

DOI:10.16781/j.0258-879x.2020.03.0325

• 短篇论著 •

磁共振影像组学主成分分析法预测直肠癌新辅助放射化学治疗效果

臧书芹, 沈 浮, 郝 强*

海军军医大学(第二军医大学)长海医院影像医学科, 上海 200433

[摘要] **目的** 探讨基于MRI高分辨率T2加权图像的影像组学主成分分析(PCA)法对直肠癌新辅助放射化学治疗效果的预测价值。**方法** 回顾性分析2018年1月1日至2018年12月31日在我院因局部进展期直肠癌接受新辅助放射化学治疗后行直肠癌根治性切除术的80例患者资料,男60例、女20例,年龄为28~74岁,平均年龄为(56.2±9.9)岁。患者行新辅助放射化学治疗前接受3.0 T MRI检查,在高分辨率T2加权图像上提取影像组学特征,再采用PCA法进行特征降维,使用降维后的特征与病理完全缓解(pCR)标签建立logistic回归分类器模型,将样本随机分为训练集与测试集进行机器学习,分别绘制ROC曲线并计算AUC及灵敏度、特异度、准确度。**结果** MRI高分辨率T2加权图像共提取到1 409个影像组学特征,PCA法重新组合并选取了前5个最能代表整个影像组学特征矩阵的新特征,分别能代表整个影像组学特征矩阵中 $9.926\ 016\ 67\times 10^{-1}$ 、 $4.854\ 545\ 00\times 10^{-3}$ 、 $2.509\ 013\ 91\times 10^{-3}$ 、 $2.489\ 032\ 30\times 10^{-5}$ 、 $7.372\ 984\ 50\times 10^{-6}$ 的信息。Logistic回归分类器模型交叉验证测试集的平均AUC为0.761(95% CI: 0.694~0.828),灵敏度为90.3%,特异度为40.0%,准确度为79.0%。**结论** 基于MRI高分辨率T2加权影像组学PCA法对直肠癌新辅助放射化学治疗的疗效具有较好的预测价值。

[关键词] 直肠肿瘤;新辅助放射化学治疗;影像组学;磁共振成像;主成分分析

[中图分类号] R 735.37

[文献标志码] A

[文章编号] 0258-879X(2020)03-0325-05

Magnetic resonance imaging-based radiomics of rectal cancer: principal component analysis to predict treatment response after neoadjuvant chemoradiotherapy

ZANG Shu-qin, SHEN Fu, HAO Qiang*

Department of Radiology, Changhai Hospital, Naval Medical University (Second Military Medical University), Shanghai 200433, China

[Abstract] **Objective** To explore the value of radiomics with principal component analysis (PCA) based on magnetic resonance imaging (MRI) high-resolution T2-weighted images for predicting treatment response after neoadjuvant chemoradiotherapy (nCRT) of rectal cancer. **Methods** Clinical data of 80 patients with locally advanced rectal cancer, who received nCRT before radical resection from Jan. 1, 2018 to Dec. 31, 2018 in our hospital, were retrospectively analyzed. There were 60 males and 20 females with a mean age of (56.2±9.9) years (range, 28-74 years). All the patients underwent 3.0 T MRI examination before nCRT. The radiomics features were extracted from the high-resolution T2-weighted images. PCA was applied to reduce the dimension. The logistic regression classifier model was built using the dimension-reduced feature and pathologic complete response label. The samples were randomly divided into training set and testing set for machine learning. ROC curves were drawn and AUC, sensitivity, specificity and accuracy were calculated. **Results** A total of 1 409 radiomics features were extracted from MRI high-resolution T2-weighted images. Five optimal features, which could best represent the overall radiomics feature matrix, were recombined and selected by PCA method. They represented the information on $9.926\ 016\ 67\times 10^{-1}$, $4.854\ 545\ 00\times 10^{-3}$, $2.509\ 013\ 91\times 10^{-3}$, $2.489\ 032\ 30\times 10^{-5}$ and $7.372\ 984\ 50\times 10^{-6}$, respectively. The mean AUC of the logistic regression classifier model was 0.761 (95% CI 0.694-0.828), sensitivity was 90.3%, specificity was 40.0%, and accuracy was 79.0%. **Conclusion** Radiomics with PCA based on MRI high-resolution T2-weighted images has good predictive value of treatment response after nCRT of rectal cancer.

