

doi: 10.3969/j.issn.1674-1242.2026.04.012

融合 CT 影像组学与深度学习特征的结直肠癌 淋巴结转移风险预测模型构建

蔡振辉, 茹晓辉, 王江利*

(郑州华卓医院 外科一病区, 河南郑州 450100)

【摘要】目的 构建并验证一种融合计算机断层扫描(computed tomography, CT)影像组学与深度学习(deep learning, DL)特征的列线图模型,用于术前预测结直肠癌(colorectal cancer, CRC)淋巴结转移(lymph node metastasis, LNM)风险。**方法** 回顾性纳入2022年1月至2025年6月于郑州华卓医院经病理确诊的116例CRC患者的临床资料。按LNM状态和T分期进行分层随机抽样后按7:3划分训练集($n=81$)与验证集($n=35$)。基于静脉期CT图像,分别提取手工影像组学特征与预训练ResNet50的DL特征。经特征筛选后,构建并比较临床模型、单一特征模型、融合模型和联合列线图模型,以曲线下面积(area under the curve, AUC)评价模型区分度,采用Hosmer-Lemeshow检验与Brier评分评估模型校准效能。**结果** 联合列线图模型在验证集中的AUC为0.823(95%CI:0.694~0.952),敏感度为85.7%,阴性预测值为89.5%,显著优于临床模型(AUC=0.701);其预测概率也与实际观测具有良好一致性(Hosmer-Lemeshow检验 $P=0.719$, Brier score=0.148)。**结论** 本研究构建的联合预测模型能有效评估CRC LNM风险,可为术前个体化决策提供可量化工具。

【关键词】 结直肠癌; 淋巴结转移; CT影像组学; 深度学习; 预测模型; 列线图

【中图分类号】 R735.3

【文献标志码】 A

文章编号: 1674-1242 (2026) 04-0063-05

Construction of a lymph node metastasis risk prediction model for colorectal cancer integrating CT radiomics and deep learning features

CAI Zhenhui, RU Xiaohui, WANG Jiangli*

(Surgical Ward 1, Zhengzhou Huazhuo Hospital, Zhengzhou 450100, Henan, China)

【Abstract】Objective To construct and validate a nomogram model that integrates computed tomography (CT) radiomics and deep learning (DL) features to predict the risk of lymph node metastasis (LNM) in colorectal cancer (CRC) before surgery. **Methods** The clinical data of 116 CRC patients who were pathologically diagnosed at Zhengzhou Huazhuo Hospital from January 2022 to June 2025 were retrospectively included. After stratified random sampling based on LNM status and T stage, the data were divided into a training set ($n=81$) and a validation set ($n=35$) according to 7 vs. 3. Based on the CT images of the venous phase, manually extracted imaging features and DL features from the pre-trained ResNet50 were respectively obtained. After feature screening, the clinical model, single feature model, fusion model and joint nomogram model were constructed and compared. The discrimination of the model was evaluated by the area under the curve (AUC), and the calibration efficacy of the model was assessed using the Hosmer-Lemeshow test and the Brier score. **Results** The AUC of the joint nomogram model in the validation set was 0.823 (95%CI: 0.694-0.952), with a sensitivity of 85.7% and a negative predictive value of 89.5%, which was significantly better than that of the clinical model (AUC=0.701), and its prediction probability was also in good agreement with the actual observations (Hosmer-Lemeshow test $P=0.719$, Brier score=0.148). **Conclusion** The joint prediction model constructed in this study can effectively assess the risk of LNM in CRC, and can provide a quantifiable tool for individualized preoperative decision-making.

【Key words】 Colorectal cancer; Lymph node metastasis; CT radiomics; Deep learning; Predictive models; Nomogram

