

本文引文格式:刘雪成,吴树剑,程晓娇,等.基于瘤体及瘤周CT组学特征的机器学习模型对直肠癌T2/T3分期的鉴别价值研究[J].右江民族医学院学报,2026,48(1):67-74.

【论著与临床报道】

基于瘤体及瘤周CT组学特征的机器学习模型 对直肠癌T2/T3分期的鉴别价值研究

刘雪成¹,吴树剑¹,程晓娇²,程正武²,赵宇¹

(1. 皖南医学院第一附属医院放射科,安徽 芜湖 241000;

2. 皖南医学院第一附属医院胃肠外科,安徽 芜湖 241000)

摘要:目的 探讨基于临床资料、增强CT常规特征、瘤体及瘤周组学特征的机器学习模型对鉴别直肠癌T2、T3分期的应用价值。方法 回顾性收集2019年12月至2022年2月期间在皖南医学院第一附属医院经病理组织学确诊的276例直肠癌病人的临床及CT影像资料,按照7:3比例随机分为训练集(193例)和验证集(83例)。使用MaZda软件在增强CT静脉期图像上提取瘤体及瘤周脂肪的影像组学特征,并用Logistic回归分析筛选T2、T3分期的影响因素。采用支持向量机(SVM)、弹性网络(EN)、朴素贝叶斯(NB)、决策树(DT)、逻辑回归(LR)5种机器学习方法建立预测模型,使用受试者工作特征曲线(ROC)分析模型的区分能力,并使用决策曲线分析(DCA)评价模型的临床实用价值。结果 瘤体及瘤周脂肪均获得6个最优CT组学特征。多因素Logistic回归分析显示,CEA、CA19-9、PLR、分化程度、肠壁毛糙、瘤体及瘤周脂肪组学标签得分为鉴别直肠癌T2、T3期分期的独立预测因素($P < 0.05$)。5种机器学习预测模型的训练集AUC分别为0.897、0.895、0.853、0.880、0.899,验证集AUC分别为0.896、0.866、0.862、0.829、0.839;SVM为最优模型。决策曲线显示:5种机器学习模型在0.1~0.9阈值概率时有较高的临床适应性。结论 基于临床资料、增强CT静脉期瘤体和瘤周影像组学特征所构建的机器学习预测模型,在术前鉴别直肠癌T2与T3分期中展现出良好的诊断性能,可为临床治疗决策的精准制定提供客观依据。

关键词:直肠肿瘤;影像组学;机器学习;TNM分期

中图分类号:R735.37

文献标识码:A

文章编号:1001-5817(2026)01-0067-08

doi:10.3969/j.issn.1001-5817.2026.01.011

Study on the discriminative value of a machine learning model based on tumor and peritumoral CT radiomics features for T2/T3 staging in rectal cancer

LIU Xuecheng¹, WU Shujian¹, CHENG Xiaojiao², CHENG Zhengwu², ZHAO Yu¹

(1. Department of Radiology, the First Affiliated Hospital of Wannan Medical College, Wuhu 241000, Anhui, China; 2. Department of Gastrointestinal Surgery, the First Affiliated Hospital of Wannan Medical College, Wuhu 241000, Anhui, China)

Abstract: **Objective** To explore the application value of a machine learning model incorporating clinical data, conventional contrast-enhanced CT features, and tumor/peritumoral radiomic features in differentiating T2 and T3 stages of rectal cancer. **Methods** A retrospective collection of clinical and CT imaging data was performed on 276 pathologically confirmed rectal cancer patients from the First Affiliated Hospital of Wannan Medical College between December 2019 and February 2022. Patients were randomly allocated into training (n

基金项目:安徽省教育厅重点项目(2022AH051248);皖南医学院中青年科研基金项目(WK2023ZQNZ53)

第一作者:刘雪成,主管技师,研究方向:CT及MR功能扫描,E-mail:1135028790@qq.com

通讯作者:赵宇,主治医师,研究方向:腹部影像诊断,E-mail:273523789@qq.com

=193) and validation ($n=83$) sets at a 7:3 ratio. Radiomic features were extracted from venous-phase enhanced CT images of both the tumor and peritumoral fat regions using MaZda software. Logistic regression (LR) analysis was then employed to identify significant influencing factors for distinguishing between T2 and T3 stages. Five machine learning algorithms, including support vector machine (SVM), elastic net (EN), naive bayes (NB), decision tree (DT), and LR, were employed to develop prediction models. Model performance was assessed using receiver operating characteristic (ROC) curves, with clinical utility evaluated through decision curve analysis (DCA). **Results** Six optimal CT radiomic features were selected for both tumor and peritumoral fat regions. Multivariate logistic regression (MLR) analysis revealed that CEA, CA19-9, PLR, differentiation grade, intestinal wall irregularity, and the radiomic signature scores from both the tumor and peritumoral fat were independent predictive factors for distinguishing between T2 and T3 stages in rectal cancer ($P < 0.05$). The training set AUC values were 0.897, 0.895, 0.853, 0.880, and 0.899 for the five models respectively, with corresponding validation set AUCs of 0.896, 0.866, 0.862, 0.829, and 0.839. SVM demonstrated superior performance. DCA showed that all five machine learning models had high clinical utility at threshold probabilities ranging from 0.1 to 0.9. **Conclusion** The machine learning prediction model integrating clinical data and venous-phase CT radiomic features of both tumor and peritumoral regions demonstrates favorable diagnostic performance for preoperative differentiation of T2/T3 rectal cancer stages, providing valuable objective evidence for precise clinical decision-making.