[Key words] rectal neoplasms; neoadjuvant chemoradiotherapy; radiomics; magnetic resonance imaging; principal component analysis

[Acad J Sec Mil Med Univ, 2020, 41(3): 325-329]

[收稿日期] 2019-08-12 [接受日期] 2019-12-23

[基金项目] 海军军医大学(第二军医大学)青年启动基金(2018QN05)。Supported by Youth Initiative Fund of Naval Medical University (Second Military Medical University) (2018QN05)。

[作者简介] 臧书芹,住院医师。E-mail: zangshuqinedri@163.com

*通信作者(Corresponding author)。Tel: 021-31162155, E-mail: haoqiang@189.cn

结直肠癌是全球最常见的消化道恶性肿瘤之一,发病率居第3位,死亡率居第2位^[1]。在我国,结直肠癌是近10年唯一发病率呈逐年上升的消化道肿瘤,其发病率、死亡率位列全部恶性肿瘤的第5位,在经济发达地区如上海的发病率已高居第2位,且局部进展期直肠癌比例高^[2]。对局部进展期直肠癌目前推荐全直肠系膜切除术(total mesorectal excision, TME)前进行新辅助放射化学治疗(neoadjuvant chemoradiotherapy, nCRT),该方案降低了术后复发率,提高了长期生存率及保肛率^[3]。文献报道,经nCRT及手术治疗后,15%~27%的局部进展期直肠癌患者在病理检查结果中显示无肿瘤细胞存活,即达到了病理完全缓解(pathologic complete response, pCR),这部分患者的长期预后明显优于有残留肿瘤细胞的患者^[4]。对局部进展期直肠癌的nCRT疗效进行准确评估及早期预测将使治疗方案更加个性化和有效化,应用影像学方法准确预测与评估直肠癌经nCRT后的转归非常迫切^[5]。

影像组学的出现和发展为我们提供了新的思路^[6]。影像组学从影像数据(CT、MRI、PET-CT等)中高通量地提取大量影像信息,通过肿瘤分割、特征提取与模型建立,对海量影像数据信息进行更高层次的挖掘、预测和分析,从而做出最准确的诊断,即将视觉影像信息转化为深层次特征进行量化研究。本研究基于高分辨率T2加权MRI,探讨影像组学模型对直肠癌nCRT疗效的预测效能。

1 资料和方法

1.1 临床资料 回顾性分析2018年1月1日至2018年12月31日在我院因局部进展期直肠癌接受nCRT后行直肠癌根治性切除术的113例患者资料。纳入标准:(1)术后病理检查证实为直肠腺癌;(2)单一病灶;(3)nCRT前行直肠高分辨率T2加权MRI检查;(4)nCRT后行根治性手术治疗,有详细病理报告结果。排除标准:(1)图像质量差,无法用于图像分割及影像组学特征提取(11例);(2)nCRT与手术间隔时间过短或过长(6例);(3)术前接受其他治疗,如靶向治疗或免疫治疗(16例)。

1.2 MRI检查方法 MRI扫描采用德国Siemens MAGNETOM Skyra 3.0 T超导型磁共振成像系

统。所有患者检查前4 h禁食。扫描前嘱患者使用1支开塞露(20 mL甘油)清洁肠道。采用腹部相控阵线圈。扫描序列:(1)常规序列包括矢状面T2加权脂肪抑制序列及横断面T1加权平扫、增强、弥散加权成像(diffusion-weighted imaging, DWI)。(2)高分辨率T2加权为斜横断面,扫描平面垂直于病灶所在肠管的长轴。扫描参数:重复时间(repetition time, TR)为4 000 ms,回波时间(time of echo, TE)为108 ms,矩形视野(field of view, FOV)为18 cm,矩阵为320×320,层厚为3 mm,层间距为0,层数为28,反转角度为150°,带宽为108 Hz/像素。(3)无脂肪抑制序列,采用全局自动校准部分并行采集(generalized autocalibrating partial parallel acquisition, GRAPPA)技术,加速因子为3,平均累加次数为2,回波链长度为16,扫描时长为250 s。

