❖生殖泌尿影像学

Radiomics models based on enhanced CT for identifying renal clear cell carcinoma and non-clear cell carcinoma

WANG Ping[1,2], PEI Xu[1,2], YIN Xiaoping[1,3*], REN Jialiang[4], GUO Jiandang[1], ZHAO Zhenzhen[1], ZHAO Yingjia[1]

(1. CT-MRI Division, Affiliated Hospital of Hebei University, Baoding 071000, China; 2. Medical College, Hebei University, Baoding 071000, China; 3. Key Laboratory of Cancer Radiotherapy and Chemotherapy Mechanism and Regulations of Hebei Province, Baoding 071000, China; 4. General Electric Pharmaceutical [Shanghai] Co., Ltd., Shanghai 210000, China)

[**Abstract**] **Objective** To establish radiomics models based on enhanced CT, and to explore the value of the models for distinguishing renal clear cell carcinoma (ccRCC) and non-clear cell renal cell carcinoma (non-ccRCC). **Methods** Totally 147 patients with ccRCC and 32 patients with non-ccRCC were randomly divided into training set (n=125) and testing set (n=54). Enhanced CT data were imported into ITK-SNAP software, and ROI was manually delineated to obtain 16 features. Random Forest (RF) model and Logistic Regression (LR) model based on features were established, respectively. ROC curve was used to observe the diagnostic efficiency of the models for ccRCC. **Results** In the training set, RF model diagnosed ccRCC with AUC of 0.96 (P<0.05) specificity of 1.00, and sensitivity of 0.83, while LR model diagnosed ccRCC with AUC of 0.96 (P<0.05), specificity of 1.00, and sensitivity of 0.83. In the testing set, RF model diagnosed ccRCC with AUC of 0.96 (P<0.05), specificity of 1.00, and sensitivity of 0.89, while LR model diagnosed ccRCC with AUC of 0.88(P<0.05), specificity of 0.90, and sensitivity of 0.77. **Conclusion** Radiomics models based on enhanced CT can be used for identifying ccRCC from non-ccRCC. RF model has higher diagnostic value than LR model.

[**Keywords**] kidney neoplasms; radiomics; tomography, X-ray computed

DOI:10.13929/j.1003-3289.201901099

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

王 平[1,2],裴 旭[1,2],殷小平[1,3*],任嘉梁[4],郭建党[1],赵珍珍[1],赵莹佳[1]

[1.河北大学附属医院 CT-MRI 室,河北 保定 071000;2.河北大学(医学院),河北 保定 071000;
3.河北省肿瘤放化疗机制与研究重点实验室,河北 保定 071000;
4.通用电气药业(上海)有限公司,上海 210000]

[**摘 要**] **目的** 建立基于增强 CT 的影像组学模型,评估其鉴别肾透明细胞癌(ccRCC)与非透明细胞癌(non-ccRCC)的应用价值。**方法** 将 147 例 ccRCC 及 32 例 non-ccRCC 患者随机分为训练集 125 例和测试集 54 例。将所有患者的增强 CT 资料导入 ITK-SNAP 软件,手动勾画 ROI,获得 16 个特征,分别建立基于特征的随机森林(RF)模型和逻辑回归(LR)模型,采用 ROC 曲线观察模型对 ccRCC 的诊断效能。**结果** 训练集 RF 模型诊断 ccRCC 的 AUC 为 0.96(P<0.05),特

[**基金项目**] 河北省研究生创新资助项目(hbu2019ss036)、保定市科技计划项目(18ZF182)。

[**第一作者**] 王平(1992—),男,山东临沂人,硕士,医师。研究方向:CT 诊断学。E-mail: wangping920403@outlook.com

[**通信作者**] 殷小平,河北大学附属医院 CT-MRI 室,071000;河北省肿瘤放化疗机制与研究重点实验室,071000。E-mail: yinxiaoping78@sina.com

