

论 著

开发并验证术前基于瘤灶及瘤周的临床放射学列线图以预测肺结节病理浸润性*

夏志颖¹ 刘子蔚² 包陈政任³
胡秋根² 张 榕^{2,*}1.海南省中医院放射科(海南海口 570203)
2.南方医科大学顺德医院(佛山市顺德区第一人民医院)放射科(广东佛山 528308)
3.南方医科大学顺德医院附属陈村医院(佛山市顺德区第一人民医院附属陈村医院)放射科(广东佛山 528313)

【摘要】目的 基于肺结节瘤灶及瘤周影像组学特征开发反向传播神经网络(BPNN)深度学习模型,结合CT影像学特征及临床信息,构建用于预测肺结节浸润性的列线图模型。方法 回顾性收集经手术切除的242例GGN患者,分为腺体前驱病变(AAH/AIS)和浸润性肺腺癌(MIA/IAC),比较两组患者影像征象、临床特征的差异性。从肺部CT图像勾画肺结节及瘤周靶区,通过单因素秩和检验及相关性分析,对影像组学特征进行筛选,采用反向传播神经网络深度学习算法构建预测模型,以受试者工作特征(ROC)曲线下面积(AUC)评估模型预测性能,在外部队列中对模型进行泛化性验证。结果 BPNN模型在训练组、内部验证组和外部验证组中的AUC分别为0.883(95%CI: 0.830-0.929)、0.854(95%CI: 0.786-0.909)和0.854(95%CI: 0.786-0.909)。经过单因素及多因素分析得出月芽征、CT值、GGN长径是预测肺结节浸润性的独立危险因素($P<0.05$),由此构建临床模型在训练组及验证组中的AUC分别为0.889(95%CI: 0.835-0.934)、0.778(95%CI: 0.668-0.879)和0.901(95%CI: 0.856-0.940)。结合BPNN模型及临床模型构建联合模型,模型在训练组和验证组中的AUC分别为0.952(95%CI: 0.920-0.977)、0.891(95%CI: 0.807-0.959)和0.939(95%CI: 0.899-0.968)。相比其他两个模型,联合模型评估肺结节病理浸润性展示了更强的模型性能。结论 基于瘤灶及瘤周放射组学特征联合临床放射学信息构建的列线图模型,表现出较好的预测性能,联合模型在外部队列中得到较好的泛化性验证,可辅助临床对肺结节诊疗提供参考文献。

【关键词】肺结节; 反向神经传播网络;
影像组学; 深度学习; 列线图

【中国分类号】R445

【文献标识码】A

【基金项目】佛山市科技计划项目
(2220001005383);
南方医科大学顺德医院科研启动项目
(SRSP2021021)

DOI:10.3969/j.issn.1672-5131.2024.06.020

Develop and Validate a Preoperative Clinical Radiological Nomogram Based on Intratumoral and Peritumoral CT Radiomics to Predict Pathology Invasiveness of Pulmonary Nodules*

XIA Zhi-ying¹, LIU Zi-wei², BAOCHEN Zheng-ren³, HU Qiu-gen², ZHANG Rong^{2,*}.

1.Department of Radiology, Hainan Traditional Chinese Medicine Hospital, Haikou 570203, Hainan Province, China

2.Department of Radiology, Shunde Hospital of Southern Medical University (The First People's Hospital of Shunde), Foshan 528308, Guangdong Province, China

3.Department of Radiology, Shunde Hospital of Southern Medical University Affiliated Chencun Hospital(the First People's Hospital of Shunde Affiliated Chencun Hospital), Foshan 528313, Guangdong Province, China

ABSTRACT

Objective To develop a backpropagation neural network (BPNN) deep learning model based on the focal and peritumoral imaging features of pulmonary nodules, and combine CT imaging features and clinical information to construct a nomogram model for predicting the infiltration of pulmonary nodules. **Methods** 242 cases of GGN were retrospectively collected and divided into glandular prodromal disease (AAH/AIS) and invasive lung adenocarcinoma (MIA/IAC). The imaging signs and clinical features of the two groups were compared. Lung nodules and peritumoral target areas were characterized from lung CT images, and the image omics features were screened through single-factor rank sum test and correlation analysis. The back propagation neural network deep learning algorithm was used to construct a prediction model, and the predictive performance of the model was evaluated by area under receiver Operating characteristic (ROC) curve (AUC), and the model was generalized to verify in external validation cohort. **Results** The AUC of BPNN model in training cohort, internal validation cohort and external validation cohort were 0.883(95%CI:0.830-0.929), 0.854(95%CI:0.786-0.909) and 0.854(95%CI:0.786-0.909), respectively. Univariate and multivariate analysis showed that crescent sign, CT value and GGN length diameter were independent risk factors for predicting pulmonary nodule invasion ($P<0.05$). The AUC of the constructed clinical model in training cohort, internal validation cohort and external validation cohort were 0.889(95%CI: 0.835-0.934), 0.778(95%CI:0.668-0.879) and 0.901(95%CI:0.856-0.940), respectively. Combined with the BPNN model and the clinical model, the AUC of the model in training cohort, internal validation cohort and external validation cohort were 0.952(95%CI:0.920-0.977), 0.891(95%CI:0.807-0.959) and 0.939(95%CI:0.899-0.968), respectively. Compared with the other two models, the combined model demonstrated a stronger model performance in evaluating pulmonary nodule pathologic infiltration. **Conclusion** The nomogram model based on tumor foci and peritumoral radiomic features combined with clinical-radiological information showed good predictive performance, and the combined model was well validated by generalization in external cohort, which could provide reference for clinical diagnosis and treatment of pulmonary nodules.

