

T2WI 序列纹理分析联合机器学习的影像组学术前预测肾透明细胞癌的病理分级

石博文¹, 叶靖^{2*}, 段绍峰³, 赵泓博¹, 刘路路⁴, 许晴⁴

1. 大连医科大学, 辽宁大连 116044; 2. 江苏省苏北人民医院影像科, 江苏扬州 225001; 3. 上海 GE 医疗公司, 上海 210000;

4. 扬州大学, 江苏扬州 225001; *通讯作者 叶靖 18051061289@163.com

【摘要】目的 探讨基于磁共振 T2 加权图像纹理分析联合机器学习的影像组学评价肾透明细胞癌 (ccRCC) 病理分级的价值。**资料与方法** 回顾性分析经手术病理证实为 ccRCC 的 63 例患者, 术前行腹部 3.0T MR 扫描, 首先使用 ITK-Snap 软件在 T2 序列图像上勾画感兴趣区, 再采用 A.K. 软件提取影像组学特征, 得到对术后病理分级最具有诊断价值的特征参数, 最后构建随机森林 (RF) 机器学习模型对最具特征参数进行分类训练和测试。**结果** 通过特征筛选后最后筛选出 5 个影像组学特征参数, 分别为 Uniformity、Compactness2、Correlation_AllDirection_offset7_SD、GLCMEntropy_AllDirection_offset7_SD 和 LongRunHighGreyLevelEmphasis_angle90_offset7。使用 RF 模型计算的准确度、敏感度、特异度及 ROC 曲线下面积分别为 85%、100%、72%、0.87。**结论** 基于 T2WI 图像纹理分析联合机器学习的影像组学对高级别和低级别肾透明细胞癌的鉴别诊断具有价值。

【关键词】 肾肿瘤; 磁共振成像; 病理学; 外科; 人工智能

【中图分类号】 R730.42; R737.11 **【DOI】** 10.3969/j.issn.1005-5185.2020.08.018

Radiomics Based on T2WI Sequence Texture Analysis Combined with Machine Learning Imaging in Predicting the Pathological Grade of Clear Cell Renal Cell Carcinoma Before Surgery

SHI Bowen¹, YE Jing^{2*}, DUAN Shaofeng³, ZHAO Hongbo¹, LIU Lulu⁴, XU Qing⁴

Department of Radiology, Northern Jiangsu People's Hospital, Yangzhou 225001, China; *Address Correspondence to: YE Jing; E-mail: 18051061289@163.com

【Abstract】Purpose To explore the value of imaging histology based on magnetic resonance T2WI texture analysis combined with machine learning to evaluate the pathological grade of clear cell renal cell carcinoma (ccRCC). **Materials and Methods** Sixty-three biopsy-confirmed ccRCC patients were retrospectively analyzed. Before operation, 3.0T MR scan of the abdomen was performed. First, ITK-Snap software was used to delineate the region of interest on the T2 sequence image, and then extracted texture feature parameters using AK software, and the most valuable diagnostic parameters for postoperative pathological grading were obtained. Finally, a random forest (RF) machine learning model was constructed to classify and tested the most characteristic parameters. **Results** After feature selection, five radiomics feature parameters were finally selected, including Uniformity, Compactness2, Correlation_AllDirection_offset7_SD, GLCMEntropy_AllDirection_offset7_SD and LongRunHighGreyLevelEmphasis_angle90_offset7. The accuracy, sensitivity, specificity and area under the ROC curve calculated by the RF model were 85%, 100%, 72% and 0.87, respectively. **Conclusion** Radiomics based on T2WI image texture analysis combined with machine learning is valuable for the differential diagnosis of high-grade and low-grade ccRCC.

【Key words】 Kidney neoplasms; Magnetic resonance imaging; Pathology, surgical; Artificial intelligence

Chinese Journal of Medical Imaging, 2020, 28 (8): 629-632

肾脏肿瘤是泌尿系统常见肿瘤, 常无明显症状, 多因体检发现, 且多为恶性^[1]。肾细胞癌是泌尿系统中恶性度较高的肿瘤, 也是最常见的肿瘤之一, 其中最常见的病理类型是肾透明细胞癌 (clear cell renal cell carcinoma, ccRCC), 约占肾脏肿瘤的 60%。2016 年, WHO 提出的新分级指南 Fuhrman 核分级^[2]在目前病理学分级系统的应用最广泛, 也是判断预后的独立指标之一, Fuhrman 分级与肿瘤的生长速度和