[**收稿日期**] 2019-01-15 [**修回日期**] 2019-07-03

异度为 1.00,敏感度 0.83;LR 模型诊断 ccRCC 的 AUC 为 0.96($P<0.05$),特异度为 1.00,敏感度为 0.83。测试集 RF 模型诊断 ccRCC 的 AUC 为 0.96($P<0.05$),特异度为 1.00,敏感度为 0.89;LR 模型诊断 ccRCC 的 AUC 为 0.88($P<0.05$),特异度为 0.90,敏感度为 0.77。**结论** 基于增强 CT 影像组学模型可用于鉴别 ccRCC 与 non-ccRCC;RF 模型诊断价值较 LR 模型更高。

[**关键词**] 肾肿瘤;影像组学;体层摄影术,X 线计算机
[**中图分类号**] R737.11;R814.42 [**文献标识码**] A [**文章编号**] 1003-3289(2019)11-1689-04

肾细胞癌(renal cell carcinoma, RCC)占成人恶性肿瘤的 2%～3%,致死率较高[1],其病理分型包括肾透明细胞癌(clear cell RCC, ccRCC)和非透明细胞癌(non-clear cell RCC, non-ccRCC),后者又分为嫌色细胞癌(chromophobe renal cell carcinoma, chRCC)、乳头状细胞癌(papillary renal cell carcinoma, pRCC)、集合管癌(collecting duct carcinoma, CDC)及未分类癌。RCC 的亚型[2-3]及病理分期[4]均会影响临床治疗及预后。肾穿刺活检是诊断 RCC 亚型的金标准,但穿刺可能导致癌细胞转移、肿瘤出血等并发症,且存在抽样误差,故对于 RCC 活检一直存在争议[5-6]。尽管影像学上 ccRCC 与 non-ccRCC 表现出不同的强化方式,但临床误判率仍较高[7]。联合应用不同影像学检查可提高诊断 RCC 亚型的正确率,但增加患者经济负担,且耗时长。影像组学是从医学图像中高通量提取定量特征(纹理特征、形状特征、强度特征及高阶特征等信息),转化为可采集的组学数据,并对这些数据进行分析,从而达到判断病变性质的目的[8]。本研究探讨基于增强 CT 的影像组学模型鉴别 ccRCC 与 non-ccRCC 的价值。

1 资料与方法

1.1 一般资料 回顾性分析 2017 年 1 月—2019 年 9 月河北大学附属医院收治的 179 例 RCC 患者的影像学资料,男 93 例,女 86 例,年龄 27～88 岁,平均(58.4±10.6)岁;肿瘤最大径 2～11 cm,平均(6.48±3.46)cm;术后病理诊断 ccRCC 147 例,non-ccRCC 32 例(pRCC 20 例、chRCC 9 例、CDC 3 例)。纳入标准:①病理检查明确诊断 RCC 及其分型;②术前 3 周接受增强 CT 皮质期扫描。排除标准:①病灶内存在直径超过 4～5 cm 的坏死、囊变;②肿瘤已发生转移;③图像质量不佳。将患者随机分为训练集和测试集。训练集 125 例,男 64 例,女 61 例,年龄 27～87 岁,平均(58.4±9.9)岁,64 例≤60 岁;non-ccRCC 22 例(pRCC 15 例、chRCC 6 例、CDC 1 例),ccRCC 103 例。测试集 54 例,男 29 例,女 25 例,年龄 30～88 岁,平均(59.7±9.1)岁;其中 non-ccRCC 10 例(pRCC 5 例、chRCC 3 例、CDC 2 例),ccRCC 44 例。

1.2 仪器与方法 采用 GE Discovery HD750 64 排 CT 扫描仪。嘱患者禁食 6～8 h 后,行肾脏平扫及增强扫描。扫描参数:层厚 5 mm,螺距 0.984:1,扫描视野 36 cm×43 cm,矩阵 512×512,管电压 100～120 kV,管电流 134～409 mA,窗宽 250～450 HU,窗位 30～50 HU。对比剂为非离子型碘佛醇,经肘静脉注射,流率 3.0～3.5 ml/s,剂量 1.0～1.2 ml/kg 体质量。皮质期、髓质期、排泄期扫描时间分别为注射对比剂后 30～35 s、50～60 s 和 180 s。