Keywords: Pulmonary Nodule; Back Propagation Neural Network; Radiomics; Deep Learning; Nomogram

由于低剂量CT的广泛普及,越来越多的肺结节被检出^[1],在肺癌筛查试验中,肺结节检出率平均为33%,而检测到的肺结节中只有1.4%被诊断为肺癌^[2]。2021年WHO将肺腺癌分为腺体前驱病变和浸润性肺腺癌两大类^[3]。这一最新分类对于准确诊断、治疗方案选择和预后评估都具有重要指导意义,有助于提高肺腺癌患者的医疗管理水平和治疗效果,因此,术前精准识别肺结节浸润性具有重要意义。

通过CT图像评估肺结节的病理性质具有挑战性,因为它们之间的形态学特征(如分叶征、毛刺征、胸膜牵拉征等)有较大的重叠性及主观性,早期预测肺结节良恶性并及时临床干预成为影像及临床医生一大挑战^[4-6]。肺结节周围实质可被认为肿瘤微环境,肿瘤微环境包括了细胞转移、炎症反应、免疫细胞浸润和微血管形成等多种成分,这些因素能够影响肺结节的生长和恶变过程^[7-8]。

机器学习和人工神经网络的模型已应用于肺癌诊断,并得到了较好的预测效果^[4,9]。深度学习(deep learning, DL)是通过多层神经网络来学习和处理数据,能够帮助计算机完成复杂的任务并提高模型的性能,并可根据现有数据自动学习并提取图像特征,适用于处理大数据量医学图像^[10],本研究引入反向传播神经网络,开发和验证基于肺结节瘤灶及瘤周区域的放射组学模型,结合临床独立危险因素构建联合模型,辅助预测肺结节浸润性,为临床术前评估肺结节以提供个性化、无创预测工具。

1 资料和方法

1.1 一般资料 回顾性收集2019年1月至2023年2月经手术后病理证实的242例肺结节患者相关临床、影像及病理资料,男性67例,女性175例,平均年龄 53.9 ± 11.3 岁。存在一下任何情况之一将被排除:肺结节 >30 mm;影像、临床及病理资料不完整;图像存在伪影;存在其他肿瘤患者。额外搜集南方医科大学顺德医院附属陈村医院患者作为外

【第一作者】夏志颖,男,主治医师,主要研究方向:胸部CT影像诊断。E-mail: 170004821@qq.com

【通讯作者】张 榕,男,主治医师,主要研究方向:胸部CT影像诊断。E-mail: zrong3500@163.com

部验证组，采用与训练组相同纳入及排除标准，回顾性搜集2018年1月-2023年1月205例肺结节患者，男性63例，女性142例，平均年龄56.2±12.33岁。

1.2 图像采集 两家医院图像采集来自两台CT扫描仪，分别为Siemens Smootom Definition AS 64排和Siemens Smootom Definition Flash 64排多层螺旋CT扫描仪。病人取仰卧位，嘱病人深吸气后屏住气进行全肺扫描。扫描范围包括全肺及两侧胸壁。扫描参数：管电压 100120KV，自动管电流，FOV 400400mm，螺距0.81.0，准直0.6×128，重建厚度0.6mm1mm。

1.3 图像分析及影像组学特征提取 由两位具有高级职称胸部影像诊断医师进行双盲法阅片，并记录肺结节的放射学特征，如大小、CT值，空泡征、空气支气管征、月牙征、瘤肺边界、分叶征、毛刺征、微血管穿行征、胸膜牵拉征，遇到分歧时，由第三位具有高级职称影像诊断医师进行阅片，协商一致方式解决。

将肺部CT平扫图像以DICOM格式导入深睿软件，并由具有10年经验的影像诊断医师采用半自动方式在图像上逐层勾画感兴趣区(volume of interesting, VOI)，确定VOI后，使用软件自动在病灶外扩3mm(VOI_{3mm})及5mm(VOI_{5mm})，勾画时尽量避开肋骨、胸膜、气管、血管、心脏等区域，使用开源PyRadiomics软件包提取每个肺结节影像组学特征。

