

◁腹部影像学▷

基于 MRI 影像组学评估局部进展期直肠癌新辅助放化疗反应：
一项两中心多设备研究贾子琪¹, 林耿斌¹, 袁文静¹, 张汉良¹, 崔亮¹, 林韵颖¹, 刘岷¹, 吴元魁², 陈维翠¹

(1. 广东省中医院(广州中医药大学第二附属医院)放射科, 广东 广州 510120;

2. 南方医科大学南方医院影像中心, 广东 广州 510515)

[摘要] 目的:探讨基于 MRI 影像组学评估局部进展期直肠癌(LARC)患者新辅助放化疗(nCRT)后病理完全缓解(pCR)的价值。方法:回顾性收集 2021 年 10 月—2024 年 1 月在南方医科大学南方医院(中心 1, n=142)和广东省中医院(中心 2, n=61)共 203 例 LARC 患者,每位患者 nCRT 后的高分辨 MRI 检查于以上 2 个中心、4 种不同机型的 3.0T MR 设备上完成。两名放射医生对 nCRT 后的 T₂WI 和弥散加权成像(DWI)图像全瘤病灶进行感兴趣体积的勾画并提取组学特征。将中心 1 患者按 4:1 的比例划分为训练集(n=113)和内部验证集(n=29),中心 2 的患者作为独立的外部测试集。采用皮尔森相关系数进行降维,采用支持向量机、线性判别分析、随机森林、逻辑回归及高斯过程共 5 种分类器构建影像组学模型,并筛选出模型性能最好的分类器。采用受试者工作特征(ROC)曲线的曲线下面积(AUC)评估影像组学模型对 pCR 的评估效能。结果:从 nCRT 后的 T₂WI 及 DWI 序列上共提取 1 888 个组学特征,经过降维及筛选后获得 12 个最有价值的组学特征(5 个特征来源于 T₂WI 序列,7 个特征来源于 DWI 序列),包括 1 个形态学特征、1 个一阶统计特征、1 个拉普拉斯高斯特征和 9 个小波特征。采用高斯过程构建影像组学模型,该模型在训练集、内部验证集、外部测试集中的 AUC 分别为 0.999、0.812 和 0.803。结论:基于 MRI 影像组学模型能较准确评估 LARC 患者 nCRT 的治疗反应,该模型通过独立的外部测试集验证,有望进一步指导临床决策。

[关键词] 直肠肿瘤;磁共振成像**[中图分类号]** R735.37;R445.2**[文献标识码]** A**[文章编号]** 1008-1062(2025)03-0194-06

DOI:10.12117/jccmi.2025.03.009

MRI radiomics for assessing treatment response of neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a two-center, multi-vendor studyJIA Zi-qi¹, LIN Geng-bin¹, YUAN Wen-jing¹, ZHANG Han-liang¹, CUI Liang¹,LIN Yun-ying¹, LIU Xian¹, WU Yuan-ku², CHEN Wei-cui¹

(1. Department of Radiology, Guangdong Provincial Hospital of Traditional Chinese Medicine(The Second Affiliated Hospital of Guangzhou University of Chinese Medicine), Guangzhou 510120, China; 2. Department of Radiology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China)

Abstract: Objective: To explore the value of MRI radiomics in evaluating pathological complete response (pCR) after neoadjuvant chemoradiotherapy(nCRT) in patients with locally advanced rectal cancer(LARC). **Methods:** Between October 2021 and January 2024, a total of 203 patients with LARC were retrospectively gathered from two medical centers: Nanfang Hospital of Southern Medical University (Center 1, n=142) and Guangdong Provincial Hospital of Traditional Chinese Medicine (Center 2, n=61). High-resolution MRI examinations following nCRT were conducted using 3.0T MR equipment from four different models across the two centers. Subsequently, two radiologists delineated the volume of interest(VOI) encompassing the entire tumor lesions on T₂WI and diffusion weighted imaging(DWI) images obtained after nCRT. Subsequently, radiomic features were extracted from these delineated VOIs for further analysis. Patients from Center 1 were stratified into a training set(n=113) and an internal validation set (n=29) at a ratio of 4:1, while the patients from Center 2 were utilized as an independent external test set. Pearson correlation coefficients was used for dimensionality reduction. Radiomics models were constructed with five classifiers, including support vector machine, linear discriminant analysis, random forest, Logistic regression, and gaussian process, and the classifier with the best model performance was selected. Evaluating the model's assessment performance was based on the area under the curve(AUC) of the receiver operating characteristic(ROC) curve. **Results:** A total of 1 888 radiomics features were extracted from T₂WI and DWI images after nCRT. After dimensionality reduction and feature selection, twelve most valuable radiomic features were obtained, comprising 5 features from the T₂WI sequence and 7 features