Key words: rectal cancer; radiomics; machine learning; TNM staging

直肠癌是全球常见的恶性肿瘤之一,2020 年新发病例约 73 万例,死亡约 34 万例,分别占癌症发病与死亡总数的 3.8% 和 3.4%^[1]。在临床治疗中,直肠癌的外科手术策略高度依赖 TNM 分期。早期(T1~2N0M0)患者通常可接受局部切除,而局部进展期(T3~4 或 N+)患者则推荐新辅助放化疗后再行手术^[2]。然而,由于肿瘤周围常伴随促纤维增生反应及纤维化改变,术前影像学对 T2 与 T3 期的鉴别存在困难,易导致分期偏差^[3]。若将 T2 期过度分期为 T3 期,可能使患者承受不必要的放化疗相关风险;反之,若将 T3 期分期不足,则可能因忽视淋巴结转移或脉管侵犯而增加术后复发概率^[4]。因此,实现术前 T2、T3 分期的精准判别对个体化治疗决策具有重要意义^[5]。

高分辨率 MRI 是当前直肠癌术前 T 分期的标准影像学手段,但其应用存在一定局限,包括扫描禁忌证、耗时较长及费用高昂。近年多中心研究表明,增强 CT 在快速全面评估原发灶及远处转移方面具有独特优势,对部分无法接受 MRI 检查的患者尤为适用^[6-7]。随着影像组学与机器学习技术的结合,为无创分期提供了新的研究方向。已有研究证实,基于 CT 影像组学特征构建的预测模型在直肠癌分期中表现出良好泛化能力^[8-9]。然而,现有研究多聚焦于瘤内特征,对瘤周微环境中反映肿瘤突破固有肌层这一关键生物学过程的信息关注不足^[10-11]。值得注意的是,肿瘤的侵袭性表现均与瘤周区域病理改变密切相关。近期研究表明,基于 CT 的瘤周影像组学特征可有效预测结直肠癌淋巴结转移^[12]、癌旁肿瘤沉积^[13]及周围神经侵

犯^[14],且瘤周与瘤内特征的联合模型均显示出优于单一特征模型的性能。这些证据共同表明,瘤周区域蕴含超越瘤体本身的深层信息,对揭示肿瘤局部侵袭行为具有重要价值。

为此,本研究拟基于增强 CT 静脉期影像,提取瘤体及瘤周脂肪区域的影像组学特征,结合常规影像学特征与临床指标,构建支持向量机(support vector machine,SVM)、弹性网络(elastic net,EN)、朴素贝叶斯(naive bayes,NB)、决策树(decision tree,DT)和逻辑回归(logistic regression,LR)5 种机器学习模型,评估其在鉴别直肠癌 T2 与 T3 分期中的效能,以期为临床术前精准分期提供可靠工具。

1 资料与方法

1.1 研究对象 本研究为回顾性研究,收集 2019 年 12 月至 2022 年 2 月于皖南医学院第一附属医院就诊的 276 例直肠癌患者的临床及 CT 影像学资料。病例纳入标准:①术前 2 周内行增强 CT 检查;②临床资料及实验室检查结果完整;③经术后病理证实为 T2 或 T3 期直肠癌。排除标准:①CT 图像质量不佳,影响评估;②有其他恶性肿瘤或既往抗肿瘤治疗史;③术后病理分期非 T2 或 T3 期。最终入组的 276 例患者中,男性 171 例(62.00%),女性 105 例(38.04%),平均年龄为(64.4±10.5)岁(范围 35~85 岁)。采用简单随机化方法,按 7:3 比例将患者分为训练集(193 例)与验证集(83 例)。训练集:T2 期 139 例、T3 期 54 例;验证集:T2 期 61 例、T3 期 22 例。本研究经皖南医学院第一附属医院伦理委员会审批通过[(2022)伦审新

第(05)号]。病理分期由两位高年资病理科医师参照美国癌症联合委员会(AJCC)第 8 版结直肠癌 TNM 分期标准^[15]独立判定。分期定义如下:T2 期指肿瘤浸润超过黏膜下层,侵犯但未穿透固有肌层;T3 期指肿瘤穿透固有肌层,侵及浆膜下层或无腹膜覆盖的直肠周围组织。

