生物医学工程·论著

融合深度学习、临床与影像组学特征的综合模型提升非小细胞肺癌EGFR突变预测效能

杨涛涛[1],陈 伟[1],陈灿灿[2],闫婉莹[2],王大为[2],熊坤林[3],孙志远[4],王显棋[1] [1]陆军军医大学(第三军医大学)第一附属医院放射科,重庆;[2]推想医疗科技股份有限公司,北京;[3]陆军军医大学大坪医院放射科,重庆;[4]东部战区总医院放射诊断科,江苏南京

[摘要] **目的** 探究胸部CT图像深度学习特征及联合临床特征、影像组学特征预测非小细胞肺癌(non-small cell lung cancer, NSCLC)表皮生长因子受体(epidermal growth factor receptor, EGFR)基因突变的应用价值。**方法** 采用病例对照研究的方法,回顾性分析2013年1月至2023年10月3家医院放射科1 070例NSCLC患者的临床及影像资料。陆军军医大学第一附属医院719例NSCLC患者按照7∶3的比例随机分成训练集(n=502)、内部验证集(n=217);东部战区总医院173例患者作为外部验证集1;陆军军医大学大坪医院178例患者作为外部验证集2。深度学习特征通过使用主干为ResNet101的2.5D卷积神经网络(convolutional neural network, CNN)结构所获得,从基于CT的图像中提取影像组学特征,确定临床风险因素以构建临床预测模型,最后将深度学习特征、临床特征、影像组学特征整合构建综合模型。四种分类模型均采用随机森林(random forest, RF)的机器学习方法进行建模。绘制校正曲线以评估综合模型的拟合优度,绘制决策曲线评估模型的临床应用价值。**结果** 深度学习模型在内部验证集、外部验证集1、外部验证集2中的曲线下面积(area under the curve, AUC)值分别为0.833 7(0.770 6~0.884 7)、0.815 1(0.741 6~0.882 8)和0.810 1(0.745 2~0.873 6),临床模型对应的AUC值分别为0.731 0(0.660 2~0.802 1)、0.746 0(0.666 4~0.824 9)和0.813 4(0.743 1~0.883 6),影像组学模型对应AUC值分别为0.762 4(0.692 4~0.825 1)、0.745 4(0.671 1~0.814 3)和0.724 7(0.639 7~0.801 6),整合后的综合模型预测效能值最优,对应AUC值分别为0.905 5(0.857 0~0.945 4)、0.832 7(0.763 3~0.896 4)和0.889 0(0.834 4~0.934 3)。决策曲线分析(decision curve analysis, DCA)表明,当阈值概率为0.15~0.85时,综合模型对EGFR的预测具有较高的净获益。**结论** 基于CT图像的深度学习特征能较好地预测NSCLC患者的EGFR突变状态,当其联合临床特征、影像组学特征构建的综合模型可适当提高其预测效能。

[关键词] 非小细胞肺癌;表皮生长因子突变;深度学习;影像组学
[中图法分类号] R319;R730.42;R734.2 [文献标志码] A

Integrative model combining deep learning, clinical and radiomic features enhances EGFR mutation prediction in non-small cell lung cancer

YANG Taotao[1], CHEN Wei[1], CHEN Cancan[2], YAN Wanying[2], WANG Dawei[2], XIONG Kunlin[3], SUN Zhiyuan[4], WANG Xianqi[1] [1]Department of Radiology, First Affiliated Hospital, Army Medical University (Third Military Medical University), Chongqing; [2]Infervision Medical Technology Co., Ltd, Beijing; [3]Department of Radiology, Daping Hospital, Army Medical University, Chongqing; [4]Department of Diagnostic Radiology, Eastern Theatre General Hospital, Nanjing, Jiangsu, China

[**Abstract**] **Objective** To evaluate the predictive value of deep learning features from chest CT

[基金项目] 重庆市科卫联合重点项目(2023ZDXM008);陆军军医大学科技创新能力提升专项项目(2023XQN31)。
[通信作者] 王显棋,E-mail:wxq9401@tmmu.edu.cn

2097-0927/Copyright © 2025, 陆军军医大学(第三军医大学)。这篇开放获取文章遵循CC BY许可协议(https://creativecommons.org/licenses/by/4.0/)
引用格式:杨涛涛,陈伟,陈灿灿,等. 融合深度学习、临床与影像组学特征的综合模型提升非小细胞肺癌EGFR突变预测效能[J]. 陆军军医大学学报,2025,47(23):2991-3001. DOI:10.16016/j.2097-0927.202509040.

Citation: YANG T T, CHEN W, CHEN C C, *et al*. Integrative model combining deep learning, clinical and radiomic features enhances EGFR mutation prediction in non-small cell lung cancer[J]. J Army Med Univ, 2025, 47(23): 2991-3001. DOI: 10.16016/j.2097-0927.202509040.

images combined with clinical and radiomics features for epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer (NSCLC). **Methods** This case-control study retrospectively analyzed clinical and imaging data of 1 070 NSCLC patients from radiology departments at three hospitals (January 2013 to October 2023). Patients were divided into: a training set (n=502) and internal validation set (n=217) *via* 7∶3 randomization of 719 cases from the First Affiliated Hospital of Army Medical University; external validation set 1 (n=173) from General Hospital of Eastern Theater Command; external validation set 2 (n=178) from Daping Hospital of Army Medical University. Deep learning features were extracted using a 2.5D convolutional neural network (CNN) with ResNet101 backbone, radiomics features were derived from CT images, and clinical risk factors were identified to construct models. An integrated model combined deep learning, clinical, and radiomics features. All four models were developed using random forest (RF) classifiers. Calibration curves assessed goodness-of-fit, and decision curve analysis (DCA) evaluated clinical utility. **Results** The deep learning model achieved AUCs of 0.833 7 (95%*CI*: 0.770 6~0.884 7), 0.815 1 (0.741 6~0.882 8), and 0.810 1 (0.745 2~0.873 6) in the internal and two external validation sets, respectively. Clinical models yielded AUCs of 0.731 0 (0.660 2~0.802 1), 0.746 0 (0.666 4~0.824 9), and 0.813 4 (0.743 1~0.883 6); radiomics models showed AUCs of 0.762 4 (0.692 4~0.825 1), 0.745 4 (0.671 1~0.814 3), and 0.724 7 (0.639 7~0.801 6). The integrated model demonstrated optimal performance with AUCs of 0.905 5 (0.857 0~0.945 4), 0.832 7 (0.763 3~0.896 4), and 0.889 0 (0.834 4~0.934 3). DCA indicated significant net benefit for EGFR prediction at threshold probabilities of 0.15~0.85 using the integrated model. **Conclusion** Deep learning features from CT images effectively predict EGFR mutation status in NSCLC. The integrated model combining deep learning, clinical, and radiomics features further enhances predictive performance.

