



doi:10.7659/j.issn.1005-6947.240611
http://dx.doi.org/10.7659/j.issn.1005-6947.240611
China Journal of General Surgery, 2025, 34(7):1410-1420.

· 专题研究 ·

MRI影像组学结合临床特征的机器学习模型对结直肠癌肝转移的预测价值

李波, 刘冠男

(河南省南阳市第一人民医院 磁共振诊断室, 河南 南阳 473000)

摘要

背景与目的: 结直肠癌肝转移 (CRCLM) 是影响患者预后的主要原因, 术前无创、精准诊断对制定治疗方案至关重要。传统临床标志物特异性有限, 本研究旨在基于多模态 MRI 影像组学特征, 结合机器学习算法, 构建预测 CRCLM 的高效模型, 并评价其临床价值。

方法: 收集 2022 年 5 月—2024 年 5 月于河南省南阳市第一人民医院行术前 MRI 检查并经病理证实的 150 例结直肠癌患者, 随机分为训练集 ($n=120$) 和验证集 ($n=30$)。其中 CRCLM 57 例, 无 CRCLM 93 例。采用单因素与多因素分析筛选 CRCLM 独立危险因素, 建立临床诊断模型。提取多模态 MRI 影像组学特征, 经 LASSO 筛选后分别构建 Logistic 回归 (LR)、支持向量机 (SVM)、随机森林 (RF) 模型, 并比较其诊断效能。建立临床及影像组学联合诊断模型, 并通过受试者操作特征和决策曲线 (DCA) 评估效能与临床获益。

结果: 癌胚抗原 ($OR=1.323$, 95% CI=1.079~1.567)、糖类抗原 19-9 ($OR=2.512$, 95% CI=1.225~3.799) 及中性粒细胞/淋巴细胞比值 ($OR=1.881$, 95% CI=1.354~2.409) 是 CRCLM 独立危险因素 (均 $P<0.05$), 以上 3 个因素构建的临床诊断模型曲线下面积 (AUC) 为 0.793。RF 模型在训练集与验证集 AUC 最高 (0.770、0.763), 优于 LR 和 SVM。基于 RF 的联合诊断模型在训练集与验证集 AUC 分别为 0.913 和 0.947, 明显优于单独临床或影像组学诊断模型, DCA 显示联合诊断模型具有最高临床净获益。

结论: RF 模型在影像组学预测中表现最佳, 其与临床特征结合的联合模型能显著提高 CRCLM 的无创诊断效能, 具备较高的临床应用价值。

关键词

结直肠肿瘤; 肿瘤转移; 肝; 多模态磁共振成像; 影像组学; 支持向量机; 随机森林

中图分类号: R735.3

The predictive value of MRI imaging omics combined with clinical features in machine learning models for colorectal cancer liver metastasis

LI Bo, LIU Guannan

(Department of Magnetic Resonance Imaging, Nanyang First People's Hospital, Nanyang, Henan 473000, China)

Abstract

Background and Aims: Colorectal cancer liver metastasis (CRCLM) is a major cause of poor

收稿日期: 2024-11-25; **修订日期:** 2025-07-18。

作者简介: 李波, 河南省南阳市第一人民医院副主任医师, 主要从事磁共振方面的研究。

通信作者: 刘冠男, Email: 18695985852@163.com

prognosis in patients with colorectal cancer. Accurate and noninvasive preoperative diagnosis is essential for treatment planning. Conventional clinical biomarkers have limited specificity. This study aimed to develop an efficient predictive model for CRCLM by integrating multimodal MRI imaging omics features with machine learning algorithms, and to evaluate its clinical value.

Methods: A total of 150 patients with colorectal cancer who underwent preoperative MRI and were pathologically confirmed at Nanyang First People's Hospital between May 2022 and May 2024 were retrospectively analyzed. Patients were randomly divided into a training set ($n=120$) and a validation set ($n=30$), including 57 cases with CRCLM and 93 cases without. Univariate and multivariate analyses were performed to identify independent risk factors for CRCLM and to construct a clinical diagnostic model. Radiomics features were extracted from multimodal MRI, and the least absolute shrinkage and selection operator (LASSO) method was used for feature selection. Logistic regression (LR), support vector machine (SVM), and random forest (RF) models were built and compared for diagnostic performance. A combined clinical-imaging omics model was further established, and its performance and clinical utility were assessed using receiver operating characteristic curves and decision curve analysis (DCA).

Results: Carcinoembryonic antigen ($OR=1.323$, 95% $CI=1.079\text{--}1.567$), carbohydrate antigen 19-9 ($OR=2.512$, 95% $CI=1.225\text{--}3.799$), and neutrophil-to-lymphocyte ratio ($OR=1.881$, 95% $CI=1.354\text{--}2.409$) were identified as independent risk factors for CRCLM (all $P<0.05$). The clinical model constructed with these three factors achieved an AUC of 0.793. Among radiomics models, the RF model demonstrated the highest AUC in both training and validation sets (0.770 and 0.763), outperforming LR and SVM. The combined RF-based model yielded AUC of 0.913 and 0.947 in the training and validation sets, respectively, significantly exceeding the performance of the clinical or imaging omics models alone. DCA confirmed the superior net clinical benefit of the combined model.

Conclusion: The RF model showed the best diagnostic performance among imaging omics models. When integrated with clinical features, the combined RF model significantly improved the noninvasive diagnostic efficacy of CRCLM and demonstrated high potential for clinical application.

Key words

Colorectal Neoplasms; Neoplasm Metastasis; Liver; Multimodal Magnetic Resonance Imaging; Imaging Omics; Support Vector Machine; Random Forests

CLC number: R735.3

尽管早期结直肠癌 (colorectal cancer, CRC) 治疗效果较好, 但仍有 30% 的患者在确诊时已发生结直肠癌肝转移 (colorectal cancer liver metastasis, CRCLM), 5 年生存率不足 12%^[1-3]。尽管无创性临床标志物预测对制定治疗方案和评估预后有帮助, 但其特异性不足^[4], 因此迫切需要一种非侵入性方法在术前准确预测 CRCLM。MRI 无辐射, 能进行多序列、多参数成像, 具有较高的软组织分辨率, 有助于评估病灶大小和位置, 是肿瘤患者术前关键的检查方法^[5]。影像组学以其非侵入性和高效性, 为影像数据中高通量特征的定量提取开辟了新途径, 被广泛结合机器学习算法用于肿瘤诊断

和预后评估^[6-7]。本研究采用多种机器学习算法开发和验证影像组学模型, 并构建基于 MRI 影像组学与临床特征的联合模型用于诊断 CRCLM, 评估其临床应用潜力。

1 资料与方法

1.1 研究对象

收集 2022 年 5 月—2024 年 5 月在河南省南阳市第一人民医院术前 1 周内完成 MRI 检查的 150 例 CRC 患者, 其中 CRCLM 57 例, 无 CRCLM 93 例。纳入标准: (1) 符合 CRC 及 CRCLM^[8]相关诊断标准,

