node metastasis, which was consistent with the actual clinical situation of the training and the internal validation set as shown in Table 1. This suggests that the models in this study may help to access LVI and lymph node metastasis. Since the positive proportion of pathological factors in the external validation set included in this study was low, the above findings need to be verified by larger external data. The heatmap generated in this study indicated that the region most closely associated with VPI status was proximal to the tumor pleura, consistent with a prior report.[24] The output of this model may thus offer value as an autonomous variable that can support clinical decision-making or serve as an anatomical determinant in prognostic models used to evaluate patients with clinical stage IA lung adenocarcinoma.

In patients with NSCLC, VPI reportedly serves as a positive prognostic factor,[5,24,35] consistent with the present results. Here, VPI was employed to stratify patients, and the ability of the DL model to predict patient DFS was assessed. In Kaplan-Meier analyses, the DL model was reliably able to separate patients with lung adenocarcinoma into groups facing a low or a high risk of mortality. The VPI predictions made by the model were also found to be independently associated with DFS in multivariate analyses. Kim et al. determined that CT semantic features of VPI were not independently predictive of DFS in patients with clinical T1 lung adenocarcinoma.[27] Morphological CT features primarily depend upon radiologist input. In an effort to more accurately and reproducibly document the association between CT image data and DFS, the present study employed a DL algorithm in order to extract latent features not otherwise represented by more traditional clinical-radiologic semantic features.

Here, the 3D-ResNet-9 architecture was used to develop a DL model capable of using CT image data to predict the VPI status of patients with clinical stage IA lung adenocarcinoma. DL model developed herein offers excellent potential as a means of accurately assessing VPI status prior to surgery. Meanwhile, the DL model was also able to identify predicted VPI in a manner independently associated with DFS risk such that the established model may aid clinicians in the context of preoperative decision-making.

Limitations of the study

This study is subject to multiple limitations. For one, the retrospective design of this study renders it susceptible to selection bias. Additionally, various CT scanners were used when evaluating these patients, and the acquisition protocols that were used were also inconsistent across

Table 6. Multivariate Cox regression analysis for disease-free survival in pathologic stage I lung adenocarcinomas

Variable	Pathologic VPI		Clinical model		DL model		Combined model	
	HR (95%CI)	p value	HR (95%CI)	p value	HR (95%CI)	p value	HR (95%CI)	p value
Smoking history (ever smoker)	1.568 (1.027, 2.395)	0.037	1.557 (1.020, 2.377)	0.040	1.553 (1.017, 2.371)	0.042	1.571 (1.027, 2.403)	0.037
CTR	4.647 (0.498, 43.38)	0.178	4.906 (0.525, 45.84)	0.163	3.807 (0.392, 36.99)	0.249	4.911 (0.527, 45.80)	0.162
Clinical T stage		0.018		0.013		0.026		0.015
cT1b	1.561 (0.514, 4.746)	0.043	1.562 (0.513, 4.754)	0.043	1.459 (0.476, 4.477)	0.049	1.552 (0.511, 4.716)	0.044
cT1c	2.778 (0.899, 8.580)	0.037	2.847 (0.923, 8.776)	0.039	2.557 (0.818, 7.997)	0.017	2.79 6(0.907, 8.621)	0.033
VPI	2.227 (0.784, 8.920)	0.037	1.126 (0.748, 1.694)	0.570	4.691 (1.712, 30.90)	0.011	4.315 (1.512, 23.38)	0.027

CI, confidence interval; DL, deep learning; HR, hazard ratio; VPI, visceral pleural invasion.

reports. Even so, we believe that the diversity of these data is representative of variations observed in clinical datasets, thus making any DL model trained using these data better suited to clinical application in a real-world setting. Lastly, the survival data used for this study were derived from a relatively small subset of cases owing to the fact that follow-up was only performed for five years. Future large-scale trials will thus be vital to validate and expand on these results.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - ○ Lead contact
  - ○ Materials availability
  - ○ Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - ○ Study participants
- METHOD DETAILS
  - ○ Clinicopathologic data collection
  - ○ CT image acquisition
  - ○ Radiological feature analyses
  - ○ Pleural CT subtypes
  - ○ CT image preprocessing
  - ○ 3D-ResNet-9 model architecture
  - ○ Model construction and validation
  - ○ Evaluation of clinical application of model
  - ○ Follow-up and survival analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108712.

