DOI:10.16016/j.2097-0927.202412113

基于机器学习算法构建胸段食管鳞癌患者化疗联合免疫治疗反应预测模型的研究

陈金城[1],张小勤[2],刘 杰[1],李童心[2],吴 毅[2],何 萍[3],吴 蔚[1] 陆军军医大学(第三军医大学)第一附属医院:[1]胸外科;[3]心脏外科,重庆;[2]陆军军医大学(第三军医大学)生物医学工程与医学影像学院数字医学教研室,重庆

[**摘要**] **目的** 应用不同机器学习算法构建食管鳞癌患者化疗联合免疫治疗反应预测模型,并筛选出最优模型。**方法** 回顾性收集2022年1月至2023年12月在陆军军医大学第一附属医院胸外科住院,行化疗联合免疫治疗的174例食管鳞癌患者,收集患者治疗前的CT及临床信息。采用简单随机抽样法将患者按7∶3比例随机分为训练集(n=122)与测试集(n=52)。提取并筛选CT影像组学特征,使用5种机器学习算法构建影像组学模型和临床-影像组学模型。在训练集中进行5折交叉验证,在测试集中采用受试者工作特征曲线(receiver operating characteristic curve,ROC)、F1分数评估预测模型性能,对于表现最佳模型使用局部可解释模型不可知的解释(local interpretable model-agnostic explanations,LIME)算法进行解释。**结果** 174例患者中有115例患者出现临床缓解(66.1%)。从患者临床信息与CT影像中筛选出1个临床特征、10个影像组学特征。基于5种机器学习算法构建的影像组学模型和临床-影像组学模型的最佳ROC曲线下面积(area under of receiver operating characteristic,AUC)分别是0.750(95%*CI*:0.616~0.883)和0.766(95%*CI*:0.647~0.895)。最优临床-影像组学模型的F1分数为0.829。基于LIME算法,最佳模型对实例样本的预测显示出可靠性。**结论** 基于机器学习算法构建的临床-影像组学预测模型性能良好,通过预测食管鳞癌患者化疗联合免疫治疗反应,为医生临床决策提供参考。

[**关键词**] 食管癌;机器学习;影像组学

[**中图法分类号**] R319;R730.5;R735.1 [**文献标志码**] A

Development of a prediction model for chemotherapy and immunotherapy response in esophageal squamous cell carcinoma patients using machine learning algorithms

CHEN Jincheng[1], ZHANG Xiaoqin[2], LIU Jie[1], LI Tongxin[2], WU Yi[2], HE Ping[3], WU Wei[1] [1]Department of Thoracic Surgery, [3]Department of Cardiac Surgery, First Affiliated Hospital, Army Medical University (Third Military Medical University), Chongqing; [2]Department of Digital Medicine, College of Biomedical Engineering and Medical Imaging, Army Medical University (Third Military Medical University), Chongqing, China

[**Abstract**] Objective To develop models for predicting response to chemotherapy combined with immunotherapy in patients with esophageal squamous carcinoma with various machine learning algorithms, and then select the optimal model. Methods A retrospective study was performed for 174 patients with esophageal squamous cell carcinoma undergoing chemotherapy combined with immunotherapy admitted in

[**基金项目**] 重庆市科卫联合医学科研项目(2022ZDXM018)

[**通信作者**] 吴 蔚,E-mail:ww@tmmu.edu.cn

何 萍,E-mail:hp@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:陈金城,张小勤,刘 杰,等. 基于机器学习算法构建胸段食管鳞癌患者化疗联合免疫治疗反应预测模型的研究[J]. 陆军军医大学学报,2025,47(06):591-601. DOI:10.16016/j.2097-0927.202412113.

Citation:CHEN J C,ZHANG X Q,LIU J,*et al*. Development of a prediction model for chemotherapy and immunotherapy response in esophageal squamous cell carcinoma patients using machine learning algorithms[J]. J Army Med Univ,2025,47(06):591-601. DOI:10.16016/j.2097-0927.202412113.

Department of Thoracic Surgery of the First Affiliated Hospital of Army Medical University from January 2022 to December 2023. The CT scans and clinical information were collected before treatment. They were randomly divided into a training set (n=122) and a testing set (n=52) in a ratio of 7∶3. CT radiomic features were extracted and selected, and then 5 machine-learning algorithms were employed to establish the prediction models, including radiomics model and clinical-radiomics model. Five-fold cross-validation was conducted on the training set, and the performance of the prediction models was evaluated on the testing set using receiver operating characteristic (ROC) curve and the F1 score. The best-performing model was further explained using local interpretable model-agnostic explanations (LIME) algorithm. **Results** Among the 174 patients, 115 (66.1%) achieved clinical remission. From the clinical information and CT images, 1 clinical features and 10 radiomic features were identified. The area under of ROC curve (AUC) for the radiomics and clinical-radiomics models was 0.750 (95%*CI*: 0.616~0.883), and 0.766 (95%*CI*: 0.637~0.895), respectively. The F1 score of the optimal clinical-radiomics model was 0.829. LIME algorithm indicated that this best model demonstrated reliability in predicting individual samples. **Conclusion** The clinical-radiomics prediction model based on machine learning algorithm performs well, and can provide a reference for doctors' clinical decision-making by predicting the response to chemotherapy combined with immunotherapy in patients with esophageal squamous cell carcinoma.