且经术后病理证实；(2)初次诊断，入组前未接受过相关治疗；(3)均进行MRI检查，且影像质量良好；(4)临床资料完整。排除标准：(1)合并其他恶性肿瘤；(2)存在心肝肾等严重功能障碍；(3)既往接受过放化疗或手术治疗；(4)MRI影像质量欠佳或临床资料缺失。按4:1比例将150例患者随机分为训练集($n=120$)和验证集($n=30$)。训练集男68例，女52例；年龄55~73 (65.25 ± 8.86)岁；左半结直肠72例，右半结肠48例；原发灶TNM分期：III期74例，IV期46例；同时性CRCLM 20例，异时性CRCLM 18例。验证集男19例，女11例；年龄54~75 (66.03 ± 9.12)岁；左半结直肠21例，右半结肠9例；原发灶TNM分期：III期19例，IV期11例；同时性CRCLM 8例，异时性CRCLM 11例。训练集与验证集一般资料比较差异无统计学意义(均 $P>0.05$)。本研究通过南阳市第一人民医院伦理委员会批准(批号：[2024-yxllz]1007-17)。

1.2 MRI检查方法

所有患者均由1名10年以上工作经验的放射科医师完成MRI检查。采用3.0T Philips Achieva MR配合腹部相控矩阵线圈完成扫描。检查前6 h患者

禁食，在扫描前进行肠道准备。上机前5~10 min注射20 mg山莨菪碱以减缓肠蠕动和肌肉张力。扫描围绕直肠兴趣区，沿病变肠管长轴垂直进行，具体参数如下：轴位T2WI序列：TR、TE、矩阵、层厚、层间距、NEX、ETL分别为5 180 ms、90 ms， 220×220 ，4 mm、0.5 mm、4、16；DWI序列：TR、TE、矩阵、层厚、层间距、b值、NEX分别为3 200 ms、55 ms， 375×300 、5 mm、2 mm、 800 s/mm^2 、8；轴位T1WI增强序列：TR、TE、矩阵、层厚分别为650 ms、20 ms， 280×340 、3 mm。增强扫描静脉团注0.1 mL/kg Gd-DTPA顺磁性对比剂。

1.3 图像处理

在Radcloud平台(北京慧影医疗科技有限公司)上，由2名经验丰富的放射科医师对患者T2WI、DWI、T1WI增强图像进行逐层手动勾画，获取3D感兴趣区容积(volume of interest, VOI)，并确保感兴趣区(region of interest, ROI)于病灶最大层面，避开周围干扰(图1-2)。审查过程中，若勾画轮廓差异超过5%，则由资深医师界定肿瘤边界。

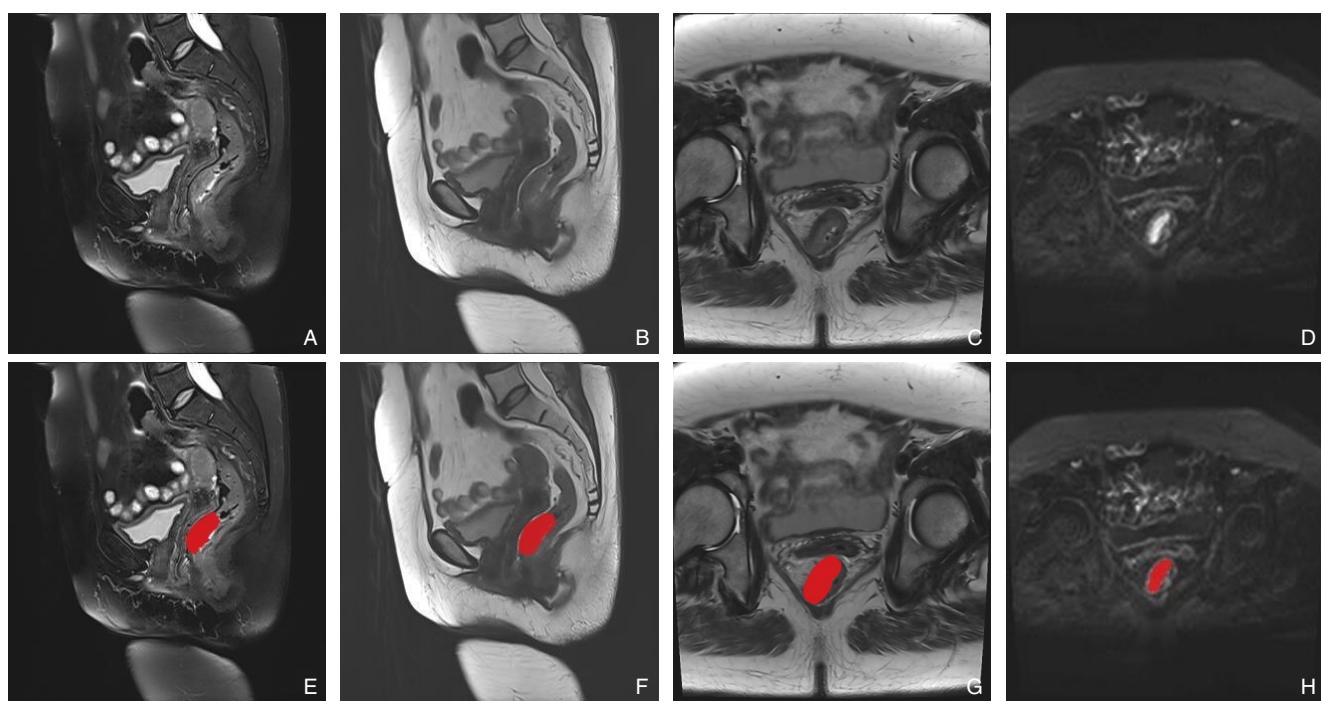


图1 CRC原发灶的多模态MRI图像及相应ROI勾画示意图
T2WI、DWI图像；E-H：分别为相应ROI勾画示意图

Figure 1 Multimodal MRI images of primary CRC lesions and corresponding ROI delineation
T2WI, sagittal T2WI, axial T2WI, and DWI images; E-H: Corresponding ROI delineation

