·直肠癌影像学专题·

直肠癌新辅助治疗效果评价——从大体影像到 AI 辅助评价体系建立

陈浩权，王屹

【摘要】 直肠癌新辅助治疗后疗效评价直接影响等待观察、手术及后续综合治疗决策。传统 MRI 可显示肿瘤退缩和纤维、黏液改变，但在残存肿瘤识别、生物学异质性刻画、远处转移和预后预测方面仍有局限。本文系统综述传统影像组学、深度学习、多模态 AI 及大语言模型辅助决策在临床完全缓解/病理完全缓解判断、免疫治疗反应、远处转移及生存预测中的研究进展，并讨论其临床转化挑战与发展方向。

【关键词】 直肠癌；新辅助治疗；磁共振成像；人工智能；影像组学；深度学习；治疗反应

【中图分类号】R735.3；R445.2 【文献标志码】A 【文章编号】1000-0313(2026)09-1013-06

DOI:10.13609/j.cnki.1000-0313.2026.09.006 开放科学(资源服务)标识码(OSID)：

直肠癌相关死亡率位居癌症相关死亡率的第二位[1]。基于高分辨率磁共振影像的形态学风险分层体系为直肠癌治疗提供了明确的决策路径[2]。影像风险因素可联合或独立于肿瘤 TNM 分期共同用于新辅助治疗方案选择、手术策略制定和随访管理[3]。术前新辅助放化疗(neoadjuvant chemoradiotherapy，nCRT)后直肠癌临床完全缓解(clinical complete response，cCR)，可实施等待观察策略以满足保留器官迫切需求[4]。尽管对于患者远处转移、生存预后以及基因免疫相关问题大体影像有提示作用，但尚不足以准确预测。因此，如何获得传统体系外肿瘤诊断及风险差异标志物是影像模型研发的核心目标。

人工智能(artificial intelligence，AI)深入挖掘影像中难以被肉眼识别的高维特征并结合人类先验知识判断治疗效果有望成为具有突破性的评价工具，推动直肠癌疗效评价从大体影像判断走向多模态生物学表型辅助决策。

影像 AI 模型判断直肠癌新辅助治疗效果

准确判断直肠癌 nCRT 后 cCR 是实施等待观察策略的关键。当前用于判断 cCR 的检查方法包括高分辨率 MRI、肛门指诊、肠镜检查及活检，必要时 PET-CT 辅助诊断。影像上，当原肿瘤区域扩散加权成像(diffusion weighted imaging，DWI)信号与周围相邻正常肠壁接近或一致，且 T_2WI 表现为正常肠壁结构或仅见 1～2 mm 低信号纤维组织时，可提示接近或达到 cCR(图 1)；但当治疗后表现为纤维化、黏液变性或混合退缩时，残存肿瘤识别仍是临床难点，成为当前迫切需要解决的问题。

近年来，基于 AI 构建综合模型判断直肠癌 nCRT 效果，特别是预测病理完全缓解(pathological complete response，pCR)和辅助判断 cCR 成为研究热点。从技术路线看，现有 AI 研究大致可分为传统影像组学、深度学习、多模态联合模型等(表 1)。传统影像组学通过人工勾画感兴趣区并提取纹理、形态及灰度等定量特征，具有样本需求较低和便于解释的优势，但受分割方式、扫描参数和特征筛选流程影响较大，模型稳定性和跨中心泛化能力有限；深度学习模型可自动学习高维影像特征，并可结合纵向 MRI 或多任务网络同时完成分割与疗效预测，更符合治疗前后动态评价的临床逻辑，但通常依赖较大样本量，且存在黑箱性和可解释性不足的问题；多模态融合模型进一步整合影像、病理、临床及分子信息，有助于刻画肿瘤生物学异质性并提高预测性能，但不同模态数据获取难度、缺失情况和标准化差异限制了其推广应用。

1.评价直肠癌新辅助治疗效果的荟萃分析

2025 年发表的一项荟萃分析结果提示基于 MRI 构建预测直肠癌 nCRT 后 pCR 的 AI 模型，其联合 T_2WI 和 DWI 较单一序列准确性高[5]；这一结果既与临床工作流程吻合，也显示多参数信息有助于提升模型稳定性。既往 Shen 等[6]发表的荟萃分析结果提示，AI 模型预测 pCR 的汇集敏感度、特异度和受试者工作特征曲线下面积(area under curve，AUC)值分别达到了 82%、84%和 90%；但研究间存在显著异质性，其中亚洲地区、低风险偏倚和深度学习的模型特异性显著高于其它；Rai 等[7]围绕 MRI 影像组学预测局部进展期直肠癌 nCRT 后 pCR 进行荟萃分析，结果提示 MRI 影像组学在 pCR 预测中具有较好的诊断潜力。其中，Shin 等[8]基于治疗后 MRI 构建 T_2WI、表观扩散系数(apparent diffusion coefficient，ADC)以及融合影像组学模型，并与影像医师基于 MRI 肿瘤退缩分级

