

基于高分辨  $T_2$ WI 的影像组学对直肠癌术前分期的应用价值

王进, 李智慧, 沈浮, 陆建平

**【摘要】** 目的:探讨基于高分辨  $T_2$ WI 影像组学在直肠癌术前 T 分期中的应用价值。方法:回顾性分析 2017 年 1 月—2018 年 12 月在上海长海医院经手术病理证实且术前行 3T 磁共振高分辨  $T_2$ WI 的直肠癌患者资料。根据病理 T 分期结果将 T1 和 T2 期患者归为未突破肌层组, T3 和 T4 期患者归为突破肌层组。在高分辨  $T_2$ WI 图像上手动勾画病灶 ROI 后提取影像组学特征,之后采用 LASSO 算法进行降维选择出对病理 T 分期有价值的特征。将所选样本按 7:3 的比例随机分为训练集与验证集进行机器学习,构建支持向量机(SVM)分类器模型,得到训练集和验证集的 ROC 曲线下面积(AUC)、敏感度、特异度及 95% 置信区间。结果:纳入病例共 154 例,男 92 例,女 62 例,平均年龄( $56.5 \pm 10.2$ ) 岁,其中 T1 期 26 例, T2 期 41 例, T3 期 76 例, T4 期 11 例。训练集 106 例(未突破肌层组 46 例,突破肌层组 60 例),验证集 48 例(未突破肌层组 21 例,突破肌层组 27 例)。LASSO 算法筛选出 5 个与 T 分期相关特征。得到 SVM 模型验证集的 AUC、95% CI、敏感度及特异度分别为 0.730、0.583~0.848、73.1%、66.7%。结论:基于高分辨  $T_2$ WI 的影像组学 SVM 模型可以判断直肠癌病灶是否突破固有肌层,对于直肠癌术前 T 分期具有较高的参考价值。

**【关键词】** 直肠肿瘤; 影像组学; 磁共振成像; T 分期

**【中图分类号】** R445.2; R735.37 **【文献标识码】** A **【文章编号】** 1000-0313(2019)11-1251-04

DOI:10.13609/j.cnki.1000-0313.2019.11.016

开放科学(资源服务)标识码(OSID):



**The value of high resolution  $T_2$ WI-based radiomics in the preoperative staging of rectal cancer** WANG Jin, LI Zhi-hui, SHEN Fu, et al. Department of Imaging Medicine, Changhai Hospital Affiliated to Naval Medical University, Shanghai 200433, China

**【Abstract】 Objective:** The purpose of this study was to explore the application value of radiomics based on high resolution magnetic resonance  $T_2$ WI in the preoperative T staging of rectal cancer. **Methods:** One hundred and fifty-four pathologically confirmed patients (92 males and 62 females, average age  $56.5 \pm 10.2$  years) with rectal cancer were analyzed retrospectively, who underwent 3-Tesla high-resolution magnetic resonance  $T_2$ WI before operation from January 2017 to December 2018. Patients of stage T1-2 were classified as the group of non-breakthrough of the muscularis propria layer; patients of stage T3-4 were classified as the group of breakthrough. The lesions were segmented to draw the outline of ROI on high resolution  $T_2$ W images. The radiomics features were extracted from the VOI, which were selected by the LASSO (least absolute shrinkage and selection operator) method to best predict the T stage. The samples were randomly divided into training set (70%) and test set (30%) for machine learning. The SVM (support vector machine) classifier model was used to obtain the receiver operating characteristic (ROC) curves and to calculate the area under the curve (AUC), 95% CI, sensitivity and specificity. **Results:** There were 26 patients with T1, 41 patients with T2, 76 patients with T3 and 11 with T4, including 106 cases of training set (46 cases of non-breakthrough and 60 cases of breakthrough) and 48 cases of test set (21 cases of non-breakthrough and 27 cases of breakthrough). Five features were extracted from high-resolution  $T_2$ WI related to the T stage by the LASSO method. For the SVM classifier model test set, the AUC (95% CI) was 0.730 (0.583~0.848), sensitivity 73.1% and specificity 66.7%. **Conclusion:** The SVM model of radiomics based on

作者单位:200433 上海,海军军医大学附属长海医院医学影像科

作者简介:王进(1989-),男,江苏盐城人,住院医师,主要从事 MR 诊断工作。

通讯作者:李智慧, E-mail: wisdom1213@163.com

基金项目:海军军医大学青年启动基金(2018QN05)

high-resolution T<sub>2</sub>WI can be used to evaluate whether the muscularis propria layer is involved, with good preoperative diagnostic performance for T stage in rectal cancer.

