

# 基于胸部 CT 影像组学特征分析预测非小细胞肺癌 ALK 基因突变的价值

曹莉, 谢玉海, 张宏杰, 庞奇, 王小雷, 高续, 宋岩

**【摘要】** 目的:探讨基于深度学习自动分割技术的胸部 CT 影像组学分析对非小细胞肺癌(NSCLC)间变性淋巴瘤激酶(ALK)基因突变的预测价值。方法:回顾性分析 2020 年 11 月—2024 年 4 月在本院经病理证实的 172 例 NSCLC 患者的临床及术前 CT 平扫资料。根据 ALK 基因检测结果,将总样本分为 ALK 阳性组(23 例)和阴性组(149 例)。采用深度学习自动分割技术基于 CT 图像进行肿瘤的三维分割,采用重采样技术(大小  $1\text{ mm}^3$ )进行图像预处理后提取肿瘤的影像组学特征,采用 Pearson 相关分析、SelectKBest 及 Logistic 回归分析筛选最优组学特征参数并构建影像组学模型。采用 ROC 曲线分析评估最优特征参数及影像组学模型的诊断效能,决策曲线分析评价模型的临床应用价值。结果:临床资料中,年龄、性别、病理 T 分期、结节部位、大小、体积、CT 值及 CT 征象在两组间的差异均无统计学意义( $P>0.05$ )。共筛选出 4 个最优组学特征,分别为 lbp-3D-m1\_firstorder\_Median、logarithm\_glcm\_Idmn、logarithm\_glcm\_InverseVariance 和 wavelet-LHH\_firstorder\_Median。四个特征参数的组间差异均具有统计学意义( $P<0.01$ ),它们判断 NSCLC 患者 ALK 突变的诊断效能分别为 0.680、0.674、0.756 和 0.682。四者构建的影像组学模型对 ALK 突变的预测效能最佳,AUC、敏感度和特异度分别为 0.868、78.26%和 83.89%。决策曲线显示影像组学模型预测 NSCLC ALK 基因突变状态具有较好的净收益。结论:联合深度学习自动分割技术的影像组学方法在无创性预测 NSCLC ALK 突变状态方面具有较高的效能。

**【关键词】** 非小细胞肺癌;间变性淋巴瘤激酶;体层摄影术,X 线计算机;影像组学;预测模型

**【中图分类号】** R814.42;R734.2 **【文献标志码】** A **【文章编号】** 1000-0313(2026)04-0404-06

DOI:10.13609/j.cnki.1000-0313.2026.04.008

开放科学(资源服务)标识码(OSID):



**Value of radiomics analysis based on chest CT images for predicting ALK gene mutations in non-small cell lung cancer** CAO Li, XIE Yu-hai, ZHANG Hong-jie, et al. Department of Radiology, Taihe People's Hospital, Taihe Hospital affiliated to Wannan Medical College, Anhui 236600, China

**【Abstract】** **Objective:** To investigate the predictive value of chest CT radiomics feature analysis based on deep learning automatic segmentation technology for anaplastic lymphoma kinase (ALK) gene mutation in non-small cell lung cancer (NSCLC) patients. **Methods:** A retrospective analysis was conducted on the clinical and preoperative plain CT image data of 172 patients with pathologically confirmed NSCLC in our hospital from November 2020 to April 2024. According to the ALK gene test results, the entire sample set was divided into the ALK-positive group (23 cases) and the ALK-negative group (149 cases). Three-dimensional segmentation of tumors based on CT images was carried out utilizing deep learning automatic segmentation technology. Image preprocessing was implemented through resampling technique (with volume of  $1\text{ mm}^3$ ), and radiomics features of the tumor were extracted. Pearson correlation, SelectKBest, and logistic regression analysis were utilized to screen the optimal radiomics feature parameters, and to construct a radiomics model. The receiver operating characteristic (ROC) curve analysis was employed to evaluate the diagnostic efficacy of the optimal feature parameters and the radiomics model, while the decision curve analysis (DCA) was utilized to assess the clinical application value of the model. **Results:** There was no statistically significant difference in age, gender, pathological T-stage, nodule location and size, volume, CT value, and CT signs between the two groups (all  $P>0.05$ ). Four optimal radiomics feature parameters were selected out, including lbp-3D-m1\_firstorder\_Median, logarithm\_glcm\_Idmn, logarithm\_glcm\_InverseVariance and wavelet-LHH\_firstorder\_Median. The difference of these features between the two groups were statistically significant (all  $P<0.01$ ), with corresponding diagnostic efficacy for ALK mutation in NSCLC patients of 0.680, 0.674, 0.756, and 0.682, respectively. The radiomics model constructed by the four factors demonstrated the best predictive efficacy for ALK mutation, with AUC, sensitivity, and specificity of 0.868, 78.26%, and 83.89%, respectively. The decision curve indicated that the radiomics radiomics model had a favorable net benefit in predicting ALK gene mutation status in NSCLC patients. **Conclusion:**