A-D: 分别为病灶矢状位压脂T2WI、矢状位T2WI、横断位

A-D: Sagittal fat-suppressed

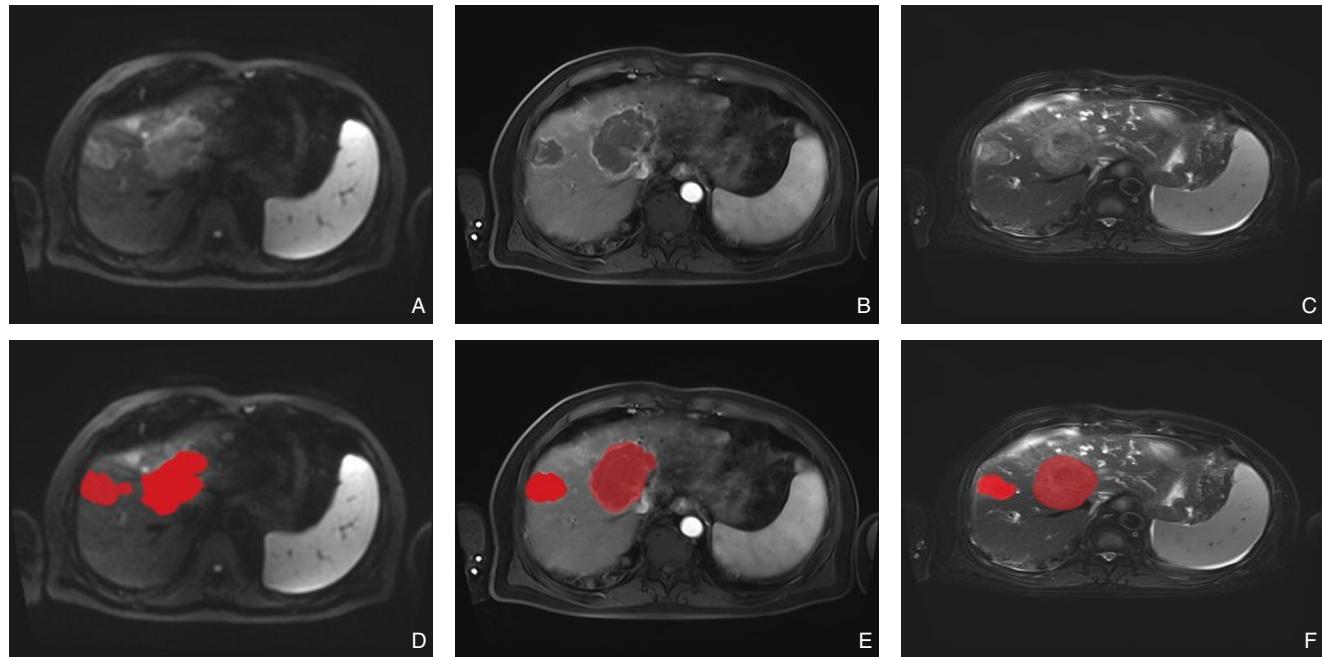


图2 CRCLM病灶的多模态MRI图像及相应ROI勾画示意图 A-C: CRCLM病灶DWI、增强横断位T1WI、横断位压脂T1WI; D-F: 相应ROI勾画示意图

Figure 2 Multimodal MRI images of CRCLM lesions and corresponding ROI delineation A-C: DWI, axial contrast-enhanced T1WI, and axial fat-suppressed T1WI images; D-F: Corresponding ROI delineation

1.4 影像组学特征筛选及模型构建

通过Radcloud平台从MRI图像中提取1 910个定量成像特征,分为三组:一阶统计特征(370个),形状和尺寸特征(40个)以及纹理特征(1 500个)。使用LASSO回归($cv=3$, $max_iter=1\ 000$)筛选出预测性最强的特征,并基于所选特征采用Logistic回归(Logistic regression, LR)、支持向量机(support vector machine, SVM)、随机森林(random forest, RF)构建预测CRCLM的影像组学模型,采用五倍交叉验证对训练集进行模型评估,并在验证组中进行验证。

1.5 统计学处理

采用SPSS 26.0分析数据。计量数据以均数 \pm 标准差($\bar{x} \pm s$)或中位数(四分位间距)[M (IQR)]表示,行 t 检验或秩和检验;计数资料以例数(百分比)[n (%)]表示,行 χ^2 检验。 $P < 0.05$ 为差异有统计学意义。绘制受试者操作特征(receiver operator characteristic, ROC)曲线,计算曲线下面积(area under curve, AUC),比较3种影像组学模型的诊断能力。采用单因素与多因素分析筛选CRCLM危险因素,据此建立临床诊断模型,并构

建临床与影像组学的联合模型。通过ROC曲线评估模型性能。采用决策曲线(decision curve analysis, DCA)评估模型的临床应用价值。

2 结果

2.1 CRCLM的影响因素分析

单因素分析结果显示,CRCLM组与无CRCLM组患者的性别、年龄、体质量指数(body mass index, BMI)、肿瘤部位、肿瘤最大径、淋巴细胞/单核细胞比值(lymphocyte to monocyte ratio, LMR)差异均无统计学意义(均 $P > 0.05$);CRCLM组患者的癌胚抗原(carcinoembryonic antigen, CEA)、糖类抗原19-9(carbohydrate antigen 19-9, CA19-9)水平及中性粒细胞/淋巴细胞比值(neutrophil to lymphocyte ratio, NLR)均高于无CRCLM组患者(均 $P < 0.05$)(表1)。多因素分析结果显示,CEA、CA19-9水平及NLR高均是CRCLM独立危险因素(均 $P < 0.05$)(表2)。据此进行临床诊断模型的构建,ROC曲线显示,CEA、CA19-9及NLR联合的临床诊断模型AUC为0.793(图3)。

表1 CRCLM影响因素的单因素分析
Table 1 Univariate analysis of factors for CRCLM

临床因素	CRCLM组(n=57)	无CRCLM组(n=93)	t/Z/χ ²	P
性别[n(%)]				
男	37(64.91)	50(53.76)	1.803	0.179
女	20(35.09)	43(46.24)		
年龄[岁,M(IQR)]	66(55~75)	68(53~73)	2.254	0.164
BMI(kg/m ² , $\bar{x} \pm s$)	22.92±3.14	23.51±2.28	1.329	0.186
肿瘤部位[n(%)]				
左半结直肠	33(57.89)	60(64.52)	0.658	0.417
右半结肠	24(42.11)	33(35.48)		
肿瘤最大径(cm, $\bar{x} \pm s$)	4.38±1.49	4.31±1.52	0.276	0.783
CEA [μg/L,M(IQR)]	5.56(2.02~23.35)	4.01(2.12~15.98)	5.527	0.012
CA19-9 [U/mL,M(IQR)]	13.56(6.63~22.49)	10.26(5.34~21.25)	6.638	0.009
NLR [M(IQR)]	2.26(1.25~3.26)	1.89(1.08~3.07)	4.498	0.024
LMR [M(IQR)]	5.08(2.45~10.02)	4.92(2.18~7.79)	2.263	0.157

表2 CRCLM影响因素的多因素分析
Table 2 Multivariate analysis of factors for CRCLM

临床因素	β	SE	Wald	P	OR(95% CI)
CEA(μg/L)	0.280	0.113	6.140	0.021	1.323(1.079~1.567)
CA19-9(U/mL)	0.921	0.252	13.357	0.009	2.512(1.225~3.799)
NLR	0.632	0.198	10.188	0.011	1.881(1.354~2.409)