[收稿日期] 2024-06-12;**[修回日期]** 2024-08-06**[作者简介]** 贾子琪(1997-),女,甘肃酒泉人,在读硕士研究生。E-mail:tyyxjiaziqu@163.com**[通信作者]** 陈维翠,广东省中医院(广州中医药大学第二附属医院)放射科,510120。E-mail:chenweicui@126.com**[基金项目]** 国家自然科学基金(82202259);广州市科技局市员联合资助项目(2023A03J0245);

广东省中医院第十三届朝阳人才项目(ZY2022YL05)。

from the DWI sequence. These features encompass 1 morphological feature, 1 first-order feature, 1 laplacian of gaussian feature and 9 wavelet features. The radiomics model was constructed using gaussian process. The model achieved AUC values of 0.999, 0.812, and 0.803 for the training set, internal validation set, and external test set, respectively. **Conclusion:** A predictive MRI-based radiomic model demonstrates promising accuracy in assessing treatment response to nCRT in patients with LARC. This model has been validated using an independent external test set, suggesting its potential utility for guiding clinical decision-making.

Key words: Rectal Neoplasms; Magnetic Resonance Imaging

结直肠癌是消化系统最常见的恶性肿瘤,与癌症相关的致死率位居第二,约 30%的病例起源于直肠^[1]。对于局部进展期直肠癌(Locally advanced rectal cancer, LARC)患者,术前新辅助放化疗(Neoadjuvant chemoradiotherapy, nCRT)加全直肠系膜切除是标准的治疗策略^[2]。nCRT 旨在缩小肿瘤,降低肿瘤分期,争取肿瘤完全切除(R0),提高保肛率及降低复发率^[3]。约 20%的患者于治疗后能达到病理完全缓解(Pathological complete response, pCR)^[4],对于这些患者,建议采取“等待观察”策略,能有效减少手术创伤和风险,提高生活质量^[5]。高分辨 MRI 是评估 nCRT 疗效的首选检查方法^[2],但由于治疗后肿瘤出现坏死、炎性反应及纤维化,导致 MRI 信号复杂,肉眼评估容易误判。因此,有必要开发新的评估方法,以准确实施“等待观察”策略。多项研究表明, MRI 影像组学在评估和预测 LARC 患者 nCRT 治疗反应方面具有潜力^[6-7]。然而,这些研究中的成像设备较单一,且缺乏外部验证,使模型难以在临床上推广^[8]。本研究回顾性收集了在 2 家医疗中心使用 2 种品牌、4 种型号的 3.0T MR 扫描仪进行扫描的 203 例 LARC 患者的资料,基于 nCRT 后的 T₂WI 及弥散加权成像(Diffusion weighted imaging, DWI)序列,探讨 MRI 影像组学在评估 LARC 患者 nCRT 后 pCR 中的价值。

1 资料与方法

1.1 研究对象

本研究获得广东省中医院伦理委员会批准(伦理号:BE2023-142),回顾性研究无需受试者签署知

情同意书。回顾性收集 2021 年 10 月—2024 年 1 月来自南方医科大学南方医院(中心 1)和广东省中医院(中心 2)的直肠癌患者。纳入标准:①首诊为中低位 LARC,即 MRI 显示肿瘤下缘距肛缘≤10 cm,临床肿瘤分期为 cT3 或 cT4 期^[9]。②接受完整的 nCRT 治疗方案。放疗方案:1 周 5 天,每天 1 次,共放疗 5 周,总辐射剂量为 45~50 Gy,分次剂量为 1.5~2.0 Gy。化疗方案共 4 种,根据患者情况选择其中 1 种:奥沙利铂+卡培他滨双药,6~8 个周期;卡培他滨单药,3~6 个周期;奥沙利铂+亚叶酸钙+5 氟尿嘧啶 3 药,3~6 个周期;5 氟尿嘧啶+亚叶酸钙,6 个周期。③完成 nCRT 后 6~8 周行全直肠系膜切除术。④nCRT 后约 2 周(术前 1 周)行高分辨 MRI 检查,包括 T₂WI 和 DWI 序列。⑤病理确诊为非黏液性直肠腺癌。排除标准:①临床及病理资料不全。②接受过免疫或靶向等其它治疗。③MRI 图像存在伪影。基于以上纳排标准,共 203 名 LARC 患者入组,男 158 例,女 45 例,年龄 24~80 岁,平均(56.31±11.56)岁。