AUTHOR CONTRIBUTIONS

Conceptualization, L.L., H.C., and S.L.; methodology, X.F.L., K.F.L., K.W.L, X.J.C., and B.Y.C.; Ssoftware, X.F.L., K.F.L., and H.J.; formal analysis, X.F.L.,K.F.L., and K.W.L; investigation, L.L., H.C., and S.L.; writing–original draft, X.F.L. and K.F.L.; writing–review and editing, K.W.L, L.L., H.C., and S.L.; visualization, K.F.L., and H.J.; supervision, L.L., H.C., and S.L.; funding acquisition, L.L. and H.C.
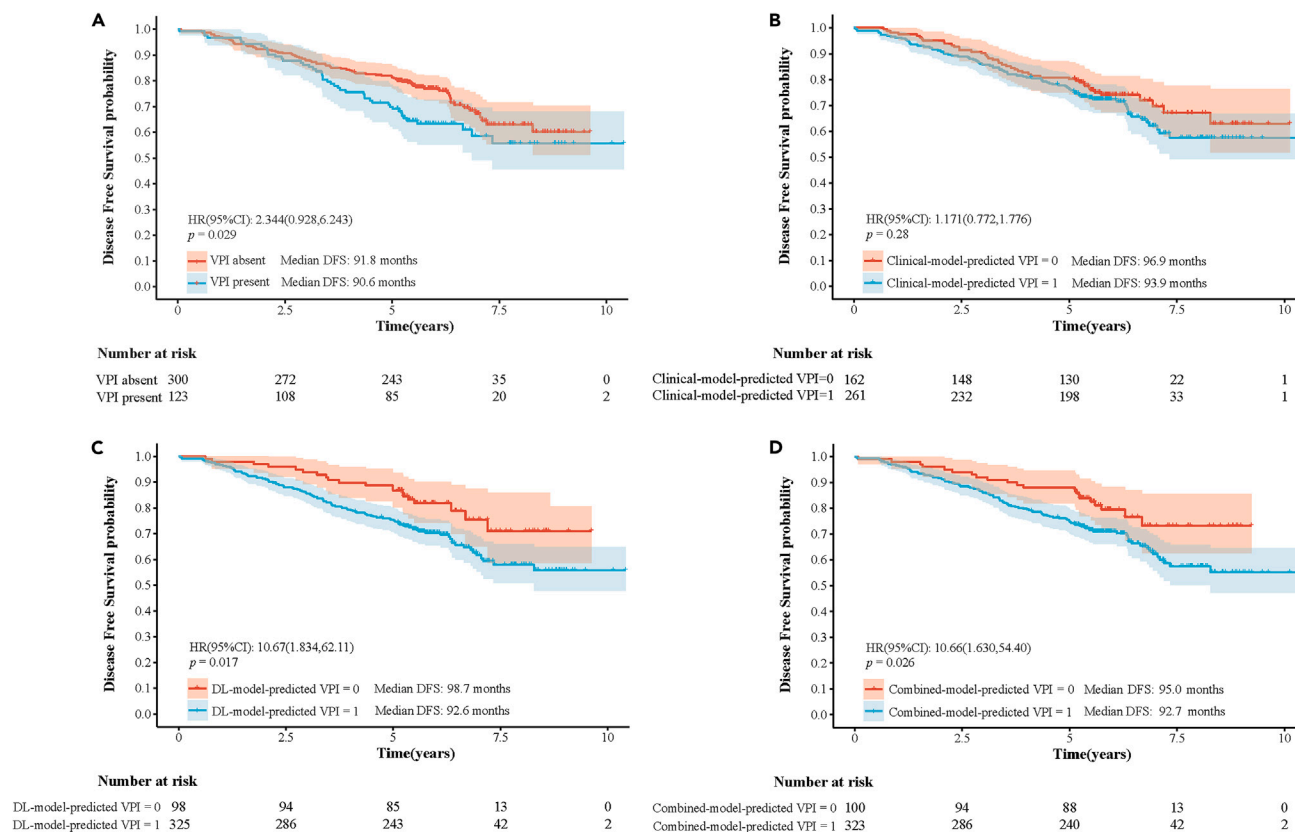
Figure 6. Prognostic value of pathological VPI status and VPI status predicted by the three models in pathologic stage I lung adenocarcinoma

(A–D) Kaplan-Meier survival curves evaluating the outcomes of pathologic stage I lung adenocarcinoma grouped according to their pathological VPI status (A), or their VPI status as predicted by the clinical model (B), DL model (C), and combined model (D). CI, confidence interval; DFS, disease-free survival; DL, deep learning; HR, hazard ratio; VPI, visceral pleural invasion.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Siegel, R.L., Miller, K.D., Fuchs, H.E., and Jemal, A. (2022). Cancer statistics, 2022. CA. Cancer J. Clin. *72*, 7–33.