作者单位:236600 安徽阜阳.太和县人民医院/皖南医学院附属太和医院放射影像科(曹莉、谢玉海、高续),心胸外科(张宏杰、庞奇、王小雷);安徽界首 236500,安徽第二医学院附属界首医院(首市人民医院)医学影像科(宋岩)

作者简介:曹莉(1984—),女,安徽阜阳人,副主任医师,主要从事影像诊断和人工智能研究工作。

通讯作者:宋岩, E-mail: jssy721@163.com

基金项目:安徽省高校科学研究重点项目(2024AH051914);北京医学奖励基金会睿影基金资助项目(YXJL-2022-0105-0116, YXJL-2023-0866-0344);皖南医学院科研项目(WK2023JXYY100)

The radiomics approach, when integrated with deep-learning automatic segmentation technique, demonstrates high effectiveness in the non-invasive prediction of ALK mutation status among NSCLC patients.

**【Key words】** Pulmonary neoplasms; Non-small cell lung cancer; Anaplastic lymphoma kinase; Tomography, X-ray computed; Radiomics; Predicting model

肺癌是我国发病率和死亡率最高的恶性肿瘤<sup>[1]</sup>,非小细胞肺癌(non-small cell lung cancer, NSCLC)是最常见的组织学类型,约占 85%<sup>[2]</sup>。在 NSCLC 中间变性淋巴瘤激酶(anaplastic lymphoma kinase, ALK)基因突变是第二常见的分子突变类型,突变率约为 4%~9%<sup>[3-4]</sup>。虽然 ALK 基因突变率较低,但 ALK 阳性患者淋巴结转移率高,疾病进展快,其复发率是 ALK 阴性患者的 2 倍<sup>[5]</sup>。然而随着酪氨酸激酶抑制剂(tyrosine kinase inhibitor, TKI)的临床应用,其可显著提高 ALK 阳性患者的无进展生存期<sup>[6-7]</sup>。因此,精准检测 ALK 突变对 NSCLC 患者临床治疗方案的选择至关重要。尽管基因检测是判定 ALK 突变状态的金标准,但其为侵入性操作且成本相对高昂,限制了其在临床上常规应用的可行性。影像组学可高通量提取与肿瘤生物学行为和微观结构变化密切相关的定量参数,并构建反应肿瘤异质性的组学模型<sup>[8-10]</sup>。因此,构建一种基于深度学习自动分割技术的无创、便捷和可靠预测 ALK 突变的模型是很有必要的。本研究旨在探讨基于深度学习自动分割的胸部 CT 影像组学模型术前预测 NSCLC 患者 ALK 基因突变的临床价值。

## 材料与方法

### 1. 一般资料

回顾性纳入 2020 年 11 月—2024 年 4 月于本院

经术后病理确诊为 NSCLC 的患者。纳入标准:①行 ALK 基因检测;②有完整的影像及临床资料;③无其它恶性肿瘤病史。排除标准:①既往行放疗化疗者;②图像伪影显著,影响结节分割。最终共纳入 172 例患者,包括男 55 例、女 117 例,年龄 28~82 岁,平均(59.58±10.57)岁。

本研究经本院伦理委员会批准(2025-16)。

### 2. CT 检查方法

胸部 CT 扫描使用 Philips Brilliance 64 或 256i 和联影 uCT760 CT 机。受试者均采用仰卧位、头先进,扫描前嘱患者吸气后屏气。扫描范围自肺尖至肺底,完整覆盖双侧全肺野。扫描参数:120 kV,管电流采用自动毫安调制技术,螺距 0.6,层厚 5 mm。进一步采用 1 mm 层厚进行薄层重建,重建间隔 0.7 mm。

### 3. 临床资料和 CT 征象分析

搜集每例患者的临床资料,包括患者基本信息(年龄、性别、吸烟史)和术后病理 T 分期。吸烟史根据患者主诉确定。

采用肺窗(窗宽 1500 HU,窗位 -400 HU)及纵隔窗(窗宽 400 HU,窗位 40 HU)对结节形态学特征的分析 and 测量,包括结节的最大径、分布、最大、最小和平均 CT 值、分叶征、毛刺征、空泡征、血管穿行征、胸膜牵拉征、支气管狭窄或闭塞。结节的最大、最小和平均 CT 值及结节最大径是使用数坤肺结节人工智能