[**Key words**] non-small cell lung cancer; epidermal growth factor receptor mutation; deep learning; radiomics

Supported by the Key Program for Science and Health Integration of Chongqing (2023ZDXM008) and the Special Project for the Improvement of Science and Technology Innovation Capability of Army Medical University (2023XQN31). Corresponding author: WANG Xianqi, E-mail: wxq9401@tmmu.edu.cn

肺癌是全球最常见的癌症,占癌症相关发病率的12.4%[1]。非小细胞肺癌(non-small cell lung cancer, NSCLC)大约占肺癌的85%[2]。NSCLC中最常见的致癌驱动基因是表皮生长因子受体(epidermal growth factor receptor, EGFR)突变,在亚洲人群中约高达40%[3]。随着精准化、个性化治疗时代的来临,酪氨酸激酶抑制剂(tyrosine kinase inhibitors, TKIs)已经成为NSCLC治疗的一线药物,可明显改善患者的生存预后[4]。较多研究表明,TKIs治疗疗效与EGFR突变状态相关,EGFR阳性患者往往表现出更长的总生存期(overall survival, OS)和无进展生存期(progression-free survival, PFS)[5-6]。因此,在NSCLC患者个体化治疗之前早期确定其EGFR突变状态至关重要。

当前,判定EGFR突变状态的金标准是对经穿刺或手术取得的肿瘤样本进行基因测序[7]。然而,组织标本的获取是一项侵入性检查,存在禁忌证的患者无法进行活检或者手术;另外,获取的组织标本只能代表肿瘤的一小部分,无法反映肿瘤的广泛异质性[8-9];最后,由于检测周期长、费用高,以及对高质量组织样本的需求,限制了其在更广泛范围内的应用[10]。因此,临床上亟须寻找非侵入性且成本效益高的EGFR突变预测手段。

计算机断层扫描(computed tomography, CT)是肺癌筛查与诊断中广泛应用的成像手段,为肺癌研究提供了一种安全、无创且高效的技术支持[11-12]。影像组学和深度学习是过去十年中新兴的图像分析方法[13-14]。影像组学可以从医学图像中提取高通量特征,并筛选出有用的特征应用于临床决策[15-16]。深度学习是人工智能(artificial intelligence, AI)领域的一种强大算法,可直接从医学图像中提取特征来预测临床结果,显著提高医学影像分析的准确性和稳定性[17-18]。基于CT肿瘤的影像组学和深度学习在预测EGFR突变方面都显示出良好的预测性能[19-20]。最新研究表明,将影像组学和深度学习相结合来预测疾病诊断和预后,可产生更好的预测性能[21-22]。

因此,本研究旨在探究深度学习特征在预测

NSCLC患者EGFR突变应用中的价值,并开发和验证一种综合模型,该模型整合了深度学习特征、影像组学特征和临床特征,以从CT图像中无创、早期地预测EGFR的突变状态。

1 资料与方法

1.1 一般资料

本研究采用病例对照研究设计方案,已获得陆军军医大学第一附属医院伦理委员会(KY2023143)、东部战区总医院伦理委员会(2022DZKY-042-01)、陆军军医大学大坪医院伦理委员会(no. 301. 2023)批准,免除患者知情同意。所有参与的医疗中心均遵循统一的纳入与排除标准。纳入标准:①经肺组织病理学检查确诊为NSCLC,同时获取基因组测序分析结果;②基线胸部CT扫描需在病理诊断前8周内完成;③具备完整的临床信息。排除标准:①检查前已存在抗肿瘤治疗史;②基线CT病灶难以被准确勾画;③图像质量差;④伴有其他原发肿瘤。经筛选纳入陆军军医大学第一附属医院(2013年1月至2023年6月)的719例NSCLC患者,并以7∶3的比例随机分为训练集(*n*=502)、内部验证集(*n*=217)。纳入东部战区总医院的173例NSCLC患者(2013年9月至2019年10月)作为外部验证集1;陆军军医大学大坪医院的178例NSCLC患者(2020年1月至2023年10月)作为外部验证集2(图1)。患者的临床资料包括性别、年龄、吸烟史、神经元特异性烯醇化酶(neuron specific enolase, NSE)、细胞角蛋白片段19(cytokeratin 19 fragment, CYFRA21-1)、癌胚抗原(carcino embryonic antigen, CEA)、病理类型、T分期、N分期、M分期和临床分期。

1.2 样本量估算

本研究为回顾性多中心病例对照研究设计,样本量估算主要基于临床预测模型开发的每变量对应阳性事件数(events per variable, EPV)原则,即"每1个预测变量对应至少10~15例阳性事件"。研究计划拟纳入预测变量40个,所需阳性病例数约400~600例,考虑我院就诊肺癌患者人群EGFR突变率(约60%),需要入组患者667~1 000例。考虑潜在数据入组约20%的剔除率,基于上限计算最少筛选患者1 200例。

1.3 CT成像方案

胸部CT扫描采用仰卧位,自胸廓入口至肺底成像,所有数据存储为DICOM格式。CT扫描和图像重建参数如下(表1)。

1.4 CT图像肿瘤分割

患者的CT图像上传到多模态科研InferScholar平

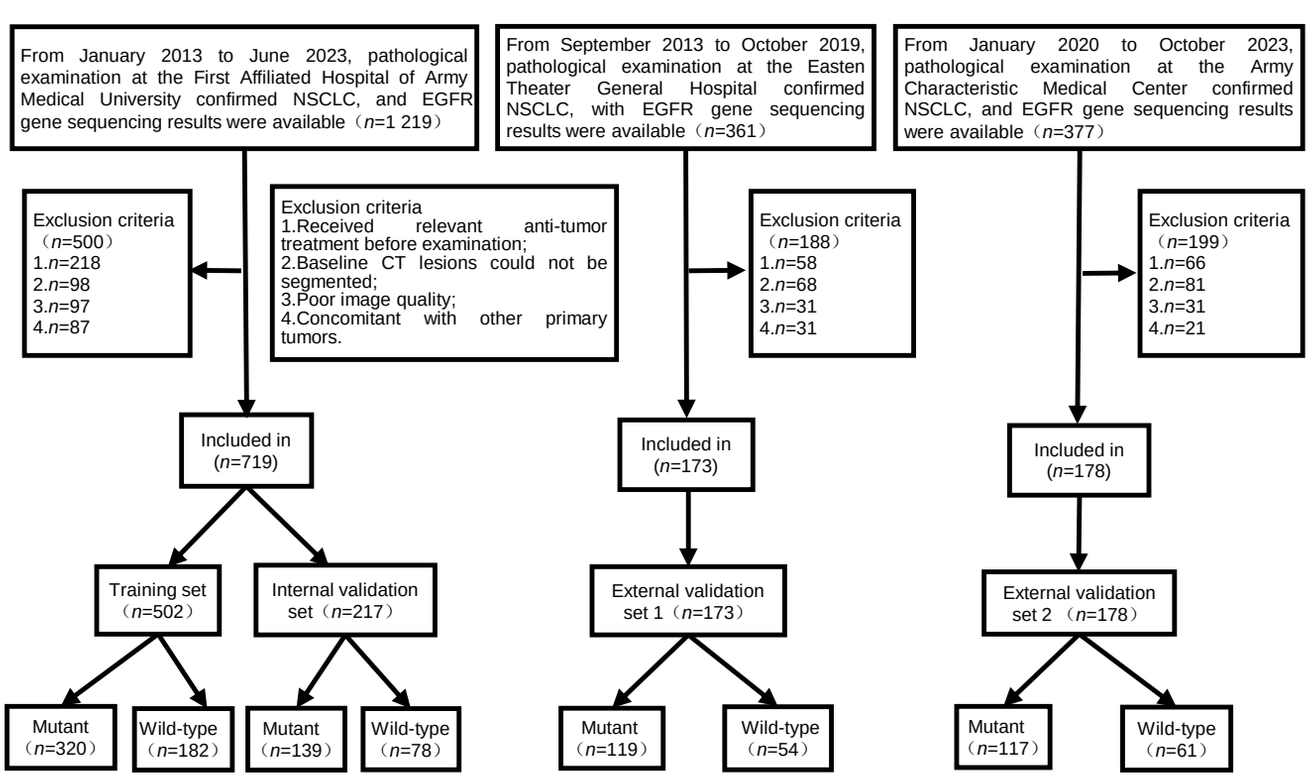


图1 患者入组流程图

Figure 1 Patient enrollment flow chart

表1 多中心医院CT扫描参数和图像重建参数

Table 1 CT scanning parameters and image reconstruction parameters in multi-center hospitals