1.4 影像组学模型构建 将南方大学顺德医院数据集以7:3的比例分为训练组和内部验证组。为了避免模型拟合过度和提高模型的泛化性，我们对提取的影像组学特征进行了由粗到细的筛选。首先，利用单因素秩和检验对组学特征进行显著性分析，以P<0.05标准进行筛选；其次，对影像组学特征进行Spearman相关性分析，阈值超过0.6定义为冗余特征进行剔除处理。联合VOI_{0mm}、VOI_{3mm}、VOI_{5mm}特征，利用单因素秩和检验对组学特征进行显著性分析，以P<0.05标准进行筛选，并构建影像组学模型。

我们设计了改良的“反向传播神经网络(back propagation neutral network, BPNN)”算法对重要影像组学特征进行筛选，并构建BPNN模型。优化原理：以最小交叉熵(cross entropy loss)为优化目标进行 n 轮迭代优化(观察迭代-误差曲线，确定n=2400)，训练误差通过“反向传播”算法反馈，以重新优化神

经元网络参数。网络架构:13-LeakyReLU-Dropout(p=0.2)-12-LeakyReLU-12-13，“输入层-隐藏层1-隐藏层2-隐藏层3-输出层”，第一层隐藏层的结果进行“BN校正-LR激活-DP失活”，第二层隐藏层的结果进行“LR激活”，第三隐藏层为FC层(全连接层)。神经元节数点：13-12-12-13-2。神经网络公式如下： $LeakyRelu=\max(0, x)+\max(0, Z) \times \min(0, x)$ 。LR: LeakyReLu模型；DP: dropout(p=0.2)，使隐藏层的20%数值随机变成0，用于防止过拟合。图1是BPNN模型迭代曲线。

1.5 临床模型构建 采用单因素Logistic分析筛选出相关性因素，再采用多因素Logistic分析筛选重要的临床-放射学因素，并构建临床模型(clinical model)。

1.6 联合模型构建及模型性能评估 采用Logistic回归算法，将重要的临床-放射学特征和放射组学模型整合，构建临床-放射组学联合模型(combined model)，并使用ROC曲线评估模型性能。

1.7 统计分析 计量资料符合正态分布采用平均值±标准差表示，计数资料使用百分比表示。使用受试者工作特征曲线下面积(AUC)，以及决策曲线分析(DCA)和校准曲线来评估模型的效能和临床价值。所有统计分析均使用R语言(4.2.1)软件进行。P<0.05为差异具有统计学意义。

2 结果

2.1 临床资料比较 本研究纳入训练组及外部验证组共447例肺结节患者，两家医院患者临床及影像信息资料基线见表1。

2.2 影像组学特征筛选和构建影像组学模型 我们根据软件对VOI肿瘤、VOI_{3mm}、VOI_{5mm}各区域共提取了2017个影像组学特征，经过从粗到细的特征筛选，共获得13个非零系数特征，对肺结节各区域分别进行LASSO回归分析，筛选特征后合并，再次相关性分析和LASSO分析，降低序列间冗余度，并构建肺结节三个区域的影像组学模型(图2)。再通过神经网络构建BPNN模型，由模型预测效能结果可知，BPNN模型表现出了更好的性能，在训练组、内部验证组和外部验证组中的AUC分别是0.883(95%CI, 0.830-0.929)、0.867(95%CI, 0.774-0.941)和0.854 (95%CI, 0.786-0.909)，因此，以BPNN模型确定为最佳影像组学模型(radiomic model)见表2。