[**Key words**] esophageal neoplasms; machine learning; radiomics

Supported by the Joint Medical Research Project of Chongqing Science and Technology Commission and Chongqing Health Commission (2022ZDXM018).
Corresponding authors: WU Wei, E-mail: ww@tmmu.edu.cn; HE Ping, E-mail: hp@tmmu.edu.cn

2024年中国国家癌症中心报告显示,食管癌在全国癌症发病人数中位居第7位,在癌症死亡人数位列第5位[1]。中国是食管癌高发地区,全球约一半食管癌发生在中国,在病理类型上,与欧美国家常见腺癌不同,我国90%病理类型为鳞癌。尽管近年来随着医疗技术的进步和国家监测体系的完善,中国食管癌的发病率及死亡率均呈下降趋势[2],但食管癌仍然是我国居民健康面临的主要威胁之一。由于早期食管癌并无明显症状,大部分患者在就诊时已到中晚期,错过了手术治疗的最佳时机,对于此类患者通常需行化疗或放化疗。

近年来,随着免疫治疗药物的研发,食管癌免疫治疗效果显著。最新指南建议,对于局部晚期食管鳞癌可行化疗联合免疫治疗[3]。然而,免疫治疗的疗效存在个体差异,只有15%~60%的患者出现缓解[4],此外,部分患者治疗过程中可能出现免疫相关的不良反应,严重时甚至可能导致肿瘤进展[5-6]。因此,在免疫治疗前,精确筛选出对免疫治疗有效的患者极其重要,目前常见检测手段为程序性死亡蛋白1(programmed cell death protein 1, PD-1)或程序性死亡配体1(Programmed cell death ligand 1, PD-L1)表达检测,但这种方法具有有创、昂贵、费时的缺点。如何精确、便捷地筛选出对免疫治疗有效的患者是免疫治疗成功的关键一步。

影像组学是将图像转换为大量高维数据的技术,能发现肉眼无法识别的特征[7]。早在2019年,有研究通过使用影像组学技术,以治疗前CT图像预测了黑色素瘤与非小细胞肺癌的免疫治疗反应[8],这是首个将影像组学用于预测癌症免疫治疗反应的研究。此外,随着人工智能技术运用于医疗,医学影像中包含的信息愈发受到关注,机器学习(machine learning, ML)作为人工智能最重要的实现方法已被广泛用于肿瘤领域。目前,机器学习方法被应用于食管癌早期筛查、疗效评估、预测预后等方面[9-12],且取得了良好的效果。然而,关于食管鳞癌化疗联合免疫治疗反应预测的研究却少有报道。因此,本研究拟从患者治疗前CT中提取影像组学特征,结合临床特征,使用机器学习算法构建多模态的预测模型,对食管鳞癌免疫治疗有效的患者进行精确筛选,指导医生进行临床决策,避免患者的过度治疗或者延误治疗。

1 资料与方法

1.1 数据来源

回顾性收集2022年1月至2023年12月于陆军军医大学第一附属医院胸外科住院化疗的174例食管鳞癌患者为研究对象,所有患者接受多西他赛(40 mg/m^2)

或紫杉醇(175 mg/m²)+铂类药物(70~75 mg/m²)化疗,同时联合免疫药物治疗,包括:信迪利单抗、帕博利珠单抗或卡瑞利珠单抗,剂量均为200 mg/次,每3周1次。纳入标准:①年龄>18岁;②病理活检证实为鳞癌;③根据第8版AJCC(American Joint Committee on Cancer)TNM分期,诊断临床分期≥T2期,任何N、M分期;④化疗联合免疫治疗前1月内进行过CT检查;⑤至少进行过2次化疗联合免疫治疗;⑥治疗结束4周后进行过CT检查。排除标准:①出现严重不良反应;②化疗前仅有外院CT图像;③既往进行过其他免疫治疗;④行治疗期间进行过放疗。本研究经陆军军医大学第一附属医院伦理委员会审查通过[批准号:(B)KY2021165,批准日期:2021年11月15日],所有患者数据进行匿名化处理后再分析。

1.1.1 患者临床数据收集 在医院信息系统(hospital information system,HIS)上收集患者治疗前的一般资料及实验室资料,包括:患者性别、年龄、吸烟史、饮酒史、肿瘤位置、临床TNM分期(第8版AJCC食管鳞癌临床TNM分期)、治疗前中性粒细胞数、中性粒细胞百分比、淋巴细胞数、淋巴细胞百分比,并通过中性粒细胞数与淋巴细胞数得到中性粒细胞与淋巴细胞比值(neutrophil-to-lymphocyte ratio,NLR)。

1.1.2 患者CT扫描 采用德国Siemens公司的Sensation 16、Somatom Definition AS 64排螺旋CT机或Somatom Definition Flash双源螺旋CT机扫描。扫描前所有患者吸气配合,以抑制呼吸运动伪影,采取头部朝前,双手放在头部上方体位。扫描参数:管电压100~120 kV,管电流为自动设置,螺距为1.2~1.5。增强扫描采用非离子型对比剂碘海醇(含碘300mg/mL),注射剂量为80~100 mL,注释速度为3.5~4.0 mL/s。图像包含116~560个层面,层厚为0.50~2.00 mm。将所有的CT图像临床数据脱敏处理后,经医院PACS系统以DICOM格式导出。收集患者治疗前后2次CT,用于疗效评估。

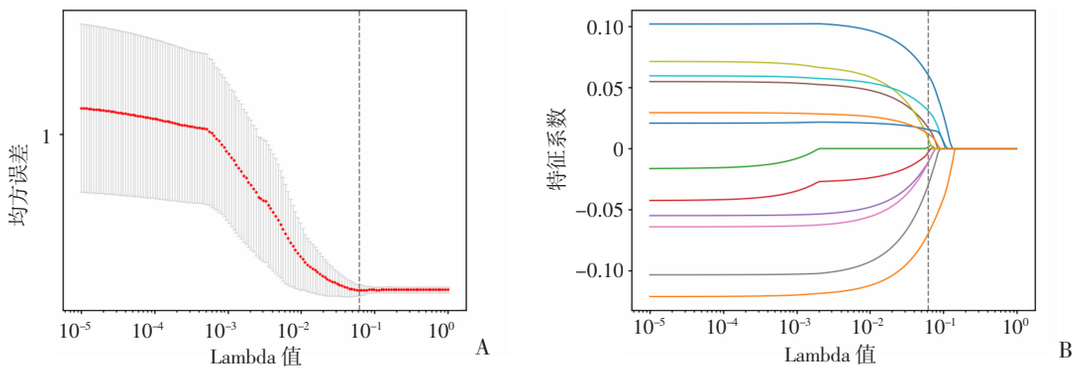
1.2 *研究方法*

1.2.1 治疗反应评估 根据实体肿瘤疗效评估标准(response evaluation crteria in solid tumors version 1.1, RECIST version1.1)[13],由2名高年资主治医师对肿瘤反应进行评估。定义临床完全缓解(clinical complete response, cCR):所有靶病灶消失,全部靶淋巴结短径<10 mm;临床部分缓解(clinical partial response, cPR):靶病灶直径之和较基线水平减少≥30%;临床肿瘤进展(clinical progressive disease, cPD):靶病灶直径之和相对增大≥20%,且直径之和绝对值增加≥5 mm;临床肿瘤稳定(clinical stable disease, cSD):靶病灶直径减少程度未达cPR,增加程度未达cPD。有反应组包括临床完全缓解与临床部分缓解,无反应组包括临床肿瘤进展与临床肿瘤稳定。当评估者间出现分歧,则由1名主任医师进行重新评估,并做出最终决定。所有患者疗效评估在化疗联合免疫治疗后,状态稳定且持续至少4周后进行。

