

• 胸部影像学 •

CT 影像组学联合细胞角蛋白 19 片段预测 EGFR 突变阳性非小细胞肺癌患者 EGFR-TKIs 的疗效

张成孟, 丁治民, 陈鹏, 刘奇峰

【摘要】 目的:探讨 CT 影像组学联合细胞角蛋白 19 片段在预测表皮生长因子受体(EGFR)突变阳性的非小细胞肺癌(NSCLC)患者行表皮生长因子受体酪氨酸激酶抑制剂(EGFR-TKIs)治疗的疗效。**方法:**回顾性搜集在本院确诊为 EGFR 突变阳性随即接受 EGFR-TKIs 治疗的 194 例 NSCLC 患者的病例资料。在 EGFR-TKIs 治疗 3 个月后进行 CT 检查来判断疗效。根据实体肿瘤疗效评价标准 1.1(RECIST 1.1), 治疗有效 121 例, 无效 73 例。采用完全随机方法将患者按 7:3 的比例分为训练集和验证集。在训练集中提取 NSCLC 病灶的组学特征, 然后使用主成分分析(PCA)、kruskal-wallis (KW)法及逻辑回归分析结合最小绝对值收敛和选择算子(LR-LASSO)对影像组学特征进行降维及影像组学模型的构建, 获得每例患者的影像组学标签值。利用临床资料、病灶的 CT 形态学特征和病理结果建立临床模型, 联合临床资料和影像组学标签建立联合模型。使用受试者工作特征曲线(ROC)下面积(AUC)评价各个模型对 EGFR-TKIs 疗效的预测效能, 应用决策曲线分析(DCA)评估模型的临床应用价值。**结果:**训练集中经 PCA、KW 方法对影像组学特征进行降维后得到 9 个影像组学特征用于模型的构建。对临床资料、病灶的 CT 形态学特征和病理结果使用单因素及多因素分析后显示仅细胞角蛋白 19 片段是疗效的独立预测因子, 以此构建临床模型; 联合模型由细胞角蛋白 19 片段结合影像组学标签进行构建。临床模型、影像组学模型及联合模型在训练集中的 AUC 分别为 0.686、0.800 和 0.836; 在验证集中的 AUC 分别为 0.666、0.774 和 0.837。DCA 显示联合模型较影像组学模型和临床模型有更好的临床净收益。**结论:**CT 影像组学联合细胞角蛋白 19 片段对预测 EGFR 突变阳性 NSCLC 患者 EGFR-TKIs 的疗效具有较高的价值。

【关键词】 影像组学; 体层摄影术, X 线计算机; 表皮生长因子受体; 非小细胞肺癌; 表皮生长因子受体酪氨酸激酶抑制剂; 预测模型

【中图分类号】 R814.48; R734.2 **【文献标志码】** A **【文章编号】** 1000-0313(2023)12-1532-07

DOI: 10.13609/j.cnki.1000-0313.2023.12.007

开放科学(资源服务)标识码(OSID):



The value of CT radiomics combined with cytokeratin-19 fragment in predicting the effectiveness of EGFR-TKIs for non-small cell lung cancer with EGFR mutation ZHANG Cheng-meng, DING Zhi-min, CHEN Peng, et al. Department of Radiology, Yijishan Hospital of Wannan Medical College, Anhui 241000, China

【Abstract】 Objective: To investigate the value of CT radiomics and cytokeratin-19 fragment in predicting the efficiency of tyrosine kinase inhibitors (TKI) in non-small cell lung cancer (NSCLC) patients with epidermal growth factor receptor (EGFR) mutations. **Methods:** A total of 194 NSCLC patients diagnosed as positive EGFR mutation and treated with EGFR-TKIs in our Hospital were retrospectively collected. The effect of the treatment was evaluated by CT images three months later. According to the response evaluation criteria in solid tumors 1.1 (RECIST 1.1), patients were categorized as effective (n=121), or ineffective (n=73). All 194 patients were divided into training set and validation set with a ratio of 7:3 by complete randomization method. In the training set, principal compo-

作者单位: 241001 安徽芜湖, 皖南医学院弋矶山医院放射科(张成孟、丁治民、刘奇峰); 313000 浙江湖州, 湖州市中心医院放射科(陈鹏)