表1 两家医院肺结节患者基线数据表

		训练组				内部验证				外部验证			
		腺体前驱病变		浸润肺腺癌	P	腺体前驱病变		浸润肺腺癌	P	腺体前驱病变		浸润肺腺癌	P
例数		169	61	108		73	26	47		205	68	137	
性别 (%)	女	125(74.0)	43(70.5)	82(75.9)	0.439	50(68.5)	17(65.4)	33(70.2)	0.671	142(69.3)	49(72.1)	93(67.9)	0.542
	男	44(26.0)	18(29.5)	26(24.1)		23(31.5)	9(34.6)	14(29.8)		63(30.7)	19(27.9)	44(32.1)	
年龄(mean (SD))		54.53(11.21)	53.11(10.21)	55.32(11.71)	0.22	52.58(11.37)	55.12(9.75)	51.17(12.05)	0.157	56.21(12.33)	54.09(10.53)	57.27(13.04)	0.082
长径(median [IQR])		8.00 [6.00, 11.00]	7.00 [5.00, 8.00]	9.00 [7.00, 12.25]	<0.001	[(腺体前驱病变/浸润肺腺癌/P)训练组]							
		8.00 [6.00, 10.00]	6.50 [6.00, 8.00]	9.00 [7.00, 13.00]	<0.001	[(腺体前驱病变/浸润肺腺癌/P)内部验证]							
		10.00 [7.00, 15.00]	7.00 [6.00, 9.00]	12.00 [9.00, 17.00]	<0.001	[(腺体前驱病变/浸润肺腺癌/P)外部验证]							
HU (median [IQR])		-484.90 [-588.60,-349.30]	-588.80 [-650.30,-478.80]	-417.90 [-539.78,-264.22]	<0.001	[(腺体前驱病变/浸润肺腺癌/P)训练组]							
		-525.70 [-667.20,-377.20]	-613.80 [-708.28,-499.32]	-480.00 [-593.75,-292.85]	0.003	[(腺体前驱病变/浸润肺腺癌/P)内部验证]							
		-458.90 [-590.80,-234.20]	-580.40 [-643.70,-477.70]	-351.60 [-540.70,-130.30]	<0.001	[(腺体前驱病变/浸润肺腺癌/P)外部验证]							
成份(%)	实性	13(7.7)	0(0.0)	13(12.0)	<0.001	3(4.1)	0(0.0)	3(6.4)	0.023	16(7.8)	1(1.5)	15(10.9)	<0.001
	混合磨玻璃	61(36.1)	13(21.3)	48(44.4)		29(39.7)	6(23.1)	23(48.9)		83(40.5)	16(23.5)	67(48.9)	
	纯磨玻璃	95(56.2)	48(78.7)	47(43.5)		41(56.2)	20(76.9)	21(44.7)		106(51.7)	51(75.0)	55(40.1)	
位置(%)	左肺下叶	21(12.4)	6(9.8)	15(13.9)	0.822	7(9.6)	1(3.8)	6(12.8)	0.667	25(12.2)	6(8.8)	19(13.9)	0.297
	左肺上叶	44(26.0)	17(27.9)	27(25.0)		16(21.9)	7(26.9)	9(19.1)		58(28.3)	21(30.9)	37(27.0)	
	右肺下叶	28(16.6)	10(16.4)	18(16.7)		20(27.4)	6(23.1)	14(29.8)		34(16.6)	7(10.3)	27(19.7)	
	右肺中叶	12(7.1)	3(4.9)	9(8.3)		5(6.8)	2(7.7)	3(6.4)		17(8.3)	6(8.8)	11(8.0)	
	右肺上叶	64(37.9)	25(41.0)	39(36.1)		25(34.2)	10(38.5)	15(31.9)		71(34.6)	28(41.2)	43(31.4)	
空气支气管征(%)	无	152(89.9)	59(96.7)	93(86.1)	0.028	66(90.4)	25(96.2)	41(87.2)	0.41	155(75.6)	65(95.6)	90(65.7)	<0.001
	有	17(10.1)	2(3.3)	15(13.9)		7(9.6)	1(3.8)	6(12.8)		50(24.4)	3(4.4)	47(34.3)	
月牙征(%)	无	123(72.8)	56(91.8)	67(62.0)	<0.001	53(72.6)	23(88.5)	30(63.8)	0.024	170(82.9)	64(94.1)	106(77.4)	0.003
	有	46(27.2)	5(8.2)	41(38.0)		20(27.4)	3(11.5)	17(36.2)		35(17.1)	4(5.9)	31(22.6)	
瘤肺边界(%)	无	42(24.9)	14(23.0)	28(25.9)	0.667	19(26.0)	9(34.6)	10(21.3)	0.214	47(22.9)	14(20.6)	33(24.1)	0.575
	有	127 (75.1)	47 (77.0)	80 (74.1)		54 (74.0)	17 (65.4)	37 (78.7)		158 (77.1)	54 (79.4)	104 (75.9)	
分叶征(%)	无	128(75.7)	60(98.4)	68(63.0)	<0.001	56(76.7)	23(88.5)	33(70.2)	0.077	127(62.0)	61(89.7)	66(48.2)	<0.001
	有	41(24.3)	1(1.6)	40(37.0)		17(23.3)	3(11.5)	14(29.8)		78(38.0)	7(10.3)	71(51.8)	
毛刺征(%)	无	131(77.5)	61(100.0)	70(64.8)	<0.001	63(86.3)	26(100.0)	37(78.7)	0.03	167(81.5)	68(100.0)	99(72.3)	<0.001
	有	38(22.5)	0(0.0)	38(35.2)		10(13.7)	0(0.0)	10(21.3)		38(18.5)	0(0.0)	38(27.7)	
微血管穿行征(%)	无	27(16.0)	15(24.6)	12(11.1)	0.019	6(8.2)	4(15.4)	2(4.3)	0.225	97(47.3)	15(22.1)	82(59.9)	<0.001
	有	142(84.0)	46(75.4)	96(88.9)		67(91.8)	22(84.6)	45(95.7)		108(52.7)	53(78.5)	55(40.1)	