1.2.2 CT图像分割与预处理 在图像分割前,将DICOM格式的CT图像导入到Amira软件中,将CT值显示范围调节为-300~500 HU之间。为了减少分割差异,由1名有着5年以上临床经验的胸外科医师,在CT图像上使用单盲法独立沿着肿瘤病灶边缘逐层手动勾画。当所有CT分割完成,由另一位有着10年以上临床经验的主任医师对分割图像进行审查、确认,对存在异议的图像协商统一意见。医师分割时对患者情况并不知情。为了尽量减少不同CT设备的扫描图像差异对特征提取的影响,在特征提取前,对图像进行统一预处理,归一化尺度设为100(即将像素值归一化到0~100)。重采样体素大小设置为[1,1,1] mm³,图像进行灰度离散化处理,带宽设置为25(灰度值分组间隔为25)。

1.2.3 影像组学特征提取与筛选 经过校准后的CT图像,使用Python工具包Pyradiomics提取三维肿瘤的影像组学特征。除了在原始图像(original)上提取肿瘤的形状、强度、纹理等特征,还使用经过小波变换(wavelet)、拉普拉斯-高斯变换(laplacian of gaussian, LoG)以及梯度变换(gradient)之后的图像提取肿瘤的高阶特征。然后,使用t检验与最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)方法对重要的特征进行筛选。保留t检验中$P<0.05$的特征进入LASSO模型。采用最小均方误差(mean square error, MSE)的5折交叉验证方法,从LASSO模型中选择最优参数lambda(λ),选取对治疗反应最具预测性的特征和权重系数(图1)。

1.2.4 模型建立与评估 采用简单随机抽样法对处理后的数据集进行拆分,其中70%进入训练集,30%作为测试集[14]。在训练数据集中,使用支持向量机(support vector machines, SVM)、朴素贝叶斯(naïve bayes, NB)、逻辑回归(logistic regression, LR)、K-近邻算法(K-nearest neighbors, KNN)、多层感知机(multilayer



A：通过五折交叉验证方法确定LASSO模型中最佳λ值的过程；B：纳入变量的LASSO系数曲线，不同色线表示特征变量的相应系数。

图1 运用LASSO回归模型进行影像组学特征筛选

perceptron，MLP)算法分别建立、影像组学模型与临床-影像组学模型，并进行5折交叉验证确定模型最佳参数。在建模型前，对所有数据进行Z-score标准化处理。在临床-影像组学模型中，将收集到的患者临床特征进行单因素Logistic与多因素Logistic分析，筛选出与治疗反应相关的特征，再构建预测模型。在测试数据集中测试模型性能，绘制受试者工作特征(receiver operating characteristic，ROC)、混淆矩阵(confusion matrix，CM)、F1分数、阳性预测值(positive predictive value，PPV)和阴性预测值(negative predictive value，NPV)对模型的预测性能和效果进行评价。

1.2.5 局部可解释模型不可知的解释(local interpretable model-agnostic explanations，LIME) LIME是一种模型不可知技术，应用于已经训练过的模型，以调查和分析输入参数与模型输出之间的关系[15]。其原理是对需要解释的黑盒模型，取一例实例样本，在其附近进行扰动生成新的样本点，并得到黑盒模型的预测值，使用新的数据集训练可解释模型，得到对黑盒模型良好的局部近似。为了使用LIME的可解释性，本研究采用Python3.9.1版本中的Lime Tabular Explainer来获取全局模型的训练数据。最终得到单个实例的可解释图，以此增强对模型的信任。

1.3 统计分析

采用SPSS 25.0软件进行数据分析。符合正态分布的计量资料以$\bar{x}\pm s$表示，组间比较采用独立样本t检验；不符合正态分布的计量资料以$M(P_{25}, P_{75})$表示，组间比较采用Mann Whitney U检验。计数资料以例(%)表示，组间比较采用χ^2检验。$P<0.05$为差异具有统计学意义。采用Python(version 3.9.1)软件进行影像组学特征提取与模型构建，绘制ROC曲线、混淆矩阵，并计算ROC曲线下面积(area under of receiver operating characteristic，AUC)、F1分数等模型评价指标。

2 结果

2.1 患者临床基线资料

共174名患者纳入研究，其中有反应组115人，无反应组59人。患者中位年龄为65岁，男性患者147人(84.5%)，女性患者27人(15.5%)。有反应组中性粒细胞数、中性粒细胞百分比与NLR均高于无反应组，无反应组淋巴细胞数与淋巴细胞百分比高于有反应组，但仅有淋巴细胞百分比具有统计学差异(P<0.05)，其余临床特征2组间未见明显统计学差异($P>0.05$)，见表1。

将患者按7∶3比例随机分为训练集(122人)与测试集(52人)后，2组间中性粒细胞百分比、淋巴细胞百分比与NLR差异具有统计学意义($P<0.05$)，其余特征未见明显统计学差异($P>0.05$)，见表2。

2.2 纳入模型的临床与影像组学特征

对于临床特征先进行单因素Logistic分析，发现中性粒细胞数、中性粒细胞百分比、淋巴细胞百分比与NLR具有统计学差异，但使用多因素Logistic分析发现仅淋巴细胞百分比具有统计学差异(OR：0.896，95%CI：0.847~0.949，P=0.001)(表3)。最终筛选出淋巴细胞百分比作为构建模型的临床特征，与影像组学特征共同构建临床-影像组学模型。