作者简介: 张成孟(1998-), 男, 安徽滁州人, 硕士研究生, 住院医师, 主要从事放射诊断工作。

通讯作者: 丁治民, E-mail: dzmdoctor@139.com

基金项目: 中国红十字基金会医学赋能-领航菁英科研项目(XM_HR_YXFN_2021_05_24); 安徽省卫生健康科研项目(AHWJ2022b044); 安徽省教育厅 2022 年度新时代育人质量工程(研究生教育)项目(2022zyxwxalk165); 皖南医学院弋矶山医院引进人才专项科研基金(YR20220219); 安徽省教育厅高等学校科学研究(自然科学类)重大项目(2023AH040253)

nent analysis (PCA), Kruskal-wallis (KW) methods and logistic regression via least absolute shrinkage and selection operator (LR-LASSO) were used to perform dimensionality reduction and model construction, then the radscore values was calculated. The clinical model was established by using clinical data, morphological features and pathological results, and the combined model was established by using clinical and radiomics labels. Area under curve (AUC) was used to evaluate the prediction efficiency of the models and the decision curve analysis (DCA) was used to evaluate the clinical application value of the models. **Results:** In the training set, nine radiomics features were extracted through PCA and KW algorithms to reduce the dimensionality of radiomics features, which were used for constructing the radiomics model. After conducting univariate and multivariate analysis on clinical data, CT morphological features of lesions, and pathological results, it was found that only the fragment of cytokeratin 19 was identified as an independent predictor. Based on this, a clinical model was constructed. Finally, the combined model was constructed by the cytokeratin 19 fragment with radscore. In the training set, the AUC of clinical model, radiomics model and combined model were 0.686, 0.800 and 0.836, and in the validation set, the AUC were 0.666, 0.774 and 0.837, respectively. DCA showed that the combined model had better benefits than the radiomics model and the clinical model. **Conclusion:** CT radiomics combined with cytokeratin-19 fragment has high value in predicting the efficacy of EGFR-TKIs in NSCLC patients with EGFR mutations.

【Key words】 Radiomics; Tomography, X-ray computed; Epidermal growth factor receptor; Non-small cell lung cancer; Epidermal growth factor receptor tyrosinase inhibitor; Predictive model

肺癌是全球第二常见的恶性肿瘤,研究表明 2020 年我国新发癌症病例数排名第一的为肺癌,约 82 万人,且肺癌死亡病例数高达 71 万^[1]。PIONEER 研究结果表明,在亚洲人群中,晚期肺腺癌患者的表皮生长因子受体(epidermal growth factor receptor, EGFR)突变的发生率高达 51.4%,且 EGFR 突变在女性和非吸烟者中更为常见,即使在经常吸烟的人群中这一比例仍有 37%^[2]。

研究表明,突变的 EGFR 是表皮生长因子受体酪氨酸酶抑制剂(epidermal growth factor receptor tyrosine kinase inhibitors, EGFR-TKIs)治疗非小细胞肺癌(non-small cell lung cancer, NSCLC)的理想靶点^[3]。近年来,随着 EGFR-TKIs 在临床的推广与使用^[4],NSCLC 患者的预后得到较大的改善^[5]。相比于化疗,EGFR-TKIs 治疗的副作用更小,患者的预后更佳。由于 EGFR-TKIs 的价格比较高,如果用药后无效,不但加重了患者的经济负担,还可能会耽误患者的治疗。因此,预测 EGFR 突变型 NSCLC 患者 EGFR-TKIs 的疗效对指导临床制订准确合理的个体化治疗方案具有相当重要意义。

材料与方法

1. 研究对象

回顾性搜集 2015 年 4 月—2022 年 6 月在本院诊断为 NSCLC 且行 EGFR 基因检测的患者的 CT 和临床资料。纳入标准:①原发病灶经活检病理证实为

NSCLC 且 EGFR 基因检测为突变阳性的患者;②有完整的临床资料(患者接受靶向治疗前的年龄、性别、肿瘤病理分型、癌胚抗原和细胞角蛋白 19 片段);③可获得患者靶向治疗前的 CT 图像;④可获得患者服用 EGFR-TKIs 3 个月后的 CT 图像。排除标准:①各种原因导致病灶边缘显示欠清,导致难以准确勾画病灶的 ROI;②肺部原发肿瘤已切除。

本研究最初将 378 例行 EGFR 基因检测的 NSCLC 患者纳入总样本,在排除 72 例 EGFR 突变阴性、20 例临床资料不全、29 例病灶边缘难以勾画、63 例服用 EGFR-TKIs 3 个月后 CT 资料缺失的患者后,最终入组 194 例。将入组患者按照 7:3 的比例随机分为训练集和验证集。

本回顾性研究获得了本院伦理委员会的批准,且豁免了患者知情同意。

2. 临床资料的选取

本研究纳入的临床指标包括性别、年龄、肿瘤病理分型、靶向治疗前、后肿瘤最大径、癌胚抗原(carcinoembryonic antigen, CEA)以及细胞角蛋白-19 片段(cytokeratin-19-fragment, CYFRA21-1)检测值。治疗前、后肿瘤最大径由 2 位工作经验分别为 5 和 15 年的影像医师在不知晓治疗结果的情况下分别在患者的胸部 CT 图像上进行测量,取两者测量结果的平均值。当 CT 平扫图像上肿瘤的边界不清时可参考胸部 CT 增强图像。以实体肿瘤疗效评价标准 1.1(response evaluation criteria in solid tumours, RECIST 1.1)为

