

基于增强 CT 的影像组学和深度学习模型预测喉癌患者术后生存期

吕芃志, 李心语, 岳华杰, 张华, 乔英, 张倩, 梁丽, 王灵杰

【摘要】 目的:探讨集成增强 CT 影像组学特征、深度学习算法及临床指标的预测模型对喉癌患者术后生存期的预测价值。方法:回顾性搜集 2016 年 1 月—2021 年 7 月两个医疗中心的 272 例喉癌患者的术前对比增强 CT(CECT)静脉期图像、临床资料及随访数据。以中心 1 的患者作为训练集($n=156$),中心 2 的患者为外部测试集($n=116$)。通过病历和电话随访,术后 2 年内每 3~6 个月一次,之后每年一次。生存期的定义为确诊至死亡或末次随访(2021 年 7 月)时间,主要观察全因死亡。使用 PyRadiomics 软件提取肿瘤全容积 ROI 的影像组学特征;并采用预训练 ResNet50 网络提取肿瘤最大截面 ROI 的深度学习(DL)特征。分别对两类特征进行归一化处理后,采用最小绝对收缩和选择算子回归分析(LASSO-Cox)筛选与患者 3 年生存期显著相关的特征,并采用逻辑回归(LR)、支持向量机(SVM)、随机森林(RF)、极限梯度提升(XGBoost)和多层感知机(MLP)五种机器学习算法(ML)构建影像组学标签(Rad-signature)和深度学习标签(DL-label),进一步结合 Cox 回归分析确定的独立临床预后因素,采用 Cox 比例风险回归分析构建联合预测模型并绘制其列线图。采用一致性指数(C-index)和 Kaplan-Meier 生存曲线分析评估模型的预测效能。结果:多因素 Cox 回归分析显示肿瘤大小($HR=1.03, P<0.05$)和临床分期($HR=2.70, P<0.05$)为喉癌术后生存期的独立预后因素。联合模型在预测喉癌术后生存期方面效能较高,C-index 值为 0.69,在外部测试集中预测 1 年、2 年、3 年生存期的 ROC 曲线下面积(AUC)分别为 0.61、0.68 和 0.72。Kaplan-Meier 生存曲线分析显示,基于约登指数对应的最大截断值能够将患者显著区分为死亡高风险组和低风险组(Log-rank 检验, $P<0.05$)。结论:基于增强 CT 影像组学特征、深度学习特征及独立临床预后因素构建的列线图模型,在喉癌术后生存期的预测中具有较高的效能,为个体化预后评估提供可靠工具。

【关键词】 喉癌;影像组学;深度学习;机器学习;体层摄影术,X 线计算机;生存期

【中图分类号】 R814.42;R739.65 **【文献标志码】** A **【文章编号】** 1000-0313(2026)04-0382-08

DOI:10.13609/j.cnki.1000-0313.2026.04.005

开放科学(资源服务)标识码(OSID):



Predicting postoperative survival in patients with laryngeal cancer using radiomics and deep learning models based on contrast-enhanced CT LV Peng-zhi, LI Xin-yu, YUE Hua-jie, et al. Department of Medical Imaging, Shanxi Medical University, Taiyuan 030001, China

【Abstract】 Objective: To investigate the predictive value of a predictive model integrating radiomics features extracted from contrast-enhanced CT (CECT) images, deep learning algorithms, and clinical indicators for postoperative survival period (SP) in patients with laryngeal cancer. **Methods:** We retrospectively collected preoperative contrast-enhanced computed tomography (CECT) venous-phase images, clinical data, and follow-up information from 272 patients with laryngeal cancer at two medical centers between January 2016 and July 2021. Patients from center A were assigned to the training cohort ($n=156$), and patients from center B served as the external test cohort ($n=116$). Follow-up was conducted through medical records and telephone interviews every 3 to 6 months during the first two years after surgery, and annually thereafter. Survival period was defined as the time from diagnosis to death or the last follow-up (July 2021), with all-cause mortality as the primary observation. Radiomics features from whole-tumor volume ROI were extracted using PyRadiomics software. Deep learning (DL) features were extracted from the ROI of the tumor's maximum cross-sectional area using a pre-trained ResNet50 network. After normalization of the two feature sets, least absolute shrinkage and selection operator (LASSO) Cox regression analysis was used to select features significantly associated with patient 3-year SP. Radiomics signature (Rad-signature) and deep learning signature (DL-label) were constructed using five machine learning algorithms, including logistic regression (LR), support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost) and multilayer perceptron (MLP). Furthermore, they were combined with independent clinical prognostic factors identified by Cox regression analysis to construct a combined prediction model using Cox proportional hazards regression analysis, and a corresponding nomogram was drawn. The predictive performance of the models was evaluated using the concordance index (C-index) and Kaplan-Meier survival curve analysis. **Results:** Multivariate Cox regression analysis revealed that tumor size [hazard ratio (HR) = 1.03, $P<0.05$] and clinical stage ($HR=2.70, P<0.05$) were independent prognostic factors for post-

作者单位:030001 山西太原,山西医科大学第一医院 CT 放射影像科

作者简介:吕芃志(1990-),男,山西临汾人,本科,主管技师,主要从事 CT 技术工作。

通讯作者:王灵杰, E-mail:4617784@qq.com

基金项目:山西省基础研究计划项目(20210302123253,202503021211277);山西省卫生健康委科研项目计划(2024056)

operative survival in patients with laryngeal cancer. The combined model demonstrated favorable predictive performance for postoperative SP, yielding a C-index of 0.69. In the external test cohort, the area under the receiver operating characteristic curve (AUC) for predicting 1-, 2-, and 3-year SP were 0.61, 0.68, and 0.72, respectively. Kaplan-Meier survival curve analysis showed that the optimal cutoff value determined by the Youden index significantly stratified patients into high-risk and low-risk groups for death (Log-rank test, $P < 0.05$). **Conclusion:** The nomogram model based on CECT radiomics features, deep learning features, and independent clinical prognostic factors demonstrates high efficacy in predicting postoperative survival period in patients with laryngeal cancer, providing a reliable tool for individualized prognostic assessment.