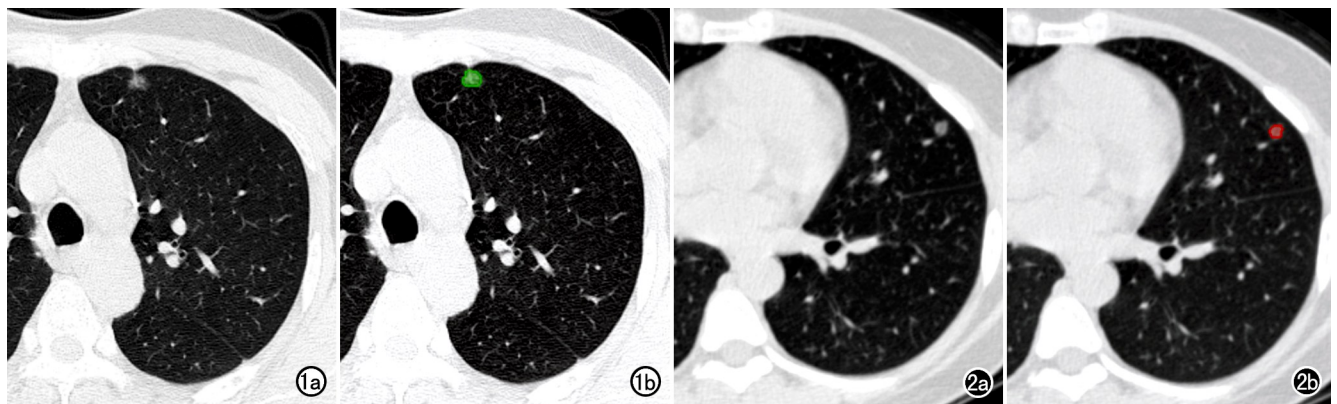


图 1 患者女,54 岁,不吸烟,体检行胸部 CT 扫描,术后病理诊断为浸润型腺癌,基因检测 ALK 基因未发生突变。a)左肺上叶磨玻璃结节,边界清晰,周围可见短毛刺和胸膜牵拉征;b)结节 ROI 自动分割示意图,提取的 4 个影像组学特征值:lbp-3D-m1\_firstorder\_Median = -2.473, logarithm\_glc\_m\_Idmn = -0.241, logarithm\_glc\_m\_InverseVariance = 1.656, wavelet-LHH\_firstorder\_Median = -2.843。图 2 患者女,38 岁,不吸烟。体检行胸部 CT 扫描,术后病理诊断为微浸润腺癌,基因检测 ALK 基因突变。a)左肺上叶磨玻璃结节,边界清晰;b)结节 ROI 自动分割示意图,提取的影像组学特征值:lbp-3D-m1\_firstorder\_Median = 1.234, logarithm\_glc\_m\_Idmn = -1.522, logarithm\_glc\_m\_InverseVariance = -0.301, wavelet-LHH\_firstorder\_Median = 3.511。

软件分析获得,支气管狭窄或闭塞、分叶征、毛刺征、空泡征、血管穿行征、胸膜牵拉征是由两位心胸组副主任医师评估确定。

#### 4. 图像分割及特征提取

肺结节图像分割及影像组学特征提取在数坤科技多模态科研平台完成。首先,采用数坤肺结节人工智能软件对结节进行自动标注(图 1~2);然后,为消除设备间分辨率不一致的问题,对所有 ROI 进行了空间标准化,重采样为体素大小为  $1.0 \text{ mm}^3$  各向同性图像;随后,系统自上述预处理后的图像中提取了肺结节的 1874 个定量特征。所提取的特征集包括一阶统计量特征 360 个、形状特征 14 个和纹理特征 1500 个,其中纹理特征包括灰度共生矩阵(gray level co-occurrence matrix, GLCM)特征 480 个、灰度区域大小矩阵(gray level size zone matrix, GLSZM)特征 320 个、灰度行程矩阵(gray level run length matrix, GLRLM)特征 320 个、灰度依赖矩阵(gray level dependence matrix, GLDM)特征 280 个、(neighbouring gray tone difference matrix, NGTDM)特征 100 个。

#### 5. 特征筛选及模型构建

采用 10 折交叉验证对影像组学特征进行标准化处理后,再对特征进行预处理(剔除特征值相同的特征)、Pearson 相关性分析( $r=0.80$ )和 SelectKBest 单变量筛选,最后采用单因素和多因素 Logistic 回归分析筛选出最优组学特征并构建 Logistic 回归预测模型。

#### 6. 统计学方法

使用 SPSS 22.0 软件进行统计学分析。符合正态分布的计量资料以均数±标准差表示,组间比较采用独立样本  $t$  检验;非正态分布的计量资料以中位数(四分位间距)表示,组间比较采用 Mann-Whitney  $U$  检验;分类资料以频数表示,组间比较采用  $\chi^2$  检验。使用 MedCalc 19.0.2 软件绘制模型的 ROC 曲线,并采用 Delong 检验比较不同模型间曲线下面积(AUC)的差异。此外,利用 R3.6.0 软件绘制模型的校准曲线和决策曲线。以  $P<0.05$  为差异有统计学意义。