1.2 仪器与方法 所有检查均采用 Siemens Somatom Definition Flash 双源 CT 完成。患者检查前需空腹 4 h,扫描时取仰卧位,头先进。扫描范围覆盖全腹部至盆腔。扫描与重建参数:管电压 120 kV,管电流 250 mA,探测器准直 128 mm×0.6 mm;图像重建采用 B30f(medium-smooth)卷积核,矩阵 512×512,视野 350 mm²×350 mm²,像素间距 0.68 mm×0.68 mm,层厚与层间距均为 5 mm。在特征提取前,对所有图像进行标准化预处理,包括统一设定窗宽 300 HU、窗位 45 HU,并进行灰度离散化(灰度级距=25)。经肘静脉以高压注射器团注非离子型对比剂碘佛醇(浓度 350 g/L),流率 2.5 mL/s,总剂量 85 mL。采用三期动态增强扫描模式:动脉期扫描时间为注射对比剂后 25~30 s、门静脉期扫描时间为注射对比剂后 55~60 s、延迟期扫描时间为注射对比剂后 160~180 s。

1.3 临床资料 基本临床资料包括性别、年龄、病理分期、肿瘤分化程度以及组织学类型。实验室指标:癌胚抗原(carcino-embryonic antigen,CEA),>5μg/L 判定为异常;糖类抗原 19-9(carbohydrate antigen 19-9,CA19-9),>40 ku/L 判定为异常;中性粒细胞-淋巴细胞比值(neutrophil-to-lymphocyte ratio,NLR);血小板-淋巴细胞比值(platelet-to-lymphocyte ratio,

PLR);淋巴细胞-单核细胞比值(lymphocyte-to-monocyte ratio,LMR)。

1.4 常规 CT 特征评估 由两位具有 10 年以上腹部影像诊断经验的放射科医师,在不知晓临床及病理结果的情况下,对所有 CT 图像进行独立盲法评估。评估的常规影像学特征包括:①肿瘤内部是否存在坏死灶;②瘤周脂肪间隙是否模糊;③肠壁边缘形态是否毛糙。若两位医师的评估结果存在分歧,则通过共同讨论达成一致诊断意见。

1.5 图像分割及特征筛选 将 276 例直肠癌患者的静脉期增强 CT 图像从 PACS 系统导出为 JPG 格式,并导入 MaZda 纹理分析软件进行预处理(包括灰度归一化)。由以上两位医师,在不知晓患者临床及病理信息的情况下,分别在瘤体及瘤周脂肪的最大横截面上独立手动勾画感兴趣区(region of interest,ROI)(见图 1)。随后,使用 MaZda 软件提取定量组学特征:①形态学特征;②一阶统计特征;③灰度共生矩阵特征;④灰度游程矩阵特征;⑤小波变换特征。使用组内相关系数(Intraclass Correlation Coefficient,ICC)评估所提取特征的一致性,保留 ICC>0.75 的稳定性特征;其次,采用最小绝对收缩和选择算子(LASSO)回归进行特征选择(在线性回归的损失函数中施加 L1 范数惩罚项,驱使对目标变量贡献较小的特征系数收缩至零,从而实现特征空间的自动稀疏化与降维),最终筛选出最优特征子集。基于筛选结果,构建影像组学评分(Radiomic score,R-score),公式为:最优特征及其权重系数的乘积与截距值相加之和,其中截距值表示所有特征取零时的基准评分。



注:A 为基于静脉期;B 为勾画瘤体 ROI;C 为沿瘤体外侧 2~3 mm 勾画瘤周脂肪 ROI。

图 1 瘤体及瘤周脂肪区域 ROI 勾画示意图

1.6 统计学方法 本研究采用 SPSS 25.0 和 R 软件(4.1.2 版)进行统计分析。符合正态分布的计量资料以($\bar{x} \pm s$)表示,组间比较采用 t 检验;不符合以中位数(四分位数间距)表示,组间比较采用 Mann-Whitney U 检验。计数资料以例数(n)和百分比(%)表示,组间比较采用 χ^2 检验。在训练集中采用单因素及多因素

Logistic 回归分析,以筛选区分直肠癌 T2、T3 分期的独立预测因素,并构建 5 种机器学习预测模型。计算受试者工作特征(receiver operating characteristic,ROC)曲线下面积评估模型的诊断性能,并通过决策曲线分析(decision curve analysis,DCA)量化模型的临床获益。采用 Hosmer-Lemeshow 检验和 Brier 分数

评估模型校准度。采用 DeLong 检验比较模型间 AUC 值差异。以 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 基线资料 训练集与验证集患者的性别、年龄、

CEA、CA19-9、NLR、PLR、LMR、分化程度、组织学类型、瘤周脂肪模糊、肠壁毛糙、病灶坏死和 TNM 分期之间的差异均无统计学意义 ($P > 0.05$), 见表 1。