【Key words】 Laryngeal cancer; Radiomics; Deep learning; Machine learning; Tomography, X-ray computed; Survival period

喉癌作为头颈部高发恶性肿瘤,全球年新发病例约 20 万例,死亡约 1 万例,其临床管理面临严峻挑战^[1]。早期喉癌病变可通过喉部分切除术、经口激光显微手术或根治性放疗等治疗方法实现功能保全,而晚期多需全喉切除联合辅助治疗,严重影响患者的生活质量^[2]。尽管诊疗技术不断进步,喉癌的预后仍不理想,亟需精准预测工具优化治疗决策^[3]。

影像组学作为新兴的医学图像分析技术,通过从 CT、PET-CT 或 MRI 数据中高通量提取定量影像特征,已展现出超越传统形态学评估的肿瘤表型解析能力^[4]。近年来,深度学习(deep learning, DL)方法在医学领域发展迅速。深度学习通过卷积神经网络(convolutional neural network, CNN)架构能自动捕获肿瘤深层语义特征,这些人类视觉无法辨别的信息已被证明与生存结局显著相关。Liao 等^[5]的研究结果显示 DL 模型在喉癌患者生存预测方面具有优越的性能,预测喉癌生存期的 C 指数值达 0.71。然而,目前对喉癌术后生存预测尚缺乏多维度整合研究。鉴于此,本研究创新性基于术前增强 CT,整合影像组学、深度学习特征及临床独立风险因素,构建生存预后风险模型并绘制其列线图,旨在为喉癌患者个体化治疗方案的制订及预后评估提供更科学的决策支持。

材料与方法

本研究为多中心回顾性研究,已获得所有参与机构的伦理审查委员会的批准,并遵循《赫尔辛基宣言》的有关规定。

1. 研究对象及临床资料

本研究为多中心回顾性队列研究,连续纳入 2016 年 1 月—2021 年 7 月在山西省肿瘤医院(中心 1)及山西医科大学第一医院(中心 2)接受手术治疗的喉癌患者。纳入标准:①术前两周内完成颈部增强 CT 检查且影像质量满足分析要求(无吞咽运动伪影等干扰);②接受根治性手术治疗(包括经口激光显微手术、喉部分切除术或全喉切除术);③经术后组织病理学检查确诊为喉鳞状细胞癌。排除标准:①术前接受过新辅助化疗或放疗;②CT 图像上存在金属伪影或呼吸运动

伪影,从而影响病灶的分割;③临床基线数据(年龄、性别、TNM 分期等)或随访记录不完整。最终纳入中心 1 的 156 例患者构成训练集,中心 2 的 116 例患者作为外部验证集。

临床数据通过医院电子病历系统提取,包括人口学特征(年龄、性别)、病理数据(T 分期、N 分期、临床分期)及治疗相关信息。对每例患者术后常规进行随访,通过门诊、住院病历系统及定期电话等方式获取患者的生存信息,随访频率为术后前 2 年内每 3~6 个月一次,之后转为每年一次。生存期的定义为确诊日期与死亡日期或末次随访日期之间的时间间隔。

2. 图像获取及预处理

所有患者行颈部 CT 增强扫描。中心 1 部分病例(39 例)使用 GE Discovery 750HD CT 扫描仪,扫描参数:100~120 kV,100~250 mA,矩阵 512×512,扫描层厚 0.625~2.500 mm;中心 1 部分病例(117 例)使用 Siemens Somatom Definition AS CT 扫描仪,扫描参数:100~120 kV,50~400 mA,矩阵 512×512,扫描层厚 0.5~3.0 mm。中心 2 病例(116 例)使用 Somatom Force CT 扫描仪,扫描参数同中心 1 的 Siemens CT 扫描仪。使用非离子型对比剂(300~370 mg I/mL),剂量 1.5 mg/kg。

所有分析均基于对比增强 CT(CECT)静脉期图像。这一选择基于以下 3 方面原因:首先,静脉期是喉癌术前评估的推荐增强期相,能提供肿瘤与周围软组织的最佳对比度,利于病灶边界的识别与分割;其次,本研究为多中心回顾性研究,各中心的扫描协议存在一定差异,为保证样本的一致性和可比性,静脉期是所有参与中心共有且质量最稳定的序列。

为消除不同 CT 机型间扫描参数的差异,对所有 CT 图像进行预处理。采用 3 次 B 样条插值算法(B-spline)将所有图像重采样至标准化体素尺寸($1 \times 1 \times 1 \text{ mm}^3$),并对体素强度值进行离散化处理。离散化参数设置:组距(Bin width)32 个单位,窗宽 350 HU,窗位 50 HU。

由 2 位具有 5 年头颈部影像诊断经验的医师,独立使用 ITK-SNAP 3.8.0 软件,采用半自动分割法在

肿瘤静脉期 CT 图像上逐层沿肿瘤边缘勾画 ROI, 并生成肿瘤全容积 ROI(图 1)。为评估肿瘤分割的可靠性, 随机抽取 30 例样本, 由上述两位医师分别独立对肿瘤区域进行勾画, 通过计算组内相关系数(intra-class correlation coefficient, ICC)评估观察者间的一致性, 以 $ICC \geq 0.85$ 作为一致性良好的判定标准。结果显示, 基于分割结果提取的影像组学特征中, 91% 的特征 $ICC > 0.85$, 提示分割结果具有良好的 consistency 与可靠性。

3. 影像组学特征提取

使用 PyRadiomics v3.1.0 软件标准化提取肿瘤的影像组学特征, 涵盖三大类核心特征: 形状特征 14 项; 一阶特征 360 项; 纹理特征 1460 项。纹理特征主要包括灰度共生矩阵(gray-level co-occurrence matrix, GLCM)特征 440 个、游程长度矩阵(gray level run length matrix, GLRLM)特征 320 个、区域大小矩阵(gray-level size zone matrix, GLSZM)特征 320 个、依赖矩阵(gray level dependence matrix, GLDM)特征 280 个及邻域灰度差分矩阵(neighborhood gray-tone difference matrix, NGTDM)100 个。