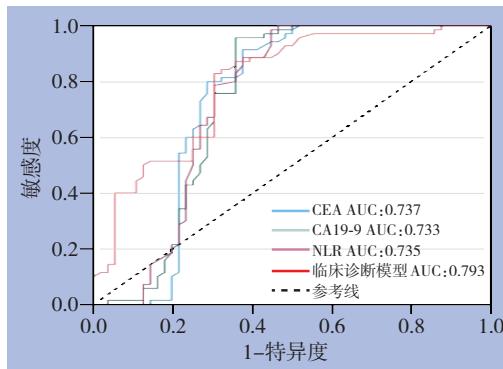


图3 CEA、CA19-9、NLR及联合预测模型的ROC曲线
Figure 3 ROC curves of CEA, CA19-9, NLR, and the combined predictive model

2.2 影像组学特征筛选

从3个序列(轴位T2WI、DWI、轴位T1WI增强序列)的ROI中共提取1 910个特征,采用包括灰度共生矩阵(gray level co-occurrence matrix, GLCM)、灰度游程长度矩阵(grey level run length matrix, GLRLM)、灰度尺寸zone矩阵(gray level size zone matrix, GLSZM)、灰度差分矩阵(grey level difference matrix, GLDM)等纹理分析算法,对图像纹理信息进行提取。通过方差阈值法筛选出1 070个,运用Select K Best法进一步选出680个,经mRMR去除冗余特征,再经LASSO回归筛选共确定10个最优特征,包括2个GLRLM、2个GLSZM、3个GLDM和3个一阶统计特征(图4)(表3)。

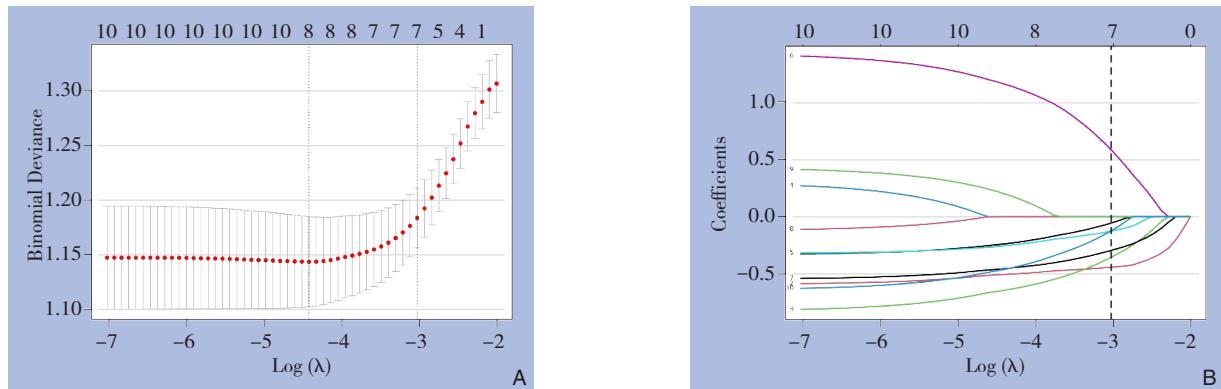


图4 LASSO回归图 A: LASSO回归交叉验证图; B: LASSO回归系数分布图
Figure 4 LASSO regression analysis A: Cross-validation curve for LASSO regression; B: Coefficient distribution map

表3 筛选出的10个最优影像组学特征及权重

Table 3 Ten selected optimal radiomics features and corresponding weights

特征名称	权重
e-THRIVE+C_exponential_grlm_RunVariance	-0.202
e-THRIVE+C_wavelet-HLL_glszm_GrayLevelNonUniformity	-0.205
T2WI_exponential_gldm_LargeDependenceEmphasis	-0.018
DWI_wavelet-HHL_gldm_DependenceVariance	-0.042
DWI_wavelet-LLH_firstrorder_Maximum	-0.056
e-THRIVE+C_square_grlm_RunVariance	0.305
T2WI_exponential_gldm_LargeDependenceLowGrayLevelEmphasis	0.058
DWI_wavelet-HLL_glszm_LngereAreaLowGrayLevelEmphasis	0.111
DWI_squareroot_firstrorder_10Percentile	0.142
e-THRIVE+C_wavelet-LHL_firstrorder_Minimum	0.051

2.3 机器学习模型构建和效能评价

基于筛选出的最优特征分别采用LR、SVM、RF构建预测CRCLM的影像组学模型,对各影像组

学模型进行ROC曲线分析,结果显示在训练集和验证集中RF模型的AUC最大,分别为0.770、0.763,故选择最佳RF模型(图5)(表4)。

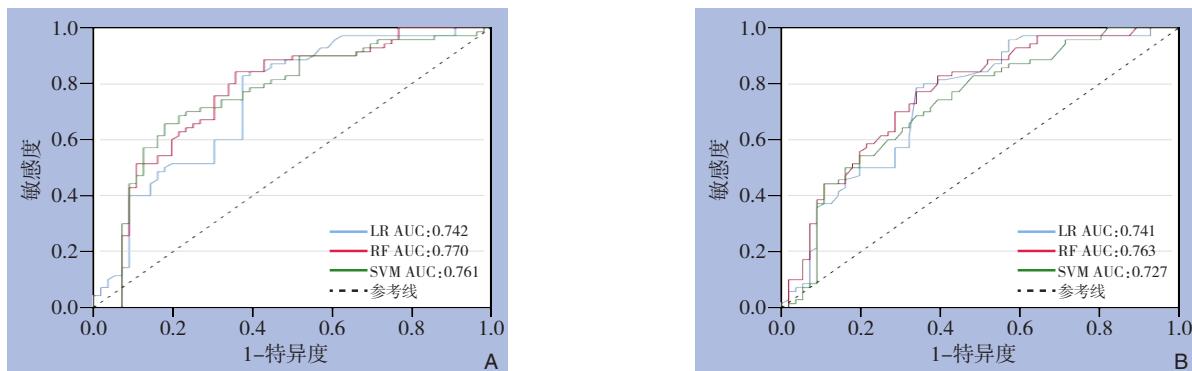


图5 LR、SVM、RF模型的ROC曲线
Figure 5 ROC curves of LR, SVM, and RF models

A: 训练集; B: 验证集
A: Training set; B: Validation set

表4 LR、SVM、RF模型的诊断效能
Table 4 Diagnostic performance of LR, SVM, and RF models

模型	AUC(95% CI)	准确度(%)	敏感度(%)	特异度(%)
LR				
训练集	0.724(0.653~0.831)	68.3	53.8	81.3
验证集	0.741(0.651~0.848)	73.7	68.3	83.3
SVM				
训练集	0.761(0.673~0.849)	71.1	55.6	84.9
验证集	0.727(0.636~0.817)	68.5	54.2	81.5
RF				
训练集	0.770(0.683~0.856)	75.7	70.9	85.3
验证集	0.763(0.677~0.848)	71.3	56.8	85.1