1.3 治疗方案与病理诊断标准

1.3.1 nCRT治疗方案 所有患者均行术前长程盆腔放射治疗,处方剂量为45~50 Gy,分25~28次,共5~5.5周完成。化学治疗于第1次盆腔放射治疗开展时实施,持续静脉滴注5-氟尿嘧啶(5-fluorouracil, 5-FU)或口服卡培他滨。nCRT结束后间隔5~12周行TME手术^[3]。

1.3.2 病理诊断 根据美国癌症联合委员会(American Joint Committee on Cancer, AJCC)第8版的分类标准^[7],将直肠癌病理肿瘤退缩分级分为0~3级。0级(完全反应):无可见的肿瘤细胞;1级(几乎完全反应):仅可见单个细胞,罕见肿瘤细胞簇;2级(部分反应):尚有肿瘤残留,也有肿瘤退缩,但退缩程度不如前者;3级(反应差或无反应):广泛的肿瘤残留,几乎没有肿瘤退缩。根据标准进行病理诊断,记录肿瘤退缩程度,包括纤维化百分比、残余肿瘤的存在和淋巴结状态。nCRT后病理分期为T0N0的患者纳入pCR组,其余患者纳入非pCR组。

1.4 影像组学方法

1.4.1 图像分割与特征提取 将高分辨率T2加权MRI的DICOM图像导入后处理平台[大数据智能分析云平台,慧影医疗科技(北京)有限公司研发],由1位放射科医师采用手动方式在高分辨率T2加权MRI图像上沿病灶边缘逐层勾画感兴趣区

(region of interest, ROI), 自动生成病灶的三维容积感兴趣区 (volume of interest, VOI), 由另1位高年资放射科医师检查勾画结果。使用上述平台提取的VOI影像组学特征如下。(1) 基于特征类: ①一阶统计量特征 (first-order statistic); ②形状特征 (shape-based); ③纹理特征 (texture)。包括灰度级共生矩阵 (gray-level co-occurrence matrix, GLCM)、灰度级长度矩阵 (gray-level run length matrix, GLRLM)、灰度级形状矩阵 (gray-level size zone matrix, GLSZM) 和领域灰度差分矩阵 (neighborhood gray-tone difference matrix, NGTDM)。(2) 基于过滤器类, 包括小波分析 (wavelet)、拉普拉斯转换 (laplacian transform)、对数特征 (logarithm)、指数特征 (exponential)、明可夫斯基函数 (Minkowski functional)、分形维数 (fractal dimension) 等。

1.4.2 特征降维及模型构建 提取影像组学特征后, 采用主成分分析 (principal component analysis, PCA) 法进行特征值降维, 使用降维后的

特征与pCR标签建立logistic回归分类器模型, 将样本随机分为训练集(80%)与测试集(20%)进行机器学习, 分别绘制ROC曲线并计算准确度。logistic回归分类器模型的测试集采用5折交叉验证 (cross-validation), 计算AUC、灵敏度及特异度。

1.5 统计学处理 应用SPSS 16.0软件进行统计学分析。呈正态分布的计量资料以 $\bar{x} \pm s$ 表示, 两组间比较采用独立样本 t 检验; 计数资料以例数和百分数表示, 两组间比较采用 χ^2 检验。检验水准(α)为0.05。

2 结果

2.1 患者一般资料 最终纳入80例患者, 男60例、女20例, 年龄为28~74岁, 平均年龄为(56.2±9.9)岁。nCRT至手术间隔时间为31~85 d, 平均(57.1±8.5) d。pCR组15例(18.75%), 非pCR组65例(81.25%)。见表1, 两组性别构成、年龄、nCRT至手术间隔时间差异均无统计学意义(P 均>0.05)。