对于影像组学特征，从化疗前CT肿瘤分割区域中，共提取出1 037个特征，为了避免大量特征造成过

表1 免疫治疗有反应组与无反应组临床数据比较［例(%),$M(P_{25}, P_{75})$］

变量	总数据(n=174)	有反应组(n=115)	无反应组(n=59)	χ^2/U	P
年龄/岁	65(58, 70)	65(58,70)	66(58,70)	0.317	0.753
性别				0.005	1.000
男	147(84.5)	97(84.3)	50(84.7)		
女	27(15.5)	18(15.7)	9(15.3)		
吸烟史				0.021	1.000
有	128(73.6)	85(73.9)	43(72.9)		
无	46(26.4)	30(26.1)	16(27.1)		
饮酒史				0.458	0.510
有	112(64.4)	72(62.6)	40(67.8)		
无	62(35.6)	43(37.4)	19(32.2)		
肿瘤位置				4.537	0.104
胸上段	50(28.7)	37(32.2)	13(22.0)		
胸中段	86(49.4)	58(50.4)	28(47.5)		
胸下段	38(21.8)	20(17.4)	18(30.5)		
临床T分期				2.060	0.378
T2	49(28.2)	36(31.3)	13(22.0)		
T3	114(65.5)	73(63.5)	41(69.5)		
T4	11(6.3)	6(5.2)	5(8.5)		
临床N分期				2.066	0.573
N0	6(3.4)	4(3.5)	2(3.4)		
N1	98(56.3)	69(60.0)	29(49.2)		
N2	61(35.1)	37(32.2)	24(40.7)		
N3	9(5.2)	5(4.3)	4(6.8)		
临床M分期				1.097	0.295
M0	165(94.8)	111(96.5)	54(91.5)		
M1	9(5.2)	4(3.5)	5(8.5)		
中性粒细胞数/($\times10^9$/L)	4.8(3.7,5.9)	4.8(4.0,6.3)	4.5(3.4,5.3)	2.780	0.005
中性粒细胞百分比	68.3(63.0,74.0)	70.1(65.1,75.2)	66.5(59.1,70.1)	3.702	0.001
淋巴细胞数/($\times10^9$/L)	1.5(1.2,1.8)	1.4(1.2,1.7)	1.6(1.3,1.9)	1.787	0.074
淋巴细胞百分比	20.9(17.1,25.9)	20.2(15.8,24.9)	23.1(19.5,30.2)	3.192	0.001
NLR	3.2(2.4,4.2)	3.5(2.6,4.6)	2.8(2.1,3.4)	3.594	0.001

模型拟合,使用t检验与LASSO模型对特征进行筛选。在t检验中,保留P<0.05的特征,共有548个特征进入LASSO模型。然后,使用LASSO模型对548个重要的特征进行筛选,最终筛选出10个影像组学特征,包括3个original特征、5个wavelet变换特征和2个LoG变换特征(表4)。

2.3 机器学习模型的建立与评估

如图2所示,在训练集中,使用5种机器学习算法分别构建影像组学模型和临床-影像组学模型,其中基于MLP算法构建模型表现最佳,影像组学模型AUC为0.971(95%*CI*:0.944~0.999),临床-影像组学模型AUC为0.968(95%*CI*:0.939~0.997)。在测试集中对模型进行评估,发现基于SVM算法构建的影像组学模型、临床-影像组学模型表现最佳,其AUC值分别为0.750(95%*CI*:0.616~0.883)与0.766(95%*CI*:0.637~0.895)。

5种机器学习算法下临床-影像组学模型的混淆矩阵如图3所示。最终筛选出最佳模型为基于SVM算法建立的临床-影像组学模型,其F1分数为0.833,PPV为0.811,NPV为0.667(表5)。

2.4 基于LIME算法的最佳模型可解释性分析

通过LIME算法,解释了基于SVM的最佳临床-影像组学模型如何对1个实例患者治疗反应进行预测(图4)。在图4A中,显示了模型预测该样本属于无反

表2 训练集与验证集临床数据比较［例(%)，M(P_{25},P_{75})］

变量	训练集(n=122)	测试集(n=52)	χ^2/U	P
年龄/岁	65(59,70)	64(57,70)	0.673	0.503
性别			0.181	0.819
男	104(85.2)	43(82.7)		
女	18(14.8)	9(17.3)		
吸烟史			2.551	0.134
有	94(77.0)	34(65.4)		
无	28(23.0)	18(34.6)		
饮酒史			1.441	0.299
有	82(67.2)	30(57.7)		
无	40(32.8)	22(42.3)		
肿瘤位置			3.906	0.143
胸上段	40(32.8)	10(19.2)		
胸中段	55(45.1)	31(59.6)		
胸下段	27(22.1)	11(21.2)		
临床T分期			0.744	0.733
T2	32(26.2)	17(32.7)		
T3	82(67.2)	32(61.5)		
T4	8(6.6)	3(5.8)		
临床N分期			4.716	0.182
N0	5(4.1)	1(1.9)		
N1	66(54.1)	32(61.5)		
N2	42(34.4)	19(36.5)		
N3	9(7.4)	0		
临床M分期			0.281	0.726
M0	115(94.3)	50(96.2)		
M1	7(5.7)	2(3.8)		
化疗联合免疫治疗/次			0.204	0.718
2	37(30.3)	14(26.9)		
>2	85(69.7)	38(73.1)		
中性粒细胞数/($\times10^9$/L)	4.8(3.8,6.3)	4.5(3.7,5.5)	1.713	0.087
中性粒细胞百分比	69.6(63.8,74.8)	66.9(61.7,71.4)	2.083	0.037
淋巴细胞数/($\times10^9$/L)	1.4(1.2,1.7)	1.5(1.2,1.8)	1.149	0.252
淋巴细胞百分比	20.6(15.8,25.7)	22.9(19.7,27.3)	1.807	0.071
NLR	3.4(2.6,4.6)	3.0(2.3,3.5)	2.439	0.014

应组和有反应组的概率,无反应组的预测概率为0.27,有反应组的预测概率为0.73,表明模型倾向于将该患者预测为治疗有反应,并展示了预测反应组与无反应组的特征及其对应值。其中预测该患者为无反应的特征包括:original_glrlm_Long RunLowGrayLevelEmphasis、LoG-sigma-3-0-mm-3D_glszm_GrayLevelNonUniformity、wavelet-HHH_firstorder_Kurtosis,而其余特征预测患者为有反应组。在图4B中,展示了各个特征对样本预测结果的贡献,红色条表示特征值增加预测无反应组的概率,绿色条表示特征值增加预测有反应组的概率。条形的长度表示特征对预测结果的影响程度。