2.4 联合模型构建与验证

将临床预测因素(CEA、CA19-9及NLR)和筛选出的10个最优影像组学特征结合,建立联合模型,使用RF建立临床模型、影像组学模型及二者联合模型,对三种模型进行ROC曲线分析,结果显示在训练集和验证集中联合模型的AUC分别

为0.913、0.947,准确度分别为0.858、0.939,敏感度分别为0.909、0.938,特异度分别为0.870、0.786,均高于临床诊断模型和影像组学模型(图6)。DCA曲线显示,联合模型在训练集和验证集中具有显著净获益(图7)。

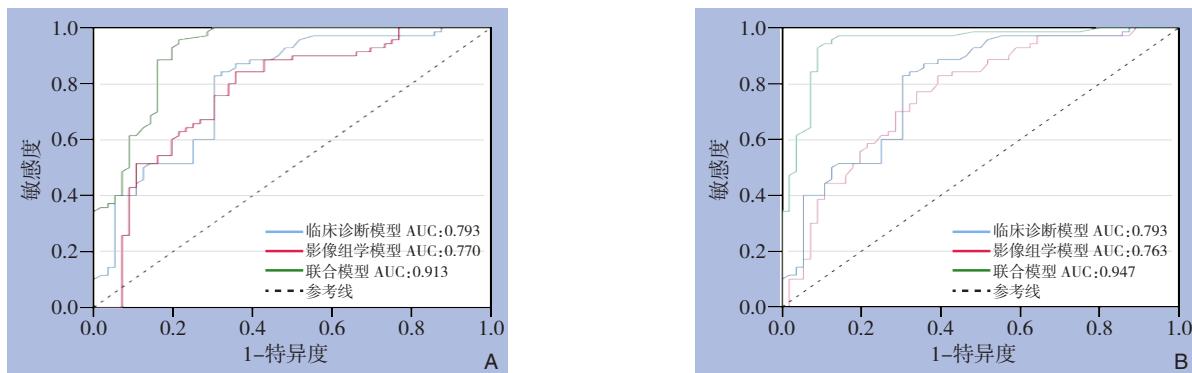


图6 RF、临床预测模型和联合模型的ROC曲线 A: 训练集; B: 验证集

Figure 6 ROC curves of RF-based radiomics model, clinical model, and combined model A: Training set; B: Validation set

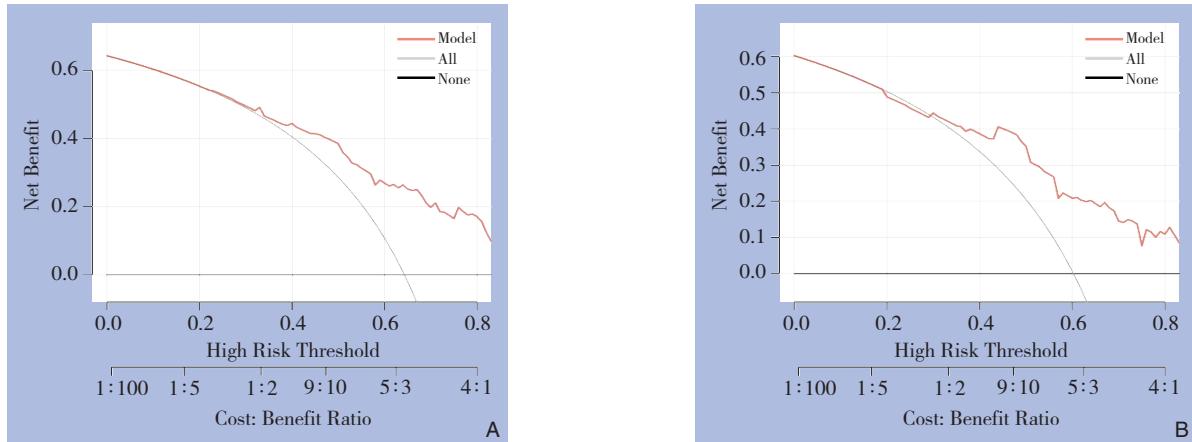


图7 联合模型的DCA曲线 A: 训练集; B: 验证集

Figure 7 Decision curve analysis (DCA) of the combined model A: Training set; B: Validation set

3 讨 论

CRC的血行转移多见于肝脏, CRCLM成为患者的主要死亡原因^[9]。据统计, 70%~90%的肝转移患者由于种种原因, 如转移肿瘤的广泛性、患者的一般状况等, 不适合接受根治性手术切除, 即便手术成功, 术后1年内的复发率也高达40%^[10]。因此术前利用MRI无创诊断CRCLM十分重要。

影像组学作为前沿技术, 凭借高效信息提取与数据分析能力, 显著提升诊断精度, 并广泛应用于肿瘤疾病的诊断、治疗、监测及预后评估^[11~13]。Xu等^[14]研究显示, 基于多参数MRI影像学特征构建的CRCLM影像组学预测模型(AUC=0.92)显示出比临床预测模型更好的(AUC=0.74)预测效能。Li等^[15]研究发现, 多参数MRI影像组学模型可有效预测直肠癌异时性肝转移, 该研究通过比较DWI、高分辨率T2WI及结合模型, 选定最佳融合模型, 结合临床特征后AUC达0.911, DCA与列线图均显示良好临床性能, 支持该模型辅助制定个体化监测与治疗策略。但上述研究多采用1种机器学习建模, 而本研究构建了3种机器学习模型, 并选定预测效果最突出的模型来构建复合模型。

受肿瘤异质性影响, 生物信息在不同区域存在差异, 多数机器学习研究仅采用二维图像作为ROI, 可能引入分析误差^[16~17], 本研究通过三维ROI分析克服了传统二维方法的肿瘤异质性局限, 使模型能够捕获更全面的生物学特征信息。经LASSO回归筛选的10个最优特征中, GLRLM特征可能反映肿瘤内部结构紊乱程度, 与癌细胞增殖导致的组织密度不均相关; GLSZM特征可能对应肿瘤微环境中血管生成或坏死区域的空间分布; GLDM特征可能表征肿瘤细胞浸润导致的密度梯度变化; 而一阶统计特征则可能体现整体肿瘤负荷和异质性程度^[18~19]。这些特征从不同角度刻画了肿瘤的生物学特性, 其组合可全面反映肿瘤微环境特征。研究结果显示, RF模型在训练集和验证集中均表现出最优预测性能(AUC值最高), 该结果与Zhang等^[20]基于SEER数据库的研究结论一致。RF模型的优势在于其集成学习机制: 通过构建多个决策树并采用投票机制, 显著提升了模型的预测准确性、泛化能力和抗过拟合性能^[21~23]。特别是在处理高维特征时, RF能够有效处理特征间的交