2. Goldstraw, P., Chansky, K., Crowley, J., Rami-Porta, R., Asamura, H., Eberhardt, W.E.E., Nicholson, A.G., Groome, P., Mitchell, A., Bolejack, V., et al. (2016). The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. J. Thorac. Oncol. *11*, 39–51.

3. Cao, C., Chandrakumar, D., Gupta, S., Yan, T.D., and Tian, D.H. (2015). Could less be more? — A systematic review and meta-analysis of sublobar resections versus lobectomy for non-small cell lung cancer

according to patient selection. Lung Cancer *89*, 121–132.

4. Huang, H., Wang, T., Hu, B., and Pan, C. (2015). Visceral Pleural Invasion Remains a Size-Independent Prognostic Factor in Stage I Non-Small Cell Lung Cancer. Ann. Thorac. Surg. *99*, 1130–1139.

5. Yoshida, J., Nagai, K., Asamura, H., Goya, T., Koshiishi, Y., Sohara, Y., Eguchi, K., Mori, M., Nakanishi, Y., Tsuchiya, R., et al. (2009). Visceral pleura invasion impact on non-small cell lung cancer patient survival: its implications for the forthcoming TNM staging based on a large-scale nation-wide database. J. Thorac. Oncol. *4*, 959–963.

6. Rami-Porta, R., Bolejack, V., Crowley, J., Ball, D., Kim, J., Lyons, G., Rice, T., Suzuki, K., Thomas, C.F., Jr., Travis, W.D., et al. (2015). The

IASLC Lung Cancer Staging Project: Proposals for the Revisions of the T Descriptors in the Forthcoming Eighth Edition of the TNM Classification for Lung Cancer. J. Thorac. Oncol. *10*, 990–1003.

7. Takizawa, H., Kondo, K., Kawakita, N., Tsuboi, M., Toba, H., Kajiura, K., Kawakami, Y., Sakiyama, S., Tangoku, A., Morishita, A., et al. (2018). Autofluorescence for the diagnosis of visceral pleural invasion in non-small-cell lung cancer. Eur. J. Cardio. Thorac. Surg. *53*, 987–992.

8. Onoda, H., Higashi, M., Murakami, T., Tao, H., Yokoyama, S., Kunihiro, Y., Kawano, R., Tanabe, M., Tanaka, N., and Matsumoto, T. (2021). Correlation between pleural tags on CT and visceral pleural invasion of peripheral lung cancer that does not appear touching

the pleural surface. Eur. Radiol. *31*, 9022–9029.

9. Hsu, J.S., Han, I.T., Tsai, T.H., Lin, S.F., Jaw, T.S., Liu, G.C., Chou, S.H., Chong, I.W., and Chen, C.Y. (2016). Pleural Tags on CT Scans to Predict Visceral Pleural Invasion of Non-Small Cell Lung Cancer That Does Not Abut the Pleura. Radiology *279*, 590–596.

10. Ahn, S.Y., Park, C.M., Jeon, Y.K., Kim, H., Lee, J.H., Hwang, E.J., and Goo, J.M. (2017). Predictive CT Features of Visceral Pleural Invasion by T1-Sized Peripheral Pulmonary Adenocarcinomas Manifesting as Subsolid Nodules. AJR Am. J. Roentgenol. *209*, 561–566.

11. Zhao, L.L., Xie, H.K., Zhang, L.P., Zha, J.Y., Zhou, F.Y., Jiang, G.N., and Chen, C. (2016). Visceral pleural invasion in lung adenocarcinoma </=3 cm with ground-glass opacity: a clinical, pathological and radiological study. J. Thorac. Dis. *8*, 1788–1797.

12. Imai, K., Minamiya, Y., Ishiyama, K., Hashimoto, M., Saito, H., Motoyama, S., Sato, Y., and Ogawa, J.i. (2013). Use of CT to evaluate pleural invasion in non-small cell lung cancer: measurement of the ratio of the interface between tumor and neighboring structures to maximum tumor diameter. Radiology *267*, 619–626.

13. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Network. *61*, 85–117.

14. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. (2017). Mastering the game of Go without human knowledge. Nature *550*, 354–359.

15. Ehteshami Bejnordi, B., Veta, M., Johannes van Diest, P., van Ginneken, B., Karssemeijer, N., Litjens, G., van der Laak, J.A.W.M., the CAMELYON16 Consortium, Hermsen, M., Manson, Q.F., et al. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer. JAMA *318*, 2199–2210.