4. 深度学习特征提取

本研究中采用标准化静脉期增强 CT 图像进行分析。沿肿瘤长轴裁剪包含最大截面的 2D 矩形 ROI, 并统一调整为 224×224 像素大小。本研究中使用预训练后的 ResNet50 系统(基于卷积神经网络架构)进行深度学习(deep learning, DL)特征的提取。首先基于 ImageNet 等大型数据集完成网络预训练, 随后通过迁移学习技术对网络进行微调。具体调整内容: 保留预训练模型的卷积基作为特征提取器, 仅对网络顶部的全连接分类层进行参数重构与随机初始化, 使其输出维度与本研究的二分类任务相匹配。在微调过程中, 所有卷积层和全连接层的参数均参与反向传播和更新, 使底层通用特征与高层语义特征均能适配医学

影像领域。微调完成后, 我们从网络的全局平均池化层提取了 2048 个维度 DL 特征向量, 作为每个样本的原始特征。随后, 基于这些特征, 采用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)回归模型进行特征筛选和融合, 最终构建出具有预测价值的深度学习标签(DL-label)。这一方法充分利用了预训练模型的特征提取能力, 同时通过特定数据集的迁移学习优化了模型对喉癌影像特征的识别性能。其中, 为提升模型的泛化能力, 集成数据增强策略: 随机旋转($\pm 15^\circ$)、平移($\pm 10\%$)及缩放(0.9~1.1 倍)。迁移学习阶段通过网格搜索优化超参数(初始学习率 0.01, 批量大小 16, 迭代 13000 步/500 周期)。

5. 特征筛选

所有提取的影像组学特征及 DL 特征分别经 z -score 标准化处理后, 执行以下 3 个步骤的特征筛选。①观察者一致性验证: 即基于 ICC 筛选可重复性高的稳定特征($ICC > 0.85$); ②消除冗余特征: 计算任意 2 个特征间的 Pearson 相关系数, 对 $|r| > 0.90$ 的强相关特征组仅保留单一代表性特征; ③判别特征优化: 采用最小绝对收缩选择算子, 经 5 折交叉验证优化正则化参数 λ , 筛选非零系数特征。为平衡训练集样本分布, 同步应用合成少数类过采样技术(synthetic minority oversampling technique)按照 1:1 的比例进行调整。

6. 模型构建

为构建最优的分类模型, 首先将筛选出的影像组学特征与 DL 特征作为输入数据, 分别导入 5 种机器学习分类器中进行训练, 包括逻辑回归(logistic regression, LR)、支持向量机(support vector machines, SVM)、随机森林(random forest, RF)、极限梯度提升(extreme gradient boosting, XGBoost)和多层感知机(multilayer perceptron, MLP), 从而得到 5 个候选诊断模型。随后, 为综合评估各模型的泛化能力和区分

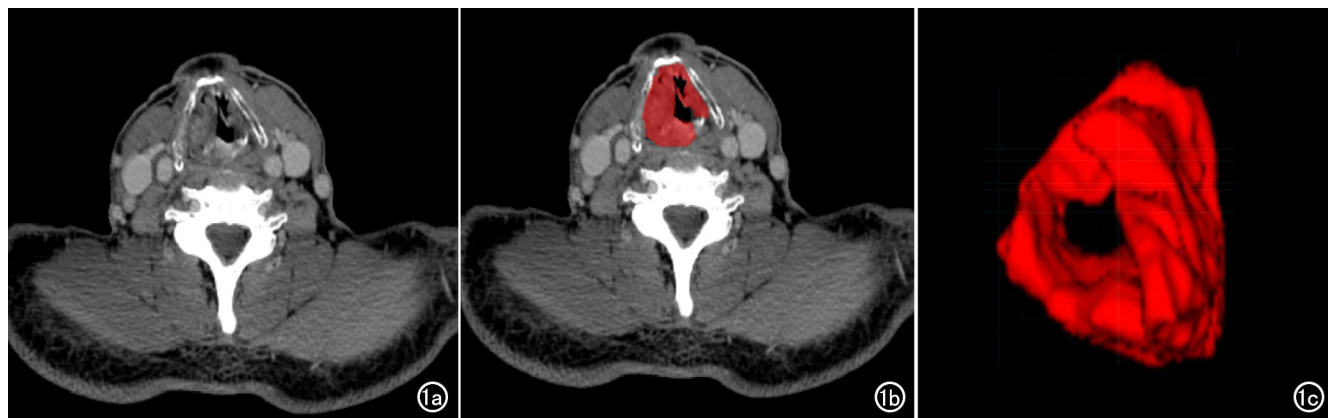


图 1 喉癌 ROI 勾画示意图。a) 静脉期 CT 图像, 显示肿瘤位于喉部, 右侧累及更显著; b) 使用 ITK-SNAP 软件, 在喉部静脉期 CT 图像上逐层沿肿瘤边缘勾画 ROI(红色区域); c) 软件自动生成肿瘤全容积 ROI。