互作用和多重共线性问题, 这在本研究的多参数特征分析中尤为重要。此外, RF对噪声数据的鲁棒性使其更适合处理医学影像数据中常见的变异^[24~25]。然而, RF模型也存在可解释性相对较弱的局限。为优化模型性能, 本研究采取了以下措施: (1)严格控制影像数据质量, 确保采集参数一致; (2)通过LASSO回归进行特征选择, 降低维度; (3)采用网格搜索进行超参数调优。这些措施有效保障了模型的稳定性和可靠性, 使其能够为临床决策提供有力支持。未来研究可结合病理学检查结果, 进一步验证这些影像组学特征与肿瘤生物学特性之间的具体关联。

临床实践中, CEA、CA19-9、NLR等指标已被证实与CRC的预后密切相关, 可作为预后评估的重要依据^[26~27]。但本研究发现单独采用这些指标进行预测时, 其AUC数值偏低, 诊断效能有限; 而仅基于影像组学特征的模型虽也有一定诊断价值, 但AUC值也在0.8以下, 诊断效能有限。相比之下, 将临床模型与影像组学模型整合构建的联合模型, 在训练集(AUC=0.913)与验证集(AUC=0.947)中均展现出更优异的诊断性能, 且DCA显示该模型在CRCLM诊断中具有较高的临床净获益, 提示联合模型能充分融合临床指标的便捷性与影像组学特征的客观性, 显著提升诊断效能。朱琦等^[28]采用多参数3D MRI技术, 通过T2WI、DWI和e-THRIVE+序列的影像组学分析, 构建了CRCLM预测模型, 该研究采用VOI勾画方法, 经特征筛选后建立的联合模型预测效能显著优于单一序列模型(训练集AUC=0.919, 验证集AUC=0.896)。卞雪莲等^[29]对63例经病理证实的CRCLM患者进行回顾性分析, 通过单因素和多因素分析筛选出T2WI信号不均、肝胆期瘤周低信号等MRI特征作为微血管侵犯(MVI)的独立预测指标, 采用Gd-EOB-DTPA增强MRI肝胆期影像组学特征, 分别构建了LR、线性判别分析和SVM三种预测模型, 其中LR模型表现最优, 最终建立的临床-影像组学联合列线图模型展现出卓越的预测能力(训练集AUC=0.970, 测试集AUC=0.917), 校准曲线和DCA分析均证实其良好的临床应用价值。与上述研究相比, 本研究不仅延续了联合模型在预测效能上的优势, 更在特征维度与模型构建逻辑上有所侧重: 一方面, 通过三维ROI勾画捕捉肿瘤全容积特征, 减少了二维图像分析可能引入的异质性误差, 使影像组学

特征更贴合肿瘤生物学本质；另一方面，在临床指标与影像组学特征的融合过程中，通过LASSO回归进行特征精选与参数优化，进一步提升了联合模型的稳定性与泛化能力。此外，本研究明确对比了单独临床模型、单独影像组学模型与联合模型的效能差异，更直观地证实了多维度信息整合对突破单一模型局限性的关键作用，为CRCLM的无创诊断提供了更具实践意义的参考范式。但本研究亦存在一定局限性：其一，作为单中心回顾性研究，样本选择可能存在偏倚，结果的外推性需多中心前瞻性研究进一步验证；其二，未纳入更多潜在的临床预后指标（如肿瘤突变负荷、循环肿瘤DNA等），可能限制了联合模型的特征广度；其三，未对不同MRI序列（如肝胆特异期成像）的单独与联合预测价值进行分层分析，有待后续研究深入探讨；其四，本研究仅采用轴位T2WI、DWI及轴位T1WI增强序列获取多模态信息，虽能反映部分肿瘤特征，但在序列多样性上仍有拓展空间。鉴于CRCLM的高度复杂性，动态对比增强（DCE）MRI等序列的补充应用值得关注—DCE-MRI可提供肿瘤微血管灌注及通透性等关键信息，对于深入剖析肿瘤生物学行为、血管生成状态及提升鉴别诊断效能具有重要价值，将其与现有序列联合运用有望进一步增强诊断模型的准确性，这也是本研究后续需完善的方向。

综上所述，RF模型在影像组学模型中表现出较强的诊断性能，其与临床特征结合的联合模型对CRCLM有良好的诊断价值。

作者贡献声明：李波负责资料收集和撰写、指导、审查和修订分析数据；刘冠男负责文献查阅实验并统计数据、整理数据和汇总。

利益冲突：所有作者均声明不存在利益冲突。

参考文献

- [1] 朱江鹏, 李光耀, 姚远, 等. 结直肠癌同时性肝肺转移的预后因素分析及预后预测模型构建[J]. 中国普通外科杂志, 2024, 33(7): 1111–1121. doi:10.7659/j.issn.1005-6947.2024.07.010.
Zhu JP, Li GY, Yao Y, et al. Prognostic factors and construction of prognostic prediction model for simultaneously diagnosed liver and lung metastases from colorectal cancer[J]. China Journal of General Surgery, 2024, 33(7):1111–1121. doi:10.7659/j.issn.1005-6947.2024.07.010.
- [2] 申洋, 管政, 梁红, 等. 基于MRI影像组学预测同时性结直肠癌肝转移灶的化疗疗效[J]. 实用放射学杂志, 2023, 39(3):416–419. doi:10.3969/j.issn.1002-1671.2023.03.018.
Shen Y, Guan Z, Liang H, et al. MRI-based radiomics signature to predict the response of chemotherapy in synchronous liver metastasis of colorectal carcinoma[J]. Journal of Practical Radiology, 2023, 39(3): 416–419. doi: 10.3969/j. issn. 1002-1671.2023.03.018.
- [3] 官新立, 郑欣, 杨新军, 等. 表观扩散系数直方图在评估直肠癌分化程度中的价值[J]. 实用放射学杂志, 2022, 38(10):1638–1641. doi:10.3969/j.issn.1002-1671.2022.10.018.
Guan XL, Zheng X, Yang XJ, et al. The value of histogram analysis of apparent diffusion coefficient in evaluating the grade of rectal cancer[J]. Journal of Practical Radiology, 2022, 38(10):1638–1641. doi:10.3969/j.issn.1002-1671.2022.10.018.
- [4] 苏璇, 王远军. 结直肠癌肝转移预测及预后的影像组学综述[J]. 中国生物医学工程学报, 2023, 42(1): 91–98. doi: 10.3969/j. issn.0258-8021.2023.01.010.
Su X, Wang YJ. Radiomics Review on Prediction and Prognosis of Liver Metastasis in Colorectal Cancer[J]. Chinese Journal of Biomedical Engineering, 2023, 42(1): 91–98. doi: 10.3969/j. issn.0258-8021.2023.01.010.
- [5] 解添淞, 翁微微, 刘伟, 等. 基于CT图像影像组学的机器学习模型预测胰腺癌免疫细胞浸润及预后的初步研究[J]. 中华放射学杂志, 2022, 56(4): 425–430. doi: 10.3760/cma. j. cn112149-20211214-01104.
Xie TS, Weng WW, Liu W, et al. A pilot study of using CT-radiomics based machine learning model in predicting immune cells infiltrating and prognosis of pancreatic cancer[J]. Chinese Journal of Radiology, 2022, 56(4): 425–430. doi: 10.3760/cma. j. cn112149-20211214-01104.
- [6] 王岩, 赵青, 孙东兰, 等. MRI影像学组预测宫颈鳞癌盆腔淋巴结转移的价值[J]. 实用放射学杂志, 2022, 38(2):273–277. doi: 10.3969/j.issn.1002-1671.2022.02.022.
Wang Y, Zhao Q, Sun DL, et al. Value of MRI-based radiomics for the preoperative prediction of pelvic lymph node metastasis of cervical squamous cell carcinoma[J]. Journal of Practical Radiology, 2022, 38(2): 273–277. doi: 10.3969/j. issn. 1002-1671.2022.02.022.
- [7] 彭乐平, 张秀玲, 施柳言, 等. CT和MRI及影像组学预测结直肠癌微卫星不稳定状态研究进展[J]. 磁共振成像, 2024, 15(6):218–223. doi:10.12015/issn.1674-8034.2024.06.035.
Peng LP, Zhang XL, Shi LY, et al. Research progress of CT and MRI with radiomics to predict microsatellite instability in colorectal cancer[J]. Chinese Journal of Magnetic Resonance