**【Key words】** Rectal neoplasms; Radiomics; Magnetic resonance imaging; T stage

直肠癌是世界范围内最常见的恶性肿瘤之一。在我国,随着人们生活习惯的改变及人口的老龄化,直肠癌发病率逐年增高,全国结直肠癌发病率及死亡率均位列第五位<sup>[1]</sup>。直肠癌患者最佳治疗方案的选择是一个非常复杂的过程,准确的术前评估是指导治疗决策的重要依据。肿瘤的精确分期可以提示患者治疗后的局部复发和远处转移风险,从而制定合理的治疗方案,达到最佳的个体化治疗策略。但目前传统影像学方法一般仅在形态学上对肿瘤进行术前评估,缺乏定量评估。基于精准医疗的要求,传统的影像医学模式亟待改变,必须与分子生物学、分子病理学结合,必须与新型治疗模式、预后判断和康复结合<sup>[2]</sup>。放射组学(radiomics)应用大量的自动化数据特征化算法,将影像兴趣区(region of interest, ROI)的影像数据转化为具有高分辨率的、可发掘的特征空间数据,具有数据维度高、定量分析等优点,可与传统影像学及分子生物学、分子病理学、信息科学相整合,采用大数据分析方法将其用于影像诊断、疗效评估、预后预测等,有望成为精准影像医学的重要基石<sup>[3-5]</sup>。因此本研究基于磁共振高分辨率 T<sub>2</sub>WI 图像,结合影像组学的方法,希望得到直肠癌术前 T 分期的影像组学特征,探讨基于高分辨率 T<sub>2</sub>WI 的影像组学方法对直肠癌特定病理特征的应用价值,并建立和验证影像组学直肠癌术前诊断模型。

## 材料与方法

### 1. 患者资料

回顾性分析 2017 年 1 月—2018 年 12 月在上海长海医院经手术病理证实且术前行直肠高分辨率 T<sub>2</sub>WI 检查的直肠癌患者的病例资料,获取原始 DICOM 数据并记录术后病理报告结果。排除标准:①图像质量差,如明显的伪影导致病灶显示欠清,无法准确勾画 VOI;②患者手术前或 MRI 检查前后接受过任何的全身或局部治疗。根据美国癌症联合委员会(American Joint Committee on Cancer, AJCC)第七版的分类标准进行病理诊断,记录患者术后病理 T 分期,将 T1~T2 期分为未突破肌层组, T3~T4 期分为突破肌层组。

### 2. 检查方法

磁共振扫描采用 Siemens MAGNETOM Skyra 3T 磁共振成像系统。所有患者检查前 4 h 禁食。检查前嘱患者使用一支开塞露(20 mL 甘油)清洁肠道。采用腹部相控阵线圈。扫描序列均采用常规直肠

MRI 及高分辨率 T<sub>2</sub>WI 序列。常规序列包括矢状面 T<sub>2</sub>WI 脂肪抑制序列,横轴面 T<sub>1</sub>WI 及增强扫描、DWI。高分辨 T<sub>2</sub>WI 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:TR 4000 ms, TE 108 ms, FOV 18 cm, 矩阵 320 × 320, 层厚 3 mm, 层间距 0, 层数 28 层, 反转角度 150°, 带宽 108 Hz/像素。无脂肪抑制, GRAPPA 并行采集模式, 加速因子为 3, 平均累加次数 2, 回波链长度 16, 扫描时长为 4 分 10 秒。

### 3. 影像组学方法

①病灶 VOI 分割和特征提取:将获取到的高分辨率 T<sub>2</sub>WI 原始 DICOM 图像导入后处理平台(大数据智能分析云平台, 慧影医疗科技有限公司, 北京), 一名放射科医师采用手动方式在 T<sub>2</sub>WI 图像上沿病变边缘逐层勾画兴趣区(region of interest, ROI), 计算机自动生成病灶的三维容积兴趣区(volume of interest, VOI), 另一名高年资放射科医师检查勾画结果。将 VOI 使用上述平台自动提取影像组学特征, 分为强度统计特征(19 个)、形状特征(15 个)、纹理特征, 其中纹理特征包括灰度共生矩阵(gray level co-occurrence matrix, GLCM)27 个、灰度游程矩阵特征(gray level run length matrix, GLRLM)16 个及灰度区域矩阵(gray level size zone matrix, GLSZM)16 个;以及对图像进行对数滤波、指数滤波、平方滤波、平方根滤波、小波变换(wavelet)等 8 种方式处理后再提取强度特征和纹理特征, 共 1029 个放射组学特征。