作者单位：100044 北京，北京大学人民医院放射科
作者简介：陈浩权(1999－)，男，广东揭阳人，博士研究生，主要从事肿瘤影像诊断工作。
通信作者：王屹，E-mail：wang_yi@hsc.pku.edu.cn

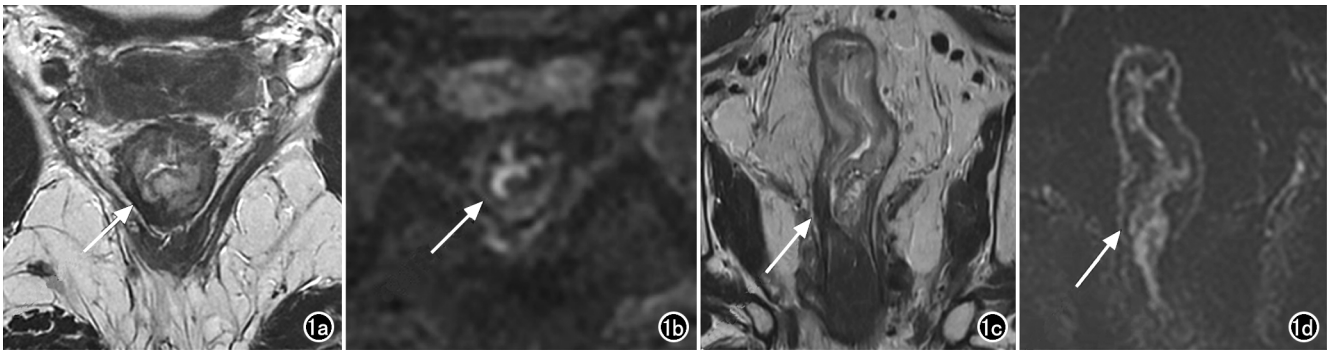

图 1 直肠癌治疗后影像学疗效评价示例图。a)高分辨率 T_2WI 轴面图像，箭头示治疗后低信号瘢痕组织；b)高分辨率 DWI 轴面图像，原病灶区未见明显异常信号，箭头示黏膜、黏膜下层和固有肌层，未见异常信号改变；c)高分辨率 T_2WI 冠状面图像，箭头示治疗后低信号瘢痕组织；d)高分辨率 DWI 冠状面图像，箭头示黏膜、黏膜下层及固有肌层未见明显异常信号。

表 1 直肠癌新辅助治疗疗效预测代表性 AI 模型比较

技术路线/研究者	数据来源	样本量	输入模态	主要性能指标	临床局限与数据偏倚
传统影像组学/Shin 等[8]	韩国单中心	898 例(训练集 592，测试集 306)	治疗后多参数 MRI	预测 pCR：T_2WI 模型 AUC 0.82；ADC 模型 AUC 0.79；融合模型 AUC 0.82	单中心回顾性设计；分割偏倚；设备偏倚；临床局限；模型特异度(68.4%)较低，假阳性率高，直接用于临床“等待观察”决策需谨慎。
免疫表型引导的可解释影像组学/Zhang 等[16]	中国多中心	149 例(免疫分型 108，训练集 19，验证集 22)	治疗前静脉期增强 CT 图像	预测新辅助抗 PD-1 治疗 pCR：训练集 AUC 0.965；独立验证集 AUC 0.904	样本量较小，泛化能力不足；设备存在选择偏倚；回顾性设计存在选择偏倚；生物学机制仍需多组学进一步验证。
纵向多任务深度学习/Jin 等[10]	中国双中心	622 例(训练集 321，验证集 301)	治疗前后配对多参数 MRI＋CEA	预测 pCR：训练集 AUC 0.95；验证集 AUC 0.92；联合 CEA 后 AUC 达 0.97	依赖精准图像配对，病灶自动分割成本高；PPV 尚不足以直接支持“等待观察”。
纵向深度学习联合单细胞组学/Cao 等[12]	中国多中心	1026 例(训练集 633，验证集 212，外部测试集 181)	治疗前后配对 T_2WI	预测 pCR：训练集 AUC 0.93；验证集 AUC 0.88；外部测试集 AUC 0.90	回顾性设计存在选择偏倚；scRNA-seq 样本量较小(n=26)；未整合多参数 MRI 信息；依赖手动 ROI 勾画。
多模态联合模型/Feng 等[25]	中国多中心	1033 例(训练集 303＋回顾验证 630，前瞻验证 100)	治疗前多参数 MRI＋HE 活检病理全切片图像	预测 pCR：训练集 AUC 0.868；回顾验证集 AUC 0.860；前瞻验证集 AUC 0.812	pCR 与非 pCR 样本比例不均衡，依赖手动感兴趣区勾画