### 结果

#### 1. 临床资料和常规 CT 征象的比较

172 例 NSCLC 患者中,ALK 突变型 23 例(突变率为 13.37%),野生型 149 例。两组间临床资料和常规 CT 征象的比较结果见表 1。ALK 野生型患者吸烟率高于突变型(25.5% vs. 4.3%),组间差异有统计学意义( $P=0.024$ ),其它临床资料和 CT 特征在两组间的差异均无统计学意义( $P>0.05$ )。

#### 2. 影像组学特征筛选及模型构建

首先通过特征预处理剔除 263 个冗余影像组学特征,然后经 Pearson 相关性分析(阈值 $\geq 0.80$ )及 SelectKBest 方法分别剔除 1302 和 289 个特征,最后共筛选出与 ALK 突变状态高度相关的 20 个最优影像组学特征。再进行单因素分析,结果见表 2。将其中

表 1 ALK 野生型与突变型 NSCLC 患者临床资料和常规 CT 征象的比较

指标	野生型(n=149)	突变型(n=23)	统计量	P 值
年龄/岁	59.52±11.05	57.52±9.52	0.820 <sup>a</sup>	0.414
性别/例			1.279 <sup>b</sup>	0.258
男性	50(33.6%)	5(21.7%)		
女性	99(66.4%)	18(78.3%)		
吸烟史/例	38(25.5%)	1(4.3%)	5.086 <sup>b</sup>	0.024
部位/例			0.469 <sup>b</sup>	0.493
右肺	89(59.7%)	12(52.2%)		
左肺	60(40.3%)	11(47.8%)		
T 分期/例			6.620 <sup>b</sup>	0.085
T1 期	125(83.9%)	19(82.6%)		
T2 期	17(11.4%)	2(8.7%)		
T3 期	7(4.7%)	1(4.3%)		
T4 期	0(0.0%)	1(4.3%)		
支气管狭窄或闭塞/例	12(8.1%)	5(21.7%)	4.165 <sup>b</sup>	0.056
分叶征/例	133(89.3%)	21(91.3%)	0.088 <sup>b</sup>	1.000
毛刺征/例	90(60.4%)	10(43.5%)	2.345 <sup>b</sup>	0.126
空泡征/例	22(14.8%)	6(26.1%)	1.863 <sup>b</sup>	0.221
血管穿行征/例	46(30.9%)	6(26.1%)	0.216 <sup>b</sup>	0.642
胸膜牵拉征/例	57(38.3%)	7(30.4%)	0.522 <sup>b</sup>	0.470
肿瘤长径/mm	16.51(11.00,23.50)	13.00(8.00,23.00)	-1.096 <sup>c</sup>	0.273
平均 CT 值/HU	-374.00(-557.00,-12.55)	-338.00(-568.10,-74.00)	-0.229 <sup>c</sup>	0.819
最大 CT 值/HU	75.00(-50.50,117.50)	-20.00(-87.00,122.00)	-0.583 <sup>c</sup>	0.560
最小 CT 值/HU	-813.00(-901.00,-716.50)	-837.00(-915.00,-697.00)	-0.349 <sup>c</sup>	0.727
结节体积/mm <sup>3</sup>	1506.00(482.0,3447.0)	631.0(230.0,3415.0)	-1.174 <sup>c</sup>	0.240

注:<sup>a</sup>为  $t$  值;<sup>b</sup>为  $\chi^2$  值;<sup>c</sup>为  $Z$  值。

$P < 0.05$  的特征纳入多因素二元 Logistic 回归分析, 结果见表 3。最终确定 4 个最优特征用于构建影像组学模型并可计算得到影像组学评分(Rad-score), 相应的计算公式:  $\text{Rad-score} = 0.964 \times \text{lbp-3D-m1\_firstorder\_Median} - 0.465 \times \text{logarithm\_glcm\_Idmn} - 1.085 \times \text{logarithm\_glcm\_InverseVariance} + 0.732 \times \text{wavelet-LHH\_firstorder\_Median} - 2.715$ 。