16. Zhao, X., Wang, X., Xia, W., Li, Q., Zhou, L., Li, Q., Zhang, R., Cai, J., Jian, J., Fan, L., et al. (2020). A cross-modal 3D deep learning for accurate lymph node metastasis prediction in clinical stage T1 lung adenocarcinoma. Lung Cancer *145*, 10–17.

17. Kim, H., Goo, J.M., Lee, K.H., Kim, Y.T., and Park, C.M. (2020). Preoperative CT-based Deep Learning Model for Predicting Disease-Free Survival in Patients with Lung Adenocarcinomas. Radiology *296*, 216–224.

18. Ohno, Y., Aoyagi, K., Yaguchi, A., Seki, S., Ueno, Y., Kishida, Y., Takenaka, D., and Yoshikawa, T. (2020). Differentiation of Benign from Malignant Pulmonary Nodules by Using a Convolutional Neural Network to Determine Volume Change at Chest CT. Radiology *296*, 432–443.

19. Feng, B., Chen, X., Chen, Y., Lu, S., Liu, K., Li, K., Liu, Z., Hao, Y., Li, Z., Zhu, Z., et al. (2020). Solitary solid pulmonary nodules: a CT-based deep learning nomogram helps differentiate tuberculosis granulomas from lung

adenocarcinomas. Eur. Radiol. *30*, 6497–6507.

20. Iizuka, S., Kawase, A., Oiwa, H., Ema, T., Shiiya, N., and Funai, K. (2019). A risk scoring system for predicting visceral pleural invasion in non-small lung cancer patients. Gen. Thorac. Cardiovasc. Surg. *67*, 876–879.

21. Fibla, J.J., Cassivi, S.D., Brunelli, A., Decker, P.A., Allen, M.S., Darling, G.E., Landreneau, R.J., and Putnam, J.B. (2012). Re-evaluation of the prognostic value of visceral pleura invasion in Stage IB non-small cell lung cancer using the prospective multicenter ACOSOG Z0030 trial data set. Lung Cancer *78*, 259–262.

22. Liu, Q.X., Deng, X.F., Zhou, D., Li, J.M., Min, J.X., and Dai, J.G. (2016). Visceral pleural invasion impacts the prognosis of non-small cell lung cancer: A meta-analysis. Eur. J. Surg. Oncol. *42*, 1707–1713.

23. Naito, Y., Goto, K., Nagai, K., Ishii, G., Nishimura, M., Yoshida, J., Hishida, T., and Nishiwaki, Y. (2010). Vascular Invasion Is a Strong Prognostic Factor After Complete Resection of Node-Negative Non-small Cell Lung Cancer. Chest *138*, 1411–1417.

24. Shimizu, K., Yoshida, J., Nagai, K., Nishimura, M., Ishii, G., Morishita, Y., and Nishiwaki, Y. (2005). Visceral pleural invasion is an invasive and aggressive indicator of non-small cell lung cancer. J. Thorac. Cardiovasc. Surg. *130*, 160–165.

25. Zhu, C.Q., Ding, K., Strumpf, D., Weir, B.A., Meyerson, M., Pennell, N., Thomas, R.K., Naoki, K., Ladd-Acosta, C., Liu, N., et al. (2010). Prognostic and Predictive Gene Signature for Adjuvant Chemotherapy in Resected Non–Small-Cell Lung Cancer. J. Clin. Oncol. *28*, 4417–4424.

26. Williams, C.D., Gajra, A., Ganti, A.K., and Kelley, M.J. (2014). Use and impact of adjuvant chemotherapy in patients with resected non-small cell lung cancer. Cancer *120*, 1939–1947.

27. Kim, H., Goo, J.M., Kim, Y.T., and Park, C.M. (2019). CT-defined Visceral Pleural Invasion in T1 Lung Adenocarcinoma: Lack of Relationship to Disease-Free Survival. Radiology *292*, 741–749.

28. Kim, H.J., Cho, J.Y., Lee, Y.J., Park, J.S., Cho, Y.J., Yoon, H.I., Chung, J.H., Cho, S., Kim, K., Lee, K.W., et al. (2019). Clinical Significance of Pleural Attachment and Indentation of Subsolid Nodule Lung Cancer. Cancer Res. Treat. *51*, 1540–1548.

29. Yang, S., Yang, L., Teng, L., Zhang, S., Cui, Y., Cao, Y., and Shi, H. (2018). Visceral pleural invasion by pulmonary adenocarcinoma ≤ 3 cm: the pathological correlation with pleural signs on computed tomography. J. Thorac. Dis. *10*, 3992–3999.