- Imaging, 2024, 15(6): 218–223. doi: [10.12015/issn.1674-8034.2024.06.035](https://doi.org/10.12015/issn.1674-8034.2024.06.035).
- [8] Cervantes A, Adam R, Roselló S, et al. Metastatic colorectal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up[J]. Ann Oncol, 2023, 34(1): 10–32. doi: [10.1016/j.annonc.2022.10.003](https://doi.org/10.1016/j.annonc.2022.10.003).
- [9] Wang Y, Zhong X, He X, et al. Liver metastasis from colorectal cancer: pathogenetic development, immune landscape of the tumour microenvironment and therapeutic approaches[J]. J Exp Clin Cancer Res, 2023, 42(1): 177. doi: [10.1186/s13046-023-02729-7](https://doi.org/10.1186/s13046-023-02729-7).
- [10] Wehrle CJ, Raj R, Aykun N, et al. Liquid biopsy by ctDNA in liver transplantation for colorectal cancer liver metastasis[J]. J Gastrointest Surg, 2023, 27(7): 1498–1509. doi: [10.1007/s11605-023-05723-8](https://doi.org/10.1007/s11605-023-05723-8).
- [11] Inchingolo R, Maino C, Cannella R, et al. Radiomics in colorectal cancer patients[J]. World J Gastroenterol, 2023, 29(19): 2888–2904. doi: [10.3748/wjg.v29.i19.2888](https://doi.org/10.3748/wjg.v29.i19.2888).
- [12] Zhang S, Yu M, Chen D, et al. Role of MRI-based radiomics in locally advanced rectal cancer (Review)[J]. Oncol Rep, 2022, 47(2): 34. doi: [10.3892/or.2021.8245](https://doi.org/10.3892/or.2021.8245).
- [13] 马小梅, 和建伟, 贾应梅, 等. 基于MRI影像组学预测结直肠KRAS基因突变的系统回顾和Meta分析[J]. 磁共振成像, 2025, 16(4): 60–69. doi: [10.12015/issn.1674-8034.2025.04.010](https://doi.org/10.12015/issn.1674-8034.2025.04.010).
Ma XM, He JW, Jia YM, et al. Assessing the efficacy of MRI radiomics for KRAS mutation prediction in colorectal cancer: insights from a systematic review and Meta-analysis[J]. Chinese Journal of Magnetic Resonance Imaging, 2025, 16(4): 60–69. doi: [10.12015/issn.1674-8034.2025.04.010](https://doi.org/10.12015/issn.1674-8034.2025.04.010).
- [14] Xu Y, Ye F, Li L, et al. MRI-based radiomics nomogram for preoperatively differentiating intrahepatic mass-forming cholangiocarcinoma from resectable colorectal liver metastases[J]. Acad Radiol, 2023, 30(9): 2010–2020. doi: [10.1016/j.acra.2023.04.030](https://doi.org/10.1016/j.acra.2023.04.030).
- [15] Li ZF, Kang LQ, Liu FH, et al. Radiomics based on preoperative rectal cancer MRI to predict the metachronous liver metastasis[J]. Abdom Radiol (NY), 2023, 48(3): 833–843. doi: [10.1007/s00261-022-03773-1](https://doi.org/10.1007/s00261-022-03773-1).
- [16] Huang H, Han L, Guo J, et al. Pretreatment MRI-based radiomics for prediction of rectal cancer outcome: a discovery and validation study[J]. Acad Radiol, 2024, 31(5): 1878–1888. doi: [10.1016/j.acra.2023.10.055](https://doi.org/10.1016/j.acra.2023.10.055).
- [17] Hou M, Zhou L, Sun J. Deep-learning-based 3D super-resolution MRI radiomics model: superior predictive performance in preoperative T-staging of rectal cancer[J]. Eur Radiol, 2023, 33(1): 1–10. doi: [10.1007/s00330-022-08952-8](https://doi.org/10.1007/s00330-022-08952-8).
- [18] 周小飞. 基于MRI的影像组学及纹理分析在直肠癌不同病理亚型及神经侵犯预测中的应用[D]. 苏州:苏州大学, 2023. doi: [10.27351/d.cnki.gszhu.2023.000670](https://doi.org/10.27351/d.cnki.gszhu.2023.000670).
Zhou XF. MRI based radiomics and texture analysis in the prediction of pathological subtypes and perineural invasion in rectal cancer[D]. Suzhou: Soochow University, 2023. doi: [10.27351/d.cnki.gszhu.2023.000670](https://doi.org/10.27351/d.cnki.gszhu.2023.000670).
- [19] 李国强. 基于磁共振T2WI纹理分析对直肠癌淋巴结转移的预测价值评估[D]. 沈阳:中国医科大学, 2023. doi: [10.27652/d.cnki.gzyku.2023.000117](https://doi.org/10.27652/d.cnki.gzyku.2023.000117).
Li GQ. Evaluation of predictive value for lymph node metastasis of rectal cancer based on T2WI texture analysis of magnetic resonance[D]. Shenyang: China Medical University, 2023. doi: [10.27652/d.cnki.gzyku.2023.000117](https://doi.org/10.27652/d.cnki.gzyku.2023.000117).
- [20] Zhang Y, Zhang Z, Wei L, et al. Construction and validation of nomograms combined with novel machine learning algorithms to predict early death of patients with metastatic colorectal cancer[J]. Front Public Health, 2022, 10: 1008137. doi: [10.3389/fpubh.2022.1008137](https://doi.org/10.3389/fpubh.2022.1008137).
- [21] Zhao FH, Fan HJ, Shan KF, et al. Predictive efficacy of a radiomics random forest model for identifying pathological subtypes of lung adenocarcinoma presenting as ground-glass nodules[J]. Front Oncol, 2022, 12: 872503. doi: [10.3389/fonc.2022.872503](https://doi.org/10.3389/fonc.2022.872503).
- [22] Xu H, Li H, Fan Y, et al. Analysis of factors influencing chemotherapy-induced peripheral neuropathy in breast cancer patients using a random forest model[J]. Breast, 2025, 81: 104457. doi: [10.1016/j.breast.2025.104457](https://doi.org/10.1016/j.breast.2025.104457).
- [23] Hu J, Szymczak S. A review on longitudinal data analysis with random forest[J]. Brief Bioinform, 2023, 24(2): bbad002. doi: [10.1093/bib/bbad002](https://doi.org/10.1093/bib/bbad002).
- [24] 张伟伟, 陈启, 翁和语, 等. 随机森林模型预测T1期结直肠癌淋巴结转移的初步研究[J]. 中华结直肠疾病电子杂志, 2024, 13(5): 389–393. doi: [10.3877/cma.j.issn.2095-3224.2024.05.006](https://doi.org/10.3877/cma.j.issn.2095-3224.2024.05.006).
Zhang WW, Chen Q, Weng HY, et al. Preliminary study on prediction of lymph node metastasis of T1 colorectal cancer by random forest model[J]. Chinese Journal of Colorectal Disease: Electronic Edition, 2024, 13(5): 389–393. doi: [10.3877/cma.j.issn.2095-3224.2024.05.006](https://doi.org/10.3877/cma.j.issn.2095-3224.2024.05.006).
- [25] Mi B, Jin Y, Zheng M, et al. Stigma, colorectal cancer knowledge and self-efficacy among colorectal cancer survivors: a cross-sectional study based on random forest analysis[J]. Eur J Oncol Nurs, 2025, 76: 102858. doi: [10.1016/j.ejon.2025.102858](https://doi.org/10.1016/j.ejon.2025.102858).
- [26] Chen B, Zhou J, Ma Y, et al. Evaluation of multiple biological indicators for the combined diagnosis of metastases from colorectal cancer-a retrospective study based on 1163 patients[J]. World J Surg Oncol, 2023, 21(1): 229. doi: [10.1186/s12957-023-03108-4](https://doi.org/10.1186/s12957-023-03108-4).