Hospital	Instrument	Tube voltage /kV	Tube current /mA	Pitch	Rotation time/s	Degree of collimation/mm	Matrix	Convolution kernel	Layer thickness /mm	Interlayer spacing/mm
The First Affiliated Hospital of Army Medical University	SIEMENS Force	120	automatic	1.200	0.50	192×0.600	512×512	Br59	1.000~3.000	1.000~3.000
	SIEMENS Flash	100	automatic	1.500	0.50	128×0.600	512×512	B80f	1.000~3.000	1.000~3.000
	SIEMENS Emotion	110	automatic	1.500	0.60	16×1.200	512×512	B90s	1.500~3.000	1.500~3.000
	SIEMENS AS	120	automatic	1.200	0.50	128×0.600	512×512	I70f	1.000~3.000	1.000~3.000
Eastern Theater General Hospital	GE Optima	120	automatic	0.984	0.50	64×0.625	512×512	standard	1.250~3.000	1.250-3.000
	SIEMENS Perspective	130	automatic	1.200	0.60	64×0.600	512×512	I50S	1.000~3.000	1.000~3.000
	SIEMENS Flash	100	automatic	1.200	0.50	128×0.600	512×512	L50f	1.000~3.000	1.000~3.000
	United-imaging UCT 760	120	automatic	1.088	0.50	128×0.625	512×512	B_SHARP_C	1.000~3.000	1.000~3.000
The Daping Hospital of Army Medical University	GE LightSpeed	100	175	0.984	0.50	64×0.625	512×512	standard	0.625~3.000	0.625~3.000
	PHILIPS Brilliance	100	150	0.763	0.33	128×0.600	512×512	standard	0.600~3.000	0.600~3.000

台(https://research.infervision.com/v2/)用于图像注释。由1名具有10年诊断经验的放射科医生在肺窗图像上手动描绘每个患者的感兴趣区域,包括整个肿瘤区域。在多发性病变的病例中,仅概述了报告EGFR检查的病变。为了确保分割的准确性和稳定性,所有轮廓标签都由1位在胸部CT方面拥有15年诊断经验的放射科医生独立审查和纠正。

1.5 多中心CT图像预处理

为应对多中心CT扫描参数不一致的问题,本研究对所有CT图像进行了统一的预处理。将所有DICOM图像重采样至统一的体素间距(1 mm×1 mm×1 mm),消除层厚和层间距差异带来的影响,对灰度值进行归一化处理(使用Z-score标准化),以减小不同设备参数不一致的影响。在影像组学特征提取中,本研究采用了PyRadiomics工具箱的默认设置,该工具已对CT图像插值和滤波进行了标准化处理,使得影像组学特征具有一定鲁棒性。使用的2.5D卷积神经网络(convolutional neural network, CNN)模型在训练过程中引入了数据增强(如旋转、缩放、灰度扰动等),以增强模型对CT扫描参数的泛化能力。

1.6 深度学习特征和影像组学特征提取

对于深度学习特征,针对每个病变在9个不同方向上共提取了4 608个CNN特征,包括轴向0°、轴向45°、轴向-45°、冠状0°、冠状45°、冠状-45°、矢状0°、矢状45°和矢状-45°。深度学习特征提取网络使用2.5D CNN结构,其中ResNet101作为从每个方向提取特征的骨干,长短期记忆网络(long short-term memory, LSTM)模块用于整合所有方向的特征。对于影像组学特征,从CT图像的感兴趣掩模中共提取了1 293个影像组学特征,包括12个形状特征、252个一阶特征、308个灰度共生矩阵特征、224个灰度大小区域矩阵特征、224个灰度游程矩阵特征、196个灰度依赖矩阵特征、70个相邻灰度色调差异矩阵特征和5个图像特征。

1.7 特征筛选与模型构建

为减少模型构建过程中的过拟合,首先使用基于信息熵的特征进行离散化方法处理影像组学及深度学习特征,然后应用LASSO回归对以上两类特征选择,进一步降低了数据中过拟合的可能性。对于临床特征,所有特征均被离散化处理后进行建模。

基于上述单一类型特征均采用随机森林(random forest, RF)的机器学习方法进行建模。选择RF的原因有如下4个:首先,RF通过Bagging集成多个决策树并引入特征随机选择机制,有效降低模型方差,在高维小样本场景下表现稳健,抗过拟合能力强;其次,RF对输入特征的尺度不敏感,适用于同时整合离散化临床变量、连续型影像组学特征及高维深度学习特征;再次,RF可提供特征重要性排序(如SHAP图),便于后续分析哪些特征对EGFR突变预测贡献最大,具有较强可解释性;最后,RF具有多个可适应不同数据样本量的超参数,如决策树数量n_estimators、决策树的最大深度max_depth、决策树内部节点划分的最小样本数量min_samples_split、决策树叶子节点的最小样本数量min_samples_leaf等,通过参数自适应方法可构建可靠的机器学习模型。

将深度学习特征和基于CT肿瘤图像的影像组学特征及临床特征进行多模态整合,形成综合模型。本研究构建了EGFR基因突变与否的二分类预测模型,因变量Y是患者EGFR基因突变型的概率(取值范围

0~1),模型预测时采用由约登指数确定的概率阈值进行划分,当Y≥阈值时,判定为阳性(即EGFR突变型),当Y<阈值时,判定为阴性(即EGFR野生型),确保因变量的定义与临床实践中的判断标准一致。

1.8 统计学分析

本研究借助SPSS 26.0软件完成数据分析。为比较EGFR突变组与野生组的差异,连续变量采用独立样本t检验分析,分类变量则通过χ^2检验或Fisher精确检验处理。模型预测性能的评估指标包括受试者工作特征(receiver operating characteristic, ROC)曲线的曲线下面积(area under curve, AUC)、准确率、灵敏度及特异度。通过绘制校正曲线来评估模型的拟合优度,并利用决策曲线分析模型的临床应用价值。P<0.05为差异具有统计学意义。