表 1 训练集与验证集临床资料和 CT 影像特征的比较

参数	训练集 ($n = 193$)	验证集 ($n = 83$)	χ^2 / z	P
性别			0.182	0.670
男	118(61.14)	53(63.86)		
女	75(38.86)	30(36.14)		
年龄/岁	64.23±10.02	64.69±11.45	0.334 ^a	0.739
CEA/($\mu\text{g} \cdot \text{L}^{-1}$)			0.579	0.447
≤ 5	98(50.78)	38(45.78)		
> 5	95(49.22)	45(54.22)		
CA19-9/($\text{ku} \cdot \text{L}^{-1}$)			1.917	0.166
≤ 40	174(90.16)	70(84.34)		
> 40	19(9.84)	13(15.66)		
NLR	2.22(1.67, 3.06)	2.07(1.62, 2.84)	0.581 ^b	0.561
PLR	128.57(91.32, 171.66)	122.22(93.68, 154.30)	0.742 ^b	0.458
LMR	4.00(3.00, 5.00)	4.20(3.22, 5.29)	0.492 ^b	0.623
分化程度			1.415	0.234
高	156(80.83)	72(86.75)		
中-低	37(19.17)	11(13.25)		
组织学类型			1.046	0.306
腺癌	181(93.78)	81(97.59)		
非腺癌	12(6.22)	2(2.41)		
瘤周脂肪模糊			0.250	0.875
是	98(50.78)	43(51.81)		
否	95(49.22)	40(48.19)		
肠壁毛糙			0.156	0.693
是	128(66.32)	53(63.86)		
否	65(33.68)	30(36.14)		
病灶坏死			0.343	0.558
有	12(6.22)	3(3.61)		
无	181(93.78)	80(96.39)		
TNM			0.063	0.802
T2	139(72.02)	61(73.49)		
T3	54(27.98)	22(26.51)		

注: a 表示符合正态分布的计量资料以 ($\bar{x} \pm s$) 表示, 组间比较采用 t 检验; b 表示不符的以中位数 (四分位数间距) 表示, 组间比较采用 Mann-Whitney U 检验。

2.2 组学标签评分 基于静脉期 CT 图像, 分别从瘤体及瘤周脂肪区域各提取 302 个影像组学特征。采用 LASSO 结合 10 折交叉验证进行特征降维与筛选, 最终瘤体区域筛选出 6 个最优特征 (3 个一阶直方图特征、2 个灰度游程矩阵特征及 1 个小波特征); 瘤周脂肪区域筛选出 6 个最优特征 (1 个一阶直方图特征、3 个灰度游程矩阵特征及 2 个小波特征) (见图 2、图 3)。基于以上最优特征分别建立瘤体及瘤周脂肪的组学标签评分 (R-score)。R-score = 截距值 + 最优特征 \times 系

数。

2.3 训练集中单因素及多因素 Logistic 分析 单因素 Logistic 回归分析显示 CEA、CA19-9、PLR、分化程度、瘤周脂肪模糊、肠壁毛糙、瘤体和瘤周脂肪 R-score 与区分 T2、T3 分期有关联 ($P < 0.05$); 机器学习模型多因素 Logistic 回归分析结果显示, CEA、CA19-9、PLR、分化程度、肠壁毛糙、瘤体和瘤周脂肪 R-score 为预测直肠癌 T2、T3 分期的独立预测因素 (见表 2)。