1.3 图像分割、特征提取与建模 将 CT 增强图像(皮质期)导入 ITK-SNAP(www.itk-snap.org)软件,由 2 名具有 8～10 年工作经验的影像学医师分别独立勾画 ROI(图 1～3)。ROI 勾画原则:①选取 PACS 系统上 DICOM 格式的图像;②距病灶皮质期边缘内侧 0～1 mm;③逐层勾画,获得病灶的 ROI 容积(volume of interest, VOI);④包含小范围坏死囊变及强化血管;⑤避开肾静脉内的瘤栓;⑥对于 non-ccRCC 皮质期,参考髓质期图像勾画 ROI。

将所有原始图像文件及勾画好的 ROI 文件导入 A.K 软件(GE Artificial Intelligence Kit)进行特征提取,包括直方图特征、形态学特征及二阶纹理特征(灰度共生矩阵、游程矩阵等),共生成 396 个特征;采用单因素方差分析及秩和检验、一般线性模型和前 10%互信息(mutual information, MI)3 种方法依次进行特征降维,最终选取 16 个具有统计学意义的特征,即 LargeAreaEmphasis、OneVoxelVolume HaralickCorrelation _ angle0 _ offset7、ClusterProminence_angle135_offset7、ShortRunEmphasis_angle0_offset4、LowGreyLevelRunEmphasis_AllDirection_offset1 _ SD、HaralickCorrelation _ angle135 _ offset4、LeastAxisLength、LowIntensityLargeAreaEmphasis、最大像素值(MaxIntensity)、直方图熵(histogramEntropy)、方差(variance)、直方图能量(histogramEnergy)、平均像素值(MeanValue)、直方图标准差(stdDeviation)和 hara 熵(haraEntroy)。分别建立线性逻辑回归(logistic regression, LR)模型及非线性随机森林(random forest,

RF)模型，保留 AUC 最大且特异度较高的模型，采用测试数据进行模型评估。

1.4　统计学分析　采用 SPSS 24.0 统计分析软件，以独立样本 t 检验比较训练集和测试集患者年龄，χ^2 检验比较性别和 ccRCC 在的训练集和测试集之间的分布。采用 ROC 曲线评价 LR、RF 模型鉴别 ccRCC 与 non-ccRCC 的效能。$P<0.05$ 为差异有统计学意义。

2　结果

训练集和测试集患者年龄（$t=2.13$，$P=0.05$）、性别（$\chi^2=0.41$，$P=0.05$）、ccRCC 在的训练集和测试集之间的分布（$\chi^2=0.02$，$P=0.90$）差异均无统计学意义。

ROC 曲线结果显示，训练集 RF 模型诊断 ccRCC 的 AUC 为 0.96（$P<0.05$），特异度为 1.00，敏感度 0.83；LR 模型诊断 ccRCC 的 AUC 为 0.96（$P<0.05$），特异度为 1.00，敏感度为 0.83。测试集 RF 模型诊断 ccRCC 的 AUC 为 0.96（$P<0.05$），特异度为 1.00，敏感度为 0.89；LR 诊断 ccRCC 的 AUC 为 0.88（$P<0.05$），特异度为 0.90，敏感度为 0.77。见表 1、图 4。

表 1　训练集、测试集中 RF、LR 模型诊断 ccRCC 的效能

组别	AUC	准确率	精确度	特异度	敏感度	F1 评分
训练集（$n=125$）						
RF	0.96	0.86	1	1	0.83	0.90
LR	0.96	0.86	1	1	0.83	0.90
测试集（$n=54$）						
RF	0.96	0.91	1	1	0.89	0.94
LR	0.88	0.80	0.97	0.90	0.77	0.86

注：精确度＝真阳性/（真阳性＋假阳性）



图 1　患者男，59 岁，pRCC 皮质期 CT 增强图像（虚线为 ROI）　**图 2**　患者女，62 岁，ccRCC 皮质期 CT 增强图像（虚线为 ROI）　**图 3**　患者男，71 岁，chRCC 皮质期 CT 增强图像（虚线为 ROI）