2 结果

2.1 患者一般基线资料

本研究从3家医疗机构共纳入1 070例经病理确诊的NSCLC患者,其中EGFR突变型695例(64.95%),野生型375例(35.05%),患者人口学特征及临床资料如下(表2)。在训练集中,EGFR突变型与野生型患者在性别、吸烟史、CYFRA21-1、病理类型及临床分期方面存在显著差异(P<0.05),女性、不吸烟者、CYFRA21-1≤5 ng/mL、腺癌以及临床分期为Ⅰ、Ⅳ期患者更容易发生EGFR突变,而年龄、CEA、NSE、T分期、N分期及M分期指标差异无统计学意义。

2.2 综合模型具有影响力的特征的SHAP值

通过对多模态数据特征进行离散化处理,将所有特征融合后建立综合模型,综合模型中的特征对EGFR突变预测有显著贡献,下图展示最具影响力的前20个特征及其对应的SHAP值(图2)。

2.3 四种分类模型的验证

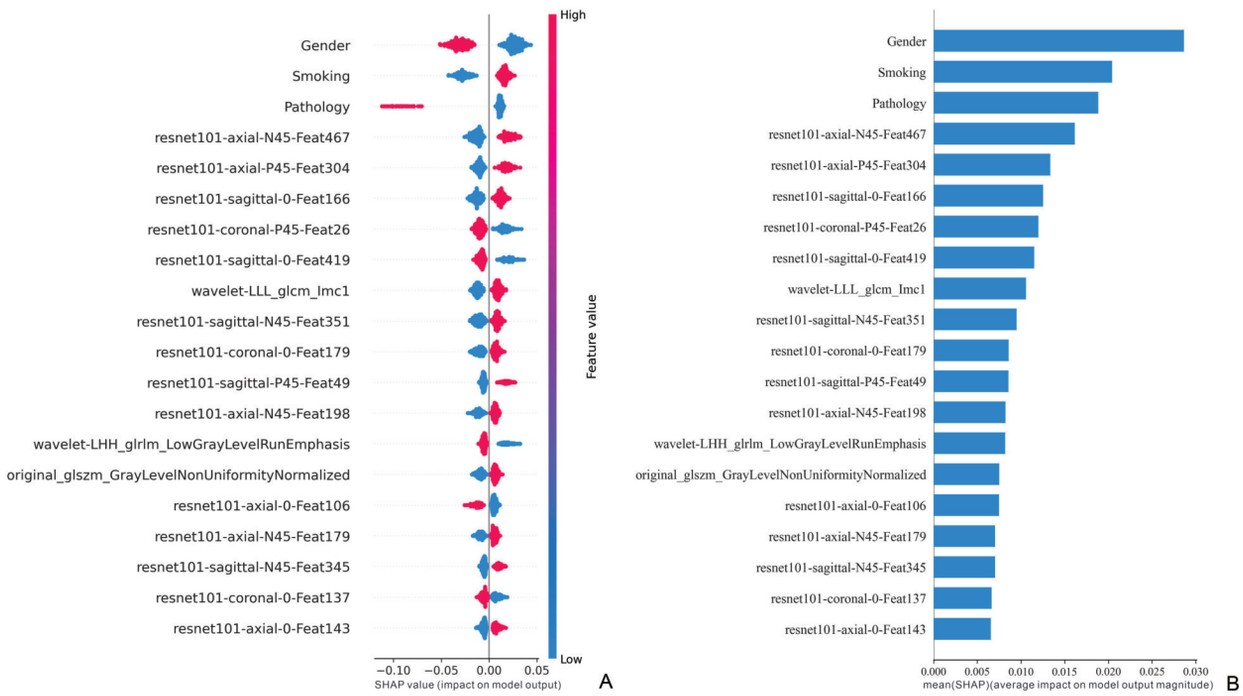
经过训练集的开发后,4种分类模型在内部验证集、外部验证集1、外部验证集2中的ROC曲线图表明深度学习模型预测效能高于影像组学模型和临床模型,将多模态数据融合后的综合模型的预测效能值得到提高,高于任何单一的预测模型(图3)。各模型的详细预测性能指标及准确性、特异性、敏感性如下(表3)。DeLong检验表明,深度学习模型与临床模型、影像组学模型、综合模型在内部验证集、外部验证集2中差异有统计学意义(P<0.05),在外部验证集1中无显著差异(P>0.05);综合模型与临床模型、影像组学模型在所有验证集中差异有统计学意义(P<0.05);临床模型与影像组学模型在所有验证集中差异无统计学意义。

表2 1 070例NSCLC患者的临床特征[例(%)]

Table 2 Clinical characteristics of 1 070 patients with NSCLC[n(%)]

Characteristics	Training set (n=502)		P	Internal validation (n=217)		P	External validation set 1 (n=173)		P	External validation set 2 (n=178)		P
	Wild-type	Mutant		Wild-type	Mutant		Wild-type	Mutant		Wild-type	Mutant	
Gender			<0.001			<0.001			<0.001			<0.001
Female	66 (36.3)	199 (62.2)		25 (32.1)	84(60.4)		19(35.2)	78(65.5)		9(14.8)	76(65.0)	
Male	116 (63.7)	121 (37.8)		53(67.9)	55(39.6)		35(64.8)	41(34.5)		52(85.2)	41(35.0)	
Smoking history			<0.001			<0.001			<0.001			<0.001
Yes	99 (54.4)	90 (28.1)		47(60.3)	36 (25.9)		26(48.1)	21(17.6)		47(77.0)	25(21.4)	
No	83 (45.6)	230 (71.9)		31(39.7)	103(74.1)		28(51.9)	98(82.4)		14(23.0)	92(78.6)	
CYFRA21-1/(ng/mL)			0.002			0.131			0.286			0.131
>5	49(26.9)	50 (15.6)		26(33.3)	23(16.5)		26(48.1)	47(39.5)		24(39.3)	33(28.2)	
≤5	133 (73.1)	270 (84.4)		52(66.7)	116(83.5)		28(51.9)	72(60.5)		37(60.7)	84(71.8)	
Pathological type			<0.001			0.637						0.637
Adenocarcinoma	144 (79.1)	317 (99.1)		64(82.1)	135 (97.1)		54(100.0)	119(100.0)		60(98.4)	116(99.1)	
Squamous carcinoma	38 (20.9)	3 (0.9)		14 (17.9)	4 (2.9)		0(0.0)	0(0.0)		1 (1.6)	1 (0.9)	
Clinical stage			0.025			0.005			0.032			0.005
Ⅰ	54 (29.7)	104 (32.5)		35(44.9)	52 (37.4)		2 (3.7)	0 (0.0)		16(26.2)	35 (29.9)	
Ⅱ	8 (4.4)	15 (4.7)		6 (7.7)	7(5.0)		2(3.7)	3 (2.5)		3 (4.9)	8 (6.8)	
Ⅲ	41 (22.5)	39 (12.2)		10 (12.8)	18 (12.9)		8 (14.8)	7(5.9)		23(37.7)	17 (14.5)	
Ⅳ	79 (43.4)	162 (50.6)		27(34.6)	62(44.6)		42 (77.8)	109 (91.6)		19(31.1)	57 (48.7)	

NSCLC: non-small cell lung cancer; CYFRA21-1: cytokeratin 19 fragment.