注：ADC 为表观扩散系数；HE 为苏木精-伊红染色；CEA 为癌胚抗原；pCR 为病理完全缓解；nCRT 为新辅助放化疗；PD-1 为程序性细胞死亡蛋白-1；scRNA-seq 为单细胞 RNA 测序；AUC 为受试者工作特征曲线下面积；PPV 为阳性预测值。

的视觉评估进行比较；测试集中，T_2WI、ADC 和融合模型的 AUC 分别为 0.82、0.79 和 0.82。周晓俞等[9]进一步采用多方位 MRI 影像组学并联合癌胚抗原(carcinoembryonic antigen，CEA)、壁外血管侵犯(extramuralvascular invasion，EMVI)等临床及常规影像指标预测 nCRT 疗效。上述研究提示定量纹理特征能够在一定程度上弥补常规视觉判断在残存肿瘤识别中的不足，但现有研究仍存在样本量偏小、单中心比例高、外部验证不足、扫描协议差异大和肿瘤分割成本高等问题。

2. 纵向 MRI 及多时间点深度学习的核心逻辑

当研究者假设肿瘤初始状态与治疗效果密切相关时，即基于 nCRT 前影像搭建 AI 预测模型。然而，影像医师需通过识别原发灶、淋巴结、系膜侵犯、EMVI、肿瘤沉积(tumor deposit，TD)等治疗前后纵向变化综合判断治疗效果。因此，直肠癌疗效评价不应仅依赖治疗前单一时间点影像，将纵向影像变化纳入深度学习模型更符合真实阅片逻辑。深度学习将预测 nCRT 效果推向端到端、多任务、多时间点的方向，即以时间为轴的纵向分析更有益于准确判断 cCR。Jin 等[10]设计 3D RP-Net 多任务深度学习模型完成肿瘤自动分割与效果预测双任务，利用治疗前后成对的纵向 MRI 数据对比学习，在两个独立训练队列和外部验证队列中取得 AUC 分别为 0.95 与 0.92 的高预测性能。当整合血清 CEA 等临床指标后 AUC 值提升至 0.97；随后，Ke 等[11]在此基础上进一步构建并验证 DeepRP-RC 系统，基于治疗前后配对 MRI 预测直肠癌治疗后 pCR，并在大样本多中心队列中进行内部和外部验证，进一步强化了纵向 MRI 深度学习模型在临床推广中的证据基础。近年来，Cao 等[12]提出的 MR-DELTA-net 模型利用纵向 MRI 构建的 Transformer 模型进一步将动态影像特征与单细胞转录组学相结合，尝试解释模型预测结果背后的免疫微环境差异。这提示未来模型不仅需要提高预测准确性，还应从单纯影像分类走向纵向影像变化与组织微环境关联，使其具备生物可解释性。

3. 免疫治疗反应的影像-病理-基因联合预测

随着错配修复功能缺陷(deficient mismatch repair，dMMR)/微卫星高度不稳定(microsatellite instability-high，MSI-H)直肠癌免疫治疗被推荐为一线方案，如何评价其治疗效果成为另一个迫切需要解决的临床问题。前期研究提示，靶向程序性细胞死亡蛋白-1(programmed cell death protein-1，PD-1)/程序性死亡配体-1(programmed cell death ligand 1，PD-L1)

途径的免疫检查点抑制剂(immune checkpoint inhibitors,ICIs)在 d-MMR/MSI-H 结直肠癌中的 pCR 率范围约 60%～100%[13,14]。然而,免疫治疗后可能出现假进展、超进展或混合进展,大体影像评价局限更为突出。Yang 等[15]基于多维 MRI 特征构建三级淋巴结构预测模型,用于无创评估直肠癌肿瘤免疫微环境,并进一步验证其对免疫治疗 pCR 和预后的预测价值,提示 MRI 影像 AI 有望从单纯形态学评价拓展至免疫表型识别。Zhang 等[16]基于 CT 影像、RNA 测序以及免疫组化染色三个维度,提取与免疫浸润最为相关的影像组学特征,构建可解释的决策树模型用于预测 PD-1 治疗后 pCR。在独立验证集中,模型预测 AUC 达到 0.904,显著优于临床或 CT 评估。影像组学特征与免疫细胞浸润程度的高度相关,实现"影像-免疫表型-临床结局"的逻辑闭环。

当前构建影像-病理-基因联合的多模态联合模型正在成为直肠癌疗效预测的重要发展方向。单模态 MRI 或 CT 影像主要反映肿瘤形态、信号和空间异质性,而病理图像、免疫组化、基因突变和转录组信息能够提供肿瘤细胞组成、免疫浸润状态和分子通路活性等更深层生物学信息。将这些信息与影像特征联合建模,有助于从影像表型进一步追溯至组织表型和分子表型,从而提高 pCR、免疫治疗反应和远处转移风险预测的可解释性,推动影像向融合基因、组织细胞和大体影像标志的生物特征驱动的诊断新范式发展。