表1 两组患者一般资料比较

分组	N	性别 n		年龄(岁), $\bar{x} \pm s$	nCRT至手术间隔(d), $\bar{x} \pm s$
		男	女		
非pCR组	65	49	16	57.1±8.8	57.5±9.1
pCR组	15	11	4	56.0±10.2	56.3±7.4
统计值		$\chi^2=0.027$		$t=0.403$	$t=0.411$
P值		0.869		0.688	0.683

pCR:病理完全缓解;nCRT:新辅助放射化学治疗

2.2 影像组学结果

2.2.1 影像组学特征 共提取到1409个组学特征, 采用PCA法进行特征值降维, 将影像组学特征重新组合为新的特征, 选取前5个最能代表整个影像组学特征矩阵的特征, 保证得到整个影

像组学特征矩阵中99%以上的信息(图1)。选取的5个特征分别能代表整个影像组学特征矩阵中 $9.926\ 016\ 67 \times 10^{-1}$ 、 $4.854\ 545\ 00 \times 10^{-3}$ 、 $2.509\ 013\ 91 \times 10^{-3}$ 、 $2.489\ 032\ 30 \times 10^{-5}$ 、 $7.372\ 984\ 50 \times 10^{-6}$ 的信息。

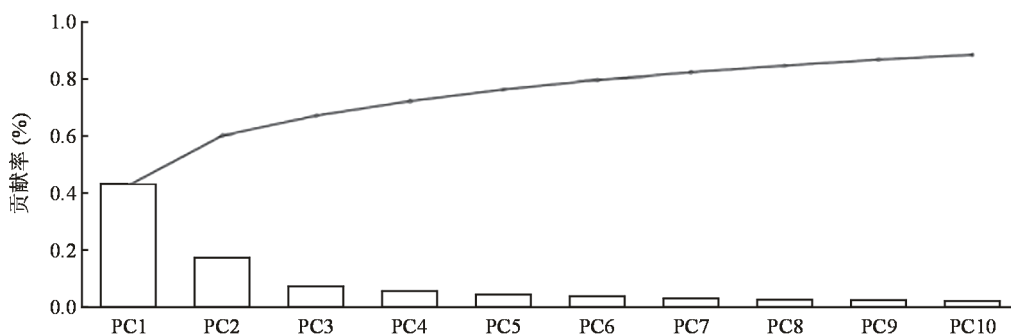


图1 主成分分析(PCA)法特征图

PC:主成分.PCA法重新组合并选取前5个最能代表整个影像组学特征矩阵的特征,该线图表示各主成分的累积贡献率

2.2.2 分类器模型预测表现 Logistic回归分类器模型训练集和5折交叉验证测试集的ROC曲线见

图2。训练集的平均AUC(95%CI)、灵敏度、特异度、准确度分别为0.678(0.443~0.883)、

61.0%、60.0%、84.0%，测试集的平均AUC (95% CI)、灵敏度、特异度、准确度分别为 0.761

(0.694~0.828)、90.3%、40.0%、79.0%。

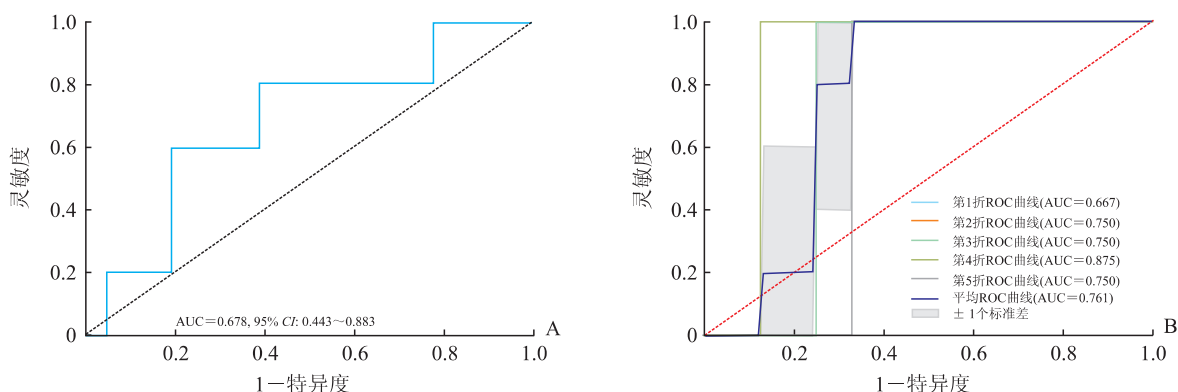


图2 训练集(A)和5折交叉验证测试集(B) ROC曲线

ROC:受试者工作特征;AUC:曲线下面积;CI:置信区间

3 讨论

术前nCRT结合TME是局部进展期直肠癌患者的标准治疗方式^[3]。nCRT疗效的预测与评估是决定下一步治疗决策的重要依据。一项meta分析指出，pCR的患者可不接受TME治疗，而进行“观察与等待(watch-and-wait)”，该类患者与手术治疗的患者相比，总生存期并无明显不同^[8]。然而pCR目前仅能在根治性手术后确认，临床判断为临床完全缓解(clinical complete response, cCR)的患者中仅有25%实际为pCR^[9]。对非pCR的患者需要及时调整治疗方案，以期取得较好的治疗效果。因此为了加强长期肿瘤患者慢性化管理的安全性，筛选有可能为pCR的cCR患者即成了关键点。术后病理学检查是判断nCRT肿瘤退缩的“金标准”，但目前临床依据的术前非病理学评估手段和标准尚缺乏共识^[10]。