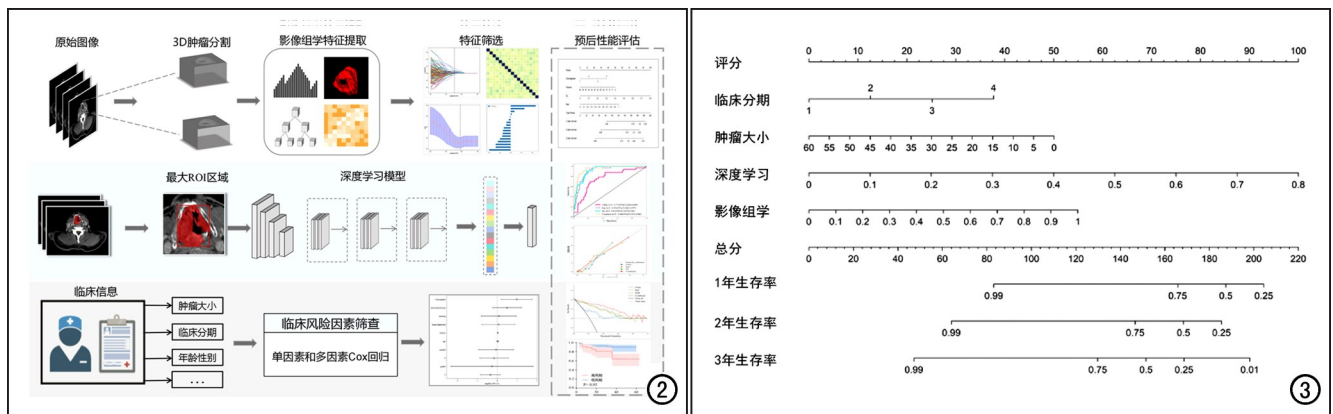


图2 本研究的工作流程图。图3 基于临床独立危险因素、影像组学标签和深度学习标签构建的联合模型的列线图。

效能,我们基于验证集绘制了各模型的ROC曲线,并计算曲线下面积(AUC)。通过比较AUC的大小,筛选出性能最优的模型作为本研究的最终分类器。

多维度特征融合与生存预测模型的构建主要采用以下几个步骤。①临床危险因素筛选:通过单因素Cox回归分析初筛影响临床生存期的独立因子($P < 0.05$),进一步使用多因素Cox回归分析确定独立危险因素;②联合预测模型构建:整合临床独立危险因素、影像组学标签及深度学习标签,构建多变量Cox比例风险回归模型;③模型的可视化及验证:采用诺莫图(nomogram)可视化1、2、3年生存率预测,并使用Kaplan-Meier曲线评估高、低风险组的生存期差异(Log-rank检验)。

本研究的影像组学分析工作流程示意图见图3。

7. 统计分析

使用SPSS 25.0和Python 3.7软件进行统计分析。连续变量采用均值±标准差的方式进行描述,正态性假设通过Shapiro-Wilk检验进行检验;符合正态分布时采用独立样本 t 检验进行组间比较,非正态分布时采用Mann-Whitney U 检验;分类变量以频率和百分比的形式表示,并使用卡方检验或Fisher精确检验(当理论频数 < 5 时)进行组间比较。采用DeLong检验比较不同模型间ROC曲线下面积(AUC)的差异。以 $P < 0.05$ (双尾)为差异有统计学意义。

结果

1. 临床资料

训练集中纳入156例患者(男149例,女7例),外部测试集中纳入116例患者(男108例,女8例)。2个数据集中存活组与死亡组基线临床特征比较结果见表1。组间比较结果显示,在训练集及外部测试集中,死亡组与存活组之间年龄、性别、吸烟史、饮酒史及原发部位的差异均无统计学意义($P > 0.05$),而肿瘤大小与手术方式的差异均有统计学意义($P < 0.05$)。

2. 影像组学和DL标签的构建

肿瘤的影像组学分析共提取了1834个影像组学特征,其中形状特征14个、一阶特征360个、纹理特征1460个。经特征筛选后,最终保留14个最佳影像组学特征,包括一阶特征4个及纹理特征10个。基于最佳影像组学特征构建的5种ML模型中,LR模型的效能最优:在训练集中的AUC为0.91(95%CI:0.86~0.96),在外部测试集中的AUC为0.71(95%CI:0.57~0.85)。

基于ResNet50模型提取了2048个DL特征。经特征筛选后,最终保留17个最佳DL特征。随后构建的5种ML模型中,MLP模型的效能最优:在训练集中的AUC达0.92(95%CI:0.88~0.96),在外部测试集中的AUC为0.69(95%CI:0.55~0.83)。

将最佳影像组学模型和DL模型输出的概率值分别定义为Rad-signature和DL-label,用于构建联合模型。

3. 联合预测模型的构建

在训练集中,采用单因素及多因素Cox回归对临床资料进行筛选,分析结果见表2。结果显示:肿瘤大小和临床分期是喉癌术后生存预后的独立危险因素($P < 0.05$)。

通过整合Rad-signature、DL-label及临床独立预后因素,构建喉癌术后生存联合模型并绘制其列线图(图3)。

各模型的预测效能详见表3和图4~6。对于单一模型,影像组学模型在训练集中预测术后患者总体生存情况的C-index=0.87,预测术后1、2、3年生存患者的AUC分别为0.90、0.88、0.91;在外部测试集中的C-index=0.68,1、2、3年的AUC分别为0.63、0.69、0.71。深度学习模型在训练集中的C-index=0.87,1、2、3年的AUC分别为0.88、0.87、0.92;在测试集中的C-index=0.66,1、2、3年的AUC分别为0.54、0.65、0.69。联合模型在训练集中预测术后患者总体生存情

表 1 2 个数据集中存活组与死亡组基线临床特征的比较

特征	训练集 (n=156)				外部测试集 (n=116)			
	存活组 (n=123)	死亡组 (n=33)	统计量	P 值	存活组 (n=92)	死亡组 (n=24)	统计量	P 值
年龄/岁	61.47±7.97	61.88±7.34	-0.265 ^a	0.79	62.29±8.35	64.29±9.65	-1.010 ^a	0.31
肿瘤大小/cm	4.75(1.68~12.02)	10.55(5.87~26.88)	-4.493 ^b	0.01	1.81(0.75~4.55)	10.0(2.31~24.80)	-3.428 ^b	0.00
性别/例			0.000 ^c	1.00			0.020 ^c	0.89
男	117(95.12%)	32(96.97%)			85(92.39%)	23(95.83%)		
女	6(4.88%)	1(3.03%)			7(7.61%)	1(4.17%)		
吸烟史/例	81(65.85%)	21(63.64%)	0.001 ^c	0.97	66(71.74%)	18(75.00%)	0.004 ^c	0.95
饮酒/例	33(26.83%)	10(30.30%)	0.031 ^c	0.86	39(42.39%)	10(41.67%)	0.000 ^c	1.00
放疗/例	42(34.15%)	17(51.52%)	2.640 ^c	0.10	21(22.83%)	11(45.83%)	3.958 ^c	0.05
T 分期/例			10.449 ^c	0.02			1.309 ^c	0.73
T1	41(33.33%)	11(33.33%)			41(44.57%)	9(37.50%)		
T2	42(34.15%)	6(18.18%)			20(21.74%)	6(25.00%)		
T3	26(21.14%)	5(15.15%)			23(25.00%)	8(33.33%)		
T4	14(11.38%)	11(33.33%)			8(8.70%)	1(4.17%)		
N 分期/例			14.819 ^c	0.00			0.413 ^c	0.81
N0	92(74.80%)	13(39.39%)			71(77.17%)	18(75.00%)		
N1	11(8.94%)	7(21.21%)			6(6.52%)	1(4.17%)		
N2	20(16.26%)	13(39.39%)			15(16.30%)	5(20.83%)		
临床分期/例			23.020 ^c	0.00			0.789 ^c	0.85
I	31(25.20%)	2(6.06%)			39(42.39%)	8(33.33%)		
II	33(26.83%)	1(3.03%)			11(11.96%)	4(16.67%)		
III	28(22.76%)	9(27.27%)			21(22.83%)	6(25.00%)		
IV	31(25.20%)	21(63.64%)			21(22.83%)	6(25.00%)		
手术方式/例			6.083 ^c	0.01			5.329 ^c	0.02
部分切	80(65.04%)	13(39.39%)			79(85.87%)	15(62.50%)		
全切	43(34.96%)	20(60.61%)			13(14.13%)	9(37.50%)		
原发部位/例			0.296 ^c	0.86			0.300 ^c	0.86
声门下	3(2.44%)	1(3.03%)			2(2.17%)	1(4.17%)		
声门型	69(56.10%)	20(60.61%)			39(42.39%)	10(41.67%)		
声门上	51(41.46%)	12(36.36%)			51(55.43%)	13(54.17%)		