影像 AI 模型辅助直肠癌风险分层

远处转移是直肠癌治疗失败和肿瘤相关死亡的独立性相关因素,如何在治疗前识别出高危转移人群是临床挑战之一[17]。MRI 所示肿瘤位置和浸润深度、EMVI、癌结节和淋巴结转移与肿瘤转移复发存在显著关联,但传统阅片主要依赖主观定性观察,难以量化肿瘤微观异质性。影像组学标签可定量反映直肠癌潜在侵袭性,并用于预测远处转移风险及辅助化疗获益,体现其在治疗前风险分层中的价值[18,19]。进一步,Zhao 等[20]整合基线多参数 MRI 影像组学、HE 活检病理组学及 CD3+/CD8+活检免疫评分,构建用于预测局部进展期直肠癌 nCRT 后远处转移风险的列线图,提示宏观影像表型与微观病理-免疫特征融合有助于提升治疗前高危转移患者的识别能力。另有研究揭示了影像组学特征与肿瘤增殖通路、免疫浸润通路之间的深层关联,且与肿瘤分子病理通路高度一致,为未来影像替代组织分子检测提供可能[21,22]。

深度学习可以从更宏观的空间模式中学习肿瘤生物学特征,其将不再局限于局部纹理或结构,而是从更高维度探究影像数据信息与肿瘤生物特征的复杂关联。因此,不断有研究应用深度学习网络探索影像或病理与患者生存内在关联关系,以直接预测远处转移和总生存期。Liu 等[23]基于 MRI 多参数组学标签和临床病理因素构建迁移学习模型,将患者分为高危和低危组;结果显示高危亚组获益于辅助治疗且远处转移风险显著下降,而低危组无明显受益,C-index 约 0.75～0.78。另有相关研究整合 T_2WI、ADC、DKI 和临床多模态信息构建深度学习端到端、多任务网络,实现自动分割、预测复发风险与无疾病进展生存,在外部验证中 C-index 超过 0.80,预测 3 年复发的 AUC 高达 0.97[24]。总体而言,AI 已经初步展示出预测无病生存和总生存超越传统临床分期和病理指标的潜力,其性能甚至可以达到临床可用于个体化随访和治疗强化决策的水平。

基于大体 MRI 影像的形态学风险分层体系为直肠癌治疗提供了明确的决策路径。影像风险信息联合或独立于肿瘤 TNM 分期共同用于新辅助治疗方案选择、手术策略制定和随访管理,并实现肿瘤转移和患者生存的预测。与传统影像风险分层相比,影像人工智能有望进一步挖掘肉眼难以识别的潜在肿瘤生物学特征,并将影像信息转化为可量化的治疗反应或预后预测指标。Feng 等[25]融合影像组学和活检病理组学特征构建了放射病理组学综合系统(radiopathomics integrated prediction system,RAPIDS)标签以预测直肠癌 nCRT 后 pCR,并尝试应用 RAPIDS 在预测结果为 pCR 后给予患者巩固化疗,如预测结果为非 pCR 则进行根治性手术,期望融合标签具备治疗决策能力。Huang 等[26]基于多中心大样本Ⅱ期结直肠癌 CT 影像,端到端深度神经网络提取多平面 CT 图像信息构建 STAR-CRC 模型,将Ⅱ期结直肠癌患者分为低风险、高风险和不确定风险;STAR-CRC 融合术后病理标志物构建 IRIS-CRC 模型再次进行风险分层。IRIS-CRC 模型多中心外部验证集中表现稳定,其预测 3 年无病生存风险的 AUC 和 C-index 均显著高于现有指南的风险标准,同时能够在不同临床亚组中维持一致的区分度。IRIS-CRC 不仅是风险预测模型,还是一个参与治疗决策的有效工具。然而,上述研究基础是否能够成为下一步随机对照试验的依据,还需进一步探讨及验证;相关模型能否作为后续随机对照试验或真实世界应用的依据,仍需进一步验证;其可解释性、错误输出风险、人机交互方式和临床部署路径也仍需持续完善。