表3 单因素与多因素Logistic回归分析筛选临床特征［例(%)］

变量	单因素分析		多因素分析	
	OR(95%*CI*)	*P*	*OR*(95%*CI*)	*P*
年龄	1.000(0.961~1.042)	0.992		
性别				
女				
男	1.031(0.432~2.460)	0.945		
吸烟史				
无				
有	0.949(0.467~1.927)	0.884		
饮酒史				
无				
有	1.257(0.647~2.442)	0.499		
肿瘤位置				
胸上段				
胸中段	2.562(1.044~6.282)	0.040		
胸下段	1.864(0.854~4.068)	0.118		
临床T分期				
T2				
T3	2.308(0.601~8.863)	0.223		
T4	1.484(0.426~5.163)	0.535		
临床N分期				
N0				
N1	1.600(0.187~13.695)	0.668		
N2	1.903(0.477~7.600)	0.362		
N3	1.233(0.301~5.060)	0.771		
临床M分期				
M0				
M1	2.569(0.663~9.955)	0.172		
中性粒细胞数/($\times10^9$/L)	1.304(1.063~1.598)	0.011		
中性粒细胞百分比	1.060(1.021~1.100)	0.002		
淋巴细胞数/($\times10^9$/L)	0.476(0.222~1.021)	0.057		
淋巴细胞百分比	0.910(0.864~0.958)	0.001	0.896(0.847~0.949)	0.001
NLR	1.314(1.063~1.625)	0.012		

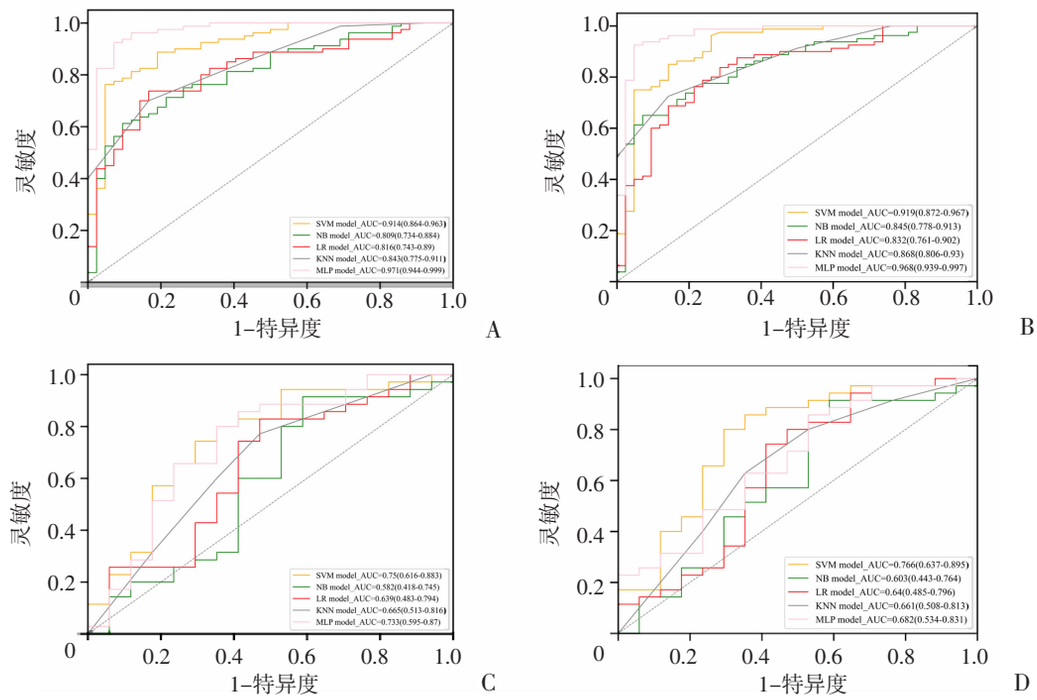
表4 筛选出来的影像组学特征及其系数

特征	权重
original_glcm_Imc2	0.000 472
original_glrlm_GrayLevelNonUniformityNormalized	−0.003 958
original_glrlm_LongRunLowGrayLevelEmphasis	−0.011 465
wavelet-LLH_firstorder_10Percentile	0.018 028
wavelet-LLH_firstorder_MeanAbsoluteDeviation	−0.011 195
wavelet-HLH_firstorder_Median	−0.030 217
wavelet-HLH_glcm-Imc1	0.008 810
wavelet-HHH_firstorder_Kurtosis	0.030 835
LoG-sigma-3-0-mm-3D_glszm_GrayLevelNonUniformity	0.015 659
LoG-sigma-3-0-mm-3D_glszm_SmallAreaEmphasis	0.012 028

3 讨论

3.1 机器学习模型预测免疫治疗反应的优势

本研究使用患者治疗前临床特征与CT影像,分别建立了影像组学模型和临床-影像组学模型,每种模型均基于5种不同的机器学习算法构建,其中基于SVM算法建立的临床-影像组学模型表现最佳。相较于传统的有创检测方法,该预测模型具有无创、可重复性、经济的优势,可协助临床医生做出最优决策。目前食管癌治疗方式仍是以手术治疗为主,化疗、放疗为辅的综合治疗。对于局部晚期食管鳞癌患者,相较于单纯手术,术前化疗联合免疫治疗能明显改善患