- [27] Manojlovic N, Savic G, Nikolic B, et al. Dynamic monitoring of carcinoembryonic antigen, CA19-9 and inflammation-based indices in patients with advanced colorectal cancer undergoing chemotherapy[J]. World J Clin Cases, 2022, 10(3): 899–918. doi: 10.12998/wjcc.v10.i3.899.
- [28] 朱琦, 陈艾琪, 邹文涛, 等. 多参数3D MRI影像组学模型预测结直肠癌肝转移及相关临床危险因素的分析[J]. 中国CT和MRI杂志, 2024, 22(2): 146–149. doi: 10.3969/j. issn. 1672-5131.2024.02.045.
- Zhu Q, Chen AQ, Zou WT, et al. Multi-parametric 3D MRI radiomics model for predicting liver metastasis of colorectal cancer and related clinical risk factors analysis[J]. Chinese Journal of CT and MRI, 2024, 22(2): 146–149. doi: 10.3969/j. issn. 1672-5131.2024.02.045.
- [29] 卞雪莲, 张培培, 孙琦, 等. 基于钆塞酸二钠增强MRI影像组学列线图模型术前预测结直肠癌肝转移瘤微血管侵犯[J]. 临床放射学杂志, 2024, 43(7): 1118–1124. doi: 10.13437/j. cnki.jcr.2024.07.012.
- Bian XL, Zhang PP, Sun Q, et al. Preoperative prediction of microvascular invasion in colorectal cancer liver metastases based on Gd-EOB-DTPA enhanced MRI radiomics nomogram model[J]. Journal of Clinical Radiology, 2024, 43(7): 1118–1124. doi: 10.13437/j.cnki.jcr.2024.07.012.

(本文编辑 姜晖)

本文引用格式:李波, 刘冠男. MRI影像组学结合临床特征的机器学习模型对结直肠癌肝转移的预测价值[J]. 中国普通外科杂志, 2025, 34(7):1410–1420. doi:10.7659/j.issn.1005-6947.240611

Cite this article as: Li B, Liu GN. The predictive value of MRI imaging omics combined with clinical features in machine learning models for colorectal cancer liver metastasis[J]. Chin J Gen Surg, 2025, 34(7):1410–1420. doi: 10.7659/j.issn.1005-6947.240611

本刊对来稿中统计学处理的有关要求

1.统计研究设计: 应交代统计研究设计的名称和主要做法。如调查设计(分为前瞻性、回顾性或横断面调查研究); 实验设计(应交代具体的设计类型, 如自身配对设计、成组设计、交叉设计、正交设计等); 临床试验设计(应交代属于第几期临床试验, 采用了何种盲法措施等)。主要做法应围绕四个基本原则(随机、对照、重复、均衡)概要说明, 尤其要交代如何控制重要非试验因素的干扰和影响。

2.资料的表达与描述: 用 $\bar{x} \pm s$ 表达近似服从正态分布的定量资料, 用 M (IQR) 表达呈偏态分布的定量资料; 用统计表时, 要合理安排纵横标目, 并将数据的含义表达清楚; 用统计图时, 所用统计图的类型应与资料性质相匹配, 并使数轴上刻度值的标法符合数学原则; 用相对数时, 分母不宜小于20, 要注意区分百分率与百分比。

3.统计分析方法的选择: 对于定量资料, 应根据所采用的设计类型、资料所具备的条件和分析目的, 选用合适的统计分析方法, 不应盲目套用 t 检验和单因素方差分析; 对于定性资料, 应根据所采用的设计类型、定性变量的性质和频数所具备条件以分析目的, 选用合适的统计分析方法, 不应盲目套用 χ^2 检验。对于回归分析, 应结合专业知识和散布图, 选用合适的回归类型, 不应盲目套用简单直线回归分析, 对具有重复实验数据的回归分析资料, 不应简单化处理; 对于多因素、多指标资料, 要在一元分析的基础上, 尽可能运用多元统计分析方法, 以便对因素之间的交互作用和多指标之间的内在联系进行全面、合理地解释和评价。

4.统计结果的解释和表达: 当 $P<0.05$ (或 $P<0.01$) 时, 应说明对比组之间的差异有统计学意义, 而不应说对比组之间具有显著性(或非常显著性)的差别; 应写明所用统计分析方法的具体名称(如: 成组设计资料的 t 检验、两因素析因设计资料的方差分析、多个均数之间两两比较的 q 检验等), 统计量的具体值(如 $t=3.45$, $\chi^2=4.68$, $F=6.79$ 等)应尽可能给出具体的 P 值(如 $P=0.023$ 8); 当涉及总体参数(如总体均数、总体率等)时, 在给出显著性检验结果的同时, 再给出95%置信区间。

中国普通外科杂志编辑部