30. Heidinger, B.H., Schwarz-Nemec, U., Anderson, K.R., de Margerie-Mellon, C., Monteiro Filho, A.C., Chen, Y., Mayerhoefer, M.E., VanderLaan, P.A., and Bankier, A.A. (2019). Visceral Pleural Invasion in Pulmonary Adenocarcinoma: Differences in CT Patterns between Solid and Subsolid Cancers. Radiol. Cardiothorac. Imaging *1*, e190071.

31. Wang, F., Lin, X., Lin, C., Huang, G., Li, M., and Zhu, L. (2023). Ability of three-dimensional 3-Tesla ultrashort echo time magnetic resonance imaging to display the morphological characteristics of pulmonary nodules: a sensitivity analysis. Quant. Imag. Med. Surg. *13*, 1792–1801.

32. Ost, D., Fein, A.M., and Feinsilver, S.H. (2003). Clinical practice. The solitary pulmonary nodule. N. Engl. J. Med. *348*, 2535–2542.

33. Gillies, R.J., Kinahan, P.E., and Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. Radiology *278*, 563–577.

34. Choi, H., Kim, H., Hong, W., Park, J., Hwang, E.J., Park, C.M., Kim, Y.T., and Goo, J.M. (2021). Prediction of visceral pleural invasion in lung cancer on CT: deep learning model achieves a radiologist-level performance with adaptive sensitivity and specificity to clinical needs. Eur. Radiol. *31*, 2866–2876.

35. Jiang, L., Liang, W., Shen, J., Chen, X., Shi, X., He, J., Yang, C., and He, J. (2015). The Impact of Visceral Pleural Invasion in Node-Negative Non-small Cell Lung Cancer. Chest *148*, 903–911.

36. Rami-Porta, R., Asamura, H., Travis, W.D., and Rusch, V.W. (2017). Lung cancer - major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA. Cancer J. Clin. *67*, 138–155.

37. Zugazagoitia, J., Enguita, A.B., Nuñez, J.A., Iglesias, L., and Ponce, S. (2014). The new IASLC/ATS/ERS lung adenocarcinoma classification from a clinical perspective: current concepts and future prospects. J. Thorac. Dis. *6* (*Suppl 5*), S526–S536.

38. Hsu, J.S., Han, I.T., Tsai, T.H., Lin, S.F., Jaw, T.S., Liu, G.C., Chou, S.H., Chong, I.W., and Chen, C.Y. (2016). Pleural tags on CT scans to predict visceral pleural invasion of non–small cell lung cancer that does not about the pleura. Radiology *279*, 590–596.

39. Ahn, S.Y., Park, C.M., Jeon, Y.K., Kim, H., Lee, J.H., Hwang, E.J., and Goo, J.M. (2017). Predictive CT features of visceral pleural invasion by T1-sized peripheral pulmonary adenocarcinomas manifesting as subsolid nodules. AJR Am. J. Roentgenol. *209*, 561–566.

40. He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition (IEEE), pp. 770–778.

41. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., and Houlsby, N. (2020). An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale.

42. Nohara, Y., Matsumoto, K., Soejima, H., and Nakashima, N. (2022). Explanation of machine learning models using shapley additive explanation and application for real data in hospital. Comput. Methods Progr. Biomed. *214*, 106584.

43. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). GradCAM: Visual Explanations from Deep Networks Via Gradient-Based Localization. In IEEE International Conference on Computer Vision (IEEE), pp. 618–626.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Python (version 3.7.0)	Python software	https://www.python.org/
PyTorch (version 1.6.0)	PyTorch software	https://pytorch.org/
CUDA toolkit (version 9.2)	CUDA toolkit software	https://developer.nvidia.com/cuda-92-download-archive
R (version 3.5.3)	R software	https://www.r-project.org
Other		
Research Data Deposit	Sun Yat-Sen University Cancer Center	https://www.researchdata.org.cn (No.RDDA2023776696)
Source code	Github	https://github.com/Johnhave/lung_nodule_cls.git

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Li Li (lil@sysucc.org.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- De-identified patient standardized data have been deposited at the Research Data Deposit public platform (No.RDDA2023776696), and DOIs are listed in the key resources table. They are available upon request if access is granted. To request access, contact Sun Yat-Sen University Cancer Center.
- All original code has been deposited at the Github and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study participants

The ethical review board of Sun Yat-sen University Cancer Center and the First Affiliated Hospital of Guangzhou Medical University (B2022-293-01), the Fifth Affiliated Hospital of Sun Yat-sen University (K107-1) approved this study, and waived the requirement for written informed consent.