A: SHAP value of the top 20 features in the comprehensive model, red indicating high feature value, blue indicating low feature value; B: Average SHAP value of the top 20 features in the comprehensive model

图2 综合模型中前20个特征的SHAP值分布图

Figure 2 SHAP value distribution diagram of the top 20 characteristics in the comprehensive model

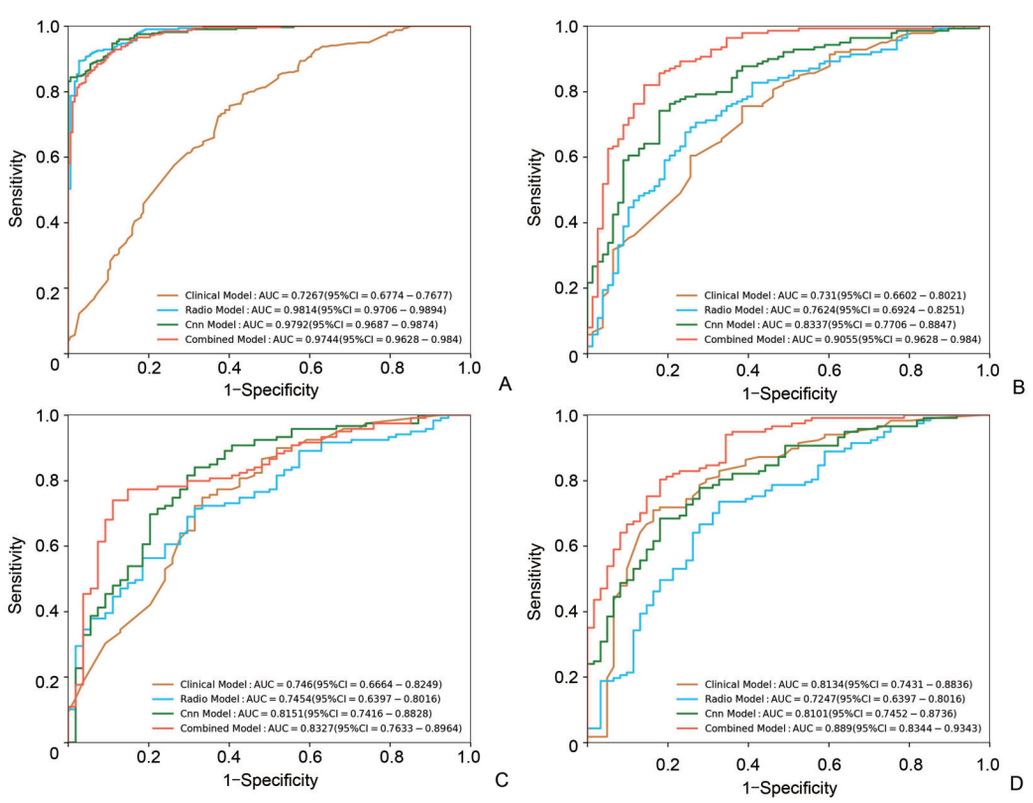


图3 4种模型分别在训练集(A)、内部验证集(B)、外部验证集1(C)、外部验证集2(D)中预测EGFR基因突变的ROC曲线

Figure 3 ROC curves of four models predict EGFR gene mutations in the training set (A), internal validation set (B), external validation set 1(C), and external validation set 2(D) respectively

表3 四种分类模型在不同数据集中的预测效能值

Table 3 Prediction power values of four classification models in different data sets

Data sets	Model	AUC(95%*CI*)	Accuracy	Sensitivity	Specificity
Training set	CNN model	0.979 2(0.968 7~0.987 4)	0.898 4	0.843 7	0.994 5
	Radiomics model	0.981 4(0.970 6~0.989 4)	0.922 3	0.893 7	0.972 5
	Clinical model	0.726 7(0.677 4~0.767 7)	0.709 2	0.790 6	0.565 9
	Combined model	0.974 4(0.962 8~0.984 0)	0.910 4	0.918 7	0.895 6
Internal validation set	CNN model	0.833 7(0.770 6~0.884 7)	0.718 9	0.661 9	0.820 5
	Radiomics model	0.762 4(0.692 4~0.825 1)	0.705 1	0.690 6	0.730 8
	Clinical model	0.731 0(0.660 2~0.802 1)	0.700 5	0.791 4	0.538 5
	Combined model	0.905 5(0.857 0~0.945 4)	0.843 3	0.877 7	0.782 1
External validation set 1	CNN model	0.815 1(0.741 6~0.882 8)	0.664 7	0.605 0	0.796 3
	Radiomics model	0.745 4(0.671 1~0.814 3)	0.630 1	0.546 2	0.814 8
	Clinical model	0.746 0(0.666 4~0.824 9)	0.751 4	0.873 9	0.481 5
	Combined model	0.832 7(0.763 3~0.896 4)	0.745 7	0.806 7	0.611 1
External validation set 2	CNN model	0.810 1(0.745 2~0.873 6)	0.612 4	0.444 4	0.934 4
	Radiomics model	0.724 7(0.639 7~0.801 6)	0.668 5	0.632 5	0.737 7
	Clinical model	0.813 4(0.743 1~0.883 6)	0.764 0	0.794 9	0.704 9
	Combined model	0.889 0(0.834 4~0.934 3)	0.758 4	0.709 4	0.852 5

AUC: area under the curve; CNN: convolutional neural network

2.4 模型的临床实用性评价

本研究的校准曲线显示综合模型在预测EGFR突变状态方面与临床实际校准度尚可(图4)。Hosmer-Lemeshow(bins=10)检验结果显示,在训练集中χ^2=152.5,P<0.001,内部验证集中χ^2=46.5,P<0.001,外部验证集1中χ^2=29.12,P<0.001,外部验证集2中χ^2=42.7,P<0.001。上述检验结果的原因可能有不同医院扫描参数差异导致影像组学特征不稳定或者深度学习特征偏移,进而影响模型在校准上出现系统性偏差;另外,虽然综合模型AUC表现优异,但RF本身不直接输出校准良好的概率估计,其预测概率基于叶节点中阳性样本比例,在训练集过拟合或特征冗余时容易产生过于极端的概率值(接近0或1),而实际事件率并未达到该水平;最后由于回顾性研究的设计,纳入患者均已完成EGFR基因检测,可能存在选择偏倚,例如,临床怀疑EGFR突变者更可能接受基因检测(如女性、非吸烟腺癌患者),导致突变比例偏高(64.95%),这种富集效应使得模型在低风险人群中的预测概率被系统性高估,进而破坏整体校准性。

临床决策曲线分析(decision curve analysis, DCA)提示综合模型具有较好的净获益,其结果显示,在阈值概率范围约0.15~0.85内,综合模型的净获益显著高于"全部治疗"和"全部不治疗"两种极端策略,尤其在临床常用决策阈值(如0.3~0.6,对应中等风险患者是否启动TKIs治疗),综合模型可带来每100例患者额外正确识别8~12例EGFR突变者的净获益,即使在外部验证集中,该优势依然保持,说明模型具有良好的泛化能力和临床转化潜力(图5)。

3 讨论

随着基因测序的发展,肺癌患者的治疗从最开始的传统化疗转向了精准化的靶向治疗。因此,在制定靶向治疗方案前实现NSCLC患者EGFR突变状态的精准无创评估,对个体化临床决策具有重要指导价值。在这项研究中,初步探究了胸部CT图像深度学习特征预测EGFR突变的能力,并进一步整合患者的临床特征、影像组学特征以构建综合模型,利用多模态信息,提高模型的预测效能。多中心结果显示:胸部CT的深度学习特征对NSCLC患者EGFR突变状态具有良好预测价值,且通过整合多模态特征的综合模型较单一特征模型预测效能更好。