影像 AI 模型辅助预测直肠癌患者生存期

总生存期是肿瘤治疗的最关键结局指标。围绕这一结局,AI研究尝试从术前影像、病理图像及临床资

表 2 直肠癌及相关结直肠癌远处转移、复发及生存风险预测代表性 AI 模型比较

技术路线/研究者	数据来源	样本量	输入模态	主要性能指标	临床局限与数据偏倚
多模态联合模型/Zhao 等[20]	中国单中心	201 例(训练集 140，测试集 61)	多参数 MRI＋HE 活检切片＋免疫组化	预测 5 年 DMFS：训练集 C-index 0.902，测试集 0.848；5 年 AUC：训练集 0.950，测试 0.872。	单中心回顾性设计存在选择偏倚，泛化能力不足；依赖手动感兴趣区勾画；活检样本存在组织异质性与取样偏倚。
多模态多任务深度学习/Liu 等[24]	中国多中心	445 例(训练集 261，验证集 37，内部测试集 75，外测试集 72)	多参数 MRI＋治疗前临床指标＋术后病理指标	预测 DFS 与 3 年 RM：内部测试集 C-index 0.92，3 年 RM 分类 AUC 0.97；外部测试集 C-index 0.81；实现端到端自动预测	回顾性设计存在选择偏倚；DKI 序列临床普及度有限；样本量仍有限，部分外部测试亚组区分度不足；终点主要为 DFS/RM，未分析 OS
多模态基础模型/Qu 等[33]	中国多中心＋公共数据集 TCGA	1679 例(训练集/内部验证集 839，外部验证集 840)	CT 图像＋病理 HE 全切片图像＋临床文本报告	预测 OS 和 PFS：训练集 OS C-index 0.812，PFS 0.901；外验集 OS C-index 0.730～0.790，PFS 0.725～0.829。	回顾性设计存在选择偏倚；队列间治疗模式差异致潜在分布偏倚；缺乏原始基因组测序数据
多模态临床决策辅助模型/Huang 等[26]	中国多中心	2992 例(训练集 1110，内部验证集 477，外部验证集 1405)	多平面 CT 图像＋临床病理因素	预测 DFS：外部验证 C-index 0.686，优于指南风险标准，辅助治疗升降阶决策	回顾性设计存在选择偏倚；未纳入 MSI 等分子标志物；临床决策价值仍需前瞻性验证；未分析 OS

注：HE 为苏木精-伊红染色；TCGA 为癌症基因组图谱；MSI 为微卫星不稳定；DMFS 为远处转移无生存期；DFS 为无病生存期；RM 为复发/转移；OS 为总生存期；PFS 为无进展生存期；AUC 为受试者工作特征曲线下面积；C-index 为一致性指数。

料中提取传统分期体系之外的预后表型，以提高患者生存风险分层能力。Jiang 等[27]基于术前 MRI 影像数据以及 Vision Transformer 构建总生存模型，在验证集中的 C-index 为 0.82；当进一步融合治疗前癌胚抗原水平后，多模态模型在验证集中的 C-index 提高至 0.86。另有研究基于 4515 例结直肠癌患者的 HE 病理切片，提取超过 1200 万个图像块构建生存预测深度学习模型，并在独立验证队列中评估其对患者结局的预测价值，提示常规病理图像中蕴含可用于预后分层的深层形态学信息[28]。同时，该研究的数据来源于临床研究，数据严谨完整，模型在训练集和外部验证集中取得相似结果。需要注意的是，总生存期受到肿瘤生物学特征、治疗决策、治疗执行情况、不良反应、患者全身状态、基础疾病以及经济、心理和家庭支持等多重因素共同影响，单一影像特征仅能反映部分因素。因此，如何从临床资料、影像和病理图像中提取更稳定、更综合的预后标志物，是直肠癌预后模型研究的重要方向。

表 2 总结了直肠癌及相关结直肠癌远处转移、复发及生存风险预测代表性模型比较。总体而言，肿瘤临床治疗到达了全面综合诊治的时代，患者总生存期不断延长，单纯预测总生存期已难以满足精准管理需求。临床实践与研究的目标逐渐转向更为精准化治疗和延长无进展生存期等更贴近临床决策的问题，并通过多模态数据融合和前瞻性验证推动预后模型向真实世界应用转化。

现存局限与未来展望

尽管直肠癌影像 AI 研究近年来持续取得突破，但其进入真实临床诊疗决策流程仍面临多维度障碍，主要包括数据和方法学偏倚与受限、临床效益验证不足、算法架构演进与可解释性、真实世界部署挑战等方面。

1. 数据和方法学层面的偏倚与局限

模型准确性不足与泛化能力受限是制约 AI 进入临床的首要障碍，其根源多与数据来源、研究设计、图像采集、病灶标注和终点定义等因素有关。现有 AI 模型多基于小样本及回顾性设计构建，既往荟萃分析提示研究样本量约为 43～1201 例[11,29]，且部分研究缺乏充分外部验证。尽管 Hu 等[30]与 Pang 等[31]率先开展了前瞻性队列验证，但针对特定临床问题，尤其是“等待观察”策略获益评估等的前瞻性随机对照试验仍罕见。与此同时，临床中真正实施该策略的患者比例有限，阳性样本稀少且组间异质性较大，进一步增加了模型过拟合风险和临床推广难度。