患者的预后密切相关, 术前准确分级对手术的治疗及预后极为重要^[3]。目前常规影像学检查对肿瘤术前病理分级的诊断价值有限, 而影像组学作为一种无创方法, 利用大量自动化数据特征化算法将感兴趣区 (ROI) 转化为可以进行发掘处理的特征空间数据, 结合机器学习, 能够对肿瘤病变内在的异质性进行定量分析, 并与病理分级进行关联^[4]。本研究拟探讨基于磁共振 T2WI 图像纹理分析联合机器学习的影像组学对术

前预测 ccRCC 病理分级的价值。

1 资料与方法

1.1 研究对象 收集 2016 年 6 月—2019 年 6 月我院 MR 诊断报告提示为肾癌并经术后病理证实为 ccRCC 的 63 例患者, 根据 Fuhrman 分级标准^[2]分为高级别 ccRCC 29 例 (III 级+IV 级), 低级别 ccRCC 34 例 (I 级+II 级), 男 43 例, 女 20 例, 年龄 27~82 岁, 平均 (58.4±11.02) 岁。纳入标准: ①手术病理证实为 ccRCC 并进行 Fuhrman 分级; ②术前进行常规 MRI 检查, 无明显影响图像纹理分析的伪影; ③病灶均为单发且初发首治。

1.2 MRI 检查 63 例患者均进行常规肾脏 MRI 平扫, 使用 GE Discovery MR750 3.0T MR 扫描仪, 患者扫描前均禁饮食 8 h, 横轴面 T2WI 采用呼吸触发脂肪抑制 PROPELLER 序列 (RT Ax fs T2 PROPELLER), 扫描范围覆盖全部肾脏, T2WI 扫描参数: TR 4800 ms, TE 74 ms, 回波链长度 28, 矩阵 320×320, 层厚 5 mm, 层间距 1.0 mm, 视野 40 cm×40 cm。

1.3 图像处理 在医院影像归档和通信系统 (PACS) 以 DICOM 格式导出纳入研究患者的肾脏 MRI 图像。将获得的 T2WI 图像导入 ITK-Snap 软件 (version 3.8.0, www.itk-snap.org), 然后勾画 ROI。由一位具有 10 年工作经验的影像诊断医师操作软件并对图像进行判读, 采用盲法以手动方式在横断位 T2WI 图像上勾画 ROI (图 1), 勾画原则: ①在病灶最大层面勾画出病灶的最大范围; ②ROI 边缘距离病灶边缘约 2 mm, 尽量避免将肾脏正常组织画入其中。由另一位高年资影像诊断医师检查勾画结果, 最后将获取的 ROI 图像以 NIFTI 格式保存。

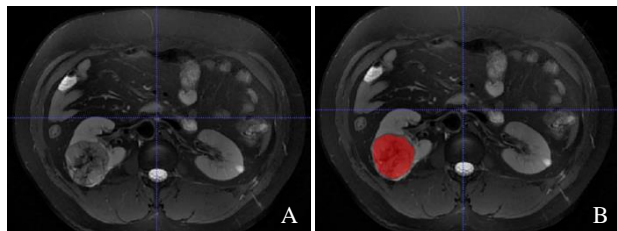


图 1 男, 42 岁。A. 横断位磁共振 T2WI 示右肾肿瘤; B. 红色区域为手动勾画感兴趣区, 术后病理证实为右肾透明细胞癌 III 级

1.4 数据采集

1.4.1 特征提取 将获取的 T2WI 原始图像和勾画完成的 ROI 图像分为高级别 ccRCC 和低级别 ccRCC, 分别导入 A.K. 软件 (Artificial Intelligent Kit), 先进

行图像预处理, 主要对异常值进行处理, 使图像标准化。然后将预处理后的图像再次导入 A.K. 软件提取病灶纹理特征, 共 396 个特征, 包括 6 类特征参数: 直方图特征、形态学特征、灰度共生矩阵参数 (grey level co-occurrence matrix, GLCM)、基于 GLCM 的 Haralick 特征参数、游程矩阵参数 (run length matrix, RLM)、灰度连通区域矩阵参数 (gray level size zone matrix, GLSZM)。