结直肠癌(colorectal cancer, CRC)为高发的消化系统恶性肿瘤,严重威胁患者生命健康^[1-2]。淋巴结

转移(lymph node metastasis, LNM)为CRC关键预后指标,可直接指导手术清扫范围及辅助治疗方案,对

收稿日期: 2025-07-22。

作者简介: 蔡振辉, 本科, 副主任医师, 研究方向: 普外胃肠。E-mail: 13603719889@163.com。

通信作者: 王江利, 本科, 副主任医师, 研究方向: 普外胃肠。E-mail: 13603711986@163.com。

其进行术前准确评估至关重要^[3]。然而,术前常规增强计算机断层扫描(computed tomography, CT)仅依据形态学特征评估 LNM,诊断敏感度有限,大量微小转移灶易被漏诊^[4]。近年来,影像组学(radiomics)与深度学习(deep learning, DL)为结直肠癌术前无创精准分期提供新思路。影像组学可高通量挖掘肿瘤异质性,在 LNM 预测中已表现出应用潜力^[5];DL 能自动提取人眼难以辨识的深层图像特征,人工智能的应用价值亦已有系统评价佐证^[6]。融合临床指标与影像特征的列线图可显著提升预测效能,在磁共振成像(magnetic resonance imaging, MRI)、超声等多模态影像中均具备应用前景^[7],新辅助治疗后的局部进展期 CRC 则进一步加大 LNM 评估难度^[8]。已有研究多简单拼接手工组学与 DL 特征,缺少针对淋巴结微环境的分层优化验证;虽提示瘤周存在潜在预测信息,但瘤内、瘤周的预测贡献尚未开展系统对比。鉴于此,本研究通过构建融合 CT 影像组学与 DL 特征的双通道模型,优化特征筛选与融合策略,并建立可解释的列线图,旨在为 CRC 术前 LNM 个体化风险评估提供依据。

1 资料与方法

1.1 一般资料

回顾性收集 2022 年 1 月至 2025 年 6 月于郑州华卓医院接受根治性切除术并经术后病理确诊的 CRC 患者的临床资料。初始通过电子病历系统共筛查 147 例,排除标准:①术前已接受新辅助放化疗;②CT 图像层厚 >3 mm,或存在明显运动伪影、金属伪影致使图像不可用;③癌胚抗原(carcinoembryonic antigen, CEA)、糖类抗原 19-9(carbohydrate antigen 19-9, CA19-9)、TNM 分期等关键临床资料缺失;④合并其他恶性肿瘤病史;⑤因肾功能不全无法完成增强 CT 扫描。最终纳入 116 例患者,其中男 67 例,女 49 例,年龄(62.4 ± 11.7)岁,年龄范围 34 ~ 85 岁。LNM 阳性 48 例(pN^+ , 41.4%), LNM 阴性 68 例($pN0$, 58.6%)。根据美国癌症联合委员会(American Joint Committee on Cancer, AJCC)第 8 版 TNM 分期标准记录肿瘤浸润深度(T category),其中 T_1 期 18 例、 T_2 期 22 例、 T_3 期 54 例、 T_4 期 22 例。所有病例按 LNM 状态(阳性/阴性)和 T 分期(T_{1-2}/T_{3-4})进行分层随机抽样,以 7:3 比例划分为训练集 81 例(LNM^+34 例、 LNM^-47 例)和验证集 35 例(LNM^+

14 例、 LNM^-21 例)。本研究经郑州华卓医院伦理委员会批准(批件号:20260615001)。

1.2 CT 图像采集与肿瘤感兴趣容积(volume of interest, VOI)勾画

所有患者术前 1 周行腹部增强 CT 扫描(Siemens SOMATOM Force)。分别于对比剂注射后延迟 25 s(动脉期)/65 ~ 70 s(静脉期)触发采集。选取静脉期图像进行后续处理。

由 2 名分别具有 5 年、10 年腹部影像诊断经验的放射科医师,采用 ITK-SNAP 3.8 软件沿原发肿瘤边界手动逐层勾画三维 VOI,同时生成瘤周 3 mm 扩展区作为第二区域,以捕获肿瘤侵入性微环境的间接信号。勾画结果存在明显分歧时,对争议层面经协商达成一致;随机抽取 20 例患者由 2 名医师独立完成二次勾画。肿瘤区和瘤周区所有影像组学特征的组内相关系数(intra-class correlation coefficient, ICC)中位数为 0.86(范围:0.78 ~ 0.94),表明勾画可重复性良好。ICC >0.75 认为一致性良好^[9-10],本研究采用 ICC ≥ 0.8 的筛选标准,以确保纳入特征的稳定性。

1.3 影像组学特征提取

使用 Python 3.9+ 搭配 PyRadiomics 3.0 开源工具库,从肿瘤 VOI 内提取标准化(Z-score)后的影像组学特征,共分 5 大类:①一阶统计量,示例特征 24 个;②形状特征(3D),示例特征 16 个;③灰度共生矩阵(gray-level co-occurrence matrix, GLCM),示例特征 $24 \times N_{angle}$;④灰度游程长度矩阵(gray-level run-length matrix, GLRLM),示例特征 16 个;⑤灰度大小区域矩阵(gray-level size zone matrix, GLSZM),示例特征 16 个。经镜像/旋转不变性去冗余后,肿瘤区与瘤周区合计初提特征 1 247 个。进一步以 ICC >0.75 筛选重现性良好的特征进入下一轮降维。