在数据层面，MRI 设备场强、厂商及扫描参数差异可引起影像特征漂移。既往荟萃分析提示，模型性能存在显著的地域与种族局限，这种数据分布偏倚与图像采集协议的异质性密切相关[6]。在病灶标注方面，既往研究多集中于原发灶或相邻肠壁，较少系统纳入直肠系膜内淋巴结、TD、EMVI 以及侧方及腹股沟淋巴结等关键区域。局部分割可能导致空间信息丢失，影响模型对肿瘤整体风险的表征能力。此外，基于内镜活检标本获取的病理组学特征或免疫评分容易受到组织异质性和取样偏倚影响，单点活检难以完全代表整个肿瘤的微观结构和免疫微环境状态。进一步地，在真实世界中，多模态数据往往并不完整，治疗方案差异、治疗前后特征演变及随访时间不一致等混杂因素也会影响模型表现。Chen 等[32]在多中心和 27433 例真实世界数据中验证 COCA 模型，提示大规模真实世界验证对于评估模型稳定性具有重要意义，但也说明随着任务复杂度和模态数量增加，数据缺失与混杂偏倚会更加突出。因此，如何在真实场景中利用不完整多模态数据，并维持模型外部验证稳定性，是当前重要技术瓶颈之一。

2. 临床效益验证的缺失与不足

目前大多数研究仍以 AUC 等区分度指标作为主要终点，但仅凭此类指标难以判断模型能否安全指导临床决策。对于直肠癌疗效评价和风险分层模型而言，除区分度外，还应关注预测概率与实际发生概率的

一致性，即模型校准度，并报告校准曲线、Brier 分数等指标，以判断模型输出风险是否可靠。只有当模型预测风险与真实风险较好匹配时，其结果才可能用于支持等待观察、追加治疗或强化随访等临床决策。

此外，模型不仅需具备预测能力，还应通过决策曲线分析（decision curve analysis，DCA）证明其在不同阈值概率下可带来临床净获益。Liu 等[19]与 Zhao 等[22]通过 DCA 评估多模态联合模型的潜在临床价值；Huang 等[26]进一步分析 IRIS-CRC 系统在治疗升降阶决策中的应用价值，该系统可帮助部分高危患者降阶以避免不必要治疗，也可提示部分低危患者需加强治疗。此类评估较单纯 AUC 更接近真实临床决策，但 DCA 仍属于模型层面的决策模拟，其对临床结局的真实改善仍需前瞻性研究进一步证实。

3. 算法架构演进与可解释性

未来直肠癌影像 AI 的发展，一方面依赖于算法架构的演进，另一方面也取决于其能否适应真实临床工作流程并满足监管要求。首先，直肠癌疗效评价具有高度复杂性，仅依赖单一高分辨率 MRI 难以全面刻画肿瘤异质性，未来模型需进一步融合影像、病理、临床检验、分子检测和基因组学等多模态信息。针对多模态数据常缺失的问题，可引入弱监督、半监督学习及缺失模态补偿策略。Qu 等[33]开发的 FLARE 模型即为处理不完整模态提供了新的思路，有望提高模型在非标准化真实场景中的可用性和鲁棒性。这提示医学基础模型和医学大模型在解决小样本、跨中心泛化和多模态融合方面具有潜力。

现有深度学习模型多聚焦于 pCR 预测、远处转移或生存分析等单一终点，而真实临床决策通常包括病灶识别、疗效评估、风险分层、预后判断和治疗获益评估等连续环节。因此，未来模型应向多任务联合建模发展，提高对复杂临床问题的综合表征能力。与此同时，当前 AI 模型的可解释性仍多停留于像素级可视化层面，如 Grad-CAM 热力图可提示模型关注区域[34]，但仍难以回答“为何这些影像表型与疗效或预后相关”的核心临床问题。更高层次的可解释性应进一步揭示影像特征背后的分子、细胞和组织机制。Cao 等[12]将 MRI 深度学习特征与单细胞 RNA 测序相映射，提示模型预测的非 pCR 倾向可能与髓系细胞富集和免疫抑制微环境相关，为验证模型预测的生物学合理性提供代表性思路。未来，影像 AI 解释应逐步构建“影像-病理-基因-微环境”的跨模态推理逻辑，使模型真正成为人类专家可理解、可信任的诊疗延伸。

4.真实世界部署与法规审批挑战

影像 AI 从实验室验证走向临床部署，还需解决工程自动化、人机交互和法规审批等问题。AI 工具不应以替代医师为目标，而需证明其作为辅助工具的临床赋能价值。Pang 等[31]的人机交互研究提示，AI 可辅助医师识别关键信息并改善诊断表现；Liu 等[24]构建的 MultiRecNet 则展示了从原始图像输入、病灶自动分割、特征提取到结局预测的端到端流程。这种工程自动化能将医生从繁琐的像素级标注中解放出来，使 AI 模型真正成为实用的临床工具。