1.4.2 特征筛选与统计学方法 使用 A.K. 软件内置 R 语言算法自动进行统计分析。首先采用 A.K. 软件提取 T2WI 图像上的 396 个影像组学特征, 使用单因素方差分析特征参数, 采用 Mann-Whitney U 检验或 t 检验, 剔除两组参数间差异无统计学意义的参数, 此步骤筛选出 89 个两组之间差异有统计学意义的影像组学特征参数。然后采用 Spearman 相关性分析进行特征去冗余, 剔除自相关性系数 > 0.9 的纹理参数, 在第一步的基础上进一步筛选出 10 个特征参数。

1.4.3 构建模型 将病例以 7 : 3 随机分为训练集与测试集, 使用降维后获得的 8 个特征参数建立随机森林 (random forest, RF) 机器学习模型, 70% 的 ROI 样本 (低级别 23 例, 高级别 20 例) 作为训练组训练机器学习模型, 30% 的样本 (低级别 11 例, 高级别 9 例) 作为测试组对模型的准确性进行测试, 计算准确度、敏感度、特异度, 并获得 ROC 曲线及曲线下面积 (AUC)。

2 结果

2.1 影像组学特征提取 采用最小绝对收缩算子算法 (least absolute shrinkage and selection operator, LASSO) 进行特征值降维, 采用交叉验证筛选各个特征在最佳 $\log(\lambda)$ 时的系数, 获得与高、低级别 ccRCC 鉴别相关的特征参数 5 个, 分为 4 类。其中, 直方图参数 1 个: uniformity; 形态学参数 1 个: Compactness 2; GLCM 2 个: Correlation_AllDirection_offset7_SD 和 GLCMEntropy_AllDirection_offset7_SD; RLM 1 个: LongRunHighGreyLevelEmphasis_angle90_offset7 (图 2)。

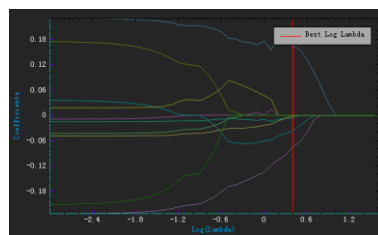


图 2 经三步筛选在最佳 $\log(\lambda)$ 时的系数 (coefficient), 获得最具特征的 5 个参数

2.2 构建 RF 机器学习模型及效能评价 使用 A.K.软件中的 RF 模型对上述 5 个影像学特征参数进行分类训练和测试,计算得训练集准确度、敏感度、特异度及 ROI

曲线下面积(AUC)均为 100%,获得训练集 ROC 曲线;测试集准确度、敏感度、特异度及 AUC 分别为 85%、100%、72%、0.87,获得测试集 ROC 曲线(图 3)。

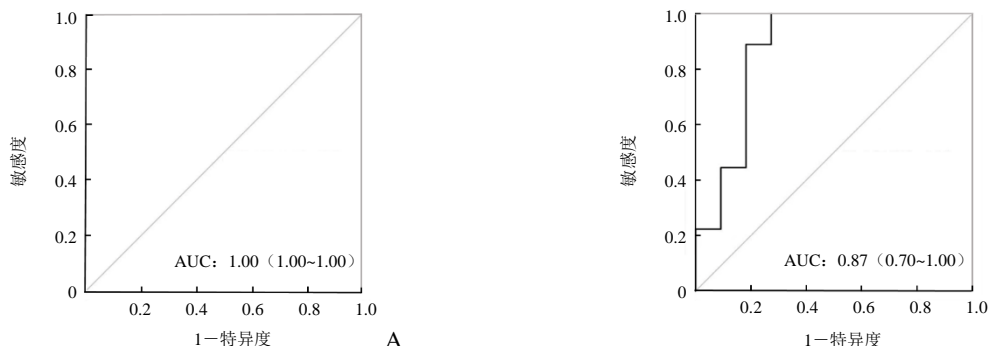


图3 筛选出的 5 个特征参数构建 RF 机器学习模型, A 为训练集 ROC 的 AUC 为 1, B 为测试集 ROC 的 AUC 为 0.87

3 讨论

临床上对不同病理级别的 ccRCC 需采用不同的手术方案,高级别 ccRCC 一般采取根治性肾切除术,但对于低级别 ccRCC,为了最大限度地保留肾单位,避免给患者带来肾功能下降及心血管事件发生的风险,通常行肾脏局部切除术或腹腔镜等创伤较小的手术^[5],因此术前明确 ccRCC 的病理分级至关重要。