1.4 深度学习特征提取

截取原始 CT 轴向切片序列中覆盖肿瘤 VOI 上下各延伸 5 层(共 20 ~ 30 层),窗位/窗宽统一设为 40/400 HU,归一化至[0, 1]后 resize 至 224×224 像素。以 ImageNet 预训练 ResNet50 为骨干网络,移除原分类层,取全局平均池化层输出(512 维)作为每张切片的深度语义特征向量。本研究采用在线数据增强策略,包括随机旋转($\pm 10^\circ$)、随机水平翻转和随机缩放(0.9 ~ 1.1 倍)。对同一病例所有切片特征取均值池化获得患者级 512 维 DL 特征。本研究仅对最后一

层自适应平均池化后的紧凑向量进行后续筛选。模型训练于配备 NVIDIA RTX 3090 GPU (24GB 显存) 的工作站进行, batch size 设置为 16, 优化器为 Adam。

1.5 特征筛选与建模

1.5.1 特征筛选策略

(1) 最小冗余最大相关算法 (minimum redundancy maximum relevance, mRMR) 在经 ICC 筛选后保留特征集运行 mRMR。针对影像组学分支, 按照互信息大小排序, 初选 30 个高相关、低冗余特征; 针对 DL 分支, 按照特征与 LNM 标签间 *t* 检验统计量的绝对值排序, 取绝对值最大的前 30 维特征。

(2) 最小绝对收缩与选择算子 (least absolute shrinkage and selection operator, LASSO) + 十折交叉验证: 以 LNM 状态为因变量, 对合并后的候选特征做 LASSO logistic 回归, 构建 L1 正则化 Logistic 回归模型。选择交叉验证误差最小时的惩罚系数 λ , 筛选得到稳定性佳、冗余度低的影像组学及 DL 特征, 构建双通道融合特征向量, 作为最终模型输入。

1.5.2 建模

本研究设计了逐步递进的模型比较框架: 以临床指标构建的模型 (M1-临床) 作为基准, 分别评估单一模态特征 (M2-影像组学、M3-DL) 的预测能力, 进而测试双模态特征融合 (M4-融合) 的效能, 最终将最具预测力的影像特征与关键临床指标融合, 构建易于临床解读的联合列线图模型 (M5-联合)。以上模型输入特征如下: (1) M1-临床: CEA (≥ 5 ng/ml)、肿瘤浸润深度 ($T_3:T_{1-2}; T_4:T_{1-2}$)、分化程度; (2) M2-影像组学: LASSO 筛选的手工特征; (3) M3-DL: LASSO 筛选的 DL 特征; (4) M4-融合: 影像组学特征 + DL 特征; (5) M5-联合: M4 的预测得分 + CEA + 肿瘤浸润深度。所有模型 (M1 ~ M5) 均采用 Logistic 回归作为统一的分类器, 模型超参数通过网格搜索 + 5 次重复十折交叉验证确定。

1.6 统计学方法

计量资料以均数 $\pm$ 标准差 ($\bar{x} \pm s$) 或 $M(Q_1, Q_3)$ 表示, 组间比较采用独立样本 *t* 检验或 Mann-Whitney *U* 检验; 计数资料以例 (%) 表示, 行 χ^2 检验或 Fisher 确切概率法。特征筛选与建模借助 R 4.3.1 与 Python 3.9 联合实现, 常规统计学分析采用 SPSS 26.0 软件完成。本研究基于多因素 Logistic 回归构建联合列线图模型 (M5), 采用交叉验证筛选最优惩罚系数, 获

取高稳定性、低冗余的影像组学与 DL 特征, 构建融合特征集用于模型输入。采用受试者操作特征 (receiver operator characteristic, ROC) 分析评价模型效能, 计算曲线下面积 (area under the curve, AUC)、敏感度、特异度、准确度、阳性/阴性预测值, 不同模型 AUC 比较采用 DeLong 检验; 采用 Hosmer-Lemeshow 检验、Brier score 评估模型校准度; 决策曲线分析 (decision curve analysis, DCA) 评估模型在不同风险阈值下的临床净获益。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 训练集与验证集的临床特征