胸膜牵拉征(%)	无	128 (75.7)	55 (90.2)	73 (67.6)	<0.001	60 (82.2)	25 (96.2)	35 (74.5)	0.046	131 (63.9)	61 (89.7)	70 (51.1)	<0.001
	有	41(24.3)	6(9.8)	35(32.4)		13(17.8)	1(3.8)	12(25.5)		74(36.1)	7(10.3)	67(48.9)	

2.3 建立临床模型及性能评估 由单因素及多因素Logistic分析可知, 月牙征、CT值、长径为独立危险因素(表3), 利用logistics算法将特征进行建模得到临床模型(clinical model), 根据ROC分析, 模型在训练组、内部验证组及外部验证组中的ACU分别是0.889 (95%CI, 0.835-0.934)、0.778 (95%CI, 0.668-0.879)和0.901 (95%CI, 0.856-0.940)。

2.4 建立联合模型及性能评估 结合BPNN模型和临床独立危险因素, 利用Logistic方法构建临床-放射组学联合模型(combined model)。在训练组、内部验证组和外部验证组中模型的AUC分别为0.952 (95%CI, 0.920-0.977)、0.891 (95%CI, 0.807-0.959)

和0.939 (95%CI, 0.899-0.968)(图2A、B), 这表明Combined model模型预测肺结节的病理浸润性比radiomic model、clinical model具有更好的诊断效能(图3, 表4)。

2.5 列线图的验证、校准及判别 为了开发一种临床上适用的方法来预测肺结节的浸润性, 我们构建了一个临床-放射组学列线图(图4A)。校准曲线(图4B)可见模型预测概率与实际概率关系良好。决策曲线DCA(图4C)用于通过计算不同概率阈值下的净收益来评估三个预测模型的效用, 结果显示在不同概率阈值下, 列线图模型实现了最高的临床净收益, 这表明列线图是预测肺结节病理浸润性的可靠临床工具。

表2 影像组学模型性能分析

	模型	AUC (95% CI)	准确度	灵敏度	特异度
训练组	瘤灶	0.795 (0.720-0.859)	0.722	0.694	0.77
	瘤周3mm	0.827 (0.754-0.892)	0.787	0.843	0.689
	瘤周5mm	0.806 (0.739-0.872)	0.746	0.769	0.705
	BPNN	0.883 (0.830-0.929)	0.74	0.602	0.984
内部验证组	瘤灶	0.687 (0.547-0.812)	0.685	0.66	0.731
	瘤周3mm	0.696 (0.573-0.825)	0.63	0.638	0.615
	瘤周5mm	0.718 (0.591-0.836)	0.63	0.66	0.577
	BPNN	0.867 (0.774-0.941)	0.699	0.532	1
外部验证组	瘤灶	0.501 (0.405-0.589)	0.429	0.372	0.544
	瘤周3mm	0.500 (0.410-0.582)	0.439	0.372	0.574
	瘤周5mm	0.535 (0.444-0.615)	0.493	0.438	0.603
	BPNN	0.854 (0.786-0.909)	0.571	0.401	0.912

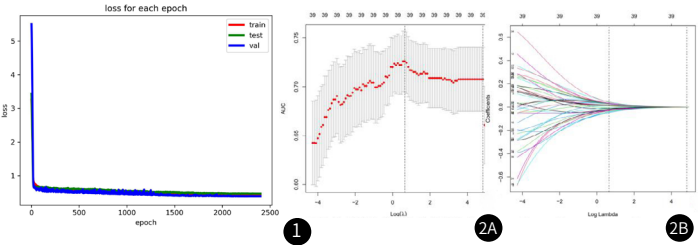


图1 BPNN模型网络迭代优化曲线; 如图1所示, 损失曲线记录了每一批损失函数值, 在50个周期后, 模型损失值迅速降至0.5且变化不再明显, 周期达到2400时, 模型得以收敛, 曲线显示模型的收敛速度比较快, 且平稳性较强, 损失值较低, 说明模型参数在经过训练之后可以很快的确定数值。

图2A-图2B 影像组学特征筛选。(2A)肺结节区域提取的放射组学特征筛选过程中的特征系数收敛图; (2B)对肺结节区域中提取的放射组学特征使用LASSO套索算法筛选出效能最好的放射组学特征集合, 纵坐标LASSO系数代表放射组学特征, 根据λ序列绘制系数剖面图, 虚线垂直线绘制在10次交叉验证中的最优λ处。