1.2 MRI 图像采集

中心 1 使用 Philips Achieva(荷兰,飞利浦医疗)3.0T MR 扫描设备,中心 2 使用 Philips Ingenia(荷兰,飞利浦医疗)、Siemens Prisma(德国,西门子医疗)、Siemens Verio(德国,西门子医疗)3.0T MR 扫描设备,对 LARC 患者进行直肠高分辨 MRI 扫描。不同中心所使用的 MR 品牌、机型及 T₂WI 和 DWI 扫描参数见表 1。

1.3 病理组织学分析

手术标本行常规 HE 染色和免疫组化染色。采

表 1 两个中心的 MR 扫描仪品牌、机型及 T₂WI 和 DWI 扫描参数

	中心 1		中心 2					
	Philips Achieva		Philips Ingenia		Siemens Prisma		Siemens Verio	
	T ₂ WI	DWI	T ₂ WI	DWI	T ₂ WI	DWI	T ₂ WI	DWI
TE(ms)	100	60	100	81	89	53	93	83
TR(ms)	3 906	2 000	3 664	3 514	8 040	4 900	6 350	5 900
EPI 因子	21	59	17	53	19	80	28	115
翻转角(°)	90	90	90	90	160	90	140	90
层厚(mm)	3	4	3	4	3	4	3	5
矩阵	316×314	120×118	228×228	108×106	320×240	140×104	320×256	192×115
视野(cm)	200×200	240×240	200×200	320×320	200×200	180×320	200×200	270×360

注:TE:回波时间;TR:重复时间;EPI 因子:平面回波成像因子。

用 2022 年美国国家综合癌症网络的肿瘤退缩分级 (Tumor regression grading, TRG) 标准对 nCRT 后直肠癌的病理反应进行评估, 在单个高倍镜视野下, 根据肿瘤区域有无肿瘤细胞残留分为 pTRG0 (无癌细胞残存)、pTRG1 (仅见小簇状或单个癌细胞残留)、pTRG2 (小部分肿瘤细胞残存, 大部分为纤维化) 和 pTRG3 (少量或无肿瘤细胞被杀死, 广泛残余癌)。将患者分为 pCR 组 (pTRG0) 和非 pCR 组 (pTRG1~3)^[2]。

1.4 手动标注肿瘤

将患者 nCRT 后的 T₂WI 及 DWI (b=1 000 s/mm²) 图像以 Dicom 格式导入 3D-Slicer 软件 (Version 4.10.2, <http://www.slicer.org>)。由 1 名标注医师和 1 名审核医师共同完成手动标注肿瘤感兴趣体积 (Volume of interesting, VOI) 的工作, 标注医师和审核医师分别具有 6 年及 12 年胃肠肿瘤诊断经验, 审

核医师对标注医师进行培训。标注原则遵循结直肠癌 CT 和 MRI 标注专家共识^[10]。T₂WI 图像的标注方法如下: 以 MR 基线图像为参考, 调整窗宽和窗位, 结合多个 MR 序列的冠状面、矢状面图像确定瘤体边界, 沿病灶边缘进行勾画, 避免包括肠腔内气体、肠壁周围脂肪等, 应尽可能包全瘤床区域的所有信号。DWI 图像的标注方法如下: 结合 MR 基线图像及 nCRT 后的多个 MR 序列识别肿瘤, 在原瘤床区域进行勾画 (图 1)。对于无肉眼可识别的肿瘤残余且怀疑 pCR 的患者, 基于原发肿瘤病灶的位置, 在相同层面和位置对正常肠壁进行勾画。标注医师完成标注后, 由审核医师对每例患者肿瘤标注进行复核、修改。最终获得的 T₂WI 及 DWI 标注数据集将用于提取影像组学特征。