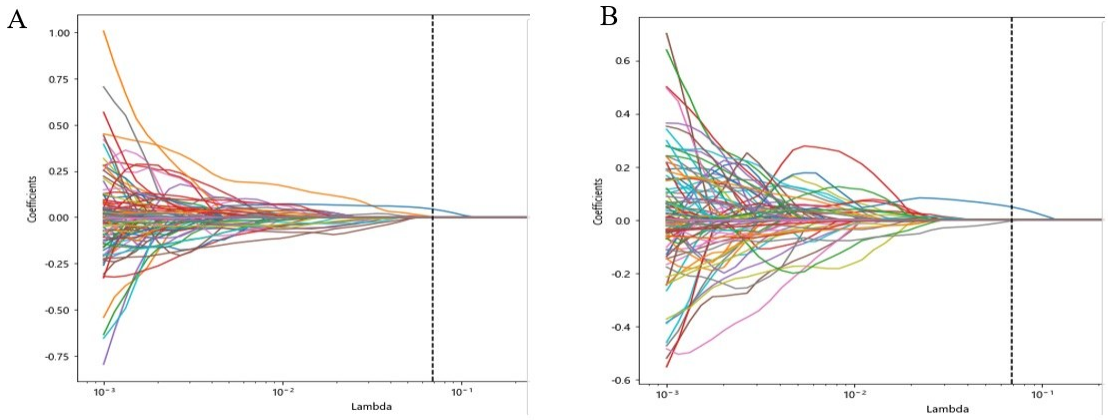


图 2 瘤体(A)及瘤周脂肪(B)特征回归系数分布

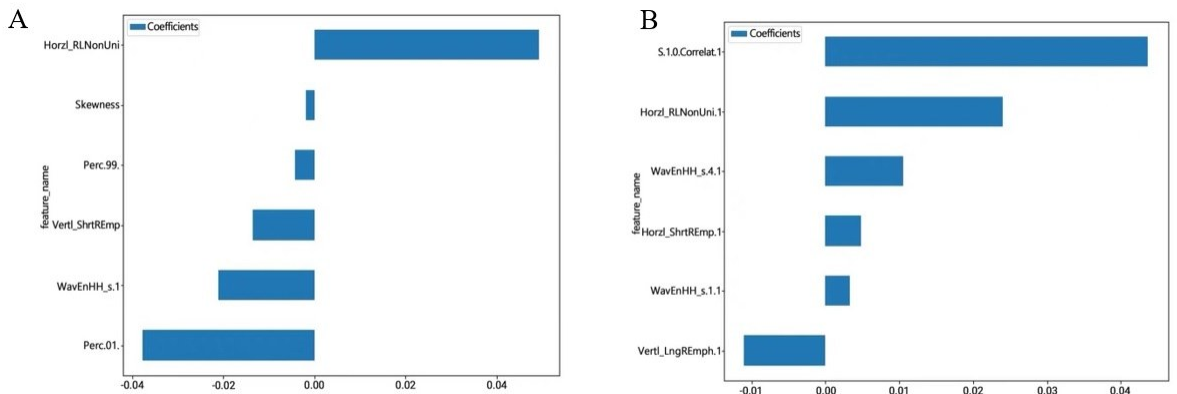


图 3 瘤体(A)及瘤周脂肪(B)最优特征及其权

表 2 训练集中单因素及多因素 Logistic 分析结果

变量	单因素 OR (95% CI)	P	多因素 OR (95% CI)	P
性别(男 vs 女)	0.719(0.372~1.390)	0.327	—	—
年龄	1.016(0.985~1.049)	0.319	—	—
CEA(>5 μg/L vs ≤5 μg/L)	3.990(1.990~8.001)	<0.001	4.085(1.628~10.247)	0.003
CA19-9(>40 ku/L vs ≤40 ku/L)	7.884(1.026~60.598)	0.047	11.132(1.122~110.470)	0.040
NLR	1.017(0.929~1.115)	0.711	—	—
PLR	1.009(1.002~1.017)	0.019	1.011(1.001~1.021)	0.040
LMR	0.949(0.819~1.100)	0.487	—	—
分化程度(中—低 vs 高)	3.892(1.037~11.584)	0.015	3.764(1.013~13.981)	0.048
组织学类型(非腺癌 vs 腺癌)	1.117(0.306~4.523)	0.813	—	—
瘤周脂肪模糊(是 vs 否)	3.854(1.942~7.646)	<0.001	1.223(0.369~4.050)	0.742
肠壁毛糙(是 vs 否)	6.693(3.359~13.335)	<0.001	7.364(2.254~24.059)	0.001
病灶坏死(有 vs 无)	4.555(0.574~36.167)	0.152	—	—
瘤体 R-score	1.690(1.257~2.239)	<0.001	2.002(1.335~3.004)	0.001
瘤周 R-score	4.967(1.838~13.420)	0.002	4.189(1.345~13.054)	0.013

2.4 模型构建与验证 根据单因素及多因素 Logistic 分析结果,基于 CEA、CA19-9、PLR、分化程度、肠壁毛糙构建临床预测模型;联合瘤体与瘤周 R-score 构建 5 种机器学习模型。在训练集和验证集中,SVM、EN、NB、DT、LR 5 种机器学习模型的训练集 AUC 分别为 0.897、0.895、0.853、0.880、0.899,验证集 AUC 分别

为 0.896、0.866、0.862、0.829、0.839,详见表 3 和图 4。DCA 曲线进一步表明所有机器学习模型在 0.1~0.9 的阈值范围内均具备良好的临床适用性(见图 5)。在验证集中,LR 模型的 Hosmer-Lemeshow 检验显示无差异($\chi^2 = 2.592, P = 0.274$),表明校准良好;其他机器学习模型,通过 Brier 分数评估其校准误差

(SVM:0.121, EN:0.125, NB:0.164, DT:0.120), 结果显示所有模型的 Brier 分数均低于 0.164, 预测概率准确性较高。

DeLong 检验显示:在训练集中, SVM、EN 和 LR 模型的 AUC 值均显著优于临床模型 ($Z = 3.119, 2.906, 2.986, P = 0.001, 0.003, 0.002$), 而 NB 与 DT 模型与临床模型无统计学意义 ($Z = 0.491, 1.131, P$

$= 0.623, 0.257$); 在验证集中, 仅 SVM 模型仍保持显著优势 ($Z = 2.111, P = 0.027$), 其余模型 (EN、NB、DT、LR) 与临床模型的差异均无统计学意义 ($Z = 1.188, 1.802, 0.552, 0.871, P = 0.234, 0.071, 0.580, 0.383$)。综合模型在训练集与验证集中的表现, SVM 确定为最优预测模型。

表 3 各模型在训练集与验证集的诊断效能

模型	AUC(95% CI)	特异度	敏感度	准确度	约登指数
训练集					
临床模型	0.839(0.781~0.897)	0.685	0.834	0.792	0.519
SVM	0.897(0.855~0.940)	0.925	0.748	0.797	0.673
EN	0.895(0.852~0.938)	0.926	0.748	0.798	0.674
NB	0.853(0.793~0.912)	0.833	0.763	0.782	0.596
DT	0.880(0.824~0.937)	0.870	0.755	0.788	0.625
LR	0.899(0.857~0.941)	0.944	0.748	0.803	0.692
验证集					
临床模型	0.783(0.672~0.896)	0.909	0.590	0.674	0.499
SVM	0.896(0.825~0.966)	0.864	0.836	0.843	0.700
EN	0.866(0.768~0.963)	0.954	0.738	0.795	0.692
NB	0.862(0.776~0.948)	0.727	0.902	0.855	0.629
DT	0.829(0.729~0.930)	0.773	0.869	0.843	0.642
LR	0.839(0.750~0.928)	0.863	0.721	0.759	0.584