影像组学方法是指对图像提取定量影像学特征,并通过分割、特征提取、模型建立等得到图像的灰度变化、体素空间关系等特征,从而实现对疾病的精准预测、诊断及预后评估等^[6-7],这不仅避免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影像图像中大量的人眼无法识别和区分的数字化信息,从而提高疾病的诊断效能,且非侵入性的影像组学检查无损伤,可重复检查^[8]。MRI 具有高分辨率、无电离辐射损失、可提供多参数功能成像等优势, T2WI 序列图像组织分辨率更高,对病灶显示更为清晰。T1WI、DWI 序列图像对病灶的大小、边缘显示不及 T2WI 序列图像。CT 扫描作为检出和诊断肾脏肿瘤的常用方法,具有一定的辐射性;CT 增强扫描需要注射对比剂,有引起对比剂过敏的风险,并增加肾功能不全患者的肾脏负担。此外,病灶碘含量值受扫描时相、血管异常等因素的干扰,容易影响高、低级别组间差异的显著性^[9]。因此本研究采用基于磁共振 T2WI 图像影像组学对高、低病理级别的 ccRCC 病变进行研究。

既往研究已证实通过影像组学诊断肾脏肿瘤性病变的可行性。王平等^[10]采用 CT 增强图像影像组学模型鉴别 ccRCC 与非透明细胞癌,显示良好的诊断效能。Juntu 等^[11]使用支持向量机分类器对 86 例良性

和 49 例恶性肾脏软组织肿瘤患者的 MRI 影像进行纹理分析,发现基于 T1WI 的纹理分析鉴别软组织肿瘤良恶性的 AUC 值(0.93)高于肉眼诊断的 AUC 值(0.90)。白旭等^[12]研究证实发生不同部位转移的 ccRCC 原发灶基于 T2WI 图像的影像组学特征差异有统计学意义。然而,利用磁共振 T2WI 图像纹理分析联合机器学习的影像组学预测 ccRCC 术前病理分级的相关研究目前鲜有报道。

本研究采用磁共振 T2WI 图像影像组学构建 RF 机器学习模型,探讨高、低病理级别 ccRCC 病变的异质性,研究采用的影像图像来自同一台 MRI 设备,手动勾画肿瘤病变区域,并进行图像预处理,以减小个体差异及采集误差,保证了提取特征的可靠性。最终筛选获得直方图参数、形态学参数、GLCM 及 RLM 4 类特征参数。直方图参数与图像像素有关,其中均匀度参数描述了图像纹理的均匀程度。GLCM 反映间隔一定距离的一对像素在既定方向上成对出现的频率,而 RLM 反映具有某灰度值的像素于既定方向上连续出现的频数,能在一定程度上反映图像的对比值,表明纹理模式中的异质性。这 4 类参数描述了图像灰度的对称性、纹理的均匀程度以及灰度分布的不确定性和随机性。而灰度均值和灰度分布的不均匀性特征又代表肿瘤的生物学侵袭性。高、低级别 ccRCC 在 T2WI 图像影像组学特征上存在差异,其原因可能是与低级别 ccRCC 相比,高级别 ccRCC 病灶存在以下特点:①成分复杂,肿瘤内常并发坏死、出血;②肿瘤新生血管丰富、迂曲,分支更不规则,血供存在差别;③肿瘤细胞的密度较大,细胞外间隙相对小;④高级别 ccRCC 细胞核较大且中重度不规则。