3.1 临床特征对EGFR突变预测能力有限

患者的临床特征是最容易取得的,也是最早用来预测EGFR突变的预测因子。前期的文献研究中,有作者利用年龄、性别、吸烟史等临床因素[23]或者CYFRA21-1、CEA等实验室检查指标[24]来预测EGFR突变。在本研究中,共纳入了1 070名患者的年龄、实验室检查、临床分期等11项临床因素,经过单因素分

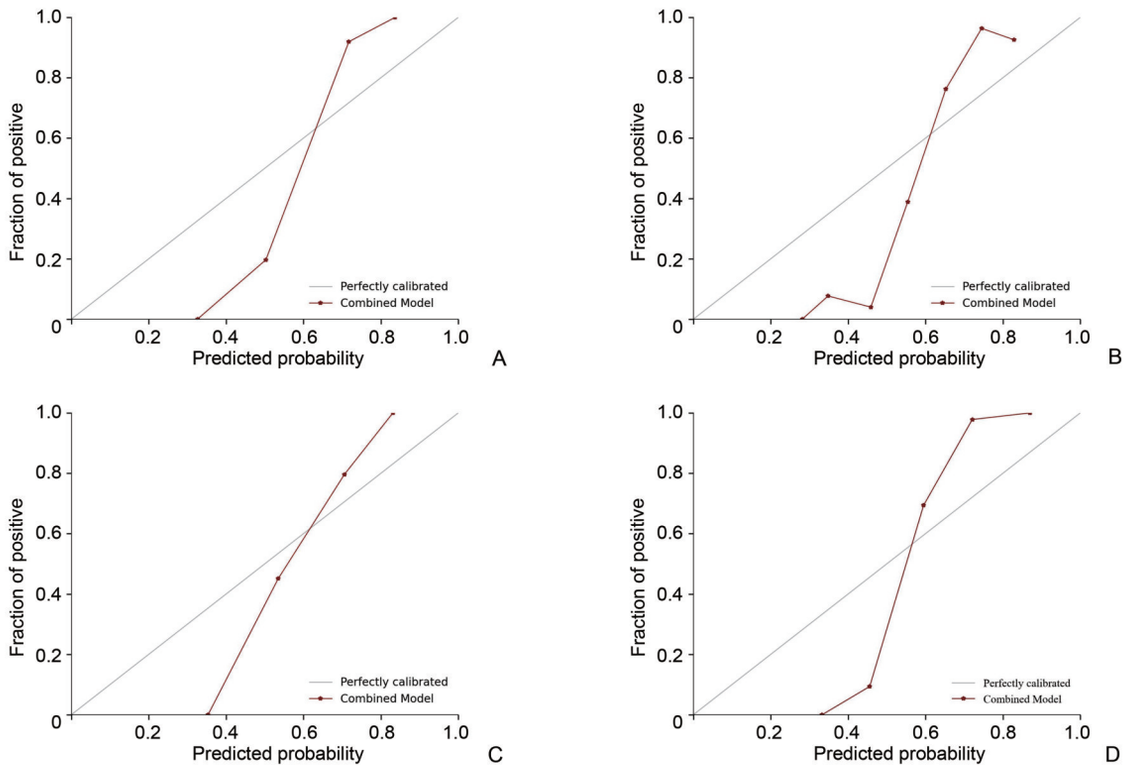


图 4 综合模型在训练集(A)、内部验证集(B)、外部验证集1(C)、外部验证集2(D)的校准曲线

Figure 4 Calibration curves of the combined model in training set (A), internal validation set (B), external validation set 1(C), and external validation set 2(D)

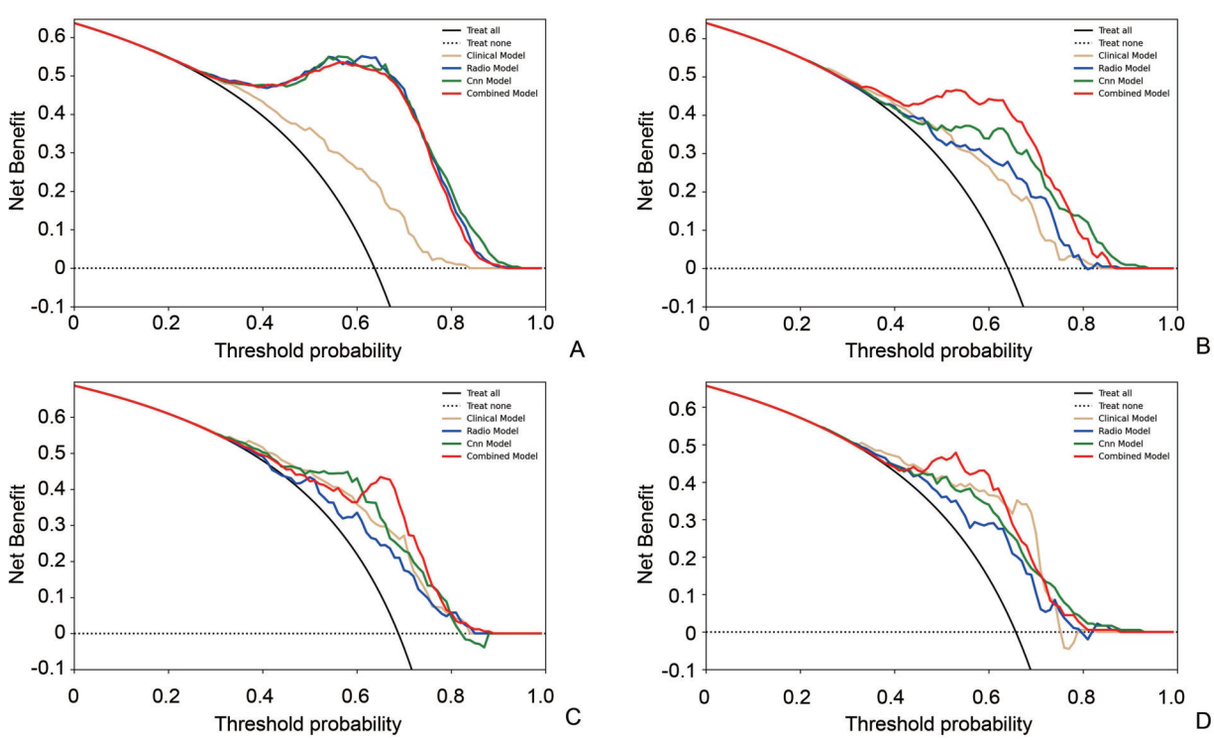


图 5 综合模型在训练集(A)、内部验证集(B)、外部验证集1(C)、外部验证集2(D)的临床决策曲线

Figure 5 DCA of the combined model in training set (A), internal validation set (B), external validation set 1(C), and external validation set 2(D)