依据对靶向治疗的效果进行评估,比较患者靶向治疗前及治疗 3 个月后 CT 图像上病灶的大小来进行判定,将完全缓解(complete response, CR)和部分缓解(partial response, PR)的患者纳入治疗有效组,病变稳定(stable disease, SD)和病变进展(pogressive disease, PD)的患者划为治疗无效组(图 1、2)。

3. CT 检查方法

所有患者在接受靶向治疗前及治疗 3 个月之后均进行了胸部 CT 检查。使用 Philips Brilliance 64 排螺旋 CT 机,患者取仰卧位,扫描参数:120 kV, 220 mA,层厚 5.0 mm,层距 5.0 mm,视野为 30 cm×30 cm。对比剂使用碘化醇(300 mg I/mL),剂量 1.5 mL/kg,注射流率 2.5 mL/s,在对比剂注射后 25 和 40~65 s 分别行动脉期和静脉期增强扫描。

4. 特征提取及预测模型的建立

将所有患者的胸部 CT 平扫图像从 PACS 系统以 DICOM 格式导入 ITK-SNAP 3.8.0 软件,并将图像进行匿名化处理。首先由一位影像科医师(具有 5 年工作经验)在胸部 CT 肺窗图像(窗宽 1500 HU,窗位 -500 HU)上沿肿瘤的边缘逐层勾画 ROI,注意避开液化和坏死区,对于部分边界欠清的病灶勾画时参考增强图像^[6]。最后,融合生成病灶的容积感兴趣区(volume of interest, VOI)。然后,将勾画了 ROI 的图像导入基于 Python 3.6.8 编写的开源组学分析平台 FeAture Explorer (FAE, v0.5.2, <https://github.com/salan668/FAE>)进行影像组学特征提取和分析^[7]。2 周后随机抽取 50 例患者的 CT 图像,由这位医师和另外一位影像科医师(具有 15 年以上工作经验)再次进行 ROI 的勾画和特征的提取。对两位医师及同一位医师前、后 2 次提取的特征进行可重复性检验,删除组内及组间相关系数(intra- and inter-class correlation coefficient, ICCs) < 0.80 的特征,然后将

筛选后的影像组学特征进行 Z-score 标准化处理,以确保数据之间的可比性。

在训练集中使用主成分分析(principal component analysis, PCA)、kruskal-wallis(KW)检验和逻辑回归分析结合最小绝对值收敛和选择算子(logistic regression via least absolute shrinkage and selection operator, LR-LASSO)对影像组学数据进行降维及模型构建,并采用 5 折法进行交叉验证,以提高模型的稳定性。通过计算每个纳入特征的值×其权重系数再计算总和得出每例患者的影像组学标签值(Radiomic score, Radscore)。

将组间比较差异有统计学意义的临床变量纳入多因素 logistics 回归分析,计算出治疗有效的独立预测因子并以此构建临床模型;然后,将有统计学意义的临床变量和影像组学标签纳入多因素 logistics 回归分析来构建联合模型。

5. 统计分析

使用 SPSS 25.0 软件和 R 4.2.1 软件进行统计分析。采用 Kolmogorov-Smirnov 检验对计量资料进行正态性检验,符合正态分布的计量资料采用均数±标准差表示,不符合正态分布的计量资料采用中位数(上、下四分位数)表示。计量资料的组间比较采用独立样本 *t* 检验(正态分布且方差齐)或 Mann-Whitney *U* 检验(偏态分布或方差不齐),计数资料的组间比较采用 χ^2 检验或 Fisher 确切概率法。采用多因素 logistic 回归分析构建预测模型并绘制诺模图,采用校正曲线评价模型的拟合优度($P > 0.05$ 表示拟合优度较好),并采用 Hosmer-Lemeshow 检验评价校正曲线的一致性。采用 AUC 评价各种模型对疗效的诊断效能,并采用 DeLong 检验比较不同模型间 AUC 的差异。采用决策曲线分析(decision curve analysis, DCA)比较模型的临床净收益。以 $P < 0.05$ 为差异有