This retrospective analysis used electronic medical records to enroll consecutive lung cancer patients who had undergone surgical resection from January 2010 – December 2021 at Sun Yat-sen University Cancer Center (Center 1) or from January 2015 – December 2019 at The First Affiliated Hospital of Guangzhou Medical University (Center 2) or January 2016 – December 2018 at Fifth Affiliated Hospital of Sun Yat-sen University (Center 3). Inclusion criteria for this study (Figure 1) included the following: (i) patients that had pathology-confirmed lung adenocarcinoma diagnoses; (ii) patients preoperatively diagnosed with clinical stage IA disease; (iii) patients that underwent non-enhanced CT imaging within two weeks prior to surgery; and (iv) patients with CT findings of pure-solid and subsolid pulmonary nodules. Overall, these criteria led to the identification of 2980 clinical stage IA lung adenocarcinoma patients without synchronous or metachronous lung cancers at these three centers. Ineligible patients were excluded based on the following criteria: (i) patients with pathological reports lacking definitive LVI status; (ii) patients that underwent preoperative neoadjuvant chemo- and/or radiotherapy; or (iii) patients lacking thin CT images (≤2 mm). Ultimately, 2077 patients met all of these criteria, including 1070, 826 and 181 from Center 1, 2 and 3, respectively. These patients of center 1 and 2 were pooled together and then separated at an 8:2 ratio via simple randomization into a training cohort (n = 1514) and an internal validation cohort (n = 382). The cohort of 181 patients from center 3 was the external validation set. The training cohort was used for model training and the selection of the best model, while model performance was evaluated with the internal and external validation cohort. 5-fold stratified cross-validation was employed during both training and two validation sets, as detailed in Figure 2. In the training cohort, 285 patients (285/1514, 18.8%) were diagnosed with VPI, while the VPI status of 73 patients (73/382, 19.1%) in the internal validation cohort and 23 patients (23/181, 12.7%) in the external validation cohort were positive (Table 1).

METHOD DETAILS

Clinicopathologic data collection

Participant clinicopathologic data including age, sex, symptoms, family history of cancer, and history of smoking were obtained from the patient electronic medical records. The subjects in this study were all of the same Asian ethnicity from China. Sex was not statistically significant in any of the groups in this study (see Table 1) and therefore had no influence on the results of this study. The 8th edition of the American Joint Committee on Cancer (AJCC) staging manual was used for clinical and pathological staging.[36] Attending pathologists were responsible for establishing the pathologic diagnoses of patients in accordance with the International Association for the Study of Lung Cancer (IASLC)/American Thoracic Society/European Respiratory Society classification system.[37] Pathologists classified VPI based on the IASLC proposal after examining hematoxylin and eosin-stained sections as follows: PL0 (lack of pleural invasion beyond the elastic layer), PL1 (invasion beyond the elastic layer), PL2 (invasion of the visceral pleural surface).[37] When lesions were found to exhibit pleural adjacency and visceral pleural involvement was uncertain, the presence of VPI was detected via elastic staining. Repeated elastic staining during the study inclusion period was not performed to reassess VPI status. No separate recordings were made for PL1 and PL2 in this analysis. Lymph node metastasis, distant metastasis, and lymphovascular invasion (LVI) incidence were also documented in this report, with LVI being defined by tumor cells detectable via microscopy that were present in the lymphatic, arterial, or venous vessels of the surrounding pulmonary tissue.

CT image acquisition

CT scans were conducted from the apex to the base of the lung. Table S8 shows the details of the scanning and reconstruction parameters.

Radiological feature analyses

Maximum tumor diameter and maximum consolidation diameter were measured using electronic calipers and an appropriate axial CT image in an image archiving and communication system in the lung window setting (window width, 1500 HU; level, −600 HU). The maximum consolidation diameter and maximum tumor diameter were then used to calculate the consolidation tumor ratio (CTR). Three radiologists with 5, 10, and 15 years of experience performing thoracic imaging-based diagnosis came to a consensus regarding nodule location (right upper lung, right middle lung, right lower lung, left upper lung, and left lower lung), density (pure-solid or subsolid), shape (regular or irregular), boundary (clear or vague), whether lobulation (defined by irregular nodule margin undulation) was present, whether spiculation (2 mm or thicker strands extending from the tumor margin to the surrounding parenchyma but not extending to the pleura) reaching the pleura), vacuole sign (<5 mm regions of intratumoral round or ovoid air attenuation), cavity (≥5 mm regions of intratumoral round or ovoid air attenuation) air bronchogram (air-filled bronchi in the tumor), vessel convergence (multiple blood vessels gathering toward the tumor), and calcification. All reviewers remained blinded to patient pathologic information and outcomes, although nodule location was provided to these reviewers.