1.5 影像组学分析

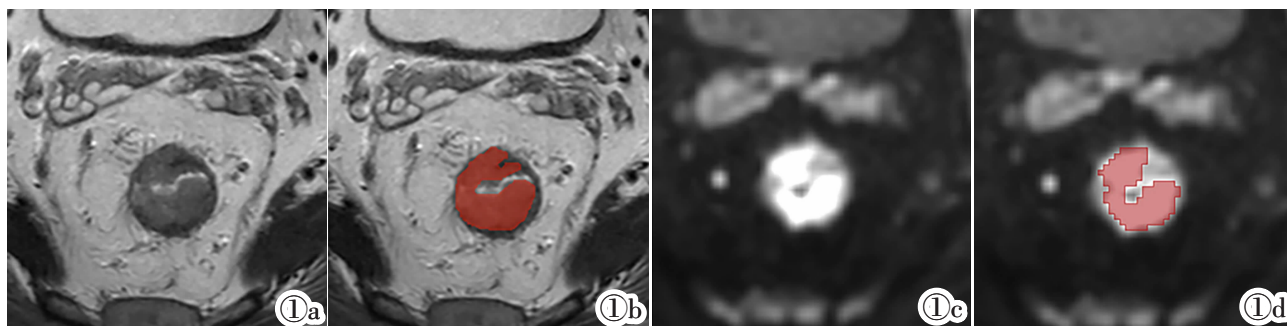


图 1 男, 32 岁, LARC 患者, 对 nCRT 反应不佳。于 T₂WI 图像 (图 1a) 上沿病变边缘标注整个肿瘤 (图 1b, 红色区域); 于 DWI 图像 (图 1c) 上沿病变边缘标注整个肿瘤 (图 1d, 红色区域)。

Figure 1. A 32-year man with LARC showed poor response to nCRT. The whole tumor (Figure 1b, red area) was marked on the edge of the lesion on the T₂WI image (Figure 1a); the tumor (Figure 1d, red area) was marked on the edge of the lesion on the DWI image (Figure 1c).

1.5.1 特征提取

使用基于 Python (Version 3.7.6) 编写的开源软件 FeAture Explorer Pro (Version 0.5.1) 进行影像组学特征提取, 包括一阶统计特征、形状特征、灰度依赖矩阵 (Gray level dependence matrix, GLDM) 特征、灰度共生矩阵 (Gray level co-occurrence matrix, GLCM) 特征、灰度距离矩阵特征 (Gray level run length matrix, GLRLM) 特征、灰度大小区 (Gray level size zone matrix, GLSZM) 特征、邻域灰度差矩阵 (Neighboring gray tone difference matrix, NGTDM) 特征、小波特征及拉普拉斯高斯特征。提取的特征使用 Combat 协调方法消除批次效应。

1.5.2 特征筛选

首先, 对数量较少的 pCR 类别样本进行过采样。其次, 由于不同特征的数值差别较大, 使用 Z-Score 对特征矩阵进行归一化处理, 将不同量级的数据缩放到相同量级的 Z-Score 分值进行比较。然后, 使用皮尔森相关系数 (Pearson correlation coefficients, PCC)

对影像组学特征进行降维处理。将任意两个特征组成一组特征对, 若该特征对的 PCC 大于 0.95, 则删除其中一个特征, 以确保每个特征相互独立。最后, 使用递归特征消除 (Recursive feature elimination, RFE) 对特征进行筛选, 通过增加或删除特定特征变量, 获得最优模型性能的特征组合。

1.5.3 模型构建

应用 5 种机器学习分类器构建影像组学模型: 支持向量机 (Support vector machine, SVM)、线性判别分析 (Linear discriminant analysis, LDA)、随机森林 (Random forest, RF)、逻辑回归 (Logistic regression, LR) 及高斯过程 (Gaussian process, GP)。将中心 1 的数据按照 4:1 的比例进行分层采样, 分为训练集 (n=113, 其中 pCR 患者 23 例, 非 pCR 患者 90 例) 和内部验证集 (n=29, 其中 pCR 患者 6 例, 非 pCR 患者 23 例), 并采用十折交叉验证以获得稳定的结果。中心 2 的数据作为独立的外部测试集 (n=61, 其中 pCR 患者 13 例, 非 pCR 患者 48 例)。采用受试者工作特

征(Receiver operating characteristic,ROC)曲线获取曲线下面积(Area under the curve,AUC),并评估模型的准确性、特异性及敏感性,以评估模型对 pCR 的预测效能。

1.6 统计学分析

采用 SPSS 26.0 及 MedCalc 20.0.22 统计软件进行数据分析。采用 Kolmogorov-Smirnov 法检验定量资料的正态分布,服从正态分布的数据以均数±标准差($\bar{x}\pm s$)表示,服从偏态分布的数据以(M(Q1, Q3))表示。训练集、内部验证集和外部测试集之间的比较采用单因素方差分析或成组的 Wilcoxon 秩和检验,计数资料的比较使用 χ^2 检验。采用 DeLong

检验比较两种组学模型 AUC 的差异。双侧 $P<0.05$ 为差异有统计学意义。

2 结果

2.1 患者的一般资料

203 名 LARC 患者中,pCR 患者有 42 例,非 pCR 患者有 161 例(pTRG1 70 例、pTRG2 74 例、pTRG3 17 例),中心 1 和中心 2 的 pCR 率分别为 20.42%(29/142)和 21.31%(13/61)。训练集、内部验证集和外部测试集的年龄、性别、化疗方案、TRG、pCR、nCRT 前后癌胚抗原(CEA)水平比较均无统计学差异(表 2)。