训练集与验证集在年龄、性别构成、肿瘤部位、肿瘤最大径、T 分期、分化程度及肿瘤标志物 (CEA、CA19-9) 水平方面比较, 差异均无统计学意义 (均 $P > 0.05$) (表 1)。作为模型预测目标的 LNM 状态组间分布比例亦无统计学差异 ($P > 0.05$)。

表 1 训练集与验证集的临床特征比较

变量	训练集 (<i>n</i> =81)	验证集 (<i>n</i> =35)	χ^2	<i>P</i>
年龄 ($\bar{x} \pm s$, 岁)	61.8 $\pm$ 12.1	63.9 $\pm$ 10.8	0.872	0.385
性别 (例)			0.031	0.861
男	47	20		
女	34	15		
肿瘤部位 (例)			0.008	0.929
结肠	46	20		
直肠	35	15		
肿瘤最大径 ($\bar{x} \pm s$, mm)	46.2 $\pm$ 18.7	48.5 $\pm$ 17.3	0.628	0.531
肿瘤浸润深度 (例)			1.874	0.599
T ₁	12	6		
T ₂	15	4		
T ₃	38	16		
T ₄	16	9		
分化程度 (例)			0.182	0.67
中-高分化	60	25		
低分化	21	10		
CEA ≥ 5 ng/ml (例)			0.294	0.587
是	39	18		
否	42	17		
CA19-9 ≥ 37 U/ml (例)			0.378	0.539
是	25	12		
否	56	23		
LNM 状态 (例)			0.107	0.744
阳性 (pN ⁺)	34	14		
阴性 (pN0)	47	21		

注: CEA: 癌胚抗原; CA19-9: 糖类抗原 19-9; LNM: 淋巴结转移。

2.2 特征筛选结果

ICC 筛选后保留稳定特征 892 个 ($ICC > 0.8$), 经 mRMR 初选 30 个 + DL 初选 30 个 → 合并后以 LASSO-logistic 回归在 $\lambda = 0.023$ (对应一标准误差) 处压缩至非零系数特征共 34 个, 其中手工影像组学特征 14 个, DL 特征 20 维。

2.3 不同模型预测效能比较

DeLong 检验显示, 验证集 M5-联合模型 AUC

(0.823) 显著高于 M1-临床模型 ($Z = 2.41, P = 0.016$); 与 M2-影像组学模型相比, M5 的 AUC 差异无统计学意义 ($Z = 1.07, P = 0.285$); 与 M3-DL 模型相比, M5-联合模型显示出一定的提升趋势 ($Z = 1.23, P = 0.219$)。M5-联合模型曲线最接近左上角, 提示其在权衡敏感度与特异度方面具有优势; M4-影像融合模型与 M5 相差不大 ($\Delta AUC = 0.018$) (表 2)。

表 2 5 种预测模型在训练集与验证集中的效能比较

模型	数据集	AUC (95%CI)	敏感度 (%)	特异度 (%)	准确度 (%)	阳性预测值 (%)	阴性预测值 (%)	Brier Score
M1-临床	训练集	0.728 (0.647 ~ 0.809)	67.6	70.2	69.1	63	74.4	0.231
	验证集	0.701 (0.562 ~ 0.840)	64.3	71.4	68.6	60	75	0.248
M2-影像组学	训练集	0.862 (0.793 ~ 0.931)	82.4	78.7	80.2	76.7	83.9	0.175
	验证集	0.779 (0.646 ~ 0.912)	78.6	76.2	77.1	73.3	80.9	0.192
M3-DL	训练集	0.849 (0.777 ~ 0.921)	79.4	80.9	80.2	78.1	82.2	0.183
	验证集	0.768 (0.630 ~ 0.906)	71.4	81	77.1	76.9	75.9	0.205
M4-影像融合	训练集	0.878 (0.812 ~ 0.944)	85.3	83	83.9	80.6	87.1	0.162
	验证集	0.805 (0.672 ~ 0.938)	85.7	76.2	80	75	88.9	0.179
M5-联合	训练集	0.891 (0.826 ~ 0.956)	88.2	85.1	86.4	83.3	89.7	0.138
	验证集	0.823 (0.694 ~ 0.952)	85.7	81	82.9	78.6	89.5	0.148