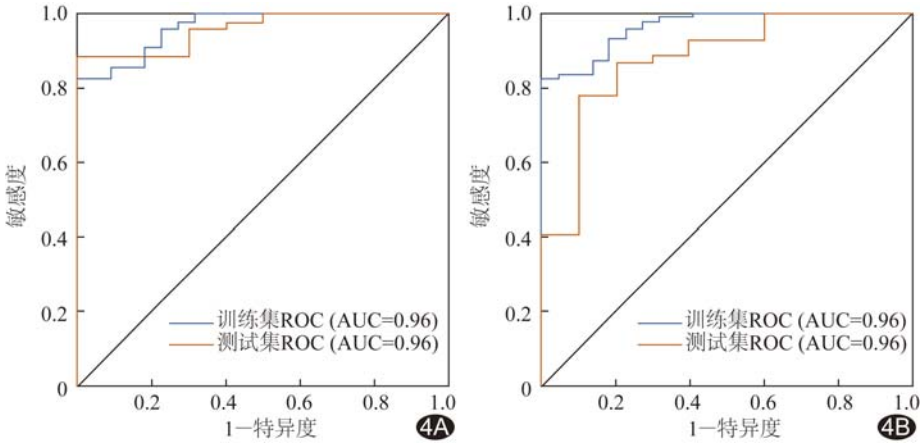


图 4　RF、LR 模型诊断 ccRCC 的 ROC 曲线　A. RF 模型的训练集和测试集 ROC 曲线；B. LR 模型的训练集和测试集 ROC 曲线。

将 RF 预测模型带入原始数据中进行预测，结果显示训练集中 ccRCC 真阳性 85 例，non-ccRCC 真阴性 22 例，测试集中 ccRCC 真阳性 39 例，non-ccRCC 真阴性 10 例。将 LR 预测模型带入原始数据中进行预测，训练集中 ccRCC 真阳性 85 例，non-ccRCC 真阴性 22 例；测试集中 ccRCC 真阳性 34 例，non-ccRCC 真阴性 9 例。

3　讨论

RCC 是肾脏最常见的恶性肿瘤，约 20% RCC 患者就诊时已发生转移，早期诊断 RCC 对于指导临床治疗意义重大[2-3]。尽管多期增强 CT 可用于鉴别 RCC 亚型[3]，但在临床工作中仍有很高的误诊率。目前采用影像组学特征鉴别 RCC 亚型的研究[9]较少，且缺乏研究 RCC 亚型之间潜在的分子基础。既往有研究[3,5-6]评估增强 CT 模式、2D 纹理和其他定量参数对 RCC 亚型的诊断效能，发现其各有不足[10]。放射组学具有可重复性，已逐渐应用于临床。本研究提取 RCC 皮质期 CT 增强图像特征建立模型，可更直观地反映肿瘤的整体特征[10]，即使选择一期进行分析，模型对鉴别 ccRCC 与 non-ccRCC 的敏感度和特异度仍较高，为传统 CT 增强所不及；尽管本组 non-ccRCC 病例数较少，但两种模型的特异度均较高，提示模型未过拟合，泛化能力较好。

Zhu 等[11]认为低强化是高级别 ccRCC 的特征。本研究于肾脏肿瘤 CT 增强皮质期勾画 ROI，利用 A.K 软件提取其影像特征（灰度分布、小波特征及纹理特征等），以反映肿瘤内部的异质性[12]。异质性是恶性肿瘤的重要特征，与其恶性生物学行为关联密切[13]，能够反映相关生长因子变化及肿瘤生长微环境的变化，而这些变化可导致肿瘤局部区域发生细胞增殖或凋亡、内部代谢活跃或抑制、肿瘤血管增加或减少等[14]；肿瘤血管生成的异质性可能导致局部有效血流灌注减少，引起该区域缺氧，导致间质静水压增高等独特的肿瘤微环境变化，最终增加肿瘤侵袭和转移的风险[15-16]。

由于本组 non-ccRCC 例数较少，可能存在建模偏差，故分别采用线性的 LR 模型和非线性的 RF 模型进行建模并对比；ROC 曲线结果显示，测试集 RF 模型诊断 ccRCC 的 AUC 达 0.96，LR 模型为 0.88。LR 模型的诊断效能不及 RF 模型，原因可能在于 RCC 亚型与特征之间呈非线性关系，而 LR 模型为线性模型，故诊断效能相对差；RF 为非线性，可更加稳定地预测非线性数据。此外，RF 模型可处理高维度数据，且毋须进行特征选择（特征子集随机），训练速度快，易采用并行化方法。Li 等[17]联合 von Hippel-Lindau（VHL）基因分别建立影像组学模型 Boruta、mRMRe，结果显示测试集在两种模型中鉴别诊断 ccRCC 与 non-ccRCC 的 AUC 分别为 0.969、0.900，与本研究的 RF 模型（AUC=0.96）相当但高于该组 LR 模型（AUC=0.88）。