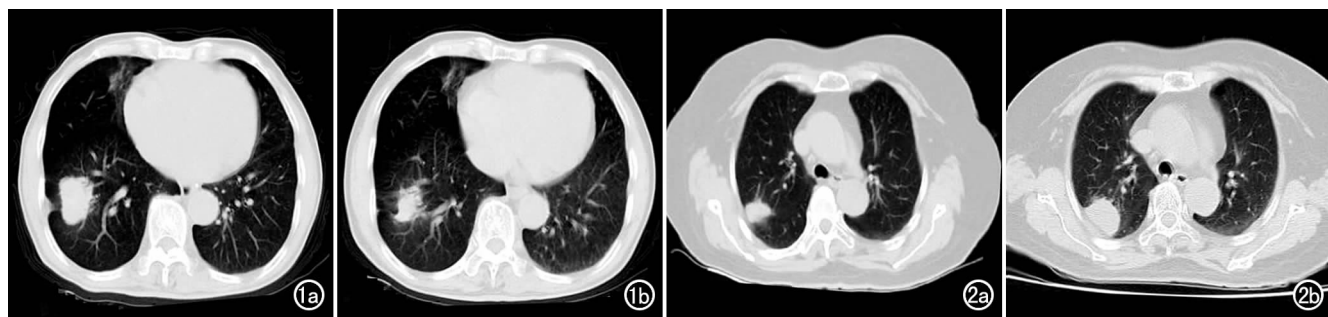


图 1 男性,76 岁。a)接受靶向治疗前的 CT 平扫肺窗图像,肿瘤最大径为 38mm;b)接受靶向治疗 3 个月后的 CT 平扫肺窗图像,肿瘤最大径为 21mm,与治疗前相比肿瘤最大径减少了约 45%,符合 PR 的标准,故该患者被纳入治疗有效组。图 2 女性,71 岁。a)接受靶向治疗前的 CT 平扫肺窗图像,肿瘤最大径为 24mm;b)接受靶向治疗 3 个月后的 CT 平扫肺窗图像,肿瘤最大径为 34mm,与治疗前相比肿瘤最大径增加了约 42%,符合 PD 的标准,故该患者被纳入治疗无效组。

统计学意义。

结 果

1. 临床资料的比较

194 例患者随机分为训练集和验证集,其中训练集 136 例,验证集 58 例,两个数据集之间以及 2 个数据集中治疗有效组与无效组之间各项临床指标、肿瘤形态学特征及病理结果的比较结果分别见表 1 和表 2。训练集与验证集之间性别、年龄、肿瘤病理分型、靶向治疗前肿瘤最大径及 CEA 和 CYFRA21-1 水平的差异均无统计学意义($P>0.05$)。在训练集和验证集中,治疗有效组与无效组之间年龄、性别、CEA、肺癌病理类型差异均无统计学意义($P>0.05$),治疗前肿瘤最大径和 CYFRA21-1 水平的差异均有统计学意义($P<0.05$)。

2. 影像组学特征提取及影像组学模型建立

自每例患者的病灶共提取到 1781 个影像组学特征,删除无效特征 196 个及组内和组间 ICC<0.80 的特征 71 个,最终剩余 1514 个特征。在训练集中通过

PCA、KW 法和 LR-LASSO 进行降维,最终筛选出 9 个影像组学特征,将其纳入回归模型中,获得每个特征相应的权重系数(表 3),计算出所有患者的 Radscore。

3. 建立临床和联合模型

将训练集中组间差异有统计学意义的临床变量纳入多因素 logistics 回归分析,结果显示 CYFRA21-1 ($OR=1.059, P=0.048$)为治疗有效的独立预测因子,以此构建临床模型。

将训练集中有统计学意义的临床病理指标(治疗前肿瘤最大径、细胞角蛋白 19 片段)及影像组学标签纳入多因素 logistic 回归分析,其中细胞角蛋白 19 片段($OR=1.058, P=0.037$)和影像组学标签($OR=3.590, P<0.001$)为独立预测因子,建立临床+影像组学联合模型。

4. 模型的诊断效能分析

采用 ROC 曲线分析各类影模型在训练集和验证集中预测 EGFR 突变阳性 NSCLC 患者 EGFR-TKIs 疗效的效能,结果见表 4、图 3。在训练集和验证集中,影像组学模型预测治疗有效的 AUC 分别为 0.800

表 1 训练集及验证集中临床资料的比较

指标	训练集(n=136)	验证集(n=58)	Z/ χ^2 值	P 值
年龄/岁	65.50(57.00,71.00)	67.50(58.00,73.50)	-1.416 ^b	0.157
性别比(男/女)	69/67	23/35	2.002 ^a	0.157
肿瘤最大径/cm	3.40(2.60,5.00)	3.85(2.85,5.13)	-0.951 ^b	0.341
CEA/ng/mL	14.68(4.36,90.16)	25.24(5.38,107.80)	-1.059 ^b	0.290
CYFRA21-1/ng/mL	3.98(2.50,10.04)	4.55(2.78,11.91)	-0.880 ^b	0.379
肺癌类型/例			— ^c	0.284
腺癌	12	57		
鳞癌	88	1		