本研究构建 RF 机器学习模型,分析基于 T2WI

图像与高、低级别 ccRCC 鉴别相关的 5 个特征参数。RF 算法是一种集成机器学习方法,它利用随机自助法重采样技术和节点随机分裂技术构建多棵决策树,通过投票得到最终分类结果^[13]。RF 算法泛化能力强,不容易过拟合,训练速度快,可并行计算,而且对样本缺失值不敏感。本文计算得 RF 模型下训练集准确度、敏感度、特异度、ROI 的 AUC 均为 1;测试集准确度为 85%、敏感度为 100%、特异度为 72%、AUC 为 0.87。本研究测试集的准确度为 85%,表明测试集能较好地模拟训练集,训练集与验证集拟合度较好,本研究构建的 RF 模型可重复性高。李小虎等^[14]对 34 例 ccRCC 患者进行基于 CT 图像的影像组学分析,构建 RF 模型鉴别低级别和高级别 ccRCC,证明影像组学结合 RF 模型对 ccRCC 术前病理分级预测有价值,AUC 为 0.88,与本研究 AUC 值相似,其敏感度为 79%,较本研究中的敏感度(100%)低,可能是由于 T2WI 组织分辨率更高,对分级的诊断价值可能高于 CT。He 等^[15]利用 CT 纹理分析联合机器学习对 227 例 ccRCC 病例进行研究,证实在提取的 556 个纹理特征中 CIF-minMSEM 参数具有很好的诊断效能,其正确率为 $(94.06 \pm 1.14)\%$,较本研究中的正确率(85%)高,尽管该研究采用 CT 图像,但其采用 MaZda (version4.6) 软件进行纹理提取和分析,该软件可能较本研究采用的 A.K.软件性能更为成熟。此外,本研究样本量较少,导致准确性略低。

本研究的局限性为:①二维的 ROI 可能会遗漏一些影像学特征,ROI 的选择为肿瘤的最大层面,影像学纹理分析仅反映该层面 ROI 的异质性,不能完全反映该肿瘤组织整体的异质性;②本研究为回顾性分析,且各病理级别组样本相对较少;③尽管影像组学为肿瘤的术前病理分类提供了较好的应用结果,但其医学解释有待进一步研究。后期将收集更多样本,并在三维 ROI 的基础上进行前瞻性研究。

总之,基于磁共振 T2WI 纹理分析联合机器学习的影像组学对鉴别诊断高、低级别 ccRCC 具有一定的价值。

参考文献

[1] 周青,陈琴. 超声造影对肾脏实性肿瘤的定性诊断价值. 中国医学影像学杂志, 2018, 26(8): 602-605, 611.
[2] Moch H, Cubilla AL, Humphrey PA, et al. The 2016 WHO classification of tumours of the urinary system and male genital organs-part a: renal, penile, and testicular tumours. Eur Urol, 2016, 70(1): 93-105.

[3] 唐彩银,李瑗,张继,等. CT 纹理分析在肾脏透明细胞癌分级的临床应用. 医学理论与实践, 2019, 32(21): 3416-3418, 3409.
[4] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer, 2012, 48(4): 441-446.
[5] Lam JS, Klatte T, Breda A. Staging of renal cell carcinoma: Current concepts. Indian J Urol, 2009, 25(4): 446-454.
[6] 孙航,李宏,张亭亭,等. Radiomics 方法研究应用进展. 肿瘤, 2017, 37(10): 1092-1099.
[7] Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and prognosis of quantitative features extracted from CT images. Transl Oncol, 2014, 7(1): 72-87.
[8] 孙钢. 放射组学的兴起及其在消化系统肿瘤中的应用. 中华消化病与影像杂志(电子版), 2017, 7(4): 145-149.
[9] 张连宇,陈雁,王绎忱,等. CT 能谱成像与 MR 扩散加权成像评价肾透明细胞癌病理分级的价值比较. 医学影像学杂志, 2018, 28(4): 627-631.
[10] 王平,裴旭,殷小平,等. 基于增强 CT 影像组学模型鉴别肾透明细胞癌与非透明细胞癌. 中国医学影像技术, 2019, 35(11): 1689-1692.
[11] Juntu J, Sijbers J, De Backer S, et al. Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. J Magn Reson Imaging, 2010, 31(3): 680-689.
[12] 白旭,王海屹,叶慧义,等. 基于 T2 加权图像的纹理分析: 不同转移部位间肾癌纹理特征的对比. 放射学实践, 2018, 33(8): 794-799.
[13] 姚登举,杨静,詹晓娟. 基于随机森林的特征选择算法. 吉林大学学报: 工学版, 2014, 44(1): 137-141.
[14] 李小虎, Cai Wenli, 裴子璐,等. 肾脏 CT 容积纹理分析及机器学习相结合的影像组学评价肾透明细胞癌病理分级的价值初探. 中华放射学杂志, 2018, 52(5): 344-348.
[15] He X, Wei Y, Zhang H, et al. Grading of clear cell renal cell carcinomas by using machine learning based on artificial neural networks and radiomic signatures extracted from multidetector computed tomography images. Acad Radiol, 2020, 27(2): 157-168.

【收稿日期】2019-12-21 【修回日期】2020-03-26

(本文编辑 刘晓萌)