②特征选择及降维:将上述获取到的全部特征采用 LASSO 算法(Least absolute shrinkage and selection operator, 即最小绝对收缩算子)进行特征值降维, 筛选各个特征在最佳平均最小误差时的变量系数(图 1), 选择对直肠癌术后 T 分期诊断有价值的特征。

③机器学习:将特征降维所获得的影像组学特征与术后 T 分期构建支持向量机(support vector machine, SVM)分类器模型。将全部样本分为训练集和验证集, 随机选取 70% 病例作为训练集进行机器学习(machine learning, ML), 并采用 5 折交叉验证方式进行验证以得出稳定的结果, 利用剩余的 30% 病例作为验证集进行验证。使用受试者工作曲线(receiver operating characteristic curve, ROC)曲线下面积(area under the curve, AUC)以及敏感度和特异度评估机器学习模型的预测能力。

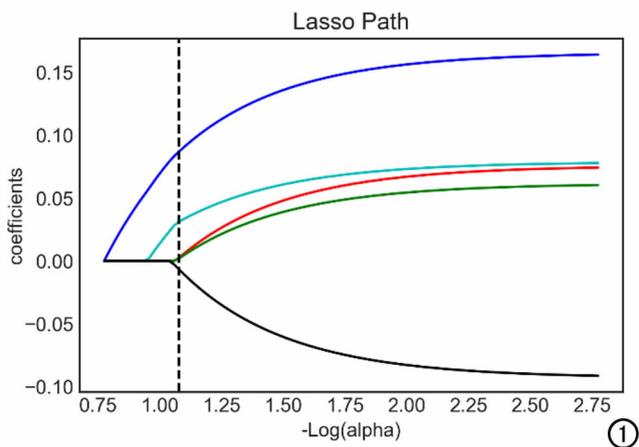


图 1 LASSO 算法系数解路径图。

## 结果

### 1. 临床资料

共纳入 154 例患者,男 92 例,女 62 例,年龄 32~78 岁,平均(56.5±10.2)岁,其中 T1 期 26 例,T2 期 41 例,T3 期 76 例,T4 期 11 例。未突破肌层组共 67 例,突破肌层组共 87 例。经组织病理证实直肠腺癌 142 例,黏液腺癌 10 例,印戒细胞癌 2 例。全部手术病例均为单发病灶,环周切缘(circumferential resection margin, CRM)均为阴性。按 7:3 的比例随机分为训练集和验证集,其中训练集 106 例(其中未突破肌层组 46 例,突破肌层组 60 例),验证集 48 例(其中未突破肌层组 21 例,突破肌层组 27 例)。

### 2. 影像组学特征

本组病例共生成的 VOI 为 154 个,共提取到 1029 个影像组学特征。通过 LASSO 算法筛选后,最终得到与 T 分期鉴别相关的特征 5 个: original\_shape\_size、waveletHLH\_firstorder\_medianvalue、waveletLLL\_glszm\_SZE、waveletLLL\_glszm\_SZLGE 和 waveletHLL\_glcm\_svarh。