注:^a 为 *t* 值;^b 为 *Z* 值;^c 为 χ^2 值。

表 2 临床特征预测喉癌患者术后生存情况的 Cox 回归分析

临床因素	单因素 Cox 分析		多因素 Cox 分析	
	HR(95%CI)	P 值	HR(95%CI)	P 值
年龄	1.00(0.96~1.05)	0.86	1.02(0.97~1.08)	0.35
性别	0.60(0.08~4.36)	0.61	0.77(0.10~6.09)	0.80
吸烟史	0.90(0.44~1.84)	0.78	1.08(0.45~2.63)	0.86
饮酒史	1.12(0.53~2.35)	0.77	0.94(0.39~2.28)	0.90
原发部位	0.86(0.45~1.63)	0.65	0.94(0.49~1.82)	0.85
肿瘤大小	1.05(1.03~1.07)	<0.05	1.03(1.01~1.06)	0.01
T 分期	1.35(0.98~1.86)	0.06	0.71(0.43~1.16)	0.17
N 分期	1.93(1.34~2.79)	<0.05	0.85(0.41~1.74)	0.65
临床分期	2.40(1.57~3.65)	<0.05	2.70(1.23~5.93)	0.01
手术方式	2.53(1.26~5.10)	0.01	1.06(0.46~2.42)	0.89
放疗	1.88(0.95~3.72)	0.07	1.64(0.80~3.38)	0.18

注:HR 为风险比(hazard ratio)。

表 3 不同模型对喉癌患者术后生存情况的预测效能

模型	1 年 AUC(95%CI)	2 年 AUC(95%CI)	3 年 AUC(95%CI)	C-index
训练集				
组学模型	0.90(0.83~0.97)	0.88(0.82~0.94)	0.91(0.86~0.96)	0.87
DL 模型	0.88(0.80~0.96)	0.87(0.81~0.93)	0.92(0.88~0.96)	0.87
临床模型	0.78(0.67~0.89)	0.75(0.64~0.87)	0.79(0.70~0.87)	0.76
联合模型	0.92(0.86~0.98)	0.90(0.84~0.95)	0.95(0.91~0.98)	0.90
外部测试集				
组学模型	0.64(0.40~0.87)	0.69(0.51~0.86)	0.71(0.57~0.85)	0.68
DL 模型	0.56(0.33~0.80)	0.65(0.46~0.83)	0.69(0.55~0.83)	0.66
临床模型	0.57(0.37~0.77)	0.60(0.42~0.78)	0.62(0.50~0.74)	0.60
联合模型	0.63(0.42~0.83)	0.68(0.52~0.84)	0.72(0.61~0.83)	0.69

况的 C-index=0.90, 预测术后 1、2、3 年生存患者的 AUC 分别为 0.92、0.90 和 0.95; 在外部测试集中, 相应结果分别为 C-index=0.69、术后 1、2、3 年 AUC 分别为 0.61、0.68、0.72。联合模型的总体效能优于单一模型。

不同模型间 AUC 比较结果显示, 联合模型在训练集及外部测试集中的 AUC 均最高(预测术后 1~3 年生存率的 AUC 在训练集中为 0.90~0.95, 在外部测试集中为 0.63~0.72), 整体预测效能优于单一影像组学模型、深度学习模型及临床模型, 表现出较好的判别