表3 单因素及多因素Logistic回归分析表

因素	Uni_OR	Uni_95%CI	Uni_P	Mul_OR	Mul_95%CI	Mul_P
年龄	1.018	(0.99~1.048)	0.219	NA	NA	NA
性别	0.757	(0.375~1.548)	0.44	NA	NA	NA
空气支气管征	4.758	(1.281~30.885)	0.043	0.224	(0.01~5.222)	0.325
HU	1.008	(1.005~1.011)	<0.001	1.007	(1.002~1.013)	0.007
毛刺征	1.01E+08	(0~9.13887371180277e+152)	0.986	NA	NA	NA
月牙征	6.854	(2.751~20.879)	<0.001	11.116	(2.521~70.294)	0.004
混合磨玻璃结节	0	(NA~3.81599155686057e+21)	0.988	NA	NA	NA
纯磨玻璃结节	0	(NA~3.81599155686057e+21)	0.987	NA	NA	NA
分叶征	35.294	(7.314~635.493)	0.001	8.051	(0.8~232.002)	0.128
长径	1.357	(1.2~1.568)	<0.001	1.284	(1.07~1.599)	0.013
微血管穿行征	2.238	(0.989~5.144)	0.054	NA	NA	NA
胸膜牵拉征	3.514	(1.545~9.118)	0.005	1.068	(0.213~4.675)	0.932
左肺上叶	0.635	(0.195~1.901)	0.429	NA	NA	NA
右肺下叶	0.72	(0.195~1.901)	0.598	NA	NA	NA
右肺中叶	1.2	(0.195~1.901)	0.825	NA	NA	NA
右肺上叶	0.624	(0.195~1.901)	0.388	NA	NA	NA
瘤肺边界	0.851	(0.399~1.755)	0.667	NA	NA	NA
空泡征	1.599	(0.683~4.077)	0.297	NA	NA	NA

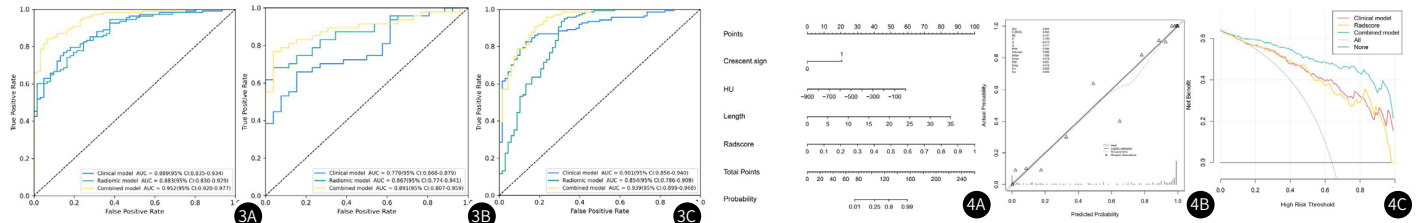


图3A-图3C 训练组、内部验证组及外部验证组模型ROC曲线; 3A: 训练组模型ROC曲线; 3B: 内部验证组模型ROC曲线; 3C: 外部验证组模型ROC曲线。
图4A-图4C 临床-影像征象联合radscore构建列线图以及诊断效能比较。4A: 临床-影像征象联合radscore构建logistics回归模型, 并采用列线图进行可视化。
4B: 预测模型校准曲线。4C: 决策曲线DCA, 黑色线(None)代表不采用任何方案进行肺结节浸润性评估时的临床收益始终为0; 红线代表采用clinical model进行肺结节浸润性评估时的临床收益; 黄色线代表采用radscore评估肺结节浸润性患者的临床收益。蓝色线代表采用combined model评估肺结节浸润性患者的临床收益。(注: Crescent sign 月牙征; Length 长径)

表4 训练组及验证组模型性能评估表

	模型	AUC (95% CI)	准确度	灵敏度	特异度
训练组	radiomic model	0.883 (0.830-0.929)	0.74	0.602	0.984
	clinical model	0.889 (0.835-0.934)	0.799	0.769	0.852
	combined model	0.952 (0.920-0.977)	0.876	0.843	0.934
内部验证组	radiomic model	0.867 (0.774-0.941)	0.699	0.532	1
	clinical model	0.778 (0.668-0.879)	0.712	0.638	0.846
	combined model	0.891 (0.807-0.959)	0.795	0.702	0.962
外部验证组	radiomic model	0.854 (0.786-0.909)	0.571	0.401	0.912
	clinical model	0.901 (0.856-0.940)	0.844	0.861	0.809
	combined model	0.939 (0.899-0.968)	0.829	0.788	0.912

3 讨 论

在这项多中心研究中，我们利用深度学习算法建立并验证了一个列线图模型，以识别术前肺结节的病理浸润性。列线图模型结合了放射组学特征及临床-放射学信息，在内部验证组及外部验证组中得到了较好的验证，模型决策曲线表明，列线图是预测肺结节患者病理浸润的可靠临床治疗决策支持工具。