表 2 203 例 LARC 患者训练集、内部验证集、外部测试集的临床病理资料比较

	中心 1(n=142)			中心 2	统计量	P 值
	训练集(n=113)	内部验证集(n=29)	外部测试集(n=61)	外部测试集(n=61)		
年龄(岁, $\bar{x}\pm s$)	56.72±10.63	55.34±11.49	56.03±13.36		0.186	0.830 ¹
性别(N,%)	男	82(72.6)	24(82.8)	52(85.2)	4.167	0.124 ²
	女	31(27.4)	5(17.2)	9(14.2)		
化疗方案(n,%)	CapeOX	81(71.7)	17(58.6)	43(70.5)	7.007	0.320 ²
	卡培他滨	15(13.3)	5(17.2)	10(16.4)		
	mFOLFOX6	12(10.6)	6(20.7)	3(4.9)		
	5-氟尿嘧啶/亚叶酸钙	5(4.4)	1(3.4)	5(8.2)		
肿瘤退缩等级(n,%)	0 级	23(20.4)	6(20.7)	13(21.3)	1.497	0.960 ²
	1 级	36(31.9)	11(37.9)	23(37.7)		
	2 级	45(39.8)	9(31.0)	20(32.8)		
	3 级	9(8.0)	3(10.3)	5(8.2)		
pCR(n,%)	是	23(20.4)	6(20.7)	13(21.3)	0.156	0.925 ²
	否	90(79.6)	23(73.9)	48(78.7)		
nCRT 前 CEA($\mu\text{g/L}$,M(Q1,Q3))	4.82(2.54,17.61)	6.02(2.53,16.77)	5.48(2.18,11.85)		1.248	0.536 ³
nCRT 后 CEA($\mu\text{g/L}$,M(Q1,Q3))	2.08(1.33,3.23)	2.30(1.46,4.01)	2.48(1.30,3.85)		0.869	0.648 ³

注:1:采用单因素方差分析;2:采用 χ^2 检验;3:采用成组的 Wilcoxon 秩和检验。CapeOX:奥沙利铂+卡培他滨双药化疗方案;mFOLFOX6:奥沙利铂+亚叶酸钙+5 氟尿嘧啶 3 药化疗方案;pCR:病理完全缓解;nCRT:新辅助放疗;CEA:癌胚抗原。

2.2 影像组学特征选择

从 T₂WI 序列和 DWI 序列中共提取出 1 888 个影像组学特征。通过比较每个特征对的相似度,删除 PCC 大于 0.95 的特征对中的一个特征,最终获得 680 个相互独立的影像组学特征。在构建影像组学模型之前,通过 RFE 递归减小特征集的规模,基于分类器选择特征,最终筛选出 12 个与 pCR 相关的影像组学特征,包括来源于 T₂WI 序列的 1 个形态学特征(Original_shape_Elongation)和 4 个小波特征(Wavelet-HHH_firstorder_Mean、Wavelet-HHH_glszm_SizeZone NonUniformityNormalized、Wavelet-HHL_firstorder_Skewness 和 Wavelet-LHL_glcm_JointEntropy),以及来源于 DWI 序列的 1 个一阶统计特征(Original_firstorder_Median)、1 个拉普拉斯高斯特征(Log-sigma-1-mm-3D_firstorder_Maximum)和 5 个小波特征(Wavelet-HHH_glszm_GrayLevelNonUniformityNormalized、Wavelet-LHH_glszm_ZoneVariance、Wavelet-LLL_firstorder_RobustMeanAbsoluteDeviation、Wavelet-HHH_

glszm_LargeAreaHighGrayLevelEmphasis 和 Wavelet-LLL_firstorder_Median)。

2.3 构建影像组学模型及模型性能评估

基于 GP、LDA、LR、RF 和 SVM 分类器构建的影像组学模型在训练集的 AUC 分别为 0.999、0.880、0.876、0.934 和 0.878,在内部验证集的 AUC 分别为 0.812、0.616、0.681、0.717 和 0.659,在外部测试集的 AUC 分别为 0.803、0.708、0.702、0.702 及 0.684。在训练集中,基于 GP 构建的影像组学模型具有最大的 AUC,两两比较差异有统计学意义(P 均 <0.05);在内部验证集和外部测试集中,基于 GP 构建的影像组学模型的 AUC 最大,但与其他 4 种分类器比较,差异无统计学意义($P>0.05$)(表 3,图 2)。