本研究的局限性：样本量较小，且 non-ccRCC 样本量更少；临床资料中无家族病史，未做组合模型；non-ccRCC 样本仅纳入 pRCC、chRCC、CDC 病例，未涉及其他病理类型 non-ccRCC；勾画 ROI 时，受部分肿块边缘不清晰或部分容积效应的影响，导致准确率降低，且病灶坏死囊变会在一定程度上影响影像组学分析结果。

总之，基于增强 CT 的影像组学模型可用于鉴别 ccRCC 与 non-ccRCC，其中 RF 模型的诊断价值更高。

[参考文献]

[1] Burton A, Martin RM, Holly J, et al. Associations of adiponectin and leptin with stage and grade of PSA-detected prostate cancer: The protect study. Cancer Causes Control, 2013, 24(2): 323-334.

[2] Cheville JC, Lohse CM, Zincke H, et al. Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma. Am J Surg Pathol, 2003, 27(5): 612-624.

[3] Young JR, Margolis D, Sauk S, et al. Clear cell renal cell carcinoma: Discrimination from other renal cell carcinoma subtypes and oncocytoma at multiphasic multidetector CT. Radiology, 2013, 267(2): 444-453.

[4] Thompson RH, Leibovich BC, Lohse CM, et al. Dynamic outcome prediction in patients with clear cell renal cell carcinoma treated with radical nephrectomy: The D-SSIGN score. J Urol, 2007, 177(2): 477-480.

[5] Raman SP, Chen Y, Schroeder JL, et al. CT texture analysis of renal masses: Pilot study using random forest classification for prediction of pathology. Acad Radiol, 2014, 21(12): 1587-1596.

[6] Feng Z, Rong P, Cao P, et al. Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol, 2018, 28(4): 1625-1633.

[7] 朱庆强，王中秋，王守安，等. 肾透明细胞癌与少见的肾癌亚型 CT 表现的对照研究. 中华泌尿外科杂志，2014，35(3)：168-173.

[8] Lambin P, Riosvelazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer, 2007, 43(8): 441-446.

[9] Yu H, Scalera J, Khalid M, et al. Texture analysis as a radiomic marker for differentiating renal tumors. Abdom Radiol, 2017, 42(10): 1-9.

[10] Gaing B, Sigmund EE, Huang WC, et al. Subtype differentiation of renal tumors using voxel-based histogram analysis of intravoxel incoherent motion parameters. Invest Radiol, 2015, 50(3): 144.

[11] Zhu YH, Wang X, Zhang J, et al. Low enhancement on multiphase contrast-enhanced CT images: An independent predictor of the presence of high tumor grade of clear cell renal cell carcinoma. AJR Am J Roentgenol, 2014, 203(3): W295-W300.

[12] 王海屹，叶慧义，马林. 定量 MR 动态增强成像的机制及其在肿瘤学方面的应用价值. 中华放射学杂志，2014，48(3)：261-264.

[13] Logothetis CJ. Re: Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. Eur Urol, 2013, 64(1): 170.

[14] Rolland Y, Bézy-Wendling J, Gestin H, et al. Analysis of texture in medical imaging. Review of the literature. Ann Radio Paris, 1995, 38(6): 315.

[15] 刘慧，王小宜，龙学颖. 基于 CT 图像纹理分析肿瘤异质性的研究进展及应用. 国际医学放射学杂志，2016，39(5)：543-548.

[16] 李颖嘉，杨莉，文戈. 肿瘤血管生成异质性及肿瘤血管生成超声分子靶向成像研究进展. 中国医学影像技术，2009，25(10)：1899-1902.

[17] Li ZC, Zhai G, Zhang J, et al. Differentiation of clear cell and non-clear cell renal cell carcinomas by all-relevant radiomics features from multiphase CT: A VHL mutation perspective. Eur Radiol, 2019, 29(8): 3996-4007.