注:CEA:癌胚抗原;CYFRA21-1:细胞角蛋白 19 片段;^a为 χ^2 值;^b为 Z 值;^c采用的是 Fisher 确切概率法,未提供统计值。

表 2 训练集及验证集中治疗有效组和无效组的临床资料、肿瘤形态学特征及病理结果的比较

项目	训练集(n=136)		Z/ χ^2 值	P 值	验证集(n=58)		Z/ χ^2 值	P 值
	有效组	无效组			有效组	无效组		
年龄/岁	66.00(58.00,70.00)	65.00(54.00,72.00)	-0.322 ^b	0.748	67.00(58.50,75.00)	68.00(55.00,72.2)	-0.465 ^b	0.642
性别比(男/女)	45/40	24/27	0.441 ^a	0.507	14/22	9/13	0.023 ^a	0.879
肿瘤最大径/cm	3.70(2.75,5.25)	3.00(2.30,4.30)	-2.525 ^b	0.012	4.3(3.53,5.38)	3.10(2.38,3.90)	-2.774 ^b	0.006
CEA/ng/mL	17.59(4.44,119.57)	10.20(4.11,80.95)	-1.000 ^b	0.317	32.67(9.11,152.34)	11.27(4.26,74.27)	-1.186 ^b	0.236
CYFRA21-1/ng/mL	5.64(3.01,14.48)	3.04(1.73,5.14)	-3.630 ^b	<0.001	5.40(3.10,19.80)	3.62(2.73,5.73)	-2.108 ^b	0.035
肺癌类型/例			— ^c	0.710			— ^c	1.000
腺癌	79	49			35	22		
鳞癌	6	2			1	0		

注:CEA:癌胚抗原;CYFRA21-1:细胞角蛋白 19 片段;^a为 χ^2 值;^b为 Z 值;^c采用的是 Fisher 确切概率法,未提供统计值。

表 3 影像组学模型中纳入的组学特征及其权重系数

特征名称	权重系数
exponential_firstorder_10Percentile	0.025
exponential_firstorder_Skewness	0.110
exponential_gldm_LargeDependenceLowGrayLevelEmphasis	0.188
exponential_glrlm_LongRunEmphasis	0.154
gradient_gldm_MaximumProbability	0.434
gradient_gldm_GrayLevelNonUniformity	-0.551
gradient_gldm_SmallDependenceEmphasis	0.652
gradient_glszm_GrayLevelNonUniformity	-0.629
gradient_ngtdm_Strength	-0.258

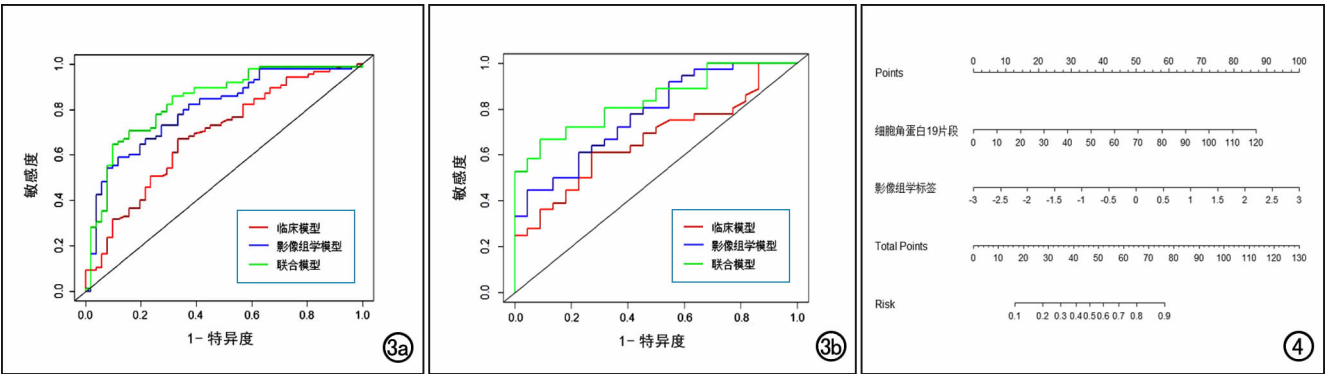


图 3 临床模型、影像组学模型及联合模型预测疗效的 ROC 曲线。a) 训练集; b) 验证集。图 4 联合模型中的 2 个变量组成的诺模图。每个变量对应应在模型中的得分, 得分相加得到总分, 再对应至下方的风险值。Points: 得分; Total Points: 总分; Risk: 风险。

(95%CI: 0.723~0.876) 和 0.774 (95%CI: 0.653~0.895); 临床模型的 AUC 分别为 0.686 和 0.666; 联合模型的 AUC 分别为 0.836 和 0.837, 优于临床模型和影像组学模型。

经 DeLong 检验, 训练集中联合模型与临床模型之间 AUC 的差异有统计学意义 ($Z = -3.136, P = 0.002$), 影像组学模型与临床模型及联合模型之间 AUC 的差异均无统计学意义 ($Z = -1.928, P = 0.054; Z = -1.958, P = 0.050$); 验证集中, 联合模型与临床模型之间 AUC 的差异有统计学意义 ($Z = -2.512, P = 0.012$), 影像组学模型与临床模型及联合模型之间 AUC 的差异均无统计学意义 ($Z = -1.109, P = 0.268; Z = -1.498, P = 0.134$)。

绘制联合模型的诺模图(图 4)。联合模型在训练集和验证集的校准曲线与理想曲线接近, 表明预测效果较好(图 5)。采用 Hosmer-Lemeshow 检验分析该模型的拟合优度, 得到一个不显著的统计学结果(训练集: $\chi^2 = 4.142, P = 0.126$, 验证集: $\chi^2 = 1.104, P = 0.576$), 表明模型未明显偏离完美的拟合。DCA 曲线显示联合模型预测 EGFR 突变阳性 NSCLC 患者 EGFR-TKIs 治疗疗效的临床净获益优于临床模型和影像组学模型(图 6)。