近些年，机器学习算法在处理大数据预测生物标志物方面得到了广泛应用，并且比传统算法具有更好的性能，传统算法常用的包括最小绝对收缩和选择算子(LASSO)和随机森林^[11-12]。LASSO算法具有降维功能，可以直接减少特征数量并进行特征选择，当特征具有较强的语义时，使用LASSO算法具有优势，后续的分析将更具可解释性。随机森林是一种集成学习算法，可以计算对最终结果贡献最大的特征。深度学习作为机器学习的一个新兴领域，近几年对在疾病预测方面也得到重大突破，相较于传统的回归分析模型具有很多优势^[10]。BPNN是一种基于神经网络的机器学习算法，通过多层神经网络的反向传播来进行训练和学习，具有非线性、自主学习、容错性良好、自我组织以及自适应性等优点。它被广泛应用于深度学习领域，用于解决图像识别、语音识别、自然语言处理等复杂的模式识别问题。它根据数据分布概率去判断结果是否异常，然后通过多层非线性分析，可将数据在高维空间中反应出可分的种群效果，这里每个移动块的数据经特征工程处理后表征为序列化信息，块内的数据被视为整体，数据间横向交叉关联，以多维并行方式构建BPNN模型，BPNN模型通过大量数据迭代训练后，在多维空间下总结出一条相对合理的分界超曲线，序列化被测队列与分界超曲线进行比对，通过其中每个元素的大小关系，计算最终异常概率^[11]。肺结节瘤周区域可以反应生物学微观信息，并且在近几年的研究中得到证实，有研究表明^[13-14]，肿瘤内和肿瘤周围特征有助于预测病理类型和淋巴结转移，并识别分子亚型，使用瘤周的放射学特征来区分肺结节的潜在恶性肿瘤具有重大价值。在Wu等^[15]人使用肺结节瘤周30mm的区域放射学特征来进行研究，可以提高模型预测性能。然而以往的研究都是基于传统人工智能算法。在本研究中，我们利用LASSO回归和BPNN算法，基于瘤灶及瘤周3mm、5mm区域分别构建影像组学模型，BPNN模型在肺结节浸润性方面表现出了更好的预测性能，有效验证了BPNN模型的准确度和可靠性。

本研究通过分析发现月牙征、CT值及长径是预测肺结节浸润性的独立危险因素。俞慧波等^[16]通过对320例肺磨玻璃结节患者进行研究，发现月牙征为病灶边缘小叶间隔不同程度地向病灶侧凹入，并且肿瘤细胞生长遇到小叶间隔的阻挡，浸润性肺腺癌对小叶间隔的牵拉较腺体前驱病变更加明显，因此，月牙征可以有效的预测肺磨玻璃结节浸润性，我们与之研究一致。肺结节的CT值及长径是反应浸润性的有效指标^[17-18]。Chu等^[19]人研究纳入了172个肺结节，包括14个AAH，59个AIS，68个MIA和31个IAC，研究结果报道，浸润性肺腺癌的概率明显高于浸润前病变(P<0.05)；Xu等^[20]对171个肺结节浸润性进行研究指出，肺结节的密度与侵袭性呈正相关。由最新第八版肺癌TNM分期可知，T分期是以肺结节长径来进行定义，相比于其他主观征象，长径更能反映出肺结节的浸润性^[21]。Horeweg^[22]等研究得出，肺结节直径<5mm发生肺癌的概率为0.4%，而肺结节直径>10mm恶性发生率为16.9%。在基于NSLT数据库的研究中，长径1019mm的肺结节的恶性概率为6%，长径<10mm的肺结节恶性概率为1.3%。Qi等^[23]报道，在肺磨玻璃结节中预测浸润性的长径临界值

是11.5mm，其敏感性为75%，特异性为91.8%。在我们的研究中，CT值及长径也是临床独立危险因素。本研究通过BPNN模型及临床独立危险因素构建联合模型预测肺结节浸润性，并得到了较好的预测价值，并以列线图将模型进行可视化，可以辅助临床评估肺结节浸润性，并提供管理策略。

本研究亦存在不足之处。首先，本研究中两个中心的患者使用三个不同的扫描仪器，由于扫描参数存在差异，可能会对研究结果产生潜在的影响。其次，勾画肺结节感兴趣区，存在人为因素和技术因素可能导致勾画误差，从而影响后续的分析。第三，不同研究对肺结节周围的定义不同，本研究中瘤周范围为3mm、5mm，后续将继续增加亚组分析。

总之，本研究基于深度学习开发了BPNN放射组学模型预测肺结节浸润性，结合临床-影像特征，构建并验证了一个列线图模型，可以有效评估肺结节的病理浸润性，并在外部数据中得到了较好验证，保证了模型的泛化性能，列线图模型可以指导临床医生提供更精准的个体化诊断和治疗建议，帮助患者获得更好的治疗效果。

参考文献

[1]张庆团,郭兴全,王家赐.低剂量CT联合血清肿瘤标志物在早期非小细胞肺癌诊断中的应用价值[J].中国CT和MRI杂志,2023,21(2):58-60.