作为可能影响治疗决策的医疗器械软件，AI 模型需满足监管机构对透明度、鲁棒性和可追溯性的要求。AI 模型的“黑箱”属性仍是其审批和推广的重要障碍。SHAP 分析（shapley additive explanations，SHAP）[35]等方法有助于量化特征贡献并增强医师信任，但仅靠算法解释仍不足以完成临床转化。Rai 等[7]的荟萃分析提示，尽管 MRI 影像组学模型总体具有较好的诊断性能，但相关研究普遍缺乏前瞻性验证、预设截断值分析和成本效益分析，方法学标准化和临床转化证据仍显不足。未来应建立标准化、自动化、可追溯的全流程质量控制体系，推动影像 AI 向合规化临床应用转化。

综上所述，当前直肠癌影像 AI 正处于向临床应用与真实世界部署过渡的临界点，但与临床落地实施间仍存鸿沟。未来需致力于建立高质量多中心数据集以验证泛化能力，推动多模态与多任务算法进阶，并实现人类专家先验知识与 AI 算法的深度融合改造。只有在数据、算法、医学知识和临床流程之间形成有效闭环，影像 AI 才能真正成为辅助医师进行直肠癌精准诊疗的可靠利器。

参考文献：

[1] GBD 2023 Cancer Collaborators. The global, regional, and national burden of cancer, 1990－2023, with forecasts to 2050: A systematic analysis for the global burden of disease study 2023[J]. Lancet, 2025, 406(10512): 1565-1586.

[2] Glynne-Jones R, Wyrwicz L, Tiret E, et al. Rectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up [J]. Ann Oncol, 2017, 28(suppl_4): iv22-iv40.

[3] Wang F, Chen G, Zhang Z, et al. The Chinese society of clinical oncology (CSCO): Clinical guidelines for the diagnosis and treatment of colorectal cancer, 2024 update[J]. Cancer Commun (Lond), 2025, 45(3): 332-379.

[4] Beets-Tan RG, Lambregts DM, Maas M, et al. Magnetic resonance imaging for the clinical management of rectal cancer patients: Recommendations from the 2012 european society of gastrointestinal and abdominal radiology (ESGAR) consensus meeting[J]. Eur Radiol, 2013, 23(9): 2522-2531.

[5] Jong BK, Yu ZH, Hsu YJ, et al. Deep learning algorithms for predicting pathological complete response in MRI of rectal cancer patients undergoing neoadjuvant chemoradiotherapy: A systematic review[J]. Int J Colorectal Dis, 2025, 40(1): 19.

[6] Shen H, Jin Z, Chen Q, et al. Image-based artificial intelligence for

the prediction of pathological complete response to neoadjuvant chemoradiotherapy in patients with rectal cancer: A systematic review and meta-analysis[J]. Radiol Med, 2024, 129(4): 598-614.

[7] Rai J, Mai DV, Drami I, et al. MRI radiomics prediction modelling for pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: A systematic review and meta-analysis[J]. Abdom Radiol, 2025, 50(11): 5103-5123.

[8] Shin J, Seo N, Baek SE, et al. MRI radiomics model predicts pathologic complete response of rectal cancer following chemoradiotherapy[J]. Radiology, 2022, 303(2): 351-358.

[9] 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.

[10] Jin C, Yu H, Ke J, et al. Predicting treatment response from longitudinal images using multi-task deep learning[J]. Nat Commun, 2021, 12(1): 1851.

[11] Ke J, Jin C, Tang J, et al. A longitudinal MRI-based artificial intelligence system to predict pathological complete response after neoadjuvant therapy in rectal cancer: A multicenter validation study[J]. Dis Colon Rectum, 2023, 66(12): e1195-e1206.

[12] Cao W, Chen H, Li J, et al. MR-DELTAnet: A longitudinal MRI-transformer model predicting pathological complete response and revealing immune microenvironment via scRNA-seq in locally advanced rectal cancer[J]. Adv Sci, 2026, 13(11): e17721.

[13] Cercek A, Lumish M, Sinopoli J, et al. PD-1 blockade in mismatch repair-deficient, locally advanced rectal cancer[J]. N Engl J Med, 2022, 386(25): 2363-2376.

[14] Williams CJM, Peddle AM, Kasi PM, et al. Neoadjuvant immunotherapy for dMMR and pMMR colorectal cancers: Therapeutic strategies and putative biomarkers of response[J]. Nat Rev Clin Oncol, 2024, 21(12): 839-851.

[15] Yang H, Wong C, Liang W, et al. A noninvasive AI model based on multi-dimensional MRI features for predicting tertiary lymphoid structures, immunotherapy response, and prognosis in rectal cancer[J]. Int J Surg, 2026, 112(2): 4317-4328.