A:影像组学模型训练集中ROC曲线;B:临床-影像组学模型训练集中ROC曲线;C:影像组学模型测试集中ROC曲线;D:临床-影像组学模型测试集中ROC曲线

图2 基于5种机器学习算法建立模型的ROC曲线

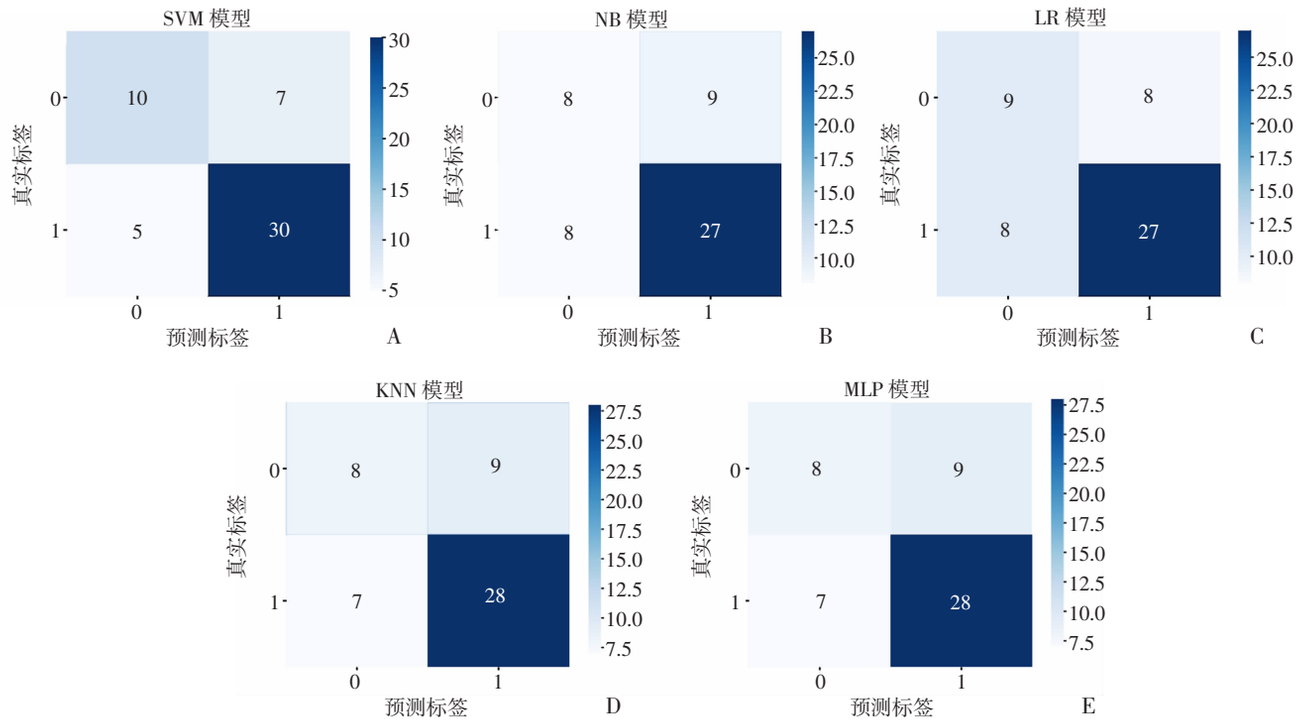
表5 测试集中基于5种机器学习的临床-影像组学模型的评价指标

模型	AUC(95%*CI*)	灵敏度	特异度	F1分数	PPV	NPV
SVM	0.766(95%CI:0.637~0.895)	0.857	0.588	0.833	0.811	0.667
NB	0.603(95%CI:0.443~0.764)	0.771	0.471	0.761	0.750	0.500
LR	0.640(95%CI:0.485~0.796)	0.771	0.529	0.771	0.771	0.529
KNN	0.661(95%CI:0.508~0.813)	0.800	0.471	0.778	0.757	0.533
MLP	0.682(95%CI:0.534~0.831)	0.800	0.471	0.778	0.757	0.533

者的预后[16]。免疫治疗作为食管鳞癌治疗的重要手段,能给患者带来巨大的临床获益,在免疫治疗过程中,医生们需要根据不同情况进行临床决策,然而,肿瘤的异质性使得这一过程变得复杂,给精准医疗带来巨大挑战[17]。机器学习模型为精准预测治疗反应提供了一种全新的视角,在以往的研究中展现出可行性。在2024年,一项关于食管鳞癌新辅助免疫治疗反应预测研究中[18],研究者使用了临床特征与影像组学特征相结合的方式构建模型,展示出了良好的预测性能。但与本研究不同,该研究仅纳入免疫治疗患者,且其主要终点事件为病理完全缓解。此类机器学习作为一种研究工具,从既往大量临床治疗数据中,发现与免疫治疗反应相关的特征,在临床上,这些特征可能并未引起医生注意。