### 3. 影像组学模型及特征的效能评估

对影像组学模型及其中包含的 4 个特征进行 ROC 曲线分析, 结果见表 4 和图 3。影像组学特征对 ALK 基因突变的诊断效能从低到高依次为  $\text{logarithm\_glcm\_Idmn}$  (AUC1 = 0.674)、 $\text{lbp-3D-m1\_firstorder\_Median}$  (AUC2 = 0.680)、 $\text{wavelet-LHH\_firstorder\_Median}$  (AUC3 = 0.682) 和  $\text{logarithm\_glcm\_InverseVariance}$  (AUC4 = 0.756)。Delong 检验显示 4 个影像组学特征间诊断效能的差异均无统计学意义 (AUC1 vs. AUC2;  $Z = 0.069, P = 0.945$ ; AUC1 vs. AUC3;  $Z = 0.122, P = 0.903$ ; AUC1 vs. AUC4;  $Z = 0.869,$

$P = 0.385$ ; AUC2 vs. AUC3;  $Z = 0.028, P = 0.978$ ; AUC2 vs. AUC4;  $Z = 1.101, P = 0.271$ ; AUC3 vs. AUC4;  $Z = 0.948, P = 0.343$ )。联合模型的 AUC 为 0.868, 高于单一特征。Delong 检验显示, 联合模型的 AUC 与 4 个独立参数间的差异均有统计学意义 ( $Z = 3.738, 2.829, 2.837, 3.592, P$  均  $< 0.01$ )。

校准曲线与理想曲线接近, 说明模型对 ALK 基因突变的预测价值与观察值具有较高的一致性 (图 4)。决策曲线分析结果表明, 在阈值概率为 4% ~ 95% 范围内, 影像组学模型在预测 NSCLC ALK 基因突变状态时, 具有良好的临床净收益 (图 5)。

## 讨论

本研究中基于 NSCLC 患者胸部 CT 平扫图像利用深度学习自动分割技术构建了一个预测 ALK 基因突变的影像组学模型, 模型的 AUC 达 0.868, 敏感度和特异度分别为 78.26% 与 83.89%, 显示出较高的预测效能。研究结果凸显了影像组学模型在 ALK 基因

表 2 ALK 野生型与突变型组之间 20 个最优影像组学特征的比较

特征参数	ALK 阴性组 (n=149)	ALK 阳性组 (n=23)	t/Z 值	P 值
gradient_gldm_SmallDependenceLowGrayLevelEmphasis	0.013(0.010, 0.017)	0.015(0.011, 0.021)	-1.788	0.074
gradient_glszm_SmallAreaLowGrayLevelEmphasis	0.052(0.034, 0.084)	0.064(0.040, 0.143)	-1.392	0.164
lbp-3D-m1_firstorder_Median	10.830(10.314, 10.830)	10.830(10.830, 11.346)	-2.959	0.003
log-sigma-1-0-mm-3D_gldm_InverseVariance	0.409(0.383, 0.434)	0.425(0.375, 0.448)	-1.064	0.287
log-sigma-1-0-mm-3D_gldm_SmallDependenceLowGrayLevelEmphasis	0.003(0.001, 0.007)	0.006(0.001, 0.009)	-1.208	0.227
logarithm_glcm_Idmn	0.985(0.967, 0.994)	0.964(0.926, 0.992)	-2.684	0.007
logarithm_glcm_InverseVariance	0.288±0.105	0.191±0.085	4.220	0.000
logarithm_ngtdm_Coarseness	0.004(0.002, 0.014)	0.007(0.001, 0.020)	-0.283	0.777
original_ngtdm_Strength	2.422(0.916, 7.278)	2.349(0.771, 4.828)	-0.074	0.941
square_glcm_JointEnergy	0.023(0.012, 0.230)	0.182(0.020, 0.846)	-2.911	0.004
square_ngtdm_Strength	1.894(0.452, 4.518)	1.904(0.513, 5.469)	-0.279	0.780
squareroot_gldm_DependenceNonUniformityNormalized	0.077(0.058, 0.130)	0.142(0.058, 0.203)	-1.833	0.067
wavelet-HHH_glcm_Idmn	0.972(0.963, 0.979)	0.966(0.956, 0.976)	-1.392	0.164
wavelet-HLL_firstorder_Skewness	-0.650(-1.252, -0.056)	-0.212(-0.945, 0.108)	-1.703	0.089
wavelet-LHH_firstorder_Median	-0.197(-0.607, 0.245)	0.163(-0.301, 1.068)	-2.805	0.005
wavelet-LHH_gldm_LargeDependenceLowGrayLevelEmphasis	1.798(0.979, 3.112)	3.137(0.639, 4.009)	-0.817	0.414
wavelet-LHH_gldm_SmallDependenceLowGrayLevelEmphasis	0.005(0.004, 0.113)	0.008(0.003, 0.016)	-1.136	0.256
wavelet-LLH_firstorder_Maximum	250.185(113.397, 379.325)	158.819(98.867, 216.828)	-1.685	0.092
wavelet-LLH_glcm_Correlation	0.192±0.140	0.122±0.102	2.298	0.023
wavelet-LLH_glrml_GrayLevelVariance	7.887(2.596, 15.792)	3.005(2.064, 5.382)	-2.333	0.020