MRI作为目前推荐度最高且常用的影像学技术，在直肠癌的术前分期及再分期中发挥了重要作用，是临床判断cCR的重要依据。MRI高分辨率T2加权序列根据形态学特征，能够判断肿瘤的浸润深度、壁外血管侵犯与否、环周切缘情况及淋巴结转移情况，是直肠癌分层治疗不可或缺的部分。但nCRT后，随着肿瘤细胞的坏死和间质纤维化等反应，单纯依靠形态学评估nCRT后有无癌灶残存较为困难，即使采用常规DWI及表面扩散系数(apparent diffusion coefficient, ADC)图也不能完全区分肿瘤组织退缩的改变或治疗后的炎性水肿及

坏死^[11]。

影像组学通过计算机技术将大量医学图像转化为可量化的数据，可实现对疾病的精准预测、诊断及预后评估等。以个体化医疗为基础的肿瘤精准治疗依赖于基因组学、蛋白质组学等组学技术^[12]，而肿瘤组织在空间和时间上均存在异质性，通过活组织检查等获取的小部分肿瘤组织标本并不能代表完全的肿瘤组织特点，因而基因组学与蛋白质组学的临床常规应用价值有限^[13]。非侵入性的影像学检查对于肿瘤的整体评估具有重要作用，具有无创性和可重复性^[14]。影像组学技术不仅规避了观察者本身对于影像特征解读的主观偏向，还能深度挖掘并整合图像中大量的人眼无法识别和区分的数字化信息，从而提高术前影像诊断效能^[15]。目前，国内外学者对结直肠肿瘤的影像组学研究主要集中在纹理特征提取、良恶性病变鉴别、术前分期、预后预测等方面。

本研究基于MRI进行影像组学研究，在高分辨率T2加权图像上勾画VOI进行特征提取，之后采用PCA法进行特征值降维，将原来众多具有一定相关性的影像组学特征重新组合成一组新的互相无关的综合指标替代原来的特征。Logistic回归算法是监督式学习(supervised learning)模式众多算法中一个简单有效的分类方法，属于判别式模型，它通过将数据拟合进一个logistic函数而预估一个事件出现的概率。其优点是分类时计算量小，速度快，实现简单，应用广泛；缺点是当特征空间很大时容易欠拟合，不能很好地处理大量多类特

征或变量。本研究利用PCA法进行降维,选取前5个最能代表整个影像组学特征矩阵的特征,极大缩减了特征量,当采用logistic回归算法时则规避了其缺点。同时,为了提高准确度,通过交叉验证进行测试,调整参数确保算法达到最优解,结果显示logistic回归分类器模型交叉验证测试集AUC为0.761,准确度为79.0%,表明logistic回归分类器模型对直肠癌nCRT后的pCR预测与评估具有诊断价值。