[16] Zhang Y, Zhang X, Zhong X, et al. Immunophenotype-guided interpretable radiomics model for predicting neoadjuvant anti-PD-1 response in stage Ⅲ-Ⅳ d-MMR/MSI-H colorectal cancer[J]. J Immunother Cancer, 2025, 13(8): e011569.

[17] Hofheinz RD, Fokas E, Benhaim L, et al. Localised rectal cancer: ESMO clinical practice guideline for diagnosis, treatment and follow-up[J]. Ann Oncol, 2025, 36(9): 1007-1024.

[18] Meng X, Xia W, Xie P, et al. Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer[J]. Eur Radiol, 2019, 29(6): 3200-3209.

[19] Liu Z, Meng X, Zhang H, et al. Predicting distant metastasis and chemotherapy benefit in locally advanced rectal cancer[J]. Nat Commun, 2020, 11(1): 4308.

[20] Zhao R, Shen W, Zhao W, et al. Integrating radiomics, pathomics, and biopsy-adapted immunoscore for predicting distant metastasis in locally advanced rectal cancer[J]. ESMO open, 2025, 10(3): 104102.

[21] Chen X, Huang Y, Wee L, et al. Integrated analysis of radiomics, RNA, and clinicopathologic phenotype reveals biological basis of prognostic risk stratification in colorectal cancer[J]. Sci Bull (Beijing), 2024, 69(23): 3666-3671.

[22] Zhao Q, Zhong H, Guan X, et al. Role of microenvironment characteristics and MRI radiomics in the risk stratification of distant metastases in rectal cancer: A diagnostic study[J]. Int J Surg, 2025, 111(1): 200-209.

[23] Liu X, Zhang D, Liu Z, et al. Deep learning radiomics-based prediction of distant metastasis in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy: A multicentre study[J]. EBioMedicine, 2021, 69: 103442.

[24] Liu Z, Meng R, Ma Q, et al. Predicting recurrence in locally advanced rectal cancer using multitask deep learning and multimodal MRI[J]. Radiol Imaging Cancer, 2025, 7(3): e240359.

[25] Feng L, Liu Z, Li C, et al. Development and validation of a radiopathomics model to predict pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: A multicentre observational study[J]. Lancet Digit Health, 2022, 4(1): e8-e17.

[26] Huang YQ, Chen XB, Cui YF, et al. Enhanced risk stratification for stage Ⅱ colorectal cancer using deep learning-based CT classifier and pathological markers to optimize adjuvant therapy decision[J]. Ann Oncol, 2025, 36(10): 1178-1189.

[27] Jiang X, Zhao H, Saldanha OL, et al. An MRI deep learning model predicts outcome in rectal cancer[J]. Radiology, 2023, 307(5): e222223.

[28] Skrede OJ, DE Raedts, Kleppe A, et al. Deep learning for prediction of colorectal cancer outcome: A discovery and validation study[J]. Lancet, 2020, 395(10221): 350-360.

[29] FU J, Zhong X, Li N, et al. Deep learning-based radiomic features for improving neoadjuvant chemoradiation response prediction in locally advanced rectal cancer[J]. Phys Med Biol, 2020, 65(7): 075001.

[30] Hu T, Gong J, Sun Y, et al. Magnetic resonance imaging-based radiomics analysis for prediction of treatment response to neoadjuvant chemoradiotherapy and clinical outcome in patients with locally advanced rectal cancer: A large multicentric and validated study[J]. MedComm, 2024, 5(7): e609.

[31] Pang X, Chen X, Zeng G, et al. Interpretable artificial intelligence in assisting treatment response prediction for locally advanced rectal cancer after neoadjuvant chemoradiotherapy: A prospective, multicenter, human-model interaction study[J]. Int J Radiat Oncol Biol Phys, 2026, 125(3): 939-948.

[32] Chen X, Qiu M Y, Zhang JP, et al. Colorectal cancer detection using noncontrast CT and deep learning: A multicenter and international cohort study[J]. Ann Oncol, 2026, 37(8): 1144-1156.

[33] Qu L, Zhang C, Hou Y, et al. Foundation model-enabled multimodal deep learning for prognostic prediction in colorectal cancer with incomplete modalities: A multi-institutional retrospective study[J]. Adv Sci, 2026, 13(17): e10931.

[34] Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep Networks via gradient-based localization[J]. Int J Comput Vis, 2017: 618-626.

[35] Nohara Y, Matsumoto K, Soejima H, et al. Explanation of machine learning models using shapley additive explanation and application for real data in hospital[J]. Comput Methods Programs Biomed, 2022, 214: 106584.

(收稿日期: 2026-06-30 修回日期: 2026-08-13)