表 3 多因素二元 Logistic 回归分析结果

特征参数	回归系数	标准误	Wald 值	P 值	OR(95%CI)值
lbp-3D-m1_firstorder_Median	0.964	0.341	7.981	0.005	2.623(1.343~5.119)
logarithm_glcm_Idmn	-0.465	0.215	4.699	0.030	0.628(0.412~0.956)
logarithm_glcm_InverseVariance	-1.085	0.325	11.151	0.001	0.338(0.179~0.639)
wavelet-LHH_firstorder_Median	0.732	0.301	5.929	0.015	2.079(1.153~3.748)
常数	-2.715				

表 4 影像组学特征及影像组学模型预测 NSCLC ALK 突变的效能

特征参数	AUC(95%CI)	最佳界值	敏感度	特异度	Z 值	P 值
lbp-3D-m1_firstorder_Median	0.680(0.605~0.749)	10.572	91.30%	41.61%	3.738	<0.001
logarithm_glcm_Idmn	0.674(0.599~0.413)	0.978	69.57%	67.11%	2.829	0.005
logarithm_glcm_InverseVariance	0.756(0.685~0.818)	0.255	86.96%	60.40%	2.837	0.005
wavelet-LHH_firstorder_Median	0.682(0.607~0.751)	0.073	60.87%	73.15%	3.592	<0.001
影像组学模型	0.868(0.808~0.914)	-1.737	78.26%	83.89%		

注: Z 值和 P 值为影像组学各特征参数与影像组学联合模型的 AUC 进行比较的 Delong 检验值。

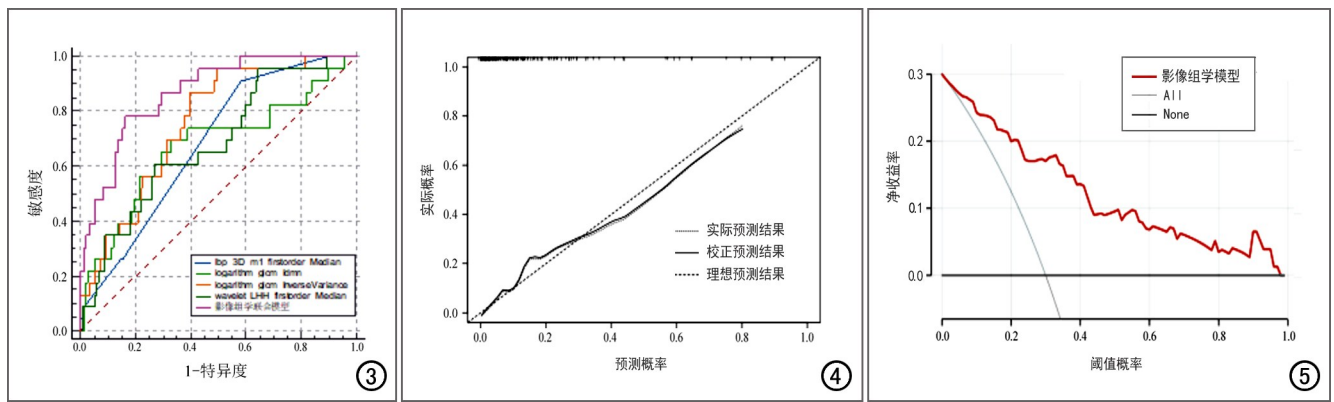


图3 四个最优影像组学特征及影像组学模型预测 ALK 基因突变的 ROC 曲线,显示模型的预测效能最高。图4 影像组学模型的决策曲线,显示组学模型在预测 ALK 基因突变状态时,具有良好的临床净收益。图5 影像组学模型的校准曲线,显示模型对 ALK 基因突变的预测值与观察值具有较高的一致性。