基于 GP 分类器构建的影像组学模型在训练集的 AUC 为 0.999,准确度为 98.23%,敏感性为 95.65%,特异性为 98.89%;在内部验证集的 AUC 为 0.812,准确度为 79.31%,敏感性为 83.33%,特异性为 60.78%;在外部验证集的 AUC 为 0.803,准确度为 81.97%,

表 3 5 种机器学习分类器在训练集、内部验证集及外部测试集的性能

分类器	训练集(n=113)		内部验证集(n=29)		外部测试集(n=61)	
	AUC(95%CI)	ACC(%)	AUC(95%CI)	ACC(%)	AUC(95%CI)	ACC(%)
GP	0.999(0.998~1.000)	99.12	0.812(0.651~0.979)	79.31	0.803(0.676~0.930)	81.97
LDA	0.880(0.806~0.955)	77.88	0.616(0.300~0.932)	51.72	0.708(0.515~0.901)	81.97
LR	0.876(0.805~0.947)	74.34	0.681(0.410~0.952)	51.72	0.702(0.513~0.891)	63.93
RF	0.934(0.886~0.982)	68.14	0.717(0.469~0.966)	48.28	0.702(0.550~0.854)	60.66
SVM	0.878(0.800~0.956)	81.42	0.659(0.364~0.955)	55.17	0.684(0.492~0.877)	77.05

注:GP:高斯过程;LDA:线性判别分析;LR:逻辑回归;RF:随机森林;SVM:支持向量机;AUC:受试者操作特征曲线下面积;95%CI:95%置信区间;ACC:准确性。

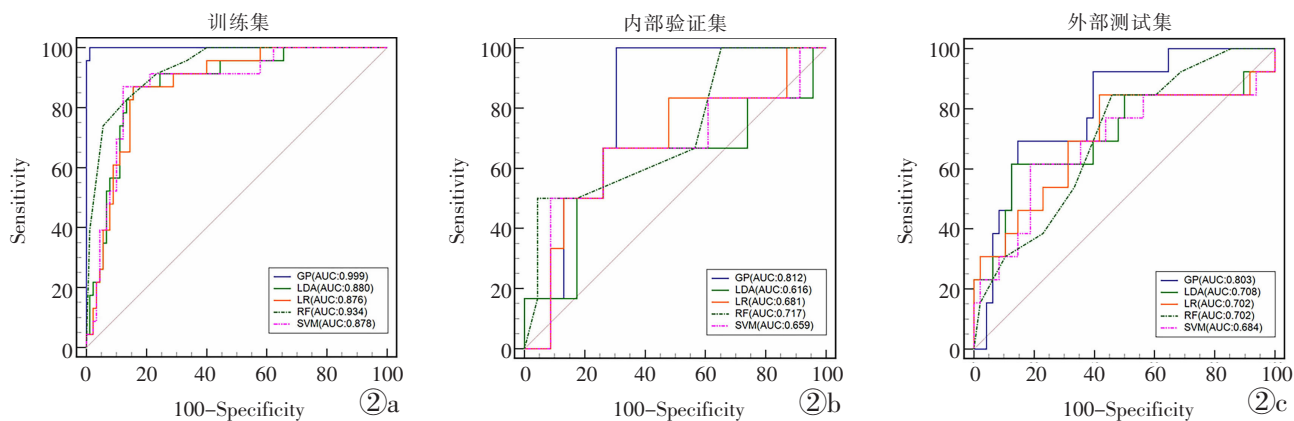


图 2 不同分类器所构建的影像组学模型训练集(图 2a)、内部验证集(图 2b)及外部测试集(图 2c)的 ROC 曲线及 AUC。GP:高斯过程;RF:随机森林;LR:逻辑回归;LDA:线性判别分析;SVM:支持向量机。

Figure 2. The receiver operating characteristic (ROC) curve and the area under the curve (AUC) of the radiomics model constructed by different classifiers in the training set(Figure 2a), internal validation set(Figure 2b) and external test set(Figure 2c). GP: Gaussian process; RF: Random forest; LR: Logistic regression; LDA: Linear discriminant analysis; SVM: Support vector machine.