Pleural CT subtypes

Multiple CT findings have been reported in the context of VPI predictions.[12,38,39] Pleural deformation has, to date, been used to establish six subgroups of patients, including: (a) Type 1 patients exhibiting direct pleural contact; (b) Type 2 patients exhibiting pleural retraction; (c) Type 3 patients with one or more linear pleural tags with thickening at the pleural end; (d) Type 4 patients with one or more cord-like pleural tags with thickening at the pleural end; (e) Type 5 patients with pleural tags but no thickening at the pleural end; and (f) Type 6 patients exhibiting no evidence of pleural contact (see Figure S2). Three thoracic radiologists also determined whether or not pleural indentation (a linear or triangular strand originating from the tumor and extending to the pleura) or pleural consolidation contact was evident. In cases where direct pleural contact was not observed, the minimum distance between the lesion and the pleura was assessed. The maximum diameter of pleural contact was measured for cases exhibiting direct pleural contact or no pleural contact. In partially solid nodules, pleural contact of the ground-glass opacity regions was not considered indicative of VPI.[26] As such, the maximum diameter of pleural consolidation contact was also analyzed, and CTR at the pleural end was calculated based on the maximum diameter of pleural consolidation contact and the maximum diameter of pleural contact.

CT image preprocessing

To develop the DL sampling, CT data were initially preprocessed in the following steps: (1) CT images were scaled and subjected to cubic spline linear interpolation to exclude the effects of voxel space variability by generating 1 × 1 × 1 mm voxels; (2) a 64 × 64 × 64 mm 3D cube subgraph was cropped from the center of the nodule for each CT scan to eliminate any interference from irrelevant regions while maximally preserving the original image information; (3) Voxel values for each sample was normalized for each extracted subgraph to values from 0 to 1, with respective upper and lower values of 400 HU and −800HU, thus retaining the maximum possible amount of nodule information.

3D-ResNet-9 model architecture

This study employed an algorithm consisting of a convolutional neural network (CNN) and a fully connected network, with the CNN utilizing the customized 3D-ResNet-9 model to enable improved dataset classification.[40] The 3D-ResNet-9 model is a residual neural network that differs from similar networks with respect to the number of layers. The architecture of the 3D-ResNet-9 model consisted of the following.

(a) Initial convolution layer: The 3D-ResNet-9 structure initially employed a 3 × 3 convolution layer with a step size of 1 that was followed by a maximum pool (maxpool) layer. This portion of the model enabled preliminary feature extraction.

(b) Residual blocks: The main components of the 3D-ResNet-9 model were formed from four residual blocks, each containing two 3D convolutional layers operated via batch normalization (BN) and ReLU activation. BN corresponds to the normalization of data feature maps such that the average value is 0 and the variance is 1. To minimize training time and resource utilization, channel numbers were decreased to 16, 32, 64, and 96. The input was then connected to the output of these convolutional layers via a skip join, enabling the network to learn the residual input-to-output mapping while avoiding the issue of exploding or disappearing gradients.

(c) Average global pool (avgpool) layer: Following the residual blocks step, 3D-ResNet-9 employs an avgpool layer rather than the fully connected layer, thereby significantly decreasing the number of model parameters required as a means of abrogating potential over-fitting. After the avgpool layer, a full connection layer was used for the final classification output.

Images were processed in various layers to enable the extraction of predictive features. The 3D-ResNet-9 algorithm eliminates the issue of vanishing gradients such that it can outperform other algorithms.

Model construction and validation

The predicted outcome in this study was VPI status, which was used with a fully connected neural network to establish a DL-based binary classification model.[41] This classifier employed a one-dimensional array input and a binary score output, with the output classification being designated with an appropriate threshold value, assigning a 1 or a 0 based on whether or not the output was greater than the threshold value, respectively. The model used a learning rate of 0.0005 and a batch size of 16. A weighted L1 norm was used for the loss function, with weights being determined based on the positive proportion. After an average of just give training cycles this model tended to converge.