讨论

影像组学是指将数字医学图像转换为可挖掘的高维数据, 其在许多疾病的鉴别诊断、预后判断及疗效评价等方面的应用日渐增多^[8-11]。影像组学近年来因其

在提供客观和可量化的成像信息方面的优势而越来越受到人们的关注。大量研究证实影像组学在良恶性疾病的鉴别及对癌症分期、病理特征、基因组突变、预后及疗效的预测等方面具有可行性和有效性^[12-15]。目前对于 NSCLC EGFR 的相关研究报道中, 大部分都停留在影像组学特征预测 EGFR 基因的某些突变方面^[16,17], 对于能否基于 CT 影像组学特征分析对 EGFR 突变阳性的 NSCLC 患者 EGFR-TKIs 的疗效进行预测却少有研究。

2003 年, Kris 等^[18]的研究结果表明对存在 EGFR 突变的肺癌患者使用 EGFR-TKIs 治疗后可以获得较好的疗效, 开启了存在 EGFR 突变的非小细胞肺癌患者靶向治疗的时代。美国国家综合癌症网络(National Comprehensive Cancer Network, NCCN)和我国制订的肺癌治疗规范中均推荐对肺癌患者应尽可能检测 EGFR 基因状态, 对 EGFR 突变阳性 NSCLC 患者推荐进行 EGFR-TKIs 治疗。但是, EGFR-TKIs 对部分患者治疗效果不佳, 所以治疗前预测 EGFR-TKIs 的有效性具有较大的价值。本研究基于胸部 CT 平扫图像, 纳入临床、病理学、影像学及影像组学特征等多参数构建预测模型, 对 EGFR 突变阳性的 NSCLC 患者接受 EGFR-TKIs 治疗的疗效进行预测, 结果显示影像组学模型优于临床模型, 联合细胞角蛋白 19 片段和影像组学构建的联合模型对疗效的预测能力进一步提升, 优于影像组学模型和临床模型。

EGFR-TKIs 对 NSCLC 患者的治疗效果受多种因素的影响, 如临床特征、组织病理、肿瘤周围浸润程

表 4 临床、影像组学和联合模型的诊断效能

模型类型	训练集 (n=136)				验证集 (n=58)			
	AUC(95%CI)	符合率	敏感度	特异度	AUC(95%CI)	符合率	敏感度	特异度
临床模型	0.686 (0.593~0.779)	66.9%	67.1%	66.7%	0.666 (0.526~0.806)	65.5%	61.1%	72.7%
影像组学模型	0.800 (0.723~0.876)	69.9%	58.8%	88.2%	0.774 (0.653~0.895)	63.8%	44.4%	95.5%
联合模型	0.836 (0.765~0.908)	75.7%	70.6%	84.3%	0.837 (0.738~0.937)	75.9%	66.7%	90.9%