[2]Huang L,Lin W,Xie D,et al.Development and validation of a preoperative CT-based radiomic nomogram to predict pathology invasiveness in patients with a solitary pulmonary nodule:a machine learning approach,multicenter,diagnostic study[J].Eur Radiol,2022,32(3):1983-1996.

[3]Nicholson A G,Tsao M S,Beasley M B,et al.The 2021 WHO classification of lung tumors:impact of advances Since 2015[J].J Thorac Oncol,2022,17(3):362-387.

[4]Kim R Y,Oke J L,Pickup L C,et al.Artificial intelligence tool for assessment of indeterminate pulmonary nodules detected with CT[J].Radiology,2022,304(3):683-691.

[5]Hsu W C,Huang P C,Pan K T,et al.Predictors of invasive adenocarcinomas among pure ground-glass nodules less than 2 cm in diameter[J].Cancers (Basel),2021,13(16).

[6]滕刚,梅建波,涂毅,等.高分辨率CT在肺腺癌病理分型中的鉴别诊断价值分析[J].中国CT和MRI杂志,2023,21(1):66-69.

[7]Beig N,Khorrami M,Alilou M,et al.Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas[J].Radiology,2019,290(3):783-792.

[8]Lin X,Jiao H,Pang Z,et al.Lung cancer and granuloma identification using a deep learning model to extract 3-dimensional radiomics features in CT imaging[J].Clin Lung Cancer,2021,22(5):e756-e766.

[9]Huang P, Park S, Yan R, et al. Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study[J]. Radiology, 2018, 286(1): 286-295.

[10]Zheng X,He B,Hu Y,et al.Diagnostic accuracy of deep learning and radiomics in lung cancer staging:a systematic review and meta-analysis[J].Front Public Health,2022,10:938113.

[11]Tu P,Li X,Cao L,et al.Machine learning and BP neural network revealed abnormal B cell infiltration predicts the survival of lung cancer patients[J].Front Oncol,2022,12:882018.

[12]Xing L,Zhang X,Zhang X,et al.Expression scoring of a small-nucleolar-RNA signature identified by machine learning serves as a prognostic predictor for head and neck cancer[J].J Cell Physiol,2020,235(11):8071-8084.

[13]Wu L,Gao C,Ye J,et al.The value of various peritumoral radiomic features in differentiating the invasiveness of adenocarcinoma manifesting as ground-glass nodules[J].Eur Radiol,2021,31(12):9030-9037.

[14]Altorki N K,Borczuk A C,Harrison S,et al.Globa evolution of the tumor microenvironment associated with progression from preinvasive invasive to invasive human lung adenocarcinoma[J].Cell Rep,2022,39(1):110639.

[15]Wu G,Woodruff H C,Shen J,et al.Diagnosis of invasive lung adenocarcinoma based on chest CT radiomic features of part-solid pulmonary nodules:a multicenter study[J].Radiology,2020,297(2):451-458.

[16]俞慧波,陈中港,李琼,等.月牙征预测纯磨玻璃结节肺腺癌浸润性的价值[J].中华放射学杂志,2021,55(4):403-408.

[17]龚海鹏,司海峰,邢金丽,等.MSCT多平面重建技术用于磨玻璃结节样肺腺癌的鉴别诊断价值研究[J].中国CT和MRI杂志,2023,21(2):52-54.

[18]孙耀耀.肺原位腺癌与微浸润腺癌的病理分化情况及超高分辨率CT鉴别效果[J].罕少疾病杂志,2023,30(4):46-47.

[19]Chu Z G,Li W J,Fu B J,et al.CT characteristics for predicting invasiveness in pulmonary pure ground-glass nodules[J].AJR Am J Roentgenol,2020,215(2):351-358.

[20]Hsu W C,Huang P C,Pan K T,et al.Predictors of invasive adenocarcinomas among pure ground-glass nodules less than 2cm in diameter[J].Cancers (Basel),2021,13(16).

[21]叶波,赵蔚.第八版国际肺癌TNM分期修订稿解读[J].中国肺癌杂志,2016,19(6):337-342.

[22]Horeweg N,van Rosmalen J,Heuvelmans M A,et al.Lung cancer probability in patients with CT-detected pulmonary nodules:a prespecified analysis of data from the NELSON trial of low-dose CT screening[J].Lancet Oncol,2014,15(12):1332-1341.

[23]Qi L,Xue K,Li C,et al.Analysis of CT morphologic features and attenuation for differentiating among transient lesions,atypical adenomatous hyperplasia,adenocarcinoma in situ,minimally invasive and invasive adenocarcinoma presenting as pure ground-glass nodules[J].Sci Rep,2019,9(1):14586.