### 3. 分类器预测表现

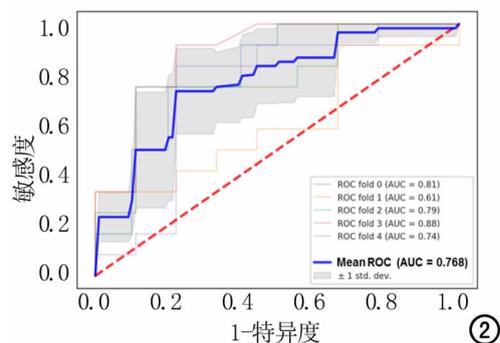


图 2 SVM 分类器训练集 ROC 曲线(5 折交叉验证)。

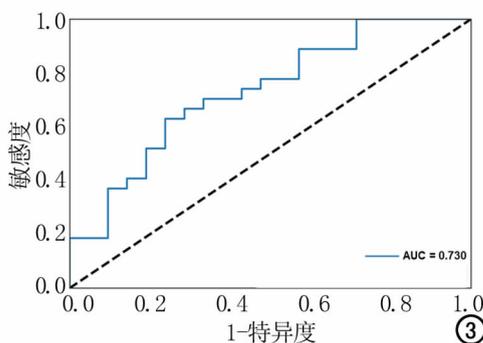


图 3 SVM 分类器验证集 ROC 曲线。

上述 5 个影像组学特征与 T 分期构建的 SVM 分类器,训练集的平均 ROC 曲线的 AUC、95% CI、敏感度及特异度为 0.768、0.696~0.860、73.3% 和 73.9% (图 2);验证集的 AUC、95% CI、敏感度及特异度为 0.730、0.583~0.848、73.1% 和 66.7% (图 3)。

## 讨论

直肠癌的术前诊断对于临床制定合理的个体化治疗决策是尤为重要的依据,依赖于精准的术前分期,包括肿瘤的浸润深度(T 分期)等反应肿瘤侵袭性及影响预后的因素。术前推荐采用 MRI 进行评估,但目前依靠传统影像学方法一般仅在形态学上进行评估,有一定的主观性,缺乏定量评估。T2 期/T3 期的分期错误是较为常见的,通常 T3 期提示较高的局部复发及远处转移风险,因此很多文献强调 T2 期与 T3 期鉴别诊断<sup>[6]</sup>。两者在 MRI 上最重要的鉴别标准是病灶是否突破固有肌层,周围系膜脂肪是否受到侵犯,而肠周线状或索条状的信号因局部纤维化或炎症也可有类似的表现,文献认为有时很难区分真正的肿瘤浸润与结缔组织增生性反应<sup>[7]</sup>,因此对 T2 期和早期 T3 期肿瘤常常很难鉴别。

现有研究证实影像组学可以为影像诊断提供新的思路 and 工具<sup>[8]</sup>,影像组学(radiomics)方法是指对 CT、MRI 和 PET 等大量医学图像高通量的提取定量影像学特征并进行分析,找到疾病的影像学标识物,从而实现对疾病的精准预测、诊断及预后评估等<sup>[4]</sup>。非侵入性的影像组学对于肿瘤的整体评估具有巨大的潜能,可对患者进行无损伤和可重复性检查<sup>[5]</sup>。目前,国内外学者对结直肠肿瘤的研究主要集中在病灶纹理特征提取、肿瘤与非肿瘤性病变的鉴别、结直肠癌术前分期、淋巴转移评估等方面<sup>[9-12]</sup>。

本研究对 154 例直肠癌病变的 T<sub>2</sub>WI 高分辨率图像进行特征提取,初始共 1029 个特征,包含两类特征。基于特征类:①一阶统计量特征(first order statistic),

描述与 ROI 内的体素强度分布有关的特征,不包含它们之间相互的空间作用。②形状特征(shape based),包括描述 VOI 大小、形态等特征,例如体积、表面积、直径、致密度、偏心度、球形度等。③纹理特征(texture),是描述体素空间分布强度等级的特征,被视为一个灰度级,是一种视觉感知的图像局部特征的综合,描述同

质性与异质性之间关系,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)。基于过滤器类:小波变换(wavelets)、拉普拉斯转换(laplacian transforms)、平方根(square root)、平方值(square)、对数特征(logarithm)、指数特征(exponential)等,表示通过不同滤波器的变换,可以表现出一些原图无法表现出的特征。

针对肿瘤是否侵犯突破肌层的判断,本研究根据术后病理结果将 T1~T2 期分为未突破肌层组, T3~T4 分为突破肌层组,经 LASSO 算法降维后最终得到与直肠癌 T 分期鉴别相关的特征 5 个。机器学习结果显示 SVM 分类器验证集的 AUC 为 0.730, 敏感度 73.1%, 特异度 66.7%, 表明该分类器有优秀的判断肿瘤浸润是否突破肌层的能力。

SVM 属于监督式学习,在对给定的数据进行逼近的精度与逼近函数的复杂性相结合来寻求最佳比例,以便取得最好的泛化能力。SVM 算法简单地说,就是升维和线性化。升维有可能会引起“维数灾难”,人们很少问津,但是 SVM 方法通过应用核函数的展开定理巧妙地解决了因低维空间向高维空间做映射而增加计算复杂性这个难题:不需要知道非线性映射的显式表达式;由于是在高维特征空间中建立线性学习机,所以与线性模型相比,不仅几乎不会增加计算的复杂性,还会在某种程度上规避了“维数灾难”,这一方面核函数的展开和计算理论功不可没。