To improve classification performance via the integration of multidimensional features, a clinical model, a DL model, and a combined model were established. And, the threshold values of above three models were 0.65, 0.93, 0.74, respectively. Eighteen features were found to be correlated with VPI status in the univariate regression analyses (Table 3), which was consistent with the results of SHAP (SHapley Additive exPlanation) analyses (see Figure S3).[42] The features ranked above contributed more weight in the model. Consolidation maximum diameter as the most important feature, which high value of SHAP made the model output tends to be positive. Although lobulated sign, CT subtype of pleural and pleural indentation were independent prediction of risk factors for VPI status in multivariate logistic regression analyses, 18 features were placed into 3D-ResNet-9 model for automatic analysis considering that individual factors still had a certain weight value through SHAP. The DL algorithm constructed a clinical model based on the different weights of these clinical radiological features. To construct the DL model, 768 DL features extracted from the 3D-ResNet-9 framework were used. The fully connected neural network contained a single hidden layer with 1024 neurons in this layer. The number of input features was used to determine the number of neurons in the input layer, while the output layer included a single neuron corresponding to the positive probability of network output.

Receiver operating characteristic (ROC) curves were used to assess the performance of the three models, and the area under the curve (AUC), sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), and F1 score values for each assessed. Calibration curves and boxplots were used to evaluate the degree of fit for these three models.

Evaluation of clinical application of model

In order to establish which CT image regions were most attributable to the predictions of VPI status made by the DL model, 3D-ResNet-9 results were visualized with heatmaps generated using gradient-weighted class activation mapping (Grad-CAM).[43] This approach employs gradients for object concept types including caption task outputs and logits of classes in the classification category, thereby generating a final convolution layer used to establish a rough location map highlighting key image areas related to prediction generation. Grad-CAM was used after 3D-ResNet-9 network training, with the 3D-ResNet-9 model having been constructed from four 3D residual neural blocks and with output activations and gradients for a third of these blocks having been used to conduct these analyses. Grad-CAM processing employed the following main steps.

(a) After inputting segmented image blocks for the two cases, a feature map was generated following the final convolution of the image block after feature extraction. The partial derivative of the feature graph was obtained to ensure the uniform size of the partial derivative matrix.

(b) Feature map weights were different for the classification of the fully connected layer. Global average pooling was performed using partial conductivity feature maps, and the weights of individual feature maps were computed through backpropagation. The number of feature maps was the length of the weight vector.

(c) To obtain weighted feature maps, each feature map was multiplied by the corresponding weight value. In three dimensions, the mean was calculated to yield a map (np. mean (axis = −1)). Any negative numbers were changed to values of 0 via Relu activation and normalization processing.

(d) Lastly, up-sampling was performed to obtain Grad-CAM, and the resultant heatmap was decreased to the image size to allow for simple image weighting.

To ascertain the clinical significance of our models, we also compared the difference of LVI, lymph node metastasis and distant metastasis between VPI-negative and -positive by the three models.

Follow-up and survival analyses

Disease-free survival (DFS) was the primary endpoint for this analysis, and was defined as the interval from surgery to the first imaging- or histological evidence-based confirmation of local or regional disease recurrence and/or distant metastasis, or all-cause death. Patients were censored at the date when their most recent chest CT scans were performed. Patients generally underwent CT-based follow-up every 6–12 months over a 2-year period postoperatively, followed by annual observation. In total, 487 patients were followed for a minimum of 5 years. Given that our study dealt with a CT-based model for prognostication in lung cancer, cases were assembled to evaluate two staging settings: preoperative prognostication at the time of clinical staging (487 patients with clinical stage IA lung adenocarcinoma) and postoperative prognostication at the time of pathological staging (355 pathologic stage I lung adenocarcinoma). Any variables exhibiting significance in univariate analyses were incorporated into a multivariate Cox regression analyses aimed at defining variables capable of independently predicting patient DFS. Kaplan-Meier survival analyses and log rank tests were used to gauge the ability of particular variables to stratify lung adenocarcinoma patients based on mortality risk.

QUANTIFICATION AND STATISTICAL ANALYSIS

Python 3.7.0. PyTorch 1.6.0, CUDA 9.2, and R 3.5.3 were used for all analyses. Continuous variables are given in the form of means $\pm$ standard deviations while categorical variables are presented as numbers (%). Unpaired t-tests, Mann-Whitney U tests, and chi-square tests were used for intergroup comparisons as appropriate. Interobserver agreement between readers 1, 2, and 3 pertaining to radiological findings was evaluated with Cohen's kappa test and intraclass correlation coefficients. ROC curves were generated with the R pROC package, and AUC values were compared to assess model predictive performance. The DeLong method was employed to compare AUC value differences. Accuracy, sensitivity, specificity, PPV, NPV, and F1 scores were incorporated into these comparative analyses. Hosmer-Lemeshow test was used to evaluated the goodness of fit of the model. Kaplan-Meier curves and log rank tests were used to assess differences in survival. A two-tailed $p < 0.05$ served as the cut-off for significance in all analyses.