析后我们发现性别、吸烟史、CYFRA21-1、组织学类型和临床分期是区分EGFR突变状态的重要因素,这与之前的研究一致[25]。然而,临床特征通常因为不同的数据集而产生不同的结果,也有文献研究结果表明,EGFR突变与患者的性别并无相关,这与之前的研究相矛盾[26]。因此,本研究将与EGFR突变常相关的临床特征离散化后纳入到模型中建模,其模型的效能值高于常规的特征筛选建模。

3.2 影像组学特征预测EGFR突变优势不明显

影像组学作为近十年迅速崛起的前沿学科,通过高通量计算技术从医学影像中挖掘定量特征,并基于特征筛选实现对肿瘤异质性的量化解析[27-29]。利用CT影像组学特征预判肺癌基因突变的研究已广泛开展。LIU等[30]从CT图像中提取了219个影像组学特征预测EGFR突变,AUC为0.647。然而,该研究受样本量较小、准确率低而限制了模型的临床应用。本研究所构建的影像组学模型不仅在内部验证集上表现出优异的预测性能,在来自两个外部医院的独立验证集中同样展现出了出色的效能。尽管影像组学模型相较于传统的临床模型没有显示出明显的优势性,然而将影像组学特征融入综合模型中时,提高了模型的预测效能值。表明影像组学特征可能拥有人眼无法识别的生物学信息,这些信息可以更加全面地反映肿瘤的空间异质性、肿瘤微环境以及基因信息等。

3.3 深度学习特征较好的预测EGFR突变

不仅影像组学在预测EGFR突变状态方面取得了应用,深度学习方法利用自动学习能力的优势在预测NSCLC患者的EGFR突变状态方面也取得了巨大的应用。深度学习方法主要通过多层次的神经网络结构从CT肿瘤图像中提取与EGFR信息直接相关的深度学习特征[31]。既往已有大量的学者利用深度学习方法从CT图像中提取特征以预测EGFR突变状态[20,31]。WANG等[32]通过深度学习方法从CT图像中来预测EGFR突变,模型的AUC值仅为0.701,且样本量较小、缺乏外部医院数据集验证,导致其临床应用价值受到限制。本研究利用2.5D CNN深度学习模型的方法从CT图像中提取深度学习特征,在内部验证集和两个外部医院的独立验证集上AUC值分别为0.833 7、0.815 1、0.810 1。深度学习模型的值无论在内部验证集还是在两个外部医院的独立验证集都是优于影像组学的。这一发现与MAHAJAN等[33]的发现一致,表明深度学习模型在预测EGFR突变方面可能优于影像组学,分析其原因可能是深度学习方法可以挖掘图像的更深层次特征。

3.4 综合模型提高预测效能

本研究提出的综合模型开创了预测EGFR突变的新模型,该模型在改善临床实践方面具有巨大的潜力。首先,该模型提供了一种快速且经济地预测基因突变的新方法。尤其是对于晚期肺癌患者,在诊断时通常症状严重急需快速准确的方法来指导治疗。综合模型的快速预测能力能够帮助患者更快地做出个性化治疗决策,从而有助于病情的改善。该多模态模型使用常规获得的临床信息、CT图像快速、准确地预测基因突变,这对无法进行基因测序的贫困地区的患者特别有益。通过提供一种更经济、更加便捷的基因突变检测方法,综合模型可以扩大贫困地区患者的治疗选择,进而改善他们的预后。其次,基于CT图像构建的综合模型可以检测肉眼无法感知的微小特征,从而通过非侵入性方式评估患者的EGFR基因突变情况,进而有效降低传统检测手段对机体造成的潜在损伤。最后,综合模型具有较高的预测效能值,能够准确预测EGFR突变,该模型无论是在内部验证集中,还是在两个外部医院的独立验证集,其模型的诊断效能值都优于单一特征模型。

本研究存在一定的不足。首先,回顾性研究不可避免存在选择偏倚,需要多中心、前瞻性研究来增强本研究的结论,并测试该模型在临床实践应用中的效果;其次,本研究所研究的人群只限于中国人,该研究的结果需要在其他种族中进行进一步验证;最后,此研究仅对EGFR是否突变进行了分析,而对其突变亚型暂未涉及,精准的亚型分型决定靶向药物敏感性,对药物选择、耐药管理到预后评估均有重要意义,未来需要进一步进行突变亚型的分析。

综上所述,本研究初步探索了基于胸部CT图像的深度学习特征在预测NSCLC患者EGFR突变状态的应用价值,并建立了一种融合深度学习特征和临床特征、影像组学特征的综合模型,该综合模型可适当提高预测效能,能够帮助临床医生无创早期识别EGFR突变状态,有助于制定患者的个体化、精准化的治疗方案,以获得最佳治疗预后。

利益冲突 本文所有作者声明不存在利益冲突

参考文献:

[1] BRAY F, LAVERSANNE M, SUNG H, *et al*. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA

Cancer J Clin, 2024, 74(3): 229-263. DOI: 10.3322/caac.21834.

[2] URAS I Z, MOLL H P, CASANOVA E. Targeting KRAS mutant non-small-cell lung cancer: past, present and future [J]. Int J Mol Sci, 2020, 21(12): 4325. DOI: 10.3390/ijms21124325.

[3] PASSARO A, MOK T, PETERS S, *et al*. Recent advances on the role of EGFR tyrosine kinase inhibitors in the management of NSCLC with uncommon, non exon 20 insertions, EGFR mutations [J]. J Thorac Oncol, 2021, 16(5): 764-773. DOI: 10.1016/j.jtho.2020.12.002.

[4] PASCHE B, GRANT S C. Non-small cell lung cancer and precision medicine: a model for the incorporation of genomic features into clinical trial design [J]. JAMA, 2014, 311(19): 1975-1976. DOI: 10.1001/jama.2014.3742.

[5] LEDUC C, MERLIO J P, BESSE B, *et al*. Clinical and molecular characteristics of non-small-cell lung cancer (NSCLC) harboring EGFR mutation: results of the nationwide French Cooperative Thoracic Intergroup (IFCT) program [J]. Ann Oncol, 2017, 28(11): 2715-2724. DOI: 10.1093/annonc/mdx404.

[6] ZHOU Q, ZHANG X C, CHEN Z H, *et al*. Relative abundance of EGFR mutations predicts benefit from gefitinib treatment for advanced non-small-cell lung cancer [J]. J Clin Oncol, 2011, 29(24): 3316-3321. DOI: 10.1200/JCO.2010.33.3757.

[7] LINDEMAN N I, CAGLE P T, BEASLEY M B, *et al*. Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors: guideline from the College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Pathology [J]. J Thorac Oncol, 2013, 8(7): 823-859. DOI: 10.1097/JTO.0b013e318290868f.

[8] AGGARWAL C, ROLFO C D, OXNARD G R, *et al*. Strategies for the successful implementation of plasma-based NSCLC genotyping in clinical practice [J]. Nat Rev Clin Oncol, 2021, 18(1): 56-62. DOI: 10.1038/s41571-020-0423-x.

[9] ZITO MARINO F, BIANCO R, ACCARDO M, *et al*. Molecular heterogeneity in lung cancer: from mechanisms of origin to clinical implications [J]. Int J Med Sci, 2019, 16(7): 981-989. DOI: 10.7150/ijms.34739.