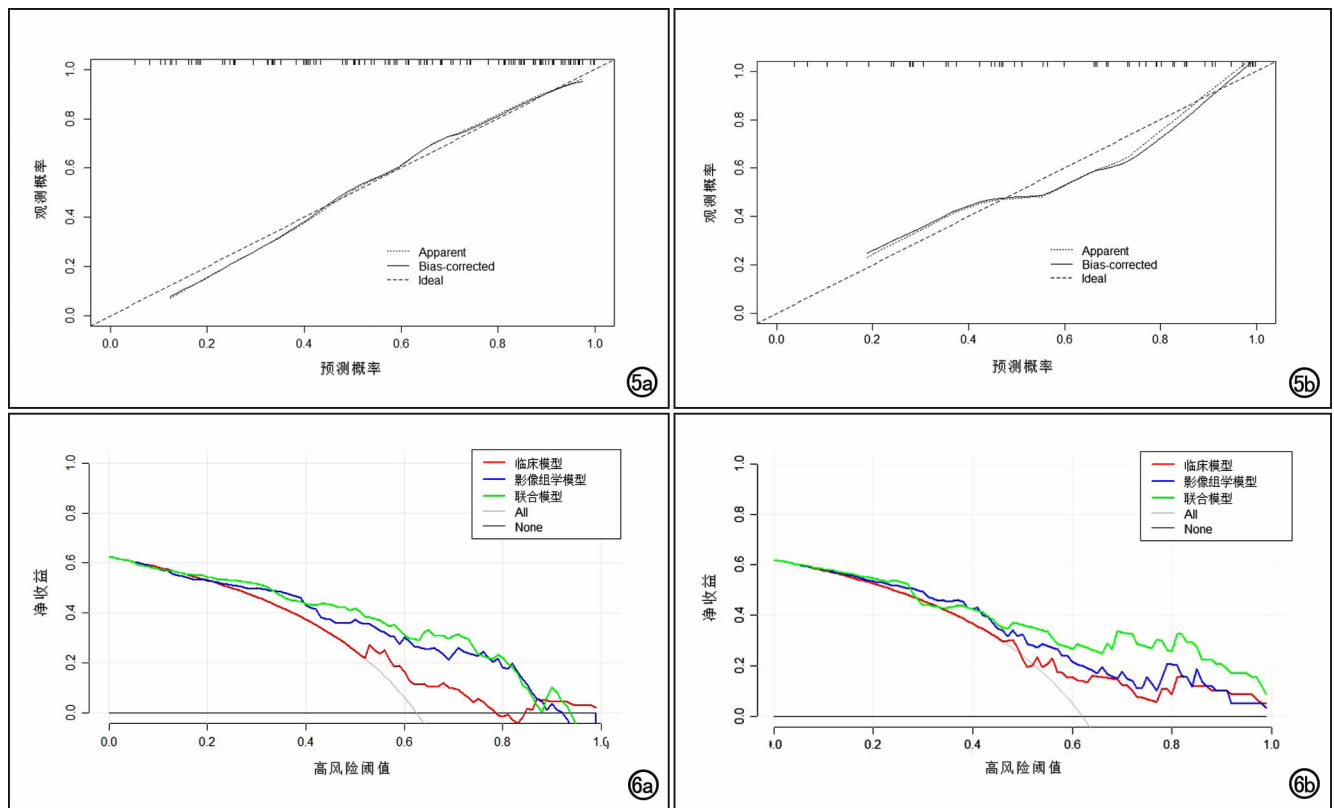


图 5 联合模型的诺模图的校准曲线。Apparent 为拟合线, Bias-corrected 为偏差校准曲线, Ideal 为参考线。a) 训练集; b) 验证集。图 6 临床模型、影像组学模型和联合模型预测 EGFR 突变阳性 NSCLC 患者 EGFR-TKIs 疗效的决策曲线, 结果显示在训练集和验证集中联合模型对预测 EGFR 突变阳性 NSCLC 患者 EGFR-TKIs 的临床价值最高, 优于临床模型和影像组学模型。All 表示所有治疗有效患者, None 表示所有治疗无效患者。a) 训练集; b) 验证集。

度和分子生物学特征等均会影响其疗效, 具体的机制也较为复杂。本研究中纳入分析的血清肿瘤标志物中的 CYFRA21-1 是诊断 NSCLC 特异性较高的指标^[19], 该指标在血清内的水平与 NSCLC 患者的预后具有较好的相关性。张明明等^[20]的研究结果表明奥希替尼治疗晚期 NSCLC 患者效果确切, 而且可以显著降低患者血清内细胞角蛋白 19 片段的水平, 进一步说明了细胞角蛋白 19 片段与肿瘤的预后可能有一定的相关性, 本研究中联合模型中纳入的细胞角蛋白 19 片段为预测 EGFR 突变阳性的 NSCLC 患者 EGFR-TKIs 疗效的独立预测因子, 治疗前细胞角蛋白 19 片段的水平对 EGFR-TKIs 的疗效有一定的预测价值。本研究基于联合模型构建了预测 NSCLC 患者疗效的诺模图, 建立了简单易行的治疗前评估疗效的方法, 医师可以在治疗前通过简单评分进行个体化的评价, 辅助临床医师在治疗前判断 EGFR 突变阳性的非小细胞肺癌患者采用 EGFR-TKIs 进行治疗的可行性。

本研究中采用 PCA 进行特征降维, PCA 常用于高维数据的降维, 在组学特征的筛选中应用广泛。Kickingeder 等^[21]利用 PCA 降维方法构建的影像组

学模型较好地预测了胶质母细胞瘤 (glioblastoma, GBM) 患者对贝伐单抗治疗的反应, AUC 达 0.90。穆建华等^[22]采用 PCA 降维方法构建了 6 种影像组学模型对脑胶质瘤的病理分级进行预测, AUC 均大于 0.8, 提示模型的诊断效能均较高, 验证了影像组学模型对胶质瘤术前分级具有较高的准确性。

本研究存在一定的局限性: ①样本数据是回顾性收集, 纳入的患者就诊时间跨度较大, 部分患者的信息不完整, 选择偏倚是不可避免的; ②传统影像学定性评价缺乏统一评价标准, ROI 的勾画可能受到勾画者主观因素的影响, 而且本研究中曾尝了 ROI 半自动及全自动勾画, 但同手动勾画相比其不能较好的识别肿瘤边界, 根据实体肿瘤疗效评价标准, 可能对肿瘤治疗的疗效判断产生一些影响; ③本研究中仅基于 CT 平扫图像进行特征的提取和分析; ④样本数较小且为单中心研究, 需要多中心大样本进一步研究验证。

综上所述, 基于临床、影像组学标签建立的联合模型对预测 EGFR 突变阳性非小细胞肺癌患者 EGFR-TKIs 的疗效具有较高的价值, 优于临床模型及影像组学模型, 对临床合理制定个体化治疗方案及科学的评

估治疗效果具有较高的价值。

参考文献：

[1] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020; GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2021, 71(3): 209-249.