目前文献报道该模型应用于不同类型肺腺癌的鉴别<sup>[13]</sup>、膀胱癌分级评估<sup>[14]</sup>以及前列腺癌的检测<sup>[15]</sup>等方面。基于 MRI 的影像组学技术,通过提取医学图像中的大量影像学特征并进行高通量数据分析,建立分类模型,不仅规避了观察者本身对于图像解读的主观偏向,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,从而提高术前直肠癌分期的诊断效能,拥有了能够量化分析肿瘤异质性的能力。

本研究的局限性主要在于:第一,影像组学特征是从手动勾画的 VOI 中提取,很难避免肠壁形变对 VOI 产生的影响,且容易出现主观错误,这可能会影响纹理特征的准确性;第二,本研究属于回顾性研究,可能存在选择偏倚,没有纳入部分未行手术无病理结果的患者,这导致建立模型的样本量仍不够大且分布欠均衡,需扩大样本量以减少数据量对模型准确性的影响;第三,本研究是利用影像组学 SVM 模型对直肠癌的初步探索,尚未对 N 分期及远处转移进行分析,今后还需进一步研究。

综上所述,基于磁共振高分辨 T<sub>2</sub>WI 的影像组学 SVM 模型有助于判断肿瘤浸润是否突破固有肌层,对直肠癌术前 T 分期的诊断具有价值,可以作为直肠癌术前评估的有效手段。

#### 参考文献:

- [1] Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015 [J]. CA Cancer J Clin, 2016, 66(2): 115-132.
- [2] 冯晓源. 精准医疗, 影像先行[J]. 中华放射学杂志, 2016, 50(1): 1-2.
- [3] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. Eur J Cancer, 2012, 48(4): 441-446.
- [4] 孙航, 李宏, 张亭亭, 等. Radiomics 方法研究应用进展[J]. 肿瘤, 2017, 37(10): 1092-1099.
- [5] 孙钢. 放射组学的兴起及其在消化系统肿瘤中的应用[J]. 中华消化病与影像杂志(电子版), 2017, 7(4): 145-149.
- [6] Nougaret S, Reinhold C, Mikhael HW, et al. The Use of MR Imaging in Treatment Planning for Patients with Rectal Carcinoma: Have You Checked the “DISTANCE” [J]. Radiology, 2013, 268: 330-344.
- [7] Beets-Tan RG, Beets GL, Vliegen RF, et al. Accuracy of magnetic resonance imaging in prediction of tumour-free resection margin in rectal cancer surgery[J]. Lancet, 2001, 357: 497-504.
- [8] 李华秀, 李振辉, 王关顺. 影像组学在消化道系统的应用进展[J]. 中国临床医学影像杂志, 2017, 28(9): 672-674, 679.
- [9] Hu P, Wang J, Zhong H, et al. Reproducibility with repeat CT in radiomics study for rectal cancer[J]. Oncotarget, 2016, 7(44): 71440-71446.
- [10] Song B, Zhang G, Lu H, et al. Volumetric texture features from higher-order images for diagnosis of colon lesions via CT colonography[J]. Int J Comput Assist Radiol Surg, 2014, 9(6): 1021-1031.
- [11] Liang C, Huang Y, He L, et al. The development and validation of a CT-based radiomics signature for the preoperative discrimination of stage I-II and stage III-IV colorectal cancer[J]. Oncotarget, 2016, 24, 7(21): 31401-31412.
- [12] Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer[J]. J Clin Oncol, 2016, 34(18): 2157-2164.
- [13] Yuan M, Zhang YD, Pu XH, et al. Comparison of a radiomic biomarker with volumetric analysis for decoding tumour phenotypes of lung adenocarcinoma with different disease-specific survival [J]. Eur Radiol, 2017, 27(11): 4857-4865.
- [14] Zhang X, Xu X, Tian Q, et al. Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging[J]. J Magn Reson Imaging, 2017, 46(5): 1281-1288.
- [15] Khalvati F, Wong A, Haider MA. Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models[J]. BMC Med Imaging, 2015, 5: 27-40.