[10] REUBEN A, ZHANG J X, CHIOU S H, *et al*. Comprehensive T cell repertoire characterization of non-small cell lung cancer [J]. Nat Commun, 2020, 11(1): 603. DOI: 10.1038/s41467-019-14273-0.

[11] DETTERBECK F C, WOODARD G A, BADER A S, *et al*. The proposed ninth edition TNM classification of lung cancer [J]. Chest, 2024, 166(4): 882-895. DOI: 10.1016/j.chest.2024.05.026.

[12] SUCCONY L, RASSL D M, BARKER A P, *et al*. Adenocarcinoma spectrum lesions of the lung: Detection, pathology and treatment strategies [J]. Cancer Treat Rev, 2021, 99: 102237. DOI: 10.1016/j.ctrv.2021.102237.

[13] LAMBIN P, RIOS-VELAZQUEZ E, LEIJENAAR R, *et al*. Radiomics: extracting more information from medical images using advanced feature analysis [J]. Eur J Cancer, 2012, 48(4): 441-446. DOI: 10.1016/j.ejca.2011.11.036.

[14] HEILBRONER S P, MIOTTO R. Deep learning in medicine [J]. Clin J Am Soc Nephrol, 2023, 18(3): 397-399. DOI: 10.2215/cjn.0000000000000080.

[15] CAI Z P, XU Z Y, CHEN Y F, *et al*. Multiparametric MRI subregion radiomics for preoperative assessment of high-risk subregions in microsatellite instability of rectal cancer patients: a multicenter study [J]. Int J Surg, 2024, 110(7): 4310-4319. DOI: 10.1097/JS9.0000000000001335.

[16] WARKENTIN M T, AL-SAWAIHEY H, LAM S, *et al*. Radiomics analysis to predict pulmonary nodule malignancy using machine learning approaches [J]. Thorax, 2024, 79(4): 307-315. DOI: 10.1136/thorax-2023-220226.

[17] SHE Y L, HE B X, WANG F, *et al*. Deep learning for predicting major pathological response to neoadjuvant chemoimmunotherapy in non-small cell lung cancer: a multicentre study [J]. EBioMedicine, 2022, 86: 104364. DOI: 10.1016/j.ebiom.2022.104364.

[18] YANG T T, WANG X Q, JIN Y, *et al*. Deep learning radiopathomics predicts targeted therapy sensitivity in EGFR-mutant lung adenocarcinoma [J]. J Transl Med, 2025, 23(1): 482. DOI: 10.1186/s12967-025-06480-9.

[19] ROSSI G, BARABINO E, FEDELI A, *et al*. Radiomic detection of EGFR mutations in NSCLC [J]. Cancer Res, 2021, 81(3): 724-731. DOI: 10.1158/0008-5472.CAN-20-0999.

[20] WANG S, SHI J Y, YE Z X, *et al*. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning [J]. Eur Respir J, 2019, 53(3): 1800986. DOI: 10.1183/13993003.00986-2018.

[21] LAN T J, KUANG S J, LIANG P S, *et al*. MRI-based deep learning and radiomics for prediction of occult cervical lymph node metastasis and prognosis in early-stage oral and oropharyngeal squamous cell carcinoma: a diagnostic study [J]. Int J Surg, 2024, 110(8): 4648-4659. DOI: 10.1097/JS9.0000000000001578.

[22] SHI S Y, LIN C X, ZHOU J, *et al*. Development and

validation of a deep learning radiomics model with clinical-radiological characteristics for the identification of occult peritoneal metastases in patients with pancreatic ductal adenocarcinoma[J]. Int J Surg, 2024, 110(5): 2669-2678. DOI:10.1097/JS9.0000000000001213.

[23] GIRARD N, SIMA C S, JACKMAN D M, *et al*. Nomogram to predict the presence of EGFR activating mutation in lung adenocarcinoma[J]. Eur Respir J, 2012, 39(2): 366-372. DOI:10.1183/09031936.00010111.

[24] DU W X, QIU T, LIU H Q, *et al*. The predictive value of serum tumor markers for EGFR mutation in non-small cell lung cancer patients with non-stage IA[J]. Heliyon, 2024, 10(9): e29605. DOI:10.1016/j.heliyon.2024.e29605.

[25] 杨涛涛, 王显棋, 陈灿灿, 等. 基于CT影像组学特征联合临床特征预测非小细胞肺癌EGFR突变状态[J]. 陆军军医大学学报, 2025, 47(8): 847-857. DOI:10.16016/j.2097-0927.202411024.
YANG T T, WANG X Q, CHEN C C, *et al*. Prediction of EGFR mutation status in non-small cell lung cancer based on CT radiomic features combined with clinical characteristics[J]. J Army Med Univ, 2025, 47(8): 847-857. DOI:10.16016/j.2097-0927.202411024.

[26] YATABE Y, KERR K M, UTOMO A, *et al*. EGFR mutation testing practices within the Asia Pacific region: results of a multicenter diagnostic survey[J]. J Thorac Oncol, 2015, 10(3): 438-445. DOI: 10.1097/JTO.0000000000000422.

[27] QI Y J, SU G H, YOU C, *et al*. Radiomics in breast cancer: Current advances and future directions[J]. Cell Rep Med, 2024, 5(9): 101719. DOI: 10.1016/j.xcrm.2024.101719.

[28] DU Y, CAI M J, ZHA H L, *et al*. Ultrasound radiomics-based nomogram to predict lymphovascular invasion in invasive breast cancer: a multicenter, retrospective study[J]. Eur Radiol, 2024, 34(1): 136-148. DOI: 10.1007/s00330-023-09995-1.

[29] 邓冰, 黄文才, 江远亮. 人工智能和影像组学在肺癌精准诊疗中的应用进展[J]. 重庆医学, 2025, 54(1): 213-218. DOI:10.3969/j.issn.1671-8348.2025.01.039.
DENG B, HUANG W C, JIANG Y L. Progress in application of artificial intelligence and radiomics in precision diagnosis and treatment of lung cancer[J]. Chongqing Med, 2025, 54(1): 213-218. DOI: 10.3969/j.issn.1671-8348.2025.01.039.

[30] LIU Y, KIM J, BALAGURUNATHAN Y, *et al*. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas[J]. Clin Lung Cancer, 2016, 17(5): 441-448.e6. DOI:10.1016/j.cllc.2016.02.001.

[31] WANG S, YU H, GAN Y C, *et al*. Mining whole-lung information by artificial intelligence for predicting EGFR genotype and targeted therapy response in lung cancer: a multicohort study[J]. Lancet Digit Health, 2022, 4(5): e309-e319. DOI:10.1016/S2589-7500(22)00024-3.

[32] WANG X Y, WU S H, REN J, *et al*. Predicting gene comutation of EGFR and TP53 by radiomics and deep learning in patients with lung adenocarcinomas[J]. J Thorac Imaging, 2025, 40(3): e0817. DOI: 10.1097/RTI.0000000000000817.

[33] MAHAJAN A, KANIA V, AGARWAL U, *et al*. Deep-learning-based predictive imaging biomarker model for EGFR mutation status in non-small cell lung cancer from CT imaging[J]. Cancers, 2024, 16(6): 1130. DOI: 10.3390/cancers16061130.

(**收稿**:2025-09-11;**修回**:2025-11-21)
(编辑 汤锦波)