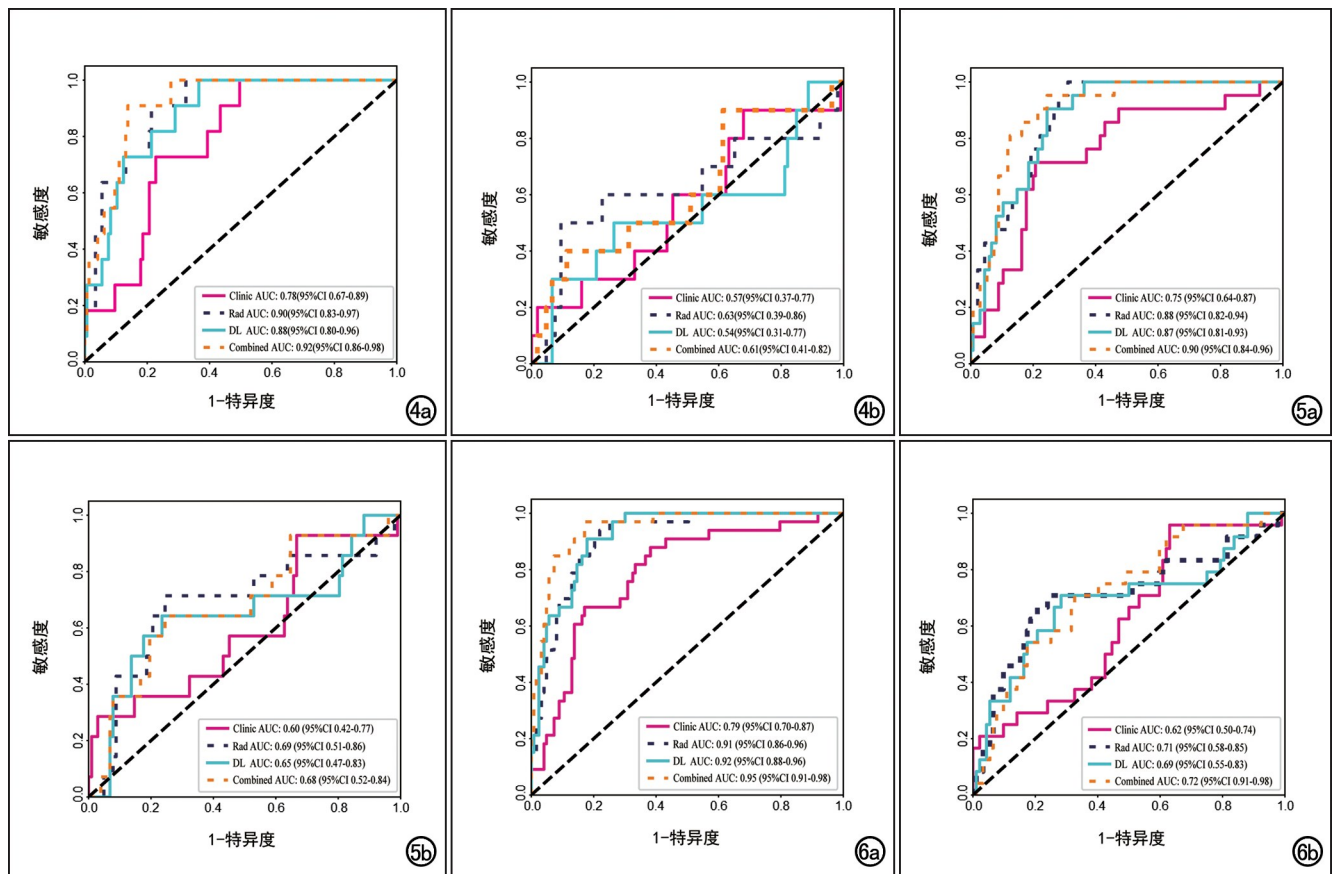


图 4 临床模型、影像组学模型、深度学习模型及联合模型预测喉癌术后 1 年生存期的 ROC 曲线。a) 训练集; b) 外部测试集。
图 5 四种模型预测喉癌术后 2 年生存期的 ROC 曲线。a) 训练集; b) 外部测试集。图 6 四种模型预测喉癌术后 3 年生存期的 ROC 曲线。a) 训练集; b) 外部测试集。

能力。KM 曲线是衡量特定时间点生存概率的常用方法^[6], 本研究中联合模型的 KM 曲线显示(图 7~9), 基于影像组学模型(Rad-signature)、深度学习模型(DL-label)及联合模型的约登指数所确定的最优截断值, 均可将患者分为高死亡风险组与低死亡风险组。三种模型中, 高风险组患者的生存概率均低于低风险组, 且组间差异均具有统计学意义(Log-rank 检验, $P < 0.05$)。

讨论

本研究中基于喉癌患者术前增强 CT 静脉期图像, 静脉期图像能显著提高肿瘤与周围组织的对比度, 清晰显示边界特征, 并稳定呈现肿瘤内部强化模式, 同时有效避免了多期相扫描可能导致的特征变异。成功开发并验证了整合影像组学特征、DL 特征及临床独立风险因素的联合预测模型能较好地应用于喉癌患者术后生存期的个体化预测。构建此模型的核心创新点在于融合了三种维度的互补信息: ①影像组学定量表征的肿瘤异质性; ②深度学习从图像中自动提取的深层语义特征; ③经过筛选得到具有独立预后价值的临床危险因素。本研究结果显示, 联合模型优于单一信

息源模型, 在外部测试集中展现出良好的预测效能, C-index 达 0.69, 预测 1、2、3 年生存率的 AUC 分别为 0.61、0.68、0.72。上述结果表明模型能够稳健地对患者进行风险分层, KM 生存曲线显示基于模型分层的两组之间的生存差异具有统计学意义($P < 0.05$)。证实了模型能为术前个体化治疗(如手术范围、辅助治疗强度)提供客观依据, 并具备临床可行性和泛化潜力。

喉癌患者的预后评估传统上高度依赖 TNM 分期系统^[7], 而原发肿瘤部位(如声门上型喉癌因淋巴结转移率高导致预后较差)也是重要的考量因素^[8]。然而, 单一或少数临床因素往往难以全面准确评估患者的预后。因此, 整合多维度信息的综合模型成为实现更精准的个体化评估的关键^[9]。因此, 本研究创新性地将影像组学特征、深度学习特征与临床危险因素融合, 构建了预测效能更优的联合模型, 且绘制其列线图进行可视化显示, 便于临床使用。

影像组学通过高通量提取定量特征, 能深入挖掘 CT 图像中肉眼难以辨别辨肿瘤表型信息, 如纹理异质性和强度分布等^[10-13]。本研究基于增强 CT 图像筛选出 14 个具有高稳定性和预测价值的影像组学特征, 覆盖形状、强度及纹理等类别。值得注意的是, 权重最