突变无创性预测方面具有较高的潜力,该方法有望成为一种新的临床辅助诊断手段。

### 1. 影像组学特征对 ALK 基因突变的预测价值

本研究最终筛选 4 个最优影像组学特征用于模型构建,分别是 lbp-3D-m1\_firstorder\_Median、wavelet-LHH\_firstorder\_Median、logarithm\_glcm\_Idmn 和 logarithm\_glcm\_InverseVariance;进一步行二元 Logistic 回归分析,结果显示前 2 个 Median 特征是 NSCLC 发生 ALK 突变的危险因素,而后面 2 个特征 (Idmn 和 InverseVariance)是保护因素。Median 是 2 个一阶特征的中位值,主要反映图像上病灶像素强度的中位值。在 CT 图像上,该值越大,表明病灶内高密度值像素的占比越大。本研究中 ALK 突变组的 Median 参数值高于野生组,作者认为肺腺癌从癌前病变发展至微浸润性和浸润性腺癌的主要病理进程是肿瘤细胞逐步突破基底膜向周围结构侵犯,在胸部 CT 上主要表现为结节内出现实性成分及其占比逐渐增加。有研究表明 ALK 突变患者胸部 CT 以实性结节或肿块为主要表现,磨玻璃成分较少<sup>[11-13]</sup>。Idmn 是 GLCM 特征中的归一化逆差矩特征参数,它主要反映图像中局部灰度值分布的均匀性和相关性,该值越低说明图像纹理越粗糙。既往有研究表明 Idmn 特征参数与肿瘤的侵袭性和复发密切相关<sup>[9,14]</sup>。本研究结果显示 ALK 突变型组的瘤体 Idmn 值低于野生型组,说明 ALK 突变型肿瘤的纹理更粗糙、异质性更高。InverseVariance 是 GLCM 特征中的逆方差特征参数,主要反映 CT 图像上病灶内灰度值局部变化的大小,该值越小,代表图像区域的灰度值分布越不均匀。本研究中 ALK 突变型组的 InverseVariance 值低于野生型组,提示与 ALK 野生型相比,突变型 NSCLC 的空间异质性更高。陈东等<sup>[15]</sup>通过对恶性软组织肉瘤的研究发现术后复发组的 InverseVariance 值小于非复发

组。吴佳琪等<sup>[16]</sup>基于胸部 CT 平扫对肺癌隐匿性淋巴结转移进行研究,也发现淋巴结转移组的 InverseVariance 值低于无转移组。既往研究表明 ALK 突变与淋巴结高转移率和肿瘤高复发率密切相关<sup>[5]</sup>。ROC 曲线分析显示 4 个影像组学特征参数对 ALK 基因突变的诊断均具有一定的诊断效能,其中以特征参数 InverseVariance 的诊断效能最高 (AUC=0.756),但各特征参数间 AUC 的差异均无统计学意义。

### 2. 影像组学模型对 ALK 基因突变的预测价值

Liu 等<sup>[17]</sup>对 2103 例 NSCLC 患者的研究发现,ALK 基因突变更多见于不吸烟患者,本研究结果也显示 ALK 突变组不吸烟率为 95.7%,高于 ALK 野生组 (74.5%),两组间差异有统计学意义。而本研究中进行多因素逻辑回归分析发现,吸烟史不是 ALK 突变的危险因素。最后,由 4 个最优组学特征构建的影像组学模型预测 NSCLC 患者 ALK 基因突变的 AUC 达 0.868,高于单一影像组学特征,且差异有统计学意义,表明联合影像组学特征可进一步提高对 NSCLC 发生 ALK 突变的诊断效能,这与贺晓天等<sup>[18]</sup>的研究结果一致。Rizzo 等<sup>[19]</sup>研究表明基于 CT 征象诊断 NSCLC 发生 ALK 突变的 AUC 为 0.649,低于本研究中基于影像组学特征及模型的诊断效能。作者分析认为常规 CT 征象只能提供一些定性或半定量参数,无法全面反映肿瘤内部结构的复杂性和异质性,而影像组学通过高通量提取的定量参数,可以从形态、密度和纹理等方面全面地反映肿瘤内部的复杂性和异质性,从而能更准确地反映肿瘤的生物行为<sup>[9,20-22]</sup>。

### 3. 本研究的局限性

虽然本研究结果表明影像组学特征与 NSCLC 发生 ALK 突变具有相关性,对 ALK 突变具有较好的诊断效能,但由于本研究是单中心研究,未设立验证集对结果进行验证,导致结果的可重复性尚待进一步证实。

此外,本研究纳入的样本中 ALK 突变率为 13.37%, 高于既往文献报道<sup>[3-4]</sup>。因此,后期需进一步扩大样本量并进行多中心验证,从而提高模型的鲁棒性。

综上所述,本研究中基于深度学习自动分割技术构建了一个胸部 CT 影像组学模型,为无创性预测 NSCLC ALK 基因突变提供了一种有效工具,其优异的预测效能对于指导临床制订个体化精准治疗策略具有重要的参考价值。