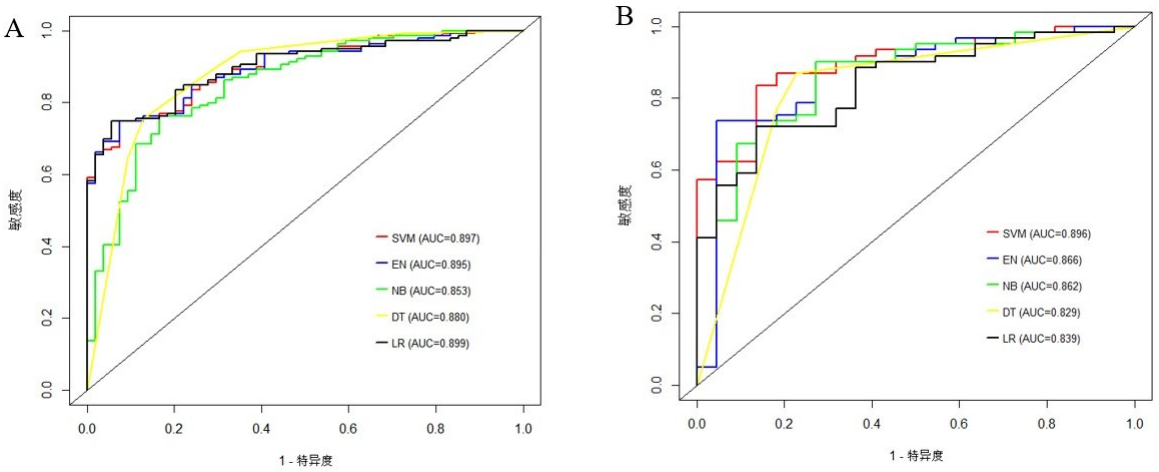


图 4 5 种机器学习模型在训练集(A)及验证集(B)的 ROC 曲线

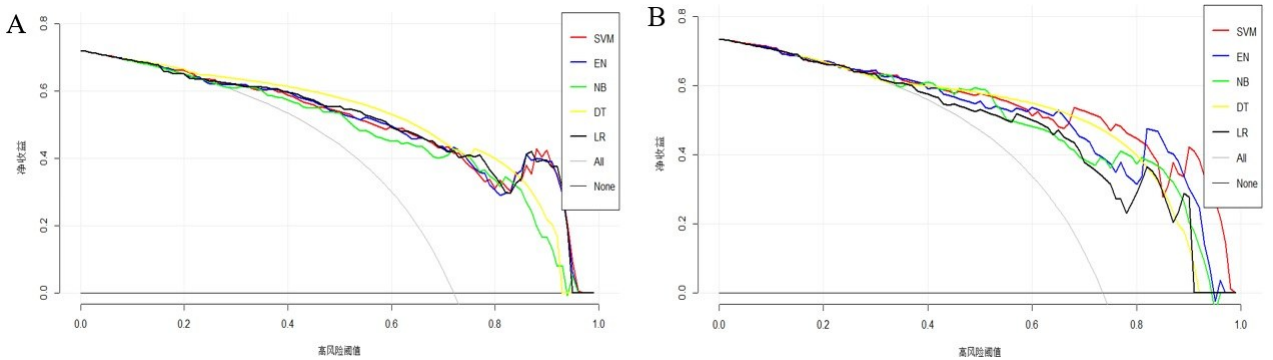


图 5 5 种机器学习模型在训练集(A)及验证集(B)的决策曲线

3 讨论

本研究通过整合瘤体及瘤周脂肪区域的 CT 影像组学特征,联合常规影像特征与临床资料,成功构建了 SVM、EN、NB、DT 和 LR 5 种机器学习算法的直肠癌 T2、T3 分期的预测模型。验证结果表明,所有模型在训练集与验证集中均表现出稳定的鉴别能力,AUC 值均不低于 0.829。其中,SVM 模型在验证集中表现尤为突出,其 AUC 达 0.896,较仅基于临床及常规特征的临床模型(AUC:0.783)提升 0.113,且 Delong 检验证实该提升具有统计学意义($P=0.027$)。这一结果充分证明了多模态特征融合与机器学习方法在提升直肠癌术前 T 分期准确性方面的应用价值与临床潜力。