敏感性为 69.23%,特异性为 85.42%。

3 讨论

“等待观察”策略是针对 nCRT 后达到 pCR 患者的一种非手术治疗观察模式,有助于保留直肠反射弧、保留直肠及肛门的器官功能、避免手术损伤和并发症风险,从而提高直肠癌患者远期生活质量和重返社会的能力^[11]。研究表明,MRI 影像组学在预测和评价 LARC 患者 nCRT 效果以及 TRG 等方面具有较高的诊断准确性^[12]。Antunes 等^[13]提取了 104 名 LARC 患者 nCRT 前的 MRI 影像组学特征构建预测 pCR 的影像组学模型,AUC 为 0.69。Cui 等^[14]利用 LARC 患者治疗前的多模态 MR 数据构建模型,该模型对 pCR 的预测效果非常显著,AUC 达到了 0.948。尽管上述研究结果令人鼓舞,但均存在单中心、缺乏外部验证及成像设备单一等不足。本研究使用开源软件 FeAture Explorer Pro 基于 T₂WI 和 DWI 图像提取影像特征,并利用多种分类器构建影像组学模型,以评估 LARC 患者 nCRT 后的 pCR 状态,研究结果表明,GP 分类器有提升影像组学模型性能的趋势。本研究的优势在于 LARC 患者的 MR 数据来源于 2 个品牌、多个机型的 MR 设备,与实际

临床诊疗过程相符合;此外,构建的影像组学模型经过独立的外部测试集验证,因此结果具有较高的可信度。

以往多项研究主要侧重于对 nCRT 前的 MR 序列进行影像组学特征的提取和模型构建^[15],而本研究则将关注点置于 LARC 患者 nCRT 后的 MR 序列上。T₂WI 序列主要提供病灶的解剖形态学信息,而 DWI 序列则是 MRI 中最常用的功能序列,反映了组织内水分子的扩散特性。因此,所获得的影像组学特征能更准确地反映肿瘤在治疗后的状态。提取的 12 个影像组学特征包括了 1 个 T₂WI 形态学特征、1 个一阶统计特征、1 个拉普拉斯高斯特征和 9 个小波特征。纹理特征有助于捕捉相邻像素间的空间关系,从而反映组织的异质性;小波特征则通过小波变换分解获得图像强度和纹理特征,使特征集中在肿瘤体积内不同频率范围内。此外,为了消除不同中心影像数据带来的批次效应,本研究采用 Combat 方法进行校正,通过参数和非参数的经验贝叶斯框架对批次效应进行了校准,以确保数据的稳健性和模型的可靠性。

本研究结果中,GP 分类器在训练集、内部验证集和外部测试集中的 AUC 分别为 0.999、0.812 和

0.803, 高于 LDA、LR、RF 和 SVM 分类器。尽管 De-long 检验结果显示, 在内部验证集和外部测试集中, 基于 GP 构建的影像组学模型的 AUC 与其他分类器相比没有统计学上的显著差异 ($P>0.05$), 但 GP 分类器呈现了提升模型诊断效能的趋势, 这种差异无统计学意义的原因可能与本研究样本量较小有关。GP 是一种非参数的监督学习方法, 它通过结合先验知识来对数据进行预测, 其中包括多元高斯到无限大小的实值变量集合的扩展^[16-17]。与大多数其他技术相比, GP 建模不仅提供平均预测, 还提供模型保真度的度量, 在小数据集和不确定性评估方面更具优势, 具有较好的平衡性^[18]。Deist 等^[19]收集了不同部位肿瘤治疗的 12 个数据集, 并比较了不同分类器的预测和评价能力, 发现 RF 和 LR 在放化疗疗效判断以及毒性预测方面表现更好。机器学习中不同分类器的性能受到数据集大小、特征空间维数和变量关系等因素的影响。此外, 数据集中的噪音和数据的线性分布性也会影响分类器的效果^[20]。

本研究存在一定的局限性。首先, 本研究为回顾性研究, 难以避免选择偏移。其次, 尽管使用了 ComBat 方法以消除批次效应, 来自两个中心的不同 MR 设备和扫描参数可能会影响影像组学特征的提取和模型的性能。第三, 本研究的总样本量和 pCR 患者样本量较少, 尽管已采用了十折交叉验证以提高模型的稳定性, 但有限的样本量仍可能影响结果的准确性。最后, 手动分割肿瘤的过程耗时费力, 因此有必要在今后的研究中利用深度学习技术进行肿瘤的自动分割, 以提高效率和准确性。

综上所述, 基于 nCRT 后 T₂WI 和 DWI 序列的影像组学模型对 LARC 患者 nCRT 后是否能达到 pCR 具有一定的预测价值, 尤其是基于 GP 分类器构建的影像组学模型展现出较高的诊断效能。该模型通过独立的外部测试集验证, 有望为 LARC 患者的精准治疗提供一种无创定量的分析方法。

[参考文献]

[1] Bray F, Laversanne 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] Benson AB, Venook AP, Al-Hawary MM, et al. Rectal Cancer, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology [J]. *J Natl Compr Canc Netw*, 2022, 20(10): 1139-1167.