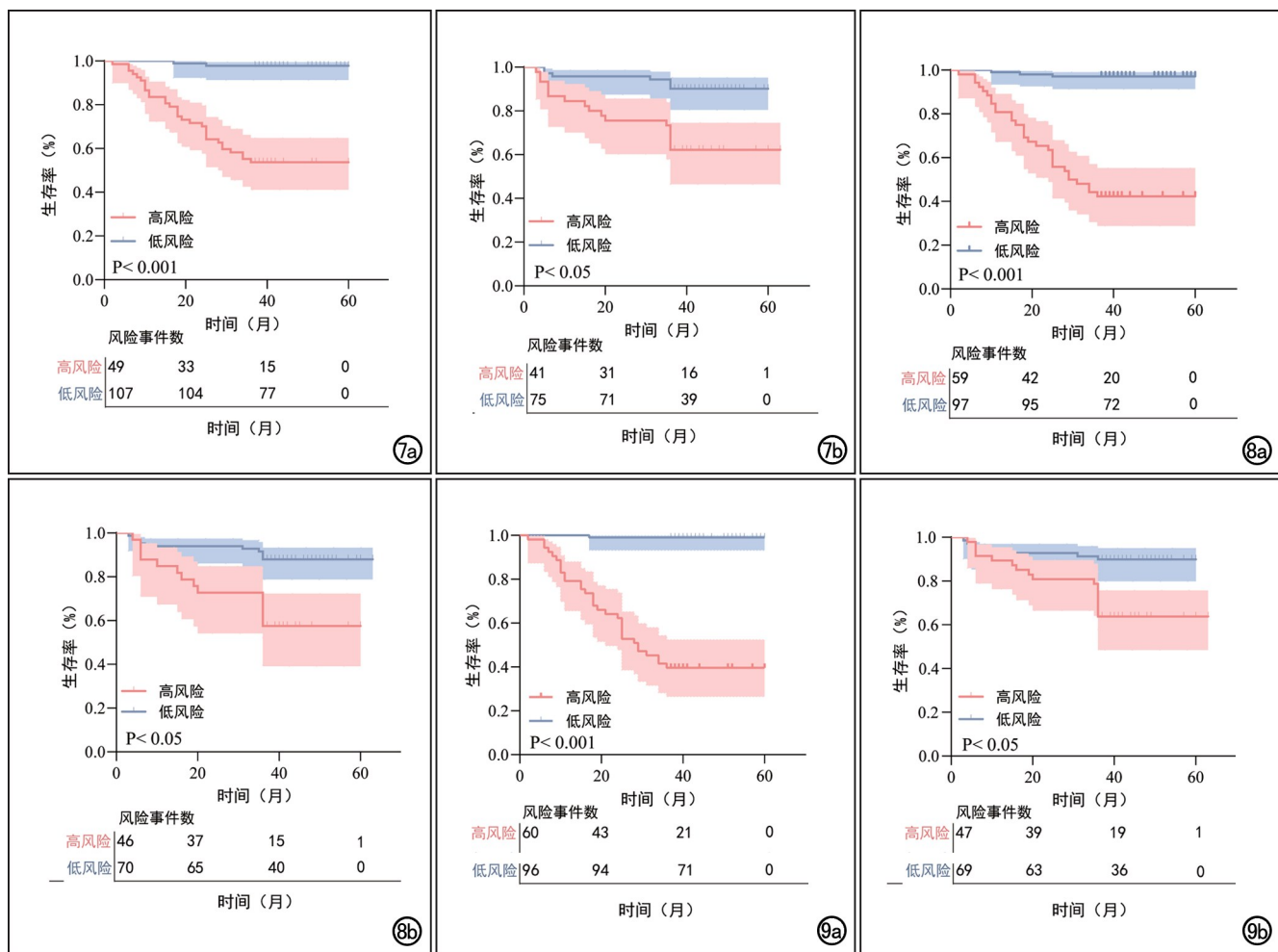


图 7 Rad-signature 的 KM 生存曲线。基于 Rad-signature 约登指数对应的最大截断值(0.037)将患者划分为高死亡风险组和低死亡风险组。a) 训练集; b) 外部测试集。图 8 DL-label 的 KM 生存曲线。基于 DL-label 约登指数对应的最大截断值(0.317)将患者划分高死亡风险组和低死亡风险组。图 9 联合模型的 KM 生存曲线。基于联合模型约登指数对应的最大截断值(0.444)将患者划分为高死亡风险组和低死亡风险组。

高的特征是主要反映肿瘤内强度峰度及纹理的特征, 这些表征异质性的指标已被多项研究证实与癌症生物学行为及预后密切相关^[14-16]。本研究基于组学特征构建的喉癌生存预后模型在外部测试集中 C-index 达 0.68, 与既往研究影像组学特征评估喉鳞状细胞癌手术切除后患者预后效能相当^[10]。

深度学习模型, 尤其是基于 CNN 架构的模型(如本研究中采用的 ResNet50), 具有强大的端到端特征学习能力, 可自动捕获涵盖从低级视觉特征至高级抽象语义等复杂肿瘤信息^[17-20]。借助 ImageNet 预训练模型开展迁移学习, 一定程度上能克服医学影像数据量相对匮乏这一局限性。本研究基于 DL 特征构建的喉癌术后生存预后模型在训练集和外部测试集中均展现出较高的预测效能(C-index 分别为 0.87 和 0.66)。DL 特征的优势在于能够更为“自然”地适配特定数据集, 并与临床结局建立关联, 进而揭示出影像信息中潜藏的、超越传统影像学特征描述的潜在信息模式^[19-21]。

本研究通过决策级融合策略构建了预后评估体系, 基于增强 CT 图像整合了影像组学特征、深度学习特征和临床风险因素三个模态的信息, 并采用机器学习分类器建立预测模型。在外部验证集中的结果显示, 联合模型的预测效能优于单一模态模型($P < 0.05$), 证实了多源数据的协同作用; 影像组学量化肿瘤异质性, 深度学习提取高维视觉特征, 临床因子提供循证预后变量。这种多维度融合策略已经被广泛应用于多种肿瘤诊断及疗效评估领域, 尤其是在胶质瘤和慢性肾病等疾病中观察到的多维度融合优势^[22-23], 进一步验证了本研究中基于增强 CT 构建的联合模型具有普适性价值。

目前, 喉癌的治疗策略正从单一方法向以根治肿瘤为核心、同时最大化保留喉功能的多学科协作模式演进^[24]。术前精准的风险分层对于指导个体化治疗决策至关重要, 如, 识别出低风险患者可能更适合喉功能保留性手术或减少辅助治疗强度, 而高风险患者则

可能需要进行更积极的综合治疗^[24-25]。本研究构建的多维度联合预测模型,为喉癌的精准确策提供了依据,在肿瘤根治与生活质量保全间寻求更优平衡。基于影像大数据和人工智能的多源信息融合策略,建立了从危险分层至预后预测的完整评估体系,提高了个体化精准诊疗水平。