[2] Shi Y, Au JS, Thongprasert S, et al. A prospective, molecular epidemiology study of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (PIONEER)[J]. Thorac Oncol, 2014, 9(2): 154-162.

[3] Zhou C, Wu YL, Chen G, et al. Erlotinib versus chemotherapy as first-line treatment for patients with advance EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomise, phase 3 study[J]. Lancet Oncol, 2011, 12(8): 735-742.

[4] 李晓琳, 谢鹏, 孟雪, 等. 表皮生长因子受体分子显像与基因突变在非小细胞肺癌中的研究进展[J]. 中华核医学与分子影像杂志, 2012, 32(5): 393-396.

[5] Recondo G, Facchinetti F, Olaussen KA, et al. Making the first move in EGFR-driven or ALK-driven NSCLC: first-generation or next-generation TKI[J]. Nat Rev Clin Oncol, 2018, 15(11): 694-708.

[6] Benelli M, Barucci A, Zoppetti N, et al. Comprehensive analysis of radiomic datasets by RadAR[J]. Cancer Res, 2020, 80(15): 3170-3174.

[7] van Griethuysen, Joost JM, Fedorov A, et al. Computational radiomics system to decode the radiographic phenotype[J/OL]. Cancer Res, 2017, 77(21): e104-e107 [2023-03-21]. <https://pubmed.ncbi.nlm.nih.gov/29092951>. DOI: 10.1158/0008-5472.CAN-17-0339.

[8] Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology[J]. Ann Oncol, 2017, 28(6): 1191-1206.

[9] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data[J]. Radiology, 2016, 278(2): 563-577.

[10] 陈俊飞, 王笑笑, 胡景卉, 等. 基于 CT 平扫影像组学模型对早期急性坏死性胰腺炎的鉴别诊断价值[J]. 放射学实践, 2023, 38(2): 177-182.

[11] 王卓, 刘世莉, 丁伟, 等. 基于 CT 影像组学结合临床影像特征预测局部晚期鼻咽癌诱导化疗疗效[J]. 放射学实践, 2023, 38(1): 20-26.

[12] Avanzo M, Stancanella J, El Naqa I. Beyond imaging: the promise of radiomics[J/OL]. Phys Med, 2017, 38: e122-e139 [2023-03-21]. <https://pubmed.ncbi.nlm.nih.gov/28595812>. DOI: 10.1016/j.ejmp.2017.05.071.

[13] Gresser E, Eva W, Piotr M, et al. Radiomics signature using manual versus automated segmentation for lymph node staging of bladder cancer[J]. Eur Urol Focus, 2023, 9(1): 145-153.

[14] Ge S, Yi XY, Jia D, et al. Application of mammography-based radiomics signature for preoperative prediction of triple-negative breast cancer[J/OL]. BMC Medical Imaging, 2022, 22(1): e166 [2022-09-14]. DOI: 10.1186/s12880-022-00875-6.

[15] Liu X, Zhang YF, Sun ZN, et al. Prediction of pelvic lymph node metastasis in prostate cancer using radiomics based on T/T₂-weighted imaging[J]. J Cent South Univ, 2022, 47(8): 1025-1036.

[16] Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning[J/OL]. Eur Respir, 2019, 53(3): e1800986 [2019-03-28]. DOI: 10.1183/13993003.00986-2018.

[17] Rios Velazquez E, Parmar C, Liu Y, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer[J]. Cancer Res, 2017, 77(14): 3922-3930.

[18] Kris MG, Natale RB, Herbst RS, et al. Efficacy of gefitinib an inhibitor of the epidermal growth factor receptor tyrosine kinase in symptomatic patients with non-small cell lung cancer: a randomized trial[J]. JAM, 2003, 290(16): 2149-2158.

[19] 吴勤如, 刘飞. 血清 TK1、CEA、CYFRA21-1 联合检测对非小细胞肺癌的诊断价值[J]. 国际检验医学杂志, 2017, 38(20): 2921-2923.

[20] 张明明, 赵云龙, 许新举. 奥希替尼治疗晚期非小细胞肺癌的疗效及其对血清癌胚抗原和细胞角蛋白 19 片段水平的影响[J]. 肿瘤基础与临床, 2022, 35(2): 173-176.

[21] Kickingeder P, Götz M, Wick A, et al. Large-scale radiomic profiling of recurrent glioblastoma identifies an imaging predictor for stratifying anti-angiogenic treatment response[J]. Clin Cancer Res, 2017, 22(23): 5765-5771.

[22] 穆建华, 张雁伟, 吴志钢. 基于常规 MRI 图像的不同影像组学模型在脑胶质瘤术前分级中的应用[J]. 磁共振成像, 2020, 11(1): 55-59.

(收稿日期: 2023-01-09 修回日期: 2023-06-19)