注: AUC: 曲线下面积; DL: 深度学习。

亚组模型分析显示, 模型 A (仅使用瘤内 VOI 特征, 分类器为 Logistic 回归) 在验证集的 AUC 为 0.782 (95%CI: 0.649 ~ 0.915); 模型 B (使用瘤内 + 瘤周 VOI 特征, 即 M2-影响组学) 在验证集的 AUC 为 0.779 (95%CI: 0.645 ~ 0.913)。二者差异无统计学意义 ($Z = 0.09, P = 0.928$), 提示本研究数据集中, 瘤周 3 mm 区域并未显著提升预测效能。

2.4 列线图模型构建与验证

2.4.1 列线图构建

将 M5-联合模型中最终入选的独立预测因子:

融合影像评分 (Rad-DL score)、CEA ≥ 5 ng/ml (是 = 1/否 = 0)、肿瘤浸润深度, 代入多元 Logistic 回归, 回归方程为: $\text{Logit}(P) = -4.872 + 2.146 \times \text{Rad-DL score} + 1.058 \times \text{CEA}_{\text{high}} + 1.324 \times T_{\text{advanced}}$ 。导出各变量回归系数并转换为列线图计分, Rad-DL 评分每增加 1 单位计 21 分, CEA ≥ 5 ng/ml 计 11 分, 肿瘤浸润深度 (T_{3-4}) 计 13 分, 总分 0 ~ 100 分对应 LNM 风险 0 ~ 95%。图 1 中虚线代表总分的读数轴, 用于将各变量得分相加后映射到对应的风险概率。以风险阈值 0.45 为 cut-off 时, 验证集约 82.9% 患者可正确分层。

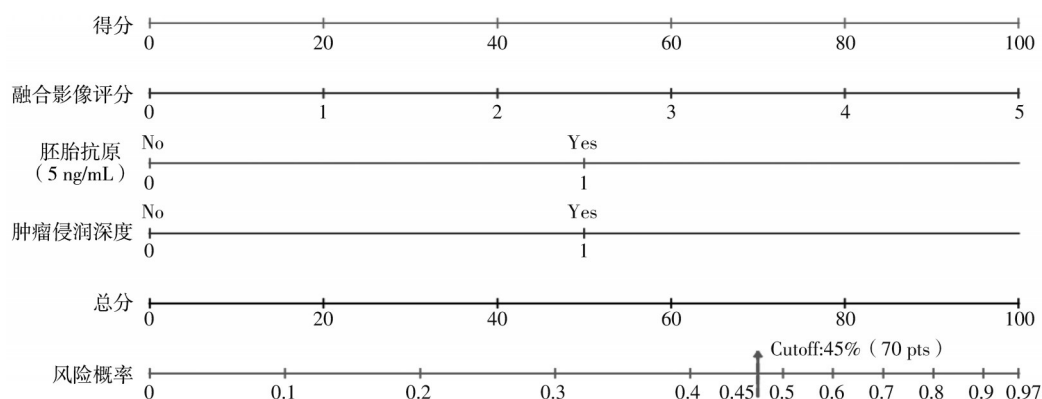


图 1 CRC LNM 风险预测列线图

注: CRC: 结直肠癌; LNM: 淋巴结转移。

2.4.2 模型验证

Hosmer-Lemeshow 检验 $\chi^2 = 5.370$, $P = 0.719$; Brier 评分为 0.148, 提示联合列线图预测概率与实际观测值一致性良好, 提示模型无显著过度乐观偏移。DCA 结果显示, 在 0.1 ~ 0.9 的风险阈值范围内, M5-联合模型相较于“全部治疗”或“全部不治疗”的策略, 均能带来更高的临床净获益。尤其在 0.3 ~ 0.6 的风险阈值区间, 其净获益最为显著。

3 讨论

术前准确评估 CRC LNM 状态对治疗决策至关重要。传统 CT 形态学评估敏感度有限。影像组学通过量化肿瘤异质性信息弥补了不足, 而 DL 能自动挖掘与转移相关的深层图像模式。本研究主要创新在于: ① 采用双源特征融合与级联降维策略, 先经 ICC 和 mRMR 筛选, 再用 LASSO 回归从融合特征中提炼出 34 个关键特征, 有效提升了小样本 ($n = 116$) 模型的稳健性; ② 引入瘤周 3 mm 区域进行分析, 并构建亚组模型进行了对比验证; ③ 构建了可解释的临床列线图, 整合了影像学评分、CEA 和肿瘤浸润深度, 便于临床快速应用。本研究严格遵守预测模型构建规范, 未将包含结局信息的 N 分期纳入自变量, 仅利用 T 分期信息, 确保了模型的科学性与独立性。