本研究仍存在以下不足:(1)本研究为回顾性研究,可能存在选择性偏倚,未来有必要进行外部验证及前瞻性多中心研究;(2)直肠癌病灶形态往往非常不规则,采用手动勾画VOI而未进行一致性检验,这可能会影响机器学习的准确性;(3)本研究的灵敏度较高,即漏诊或低估pCR的可能性降低,但是特异度不高,意味着误诊或高估pCR的可能性较高,这可能与样本量较小有关,今后需扩大样本量以减少数据量对模型的影响。

综上所述,基于MRI高分辨率T2加权图像的影像组学可用于预测直肠癌nCRT后的pCR状态,对临床选择个体化治疗方案有重要价值。

[参考文献]

- [1] BRAY F, FERLAY J, SOERJOMATARAM I, SIEGEL R L, TORRE L A, JEMAL A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68: 394-424.
- [2] 王磊,刘志华,汪建平.我国结直肠癌的诊疗现状[J]. *中华实验外科杂志*, 2015, 32: 677-679.
- [3] BENSON A B, VENOOK A P, AL-HAWARY M M, CEDERQUIST L, CHEN Y J, CIOMBOR K K, et al. Rectal cancer, version 2.2018, NCCN clinical practice guidelines in oncology[J]. *J Natl Compr Canc Netw*, 2018, 16: 874-901.
- [4] MAAS M, NELEMANS P J, VALENTINI V, DAS P, RÖDEL C, KUO L J, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data[J]. *Lancet Oncol*, 2010, 11: 835-844.
- [5] 伍兵,方鑫.直肠癌新辅助治疗后的MR评估[J]. *中国普外基础与临床杂志*, 2018, 25: 1299-1303.
- [6] 黄冠,蒲红,尹芳艳,李小雪,朱浩.影像组学的临床应用与进展[J]. *实用医院临床杂志*, 2019, 16: 193-197.
- [7] AMIN M B, GREENE F L, EDGE S B, COMPTON C C, GERSHENWALD J E, BROOKLAND R K, et al. The eighth edition AJCC cancer staging manual: continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging[J]. *CA Cancer J Clin*, 2017, 67: 93-99.
- [8] DOSSA F, CHESNEY T R, ACUNA S A, BAXTER N N. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete response following neoadjuvant chemoradiation: a systematic review and meta-analysis[J]. *Lancet Gastroenterol Hepatol*, 2017, 2: 501-513.
- [9] HIOTIS S P, WEBER S M, COHEN A M, MINSKY B D, PATY P B, GUILLEM J G, et al. Assessing the predictive value of clinical complete response to neoadjuvant therapy for rectal cancer: an analysis of 488 patients[J/OL]. *Ann Chir*, 2002, 127: 492. doi: 10.1016/S0003-3944(02)00810-6.
- [10] SMITH J J, CHOW O S, GOLLUB M J, NASH G M, TEMPLE L K, WEISER M R, et al. Organ preservation in rectal adenocarcinoma: a phase II randomized controlled trial evaluating 3-year disease-free survival in patients with locally advanced rectal cancer treated with chemoradiation plus induction or consolidation chemotherapy, and total mesorectal excision or nonoperative management[J/OL]. *BMC Cancer*, 2015, 15: 767. doi: 10.1186/s12885-015-1632-z.
- [11] 张胜潮,汪嵘,郭琴,李耀平,田利军.高分辨率磁共振成像在局部进展期直肠癌新辅助治疗后的应用价值[J]. *中国药物与临床*, 2019, 19: 724-725.
- [12] PARMAR C, LEIJENAAR R T, GROSSMANN P, RIOS VELAZQUEZ E, BUSSINK J, RIETVELD D, et al. Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer[J/OL]. *Sci Rep*, 2015, 5: 11044. doi: 10.1038/srep11044.
- [13] LAMBIN P, RIOS-VELAZQUEZ E, LEIJENAAR R, CARVALHO S, VAN STIPHOUT R G, GRANTON P, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. *Eur J Cancer*, 2012, 48: 441-446.
- [14] 孙钢.放射组学的兴起及其在消化系统肿瘤中的应用[J/CD]. *中华消化病与影像杂志(电子版)*, 2017, 7: 145-149.
- [15] 冯晓源.精准医疗,影像先行[J]. *中华放射学杂志*, 2016, 50: 1-2.

[本文编辑] 杨亚红