本研究存在一定的局限性:①系回顾性分析且研究队列的样本量相对较小,可能存在一定的选择偏倚,尚需大样本量多中心前瞻性研究进一步验证本研究的结论;②由于多中心扫描协议存在差异,本研究中仅采用静脉期增强 CT 图像进行分析,可能存在影像信息维度不足的局限性;③构建模型时未纳入遗传学和病理数据等潜在的预后相关因素,未来可整合更多信息以提升预测能力。

综上,基于增强 CT 联合影像组学、深度学习特征和临床危险因素构建的联合预测模型,有助于提升喉癌患者个体化诊疗水平及生存期预测准确性,从而改善患者的生存获益。

参考文献:

[1] Chen X, Yu Q, Peng J, et al. A combined model integrating radiomics and deep learning based on contrast-enhanced CT for preoperative staging of laryngeal carcinoma [J]. *Acad Radiol*, 2023, 30(12):3022-3031.

[2] Shen Y, Huang Q, Zhang Y, et al. A novel signature derived from metabolism-related genes GPT and SMS to predict prognosis of laryngeal squamous cell carcinoma[J/OL]. *Cancer Cell Int*, 2022, 22(1):e226[2022 Jul 8]. DOI:10.1186/s12935-022-02647-2.

[3] Liu Q, Liu S, Mao Y, et al. Machine learning model to preoperatively predict T2/T3 staging of laryngeal and hypopharyngeal cancer based on the CT radiomic signature[J]. *Eur Radiol*, 2024, 34(8):5349-5359.

[4] Chen Q, Zhang L, Liu S, et al. Radiomics in precision medicine for gastric cancer: opportunities and challenges[J]. *Eur Radiol*, 2022, 32(9):5852-5868.

[5] Liao F, Wang W, Wang J. A deep learning-based model predicts survival for patients with laryngeal squamous cell carcinoma: a large population-based study [J]. *Eur Arch Otorhinolaryngol*, 2023, 280(2):789-795.

[6] Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate[J]. *Int J Ayurveda Res*, 2010, 1(4):274-278.

[7] 周剑,刘欢,张惠紫,等. PKM2 与头颈部鳞状细胞癌淋巴结转移及临床分期相关性的 Meta 分析[J]. *临床与实验病理学杂志*, 2019, 35(6):711-713.

[8] Zheng Y, Qiu B, Liu S, et al. A transformer-based deep learning model for early prediction of lymph node metastasis in locally advanced gastric cancer after neoadjuvant chemotherapy using pre-treatment CT images [J/OL]. *EClinical Medicine*, 2024, 75:e102805, September 2024, DOI:10.1016/j.eclinm.2024.102805

[9] Zwanenburg A, Vallières M, Abdalah MA, et al. The image bio-

marker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping[J]. *Radiology*, 2020, 295(2):328-338.

[10] Chen L, Wang H, Zeng H, et al. Evaluation of CT-based radiomics signature and nomogram as prognostic markers in patients with laryngeal squamous cell carcinoma[J/OL]. *Cancer Imaging*, 2020, 20(1):e28[2020 Apr 22]. DOI:10.1186/s40644-020-00310-5.

[11] 谢凯,孙鸿飞,林涛,等.影像组学中特征提取研究进展[J]. *中国医学影像技术*, 2017, 33(12):1792-1796.

[12] Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach[J/OL]. *Nat Commun*, 2014, 5:e400603[3 June 2014]. DOI:10.1038/ncomms5644.

[13] 周澜,郭建峰,陈进兰,等.基于增强 CT 影像组学联合机器学习模型术前预测甲状腺乳头状癌 BRAFV600E 基因突变的价值[J]. *放射学实践*, 2025, 40(11):1396-1402.

[14] 于昕冉,冯兵.基于 CT 影像组学预测非肌层浸润性膀胱癌术后复发的临床价值[J]. *临床放射学杂志*, 2024, 43(5):788-793.

[15] Feng F, Jiang F, Liu Y, et al. Radiomics analysis of dual-layer spectral-detector CT-derived iodine maps for predicting tumor deposits in colorectal cancer[J]. *Eur Radiol*, 2025, 35(1):105-116.

[16] Liu Y, Wang Y, Hu X, et al. Multimodality deep learning radiomics predicts pathological response after neoadjuvant chemoradiotherapy for esophageal squamous cell carcinoma[J/OL]. *Insights Imaging*, 2024, 15(1):e277[2024 Nov 15]. DOI:10.1186/s13244-024-01851-0.

[17] 田娟秀,刘国才,谷珊珊,等.医学图像分析深度学习研究方法研究与挑战[J]. *自动化学报*, 2018, 44(3):401-424.

[18] 王旷怡,胡秀枋.深度学习在胰腺医学影像中的应用综述[J]. *软件导刊*, 2022, 21(3):249-252.

[19] 张文娟.基于 CT 影像组学和深度学习的进展期胃癌预后预测研究[D].兰州:兰州大学,2021.

[20] 牛冉冉,陈基明,任梦婷,等.基于 CT 影像组学的深度学习模型预测胃癌隐匿性腹膜转移的价值[J]. *放射学实践*, 2025, 40(4):485-492.

[21] 李鑫,陈金明,李志兴,等.基于增强 CT 深度学习技术预测肝细胞癌术后早期复发[J]. *放射学实践*, 2025, 40(8):1002-1010.

[22] 赵杉,闫子康,杨骏骏,等.影像组学及深度学习联合血液炎症指标预测胶质瘤预后的价值[J]. *磁共振成像*, 2024, 15(1):88-94, 100.

[23] Tian S, Yu Y, Shi K, et al. Deep learning radiomics based on ultrasound images for the assisted diagnosis of chronic kidney disease[J]. *Nephrology (Carlton)*, 2024, 29(11):748-757.

[24] 乐慧君,陈思宇,李芸,等.喉癌诊疗策略及进展[J]. *临床耳鼻咽喉头颈外科杂志*, 2019, 33(11):1017-1021.

[25] Shi S, Lin C, 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.

(收稿日期:2025-01-27 修回日期:2025-07-13)