#### 参考文献:

- [1] 周昌明,王泽洲,郑莹.2023 年美国癌症数据解读及对中国癌症防治的启示[J].中国癌症杂志,2023,33(2):117-125.
- [2] 中国医师协会肿瘤医师分会,中国医疗保健国际交流促进会肿瘤内科学分会.间变性淋巴瘤激酶酪氨酸激酶抑制剂治疗晚期非小细胞肺癌中国专家建议(2024 版)[J].中华医学杂志,2024,104(7):473-485.
- [3] 孙伟臻,周宇桁,林耀彬,等.ALK 融合阳性非小细胞肺癌新辅助靶向治疗的进展[J].中国肺癌杂志,2024,27(11):849-854.
- [4] 吴丹,李静,姚梅宏,等.非小细胞肺癌表皮生长因子受体、间变性淋巴瘤激酶、ROS1 基因突变及突变共存的临床病理学意义[J].中华病理学杂志,2021,50(3):251-253.
- [5] 宋晓宇,郝俊梅.肺腺癌基因突变的研究进展[J].临床与病理杂志,2024,44(3):434-441.
- [6] Tan X, Li Y, Wang S, et al. Predicting EGFR mutation, ALK rearrangement, and uncommon EGFR mutation in NSCLC patients by driverless artificial intelligence: a cohort study[J]. Respir Res, 2022, 23(1):132-144.
- [7] Du X, Shao Y, Qin HF, et al. ALK-rearrangement in non-small-cell lung cancer (NSCLC)[J]. Thorac Cancer, 2018, 9(4):423-430.
- [8] 田琪,冯峰,陈巧玲,等.CT 影像组学列线图评估非小细胞肺癌程序性死亡受体 1 表达[J].中国医学影像技术,2023,39(4):543-548.
- [9] 谢玉海,韩剑剑,马培旗,等.基于多中心数字化乳腺 X 线摄影影像组学预测乳腺癌人表皮生长因子受体-2 过表达[J].中国医学影像技术,2023,39(3):365-369.
- [10] 胡奎,张炯,白彪胜,等.CT 影像组学模型预测靶向药物治疗表皮生长因子受体基因突变非小细胞肺癌患者预后[J].中国医学影像技术,2022,38(10):1491-1495.
- [11] Choi CM, Kim MY, Hwang HJ, et al. Advanced adenocarcinoma of the lung: comparison of CT characteristics of patients with anaplastic lymphoma kinase gene rearrangement and those with epidermal growth factor receptor mutation[J]. Radiology, 2015, 275(1):272-279.
- [12] Park J, Kobayashi Y, Urayama KY, et al. Imaging characteristics of driver mutations in EGFR, KRAS, and ALK among treatment-naïve patients with advanced lung adenocarcinoma[J/OL]. PLoS One, 2016, 11(8): e0161081 [12 August 2016]. DOI: 10.1371/journal.pone.0161081.
- [13] Zhou JY, Zheng J, Yu ZF, et al. Comparative analysis of clinicoradiologic characteristics of lung adenocarcinomas with ALK rearrangements or EGFR mutations[J]. Eur Radiol, 2015, 25(5): 1257-1266.
- [14] Granata V, Fusco R, Setola SV, et al. CT-Based radiomics analysis to predict histopathological outcomes following liver resection in colorectal liver metastases[J/OL]. Cancers (Basel), 2022, 14(7): e1648 [24 March 2022]. DOI: 10.3390/cancers14071648.
- [15] 陈东,史彬,贾好东,等.基于多序列 3.0T MRI 的纹理分析预测恶性软组织肉瘤术后复发的价值[J].临床放射学杂志,2023,42(7):1182-1188.
- [16] 吴佳琪,倪炯,王培军.基于胸部 CT 平扫的影像组学模型预测肺癌隐匿性淋巴结转移的临床价值[J].中国医学计算机成像杂志,2023,29(3):266-270.
- [17] Liu Y, Ye X, Yu Y, et al. Prognostic significance of anaplastic lymphoma kinase rearrangement in patients with completely resected lung adenocarcinoma[J]. J Thorac Dis, 2019, 11(10): 4258-4270.
- [18] 贺晓天,张文娟,曹阳阳,等.CT 纹理分析评估非小细胞肺癌 ALK 基因状态的价值[J].中国医学计算机成像杂志,2024,30(3):312-316.
- [19] Rizzo S, Petrella F, Buscarino V, et al. CT radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer[J]. Eur Radiol, 2016, 26(1):32-42.
- [20] 王传彬,李翠平,曹锋,等.基于 CT 临床放射组学列线图与深度学习鉴别非典型肺错构瘤和肺腺癌[J].安徽医科大学学报,2024,59(2):344-350.
- [21] 高续,庞奇,张宏杰,等.基于深度学习自动分割技术的胸部 CT 影像组学模型预测非小细胞肺癌 EGFR 基因突变[J].放射学实践,2025,40(5):573-578.
- [22] 王小雷,哈传传,谢玉海,等.基于钼靶影像组学列线图预测乳腺癌脉管浸润状态[J].放射学实践,2025,40(5):579-585.

(收稿日期:2025-10-30 修回日期:2026-01-11)