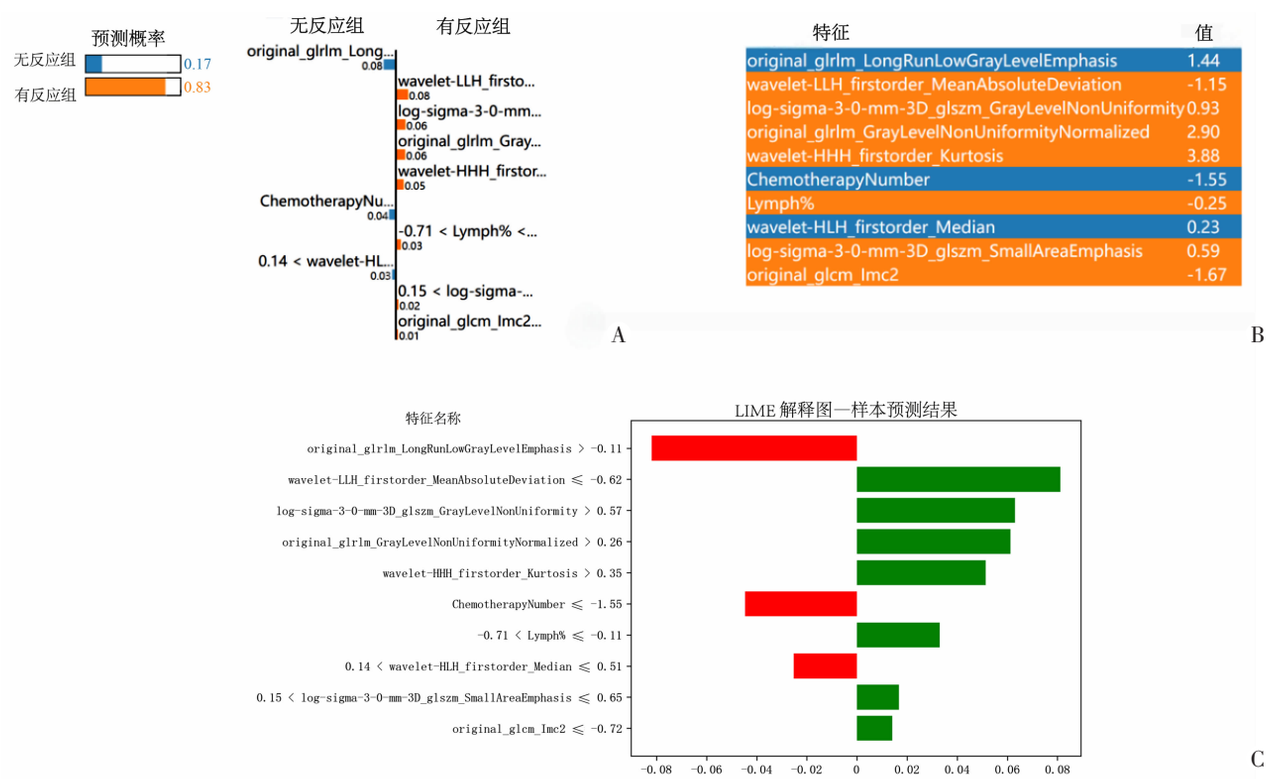
3.2 模型重要特征探讨

在本研究中,经过多因素Logistic回归分析发现淋巴细胞百分比与肿瘤缓解呈负相关。然而,一项研究显示,淋巴细胞,尤其是$CD4^+/CD8^+$ T细胞,具有影响癌症生长或转移的强效抗癌活性,在新辅助治疗期间的高淋巴细胞水平似乎与高pCR水平相关[19]。与本研究不同,该研究以pCR作为治疗反应评估标准,而本研究使用实体肿瘤疗效评估标准(RECIST version1.1)对肿瘤进行评估,这可能是导致两者差异的原因。尽管在将淋巴细胞百分比纳入模型中后,临床-影像组学模型性能较影像模型有所提升,但淋巴细胞对治疗反应的潜在生物学机制仍有待进一步揭示。此外,NLR作为全身性炎症的指标,可以反映炎症的微环境,既往的研究也发现了炎症因子与肿瘤之间关系密切[20],但本研究并未发现其与免疫治疗反映、相



A~E:分别为基于SVM、NB、LR、KNN、MLP算法建立的临床-影像组学模型的混淆矩阵

图3 基于5种机器学习算法建立的临床-影像组学模型的混淆矩阵



A、B:单个实例的LIME可解释性;C:LIME局部解释图

图4 通过LIME算法对最佳SVM模型预测过程进行解释

关。想要了解淋巴细胞、炎症因子与肿瘤反应关系,还需要更多前瞻性的研究。

关于影像组学特征,除了原始图像特征,小波变换、LoG变换、梯度变换后的图像各有其特点与优势。一些研究表明,小波变换能够有效地去除CT图像中的横纹噪声,对图像分析预测有明显优势,被广泛运用于影像组学研究中。本研究筛选了10个影像组学特征,其中有5个特征来自小波变换图像,表明小波变换在特征提取中的重要作用。在影像组学中,一阶统计量(first-order)特征描述由掩膜定义的图像区域内体素强度的分布,其具有良好的重复性[21],能够增强数据的可靠性。在本研究影像组学特征中出现4个一阶统计量特征,表明其构建的模型具有较好的可靠性。

3.3 机器学习模型在食管癌中运用

随着人工智能的发展,机器学习技术被广泛运用于医疗领域,在癌症诊断、预后评估和疗效预测中显示出专家级的能力[22]。在食管癌研究中,HU等[23]使用CT影像组学特征训练机器学习模型,以预测患者新辅助放化疗后病理完全反应情况,其准确率达到77.1%;WANG等[24]将CT影像组学特征与临床特征相结合,使用机器学习算法构建模型预测淋巴结转移情况。然而,机器学习模型的黑盒效应在很大程度上使其可信度降低,当临床医生对模型决策过程完全不可见时,阻碍了模型的临床应用。因此,在构建模型时,有必要对模型的可解释性进行探索,增强模型开发者与使用者的信心。本研究使用LIME算法对最佳临床-影像组学模型进行解释,展示了CT影像组学特征与临床特征对模型预测的具体影响,并对重要特征进行排序,极大提高了模型的可信度。LIME算法显示,临床-影像组学预测模型的11个重要特征中,前10个为影像组学特征,最后1个为临床特征,其中淋巴细胞百分比对模型产生正向影响。这些特征对模型的贡献程度从高到低排序,为临床-影像组学模型做出的每个预测提供了可理解的解释,从而显著提高了模型的可信度与临床适用性。

3.4 研究局限性

本研究存在一定的局限性:①首先本研究属于回顾性研究,数据存在不平衡的问题,且缺少外部验证;②本研究为单中心的研究,样本量总体偏少,为了使研究结果更加真实可靠,后续需纳入其他中心数据,以增强模型的泛化能力;③影像组学的生物学可解释性仍然是一个重大挑战,需要进一步探索揭示影像组学与基因组学、蛋白质组学等多组学整合之间的潜在联系。

综上所述,本研究基于5种机器学习算法构建模型,预测食管鳞癌化疗联合免疫治疗反应,其中基于SVM算法的临床-影像组学模型预测性能最好,预测模型具有较高可解释性。通过无创预测模型,可为临床决策与精确治疗提供帮助。

利益冲突 本文所有作者声明不存在利益冲突

参考文献:

[1] HAN B F, ZHENG R S, ZENG H M, *et al*. Cancer incidence and mortality in China, 2022[J]. J Natl Cancer Cent, 2024, 4(1): 47-53. DOI: 10.1016/j.jncc.2024.01.006.

[2] HE Y T, LI D J, SHAN B E, *et al*. Incidence and mortality of esophagus cancer in China, 2008 - 2012[J]. Chin J Cancer Res, 2019, 31(3): 426-434. DOI: 10.21147/j.issn.1000-9604.2019.03.04.

[3] SHAH M A, HOFSTETTER W L, KENNEDY E B, *et al*. Immunotherapy in patients with locally advanced esophageal carcinoma: ASCO treatment of locally advanced esophageal carcinoma guideline rapid recommendation update[J]. J Clin Oncol, 2021, 39(28): 3182-3184. DOI: 10.1200/JCO.21.01831.