[3] He L, Xiao J, Zheng P, et al. Lymph node regression grading of locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy [J]. *World J Gastrointest Oncol*, 2022, 14 (8): 1429-1445.

[4] Stanzione A, Verde F, Romeo V, et al. Radiomics and machine

learning applications in rectal cancer: Current update and future perspectives [J]. *World J Gastroenterol*, 2021, 27(32): 5306-5321.

[5] 中国直肠癌新辅助治疗后等待观察数据库研究协作组, 中国医师协会外科医师分会中国医师协会肛肠医师分会, 中华医学会外科学分会结直肠外科学组, 等. 直肠癌新辅助治疗后等待观察策略专家共识 (2020 版) [J]. *中华胃肠外科杂志*, 2020, 23 (1): 1-9.

[6] 胡鸿博, 赵升, 姜昊, 等. 基于原发癌灶 T₂WI 的影像组学特征预测局部进展期直肠癌新辅助治疗疗效及多种分类器效能比较 [J]. *磁共振成像*, 2023, 14(11): 77-83.

[7] 周晓俞, 杨筠, 刘学焕, 等. 多方位 MRI 影像组学预测局部进展期直肠癌新辅助放化疗疗效 [J]. *放射学实践*, 2024, 39(2): 218-226.

[8] Miranda J, Horvat N, Araujo-Filho JAB, et al. The Role of Radiomics in Rectal Cancer [J]. *J Gastrointest Cancer*, 2023, 54(4): 1158-1180.

[9] 国家卫生健康委员会医政司中华医学会肿瘤学分会. 国家卫健委中国结直肠癌诊疗规范 (2023 版) [J]. *中国实用外科杂志*, 2023, 43 (6): 602-630.

[10] 中华医学会放射学分会医学影像大数据与人工智能工作委员会, 中华医学会放射学分会腹部学组, 中华医学会放射学分会磁共振学组. 结直肠癌 CT 和 MRI 标注专家共识 (2020) [J]. *中华放射学杂志*, 2021, 55(2): 111-116.

[11] 中国直肠癌新辅助治疗后等待观察数据库研究协作组, 中华医学会外科学分会结直肠外科学组, 中国医师协会结直肠肿瘤医师专业委员会, 等. 直肠癌新辅助治疗后等待观察策略中国专家共识 (2024 版) [J]. *中华胃肠外科杂志*, 2024, 27(4): 301-305.

[12] 梁程, 安小霞, 李瑞, 等. MRI 影像组学预测局部进展期直肠癌新辅助放化疗疗效的研究进展 [J]. *磁共振成像*, 2024, 15(1): 224-228.

[13] Antunes JT, Ofshteyn A, Bera K, et al. Radiomic Features of Primary Rectal Cancers on Baseline T₂-Weighted MRI Are Associated With Pathologic Complete Response to Neoadjuvant Chemoradiation: A Multisite Study [J]. *J Magn Reson Imaging*, 2020, 52(5): 1531-1541.

[14] Cui Y, Yang X, Shi Z, et al. Radiomics analysis of multiparametric MRI for prediction of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer [J]. *Eur Radiol*, 2019, 29(3): 1211-1220.

[15] 秦思源, 陆思懿, 王奇政, 等. 不同 MRI 影像组学方法预测直肠癌新辅助治疗病理完全缓解的价值比较 [J]. *磁共振成像*, 2022, 13(11): 82-87.

[16] Rasmussen CE. Gaussian processes in machine learning [M]// Bousquet O, Vonluxburg U, Ratsch G. *Machine Learning Summer School Conference*, 2003, 2004: 63-71.

[17] Seeger M. Gaussian processes for machine learning [J]. *Int J Neural Syst*, 2004, 14(2): 69-106.

[18] Kapoor A, Grauman K, Urtasun R, et al. Gaussian Processes for Object Categorization [J]. *Int J Comput Vis*, 2010, 88 (2): 169-188.

[19] Deist TM, Dankers FJWM, Valdes G, et al. Erratum: "Machine learning algorithms for outcome prediction in (chemo)radiotherapy: An empirical comparison of classifiers" [J]. *Med Phys*, 2018, 45(7): 3449-3459.

[20] Sharma K, Cerezo M, Holmes Z, et al. Reformulation of the No-Free-Lunch Theorem for Entangled Datasets [J]. *Phys Rev Lett*, 2022, 128(7): 070501.