结果验证显示, 联合模型在验证集中 AUC 为 0.823, 敏感度达 85.7%, 阴性预测值为 89.5%。表明该模型更适用于排除低风险患者, 从而可能避免不必要的过度治疗, 与其作为辅助决策工具的定位相符。模型效能解读方面, 本研究从三个层面进行了分析: 首先是区分能力, M5 模型的 AUC 为 0.823, 优于传统临床模型, 表明其能较好地区分 LNM 阳性和阴性患者; 其次是校准性能, Hosmer-Lemeshow 检验和 Brier 评分均表明模型的预测概率与实际发生概率高度一致; 最后是临床获益, DCA 证实, 在多个风险阈值下, 使用该模型指导临床决策均可带来净获益。

本研究亚组分析发现, 加入瘤周 3 mm 区域并未显著提升模型效能。其可能原因为: 瘤周区域的特征与瘤内特征高度相关, 或在当前特征提取和筛选策略下, 其增量信息有限。该结论与既往报道相一致^[11]。提示瘤周区域的预测价值可能因肿瘤类型、成像参数和分析方法而异, 值得进一步探讨。

综上, 基于术前静脉期增强 CT, 融合影像组学

特征与 DL 特征构建的双通道预测模型, 能以 0.823 的验证集 AUC 对 CRC LNM 风险进行有效分层。整合 CEA 与 T 分期后输出的可视化列线图具备良好的校准度, 可作为术前无创分期决策的客观辅助工具。但本研究存在一定局限性: 作为中心回顾性设计, 样本量有限, 统计效能和模型泛化能力有待多中心、大样本数据进一步验证; 研究未纳入能谱等多参数影像信息, DL 特征可视化与机制解读尚未充分; 肿瘤 VOI 人工勾画存在一定主观偏差, 后续可探索半自动或全自动分割方法。

参考文献

- [1] BRAY F, LAVERANNE 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.
- [2] MORGAN E, ARNOLD M, GINI A, *et al.* Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN[J]. *Gut*, 2023, 72(2):338-344.
- [3] 伍雯, 张若昕, 翁俊勇, 等. 探索阳性淋巴结比率在 ypⅢ期结直肠癌患者中的预后价值及预测模型的建立[J]. *中国癌症杂志*, 2024, 34 (9): 873-880.
- [4] NERAD E, LAHAYE M J, MAAS M, *et al.* Diagnostic accuracy of CT for local staging of colon cancer: a systematic review and meta-analysis[J]. *AJR Am J Roentgenol*, 2016, 207(5):984-995.
- [5] 刘家璇, 卢希, 杜莹, 等. 基于增强 CT 构建影像组学模型预测直肠癌患者治疗前区域淋巴结转移[J]. *实用放射学杂志*, 2023, 39 (12): 1985-1989.
- [6] 彭琳, 王冬青, 庄子健, 等. MRI 影像组学及深度学习在直肠癌中的研究进展[J]. *中国 CT 和 MRI 杂志*, 2024, 22 (5): 177-180.
- [7] CHENG Y, YU Q, MENG W, *et al.* Clinico-radiologic nomogram using multiphase CT to predict lymph node metastasis in colon cancer[J]. *Mol Imaging Biol*, 2022, 24(5): 798-806.
- [8] 朱海涛, 朱汇慈, 李晓婷, 等. 基于淋巴结影像学特征的联合预测模型评价局部进展期直肠癌患者行新辅助化疗后淋巴结转移的应用价值[J]. *中华消化外科杂志*, 2025, 24 (6): 769-776.
- [9] KOO T K, LI M Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research[J]. *J Chiropr Med*, 2016, 15(2):155-163.
- [10] LAMBIN P, LEIJENAR R T H, DEIST T M, *et al.* Radiomics: the bridge between medical imaging and personalized medicine[J]. *Nat Rev Clin Oncol*, 2017, 14(12):749-762.
- [11] KORSAVIDOU HULT N, TARAI S, HAMMARSTRÖM K, *et al.* Inclusion of tumor periphery in radiomics analysis of magnetic resonance images does not improve predictions of preoperative therapy response in patients with rectal cancer[J]. *Abdom Radiol (NY)*, 2026, 51(3):1116-1128.