[4] DAS S, JOHNSON D B. Immune-related adverse events and anti-tumor efficacy of immune checkpoint inhibitors[J]. J Immunother Cancer, 2019, 7(1): 306. DOI: 10.1186/s40425-019-0805-8.

[5] YANG Y N, ZHANG J F, MENG H X, *et al*. Neoadjuvant camrelizumab combined with paclitaxel and nedaplatin for locally advanced esophageal squamous cell carcinoma: a single-arm phase 2 study (cohort study)[J]. Int J Surg, 2024, 110(3): 1430-1440. DOI: 10.1097/JS9.0000000000000978.

[6] WANG Z X, CUI C X, YAO J, *et al*. Toripalimab plus chemotherapy in treatment-naïve, advanced esophageal squamous cell carcinoma (JUPITER-06): a multi-center phase 3 trial[J]. Cancer Cell, 2022, 40(3): 277-288.e3. DOI: 10.1016/j.ccell.2022.02.007.

[7] GILLIES R J, KINAHAN P E, HRICAK H. Radiomics: images are more than pictures, they are data[J]. Radiology, 2016, 278(2): 563-577. DOI: 10.1148/radiol.2015151169.

[8] TREBESCHI S, DRAGO S G, BIRKBAK N J, *et al*. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers[J]. Ann Oncol, 2019, 30(6): 998-1004. DOI: 10.1093/annonc/mdz108.

[9] HOSSEINI F, ASADI F, EMAMI H, *et al*. Machine learning applications for early detection of esophageal cancer: a systematic review[J]. BMC Med Inform Decis Mak, 2023, 23

(1): 124. DOI: 10.1186/s12911-023-02235-y.

[10] LI L R, QIN Z H, BO J, *et al*. Machine learning-based radiomics prognostic model for patients with proximal esophageal cancer after definitive chemoradiotherapy [J]. Insights Imaging, 2024, 15(1): 284. DOI: 10.1186/s13244-024-01853-y.

[11] CUI Y B, LI Z J, XIANG M Y, *et al*. Machine learning models predict overall survival and progression free survival of non-surgical esophageal cancer patients with chemoradiotherapy based on CT image radiomics signatures [J]. Radiat Oncol, 2022, 17(1): 212. DOI: 10.1186/s13014-022-02186-0.

[12] SHI Z W, ZHANG Z, LIU Z Y, *et al*. Methodological quality of machine learning-based quantitative imaging analysis studies in esophageal cancer: a systematic review of clinical outcome prediction after concurrent chemoradiotherapy[J]. Eur J Nucl Med Mol Imaging, 2022, 49(8): 2462-2481. DOI: 10.1007/s00259-021-05658-9.

[13] EISENHAUER E A, THERASSE P, BOGAERTS J, *et al*. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)[J]. Eur J Cancer, 2009, 45(2): 228-247. DOI: 10.1016/j.ejca.2008.10.026.

[14] CUI J F, LI L, LIU N, *et al*. Model integrating CT-based radiomics and genomics for survival prediction in esophageal cancer patients receiving definitive chemoradiotherapy [J]. Biomark Res, 2023, 11(1): 44. DOI: 10.1186/s40364-023-00480-x.

[15] GRAMEGNA A, GIUDICI P. SHAP and LIME: an evaluation of discriminative power in credit risk [J]. Front Artif Intell, 2021, 4: 752558. DOI: 10.3389/frai.2021.752558.

[16] GE F, HUO Z Y, CAI X Y, *et al*. Evaluation of clinical and safety outcomes of neoadjuvant immunotherapy combined with chemotherapy for patients with resectable esophageal cancer: a systematic review and meta-analysis [J]. JAMA Netw Open, 2022, 5(11): e2239778. DOI: 10.1001/jamanetworkopen.2022.39778.

[17] LI J Y, LI L, YOU P M, *et al*. Towards artificial intelligence to multi-omics characterization of tumor heterogeneity in esophageal cancer [J]. Semin Cancer Biol, 2023, 91: 35-49. DOI: 10.1016/j.semcancer.2023.02.009.

[18] WANG J L, TANG L S, ZHONG X, *et al*. A machine learning radiomics based on enhanced computed tomography to predict neoadjuvant immunotherapy for resectable esophageal squamous cell carcinoma [J]. Front Immunol, 2024, 15: 1405146. DOI: 10.3389/fimmu.2024.1405146.

[19] FANG P, JIANG W, DAVULURI R, *et al*. High lymphocyte count during neoadjuvant chemoradiotherapy is associated with improved pathologic complete response in esophageal cancer [J]. Radiother Oncol, 2018, 128(3): 584-590. DOI: 10.1016/j.radonc.2018.02.025.

[20] FENG J F, WANG L, YANG X, *et al*. Pretreatment pan-immune-inflammation value (PIV) in predicting therapeutic response and clinical outcomes of neoadjuvant immunochemotherapy for esophageal squamous cell carcinoma [J]. Ann Surg Oncol, 2024, 31(1): 272-283. DOI: 10.1245/s10434-023-14430-2.

[21] PARK J E, HAM S, KIM H S, *et al*. Diffusion and perfusion MRI radiomics obtained from deep learning segmentation provides reproducible and comparable diagnostic model to human in post-treatment glioblastoma [J]. Eur Radiol, 2021, 31(5): 3127-3137. DOI: 10.1007/s00330-020-07414-3.

[22] CHEN Z H, LIN L, WU C F, *et al*. Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine [J]. Cancer Commun (Lond), 2021, 41(11): 1100-1115. DOI: 10.1002/cac2.12215.

[23] HU Y H, XIE C Y, YANG H, *et al*. Computed tomography-based deep-learning prediction of neoadjuvant chemoradiotherapy treatment response in esophageal squamous cell carcinoma [J]. Radiother Oncol, 2021, 154: 6-13. DOI: 10.1016/j.radonc.2020.09.014.

[24] WANG J, ZHU X, ZENG J, *et al*. Using clinical and radiomic feature-based machine learning models to predict pathological complete response in patients with esophageal squamous cell carcinoma receiving neoadjuvant chemoradiation [J]. Eur Radiol, 2023, 33(12): 8554-8563. DOI: 10.1007/s00330-023-09884-7.

(收稿:2024-12-24;修回:2025-01-19)

(编辑 李 玥)