王进等^[16]基于高分辨率 T2WI 的 SVM 模型判断直肠癌术前 T 分期的 AUC 为 0.730,其诊断效能低于本研究,这可能与其仅依据常规影像特征及临床资料而未纳入组学特征信息相关。MA X L 等^[17]采用 6 种机器学习方法,其中 SVM 模型的 AUC 为 0.862,略低于本研究的 SVM 表现(AUC:0.896)。这一差异可能源于本研究不仅提取了瘤体区域特征,还整合了瘤周脂肪区域的影像组学信息,提示瘤周微环境的变化对判断肿瘤是否穿透固有肌层具有重要病理学意义。具体而言,T3 期肿瘤穿透固有肌层后常侵犯周围脂肪组织,引起局部炎症、纤维化及组织结构紊乱,这些病理改变可在影像组学特征中得以体现。需要说明的是,本研究在瘤周脂肪特征提取中采用二维单层面勾画,且图像层厚为 5 mm,可能对早期细微浸润变化的捕捉存在一定局限。此外,本研究选择增强 CT 静脉期图像进行组学分析,是因为静脉期肿瘤实质强化达到峰值且趋于均匀,能够最佳地显示肿瘤与周围正常组织的对比差异,从而可能提取出更稳定、更有鉴别力的特征^[18]。

本研究从瘤体与瘤周脂肪区域各提取 302 个影像组学特征,经 LASSO 降维后分别筛选出 6 个关键特征。瘤体区域特征:3 个一阶直方图特征、2 个灰度游程矩阵特征和 1 个小波变换特征;瘤周脂肪区域:1 个一阶直方图特征、3 个灰度游程矩阵特征和 2 个小波变换特征。从病理生理机制角度分析,这些特征与肿瘤生物学行为密切相关:一阶直方图反映组织密度分布,T3 期肿瘤因浸润性生长表现出更高的异质性;灰度游程矩阵特征表征图像纹理的均匀性与复杂性^[19],T3 期因组织成分混杂而呈现更不均匀的灰度分布;小波变换特征捕捉多尺度空间结构信息^[20],T3 期因边界模糊及周围组织浸润而显示出更复杂的频率特征。邓波等^[21]基于直肠癌系膜脂肪影像组学构建的 SVM 模型验证集 AUC 为 0.811,低于本研究的 0.896,这一差异进一步证实联合瘤体与瘤周特征能够更全面地反

映 T3 期肿瘤突破固有肌层后的病理改变,从而提高分期的准确性。

在 5 种机器学习模型中,SVM 表现最优且最为稳定,这可能得益于其处理高维特征空间的优势及较强的泛化能力。EN 和 LR 在训练集中表现与 SVM 接近(AUC 分别为 0.895 和 0.899),但在验证集中 AUC 下降较为明显(分别降至 0.866 和 0.839),提示存在一定过拟合风险,未来研究可通过扩大样本量或增强正则化策略予以改进。NB 模型虽结构简单,但表现出良好的稳健性(训练集与验证集 AUC 分别为 0.853 和 0.862),适合临床快速筛查场景;DT 模型虽具有良好的可解释性,但验证集 AUC 相对较低(0.829),可能受限于其对复杂非线性关系及特征交互作用的捕捉能力。

常规 CT 特征分析显示,T3 期患者肠壁外缘毛糙比例显著高于 T2 期,且为独立预测因素,可能与 T3 期肿瘤穿透肌层后浸润周围脂肪、引发纤维增生及条索状高密度影相关^[22]。该结果与姜军等^[23]报道存在差异,可能源于其研究中 T2 期样本量偏少(14 例)所导致的统计偏差。临床资料分析进一步确定了 PLR 与 CEA 为独立预测因素,PLR 作为全身炎症反应的标志物,其升高可能反映了肿瘤侵袭所伴随的局部及系统性免疫微环境改变;而 CEA 水平则直接关联于肿瘤负荷。二者联合为术前无创评估肿瘤侵袭程度提供了互补信息^[24]。此外,肿瘤分化程度也被证实与分期显著相关,中一低分化肿瘤因其更强的生物学侵袭性,更易出现固有肌层的穿透^[25],这一发现对术前治疗策略的制定具有重要提示价值。

本研究存在以下局限性:首先,仅采用增强 CT 静脉期图像进行分析,未能评估其他扫描时相的潜在价值;其次,采用二维 ROI 勾画方式可能无法全面反映肿瘤异质性;第三,数据集中 T2 与 T3 期样本分布不均衡且总量有限,可能影响模型稳定性;此外,缺乏外部验证队列和标准化图像预处理流程也限制了结果的普适性,后续研究将开展三维全瘤分割、标准化图像预处理,并通过多中心合作进行外部验证,以提升模型的可靠性与临床转化价值。

综上所述,融合临床指标、常规 CT 特征及瘤体一瘤周影像组学特征的机器学习模型可显著提升直肠癌术前 T2、T3 分期鉴别效能,其中 SVM 模型表现最佳,有望为个体化治疗决策提供可靠依据。

参考文献:

- [1] SUNG H, FERLAY J, SIEGEL R L, 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] BENSON A B, VENOOK A P, AL-HAWARY M M, 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] 胡飞翔, 岳亚丽, 彭卫军, 等. DWI 联合 T2WI 在鉴别 T2 和 T3 期直肠癌术前分期中的应用价值 [J]. 放射学实践, 2021, 36(4): 507-513.
- [4] 国家卫生健康委员会医政司, 中华医学会肿瘤学分会, 顾晋, 等. 国家卫健委中国结直肠癌诊疗规范(2023 版) [J]. 2023(6): 602-630.
- [5] DAHLBÄCK C, KORSBAKKE K, ALSHIBIBY BERGMAN T, et al. Accuracy of magnetic resonance imaging staging of tumour and nodal stage in rectal cancer treated by primary surgery: a population-based study [J]. Colorectal Dis, 2022, 24(9): 1047-1053.
- [6] 黄必贵, 农欣欣, 邹清艺, 等. 影像组学联合临床特征对非小细胞肺癌 EGFR 突变的预测价值 [J]. 右江民族医学院学报, 2024, 46(2): 192-199.
- [7] SUN Z H. Cardiovascular computed tomography in cardiovascular disease: an overview of its applications from diagnosis to prediction [J]. J Geriatr Cardiol, 2024, 21(5): 550-576.
- [8] BELLINI D, CARBONE I, RENGO M, et al. Performance of machine learning and texture analysis for predicting response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer with 3T MRI [J]. Tomography, 2022, 8(4): 2059-2072.
- [9] FENG Y R, GONG J, HU T D, et al. Radiomics for predicting survival in patients with locally advanced rectal cancer: a systematic review and meta-analysis [J]. Quant Imaging Med Surg, 2023, 13(12): 8395-8412.
- [10] LIU Z Y, ZHANG J S, WANG H X, et al. MRI-based radiomics feature combined with tumor markers to predict TN staging of rectal cancer [J]. J Rob Surg, 2024, 18(1): 229.
- [11] ZHU W L, XU Y S, ZHU H L, et al. Multi-parametric MRI radiomics for predicting response to neoadjuvant therapy in patients with locally advanced rectal cancer [J]. Jpn J Radiol, 2024, 42(12): 1448-1457.
- [12] 董丽, 公绪涛, 李靖, 等. 直肠癌瘤内及瘤周 CT 影像组学联合临床特征对淋巴结转移的预测 [J]. 锦州医科大学学报, 2025, 46(4): 54-59.
- [13] 刘燕, 罗锦文, 刘艳丽, 等. CT 增强影像组学模型对结直肠癌旁肿瘤沉积的诊断价值 [J]. 分子影像学杂志, 2025, 48(4): 484-491.
- [14] 岳思宇, 储伟, 童鹏飞, 等. 瘤内及瘤周双能 CT 影像组学联合临床参数预测直肠癌周围神经侵犯 [J]. 影像诊断与介入放射学, 2025, 34(3): 178-183.
- [15] VANDERWALDE N, GROTHEY A. Personalizing treatment for rectal cancer: total neoadjuvant therapy is leading the way [J]. JAMA Netw Open, 2020, 3(12): e2030508.
- [16] 王进, 李智慧, 沈浮, 等. 基于高分辨 T2WI 的影像组学对直肠癌术前分期的应用价值 [J]. 放射学实践, 2019, 34(11): 1251-1254.
- [17] MA X L, SHEN F, JIA Y, et al. MRI-based radiomics of rectal cancer: preoperative assessment of the pathological features [J]. BMC Med Imaging, 2019, 19(1): 86.
- [18] MACHADO L, ALBERGE L, PHILIPPE H, et al. A promptable CT foundation model for solid tumor evaluation [J]. NPJ Precis Oncol, 2025, 9(1): 121.
- [19] 吴树剑, 俞咏梅, 范莉芳, 等. MSCT 影像组学结合机器学习预测直径 2~5 cm 胃胃肠间质瘤危险度分级研究 [J]. 中国实用外科杂志, 2022, 42(12): 1401-1407.
- [20] 董诚, 李健, 郑迎梅, 等. 基于 MRI 征象及影像组学的列线图预测腮腺良性与恶性肿瘤的价值 [J]. 2022(2): 149-155.
- [21] 邓波, 杨严伟, 刘原庆, 等. 基于直肠癌系膜脂肪 MRI 影像组学构建的机器学习模型术前鉴别 T2、T3 期直肠癌 [J]. 现代肿瘤医学, 2023, 31(20): 3822-3827.
- [22] XU L P, ZHANG Z Y, QIN Q, et al. Assessment of T and N staging with MRI(3)T in lower and middle rectal cancer and impact on clinical strategy [J]. J Int Med Res, 2020, 48(6): 300060520928685.
- [23] 姜军, 李颖, 蒋力明, 等. 64 排螺旋 CT 中“管壁僵硬”和“外膜毛糙”征象对结直肠癌术前分期的价值 [J]. 中国介入影像与治疗学, 2010, 7(1): 5-9.
- [24] CAI Z R, HUANG L Y, CHEN Y F, et al. CEA decline predicts tumor regression and prognosis in locally advanced rectal cancer patients with elevated baseline CEA [J]. J Cancer, 2020, 11(22): 6565-6570.
- [25] 金兴兴, 王帅娜, 王红霞, 等. 磁共振分数阶微积分扩散成像在直肠癌分化程度评估中的价值 [J]. 临床放射学杂志, 2025, 44(5): 902-906.

收稿日期: 2025-11-07; 修回日期: 2025-12-09

(本文编辑 覃黎黎)