

临床研究

DOI: 10.13406/j.cnki.cyx.002175

基于 DCE-MRI 的影像组学在预测乳腺癌 Luminal 型和非 Luminal 型中的价值

王世科, 孙 冬, 郭大静, 刘 佳, 陈林丽
(重庆医科大学附属第二医院放射科, 重庆 400010)

【摘要】目的:探讨基于动态增强 MRI(dynamic contrast enhanced magnetic resonance imaging, DCE-MRI)的影像组学在预测乳腺癌 Luminal 型和非 Luminal 型中的价值。**方法:**回顾性分析行乳腺 DCE-MRI 扫描,并经手术病理证实为乳腺癌的 42 例患者,采用手动分割法在 DCE-MRI 强化最明显的图像上逐层勾画病灶提取影像组学特征,用最小绝对收缩算子(least absolute shrinkage and selection operator, LASSO)筛选最优影像特征,根据雌激素受体(estrogen receptor, ER)和孕激素受体(progesterone receptor, PR)的表达情况将所有病例分为 Luminal 型和非 Luminal 型,构建基于逻辑回归(logistic regression, LR)、随机森林(Random forest, RF)、K 近邻(K nearest neighbor, KNN)和支持向量机(support vector machine, SVM)的分类器,并用 5 折交叉验证法对预测性能进行验证。**结果:**最终选出 4 个最优特征值,包括 1 个形状特征和 3 个小波特征,其中 Original_Shape_Maximum 2D Diameter Row、Wavelet-LLH_GLCM_Idn 和 Wavelet-HHH_GLCM_Correlation 三个特征值与 ER 和 PR 的表达情况均具有显著相关性,基于 RF 的分类器预测乳腺癌 Luminal 型和非 Luminal 型的性能最佳,其敏感度、特异度、准确度及曲线下面积(area under curve, AUC)分别为 0.838、0.900、0.853 和 0.876。**结论:**基于 DCE-MRI 的影像组学在预测乳腺癌 Luminal 型和非 Luminal 型中具有较好的应用价值,基于 RF 的分类器的预测性能最佳。

【关键词】影像组学;动态增强 MRI;乳腺癌;分子分型

【中图分类号】R445.2

【文献标志码】A

【收稿日期】2019-03-08

Value of radiomics based on dynamic contrast-enhanced magnetic resonance imaging in predicting luminal and non-luminal breast cancer

Wang Shike, Sun Dong, Guo Dajing, Liu Jia, Chen Linli

(Department of Radiology, the Second Affiliated Hospital of Chongqing Medical University)

【Abstract】Objective: To investigate the value of radiomics based on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) in predicting luminal and non-luminal breast cancer. **Methods:** A retrospective analysis was performed for the clinical data of 42 patients who underwent breast DCE-MRI and had pathologically confirmed breast cancer. The lesions were outlined by manual segmentation on the images with the strongest enhancement of DCE-MRI to determine radiomic features. Then the least absolute shrinkage and selection operator was used to select the most important features. All patients were divided into luminal group and non-luminal group according to the expression of estrogen receptor (ER) and progesterone receptor (PR). Classifiers were constructed based on logistic regression (LR), random forest (RF), K nearest neighbor (KNN), and support vector machine (SVM), and the 5-fold cross validation method was used to verify prediction performance. **Results:** A total of 4 optimal features were screened out, consisting of 1 shape feature and 3 wavelet features, among which Original_Shape_Maximum 2D Diameter Row, Wavelet-LLH_GLCM_Idn, and Wavelet-HHH_GLCM_Correlation had significant correlations with ER and PR levels. The RF-based classifier had the best performance in predicting luminal and non-luminal breast cancer, with a sensitivity of 0.838, a specificity of 0.900, an accuracy of 0.853, and an

area under the ROC curve of 0.876. **Conclusion:** Radiomics based on DCE-MRI has a good value in predicting luminal and non-luminal breast cancer, and the RF-based classifier has the best performance.

【Key words】radiomics; dynamic contrast-enhanced magnetic resonance imaging; breast cancer; molecular subtype

作者简介:王世科, Email: wangsk@cqmu.edu.cn,

研究方向:磁共振成像技术。

通信作者:孙 冬, Email: 303704@hospital.cqmu.edu.cn。

基金项目:重庆市教委科学技术研究项目(青年项目)资助项目(编号: KJQN201800405)。

优先出版: <http://kns.cnki.net/kcms/detail/50.1046.R.20190523.1500.004.html>
(2019-05-24)

乳腺癌是女性最常见的恶性肿瘤,随着社会发展、饮食结构和生活方式的改变等原因,乳腺癌的发病率和死亡率呈现逐年上升的趋势^[1-2]。乳腺癌根据激素受体表达不同可分为 Luminal 型和非 Luminal 型,不同的分子亚型具有不同的生物学行为,对治疗的反应和预后也存在显著差异^[3]。临床上,无创性地早期定性分析乳腺癌的分子分型是近年来的研究热点。影像组学通过高通量提取定量图像数学结构特征,可实现肿瘤分割、特征提取与模型建立,对海量影像数据信息进行深层次的挖掘、预测和分析来进行辅助诊断^[4]。本研究回顾性分析了 42 例乳腺癌患者的基于动态增强 MRI(dynamic contrast enhanced magnetic resonance imaging,DCE-MRI)图像的影像组学特征,探讨影像组学特征在预测 Luminal 型和非 Luminal 型乳腺癌中的价值。

1 资料与方法

1.1 研究对象

回顾性分析 2013 年 3 月至 2014 年 6 月的 58 例乳腺癌患者的 DCE-MRI 图像,纳入标准:①经手术病理证实为乳腺癌的患者;②MRI 检查前未进行过穿刺活检、手术切除或化疗等治疗;③病理资料完整,具有明确的雌激素受体(estrogen receptor,ER)、孕激素受体(progesterone receptor,PR)和人表皮生长因子受体-2(human epidermal growth factor receptor-2,HER2)检查结果。排除标准:图像质量差,不能满足分析要求。

1.2 MRI 图像采集

所有病人使用 3.0 T 磁共振扫描仪(Discovery MR750,美国 GE 公司),俯卧位、脚先进,使用乳腺专用相控阵线圈采集信号,行轴位 T₁WI、脂肪抑脂 T₂WI、DWI 和三维乳腺容积成像(3D volume imaged breast assessment,3D VIBRANT)动态增强扫描。3D VIBRANT 扫描参数如下:TR 4.5 ms,TE 2.1 ms,翻转角 10°,矩阵 384 × 256,层厚 1.2 mm;先行一期平扫,然后经肘静脉快速团注对比剂 0.1 mmol/kg(欧乃影,美国 GE 公司),注入对比剂后再连续扫描 6 期动态增强,每期扫描时间 58 s,总扫描时间为 6 min 46 s。

1.3 免疫组化检测及分子分型判定

所有患者均进行了乳腺肿瘤切除术,术后病理标本进行免疫组织化学检测,采用 SP 法检测患者的 ER、PR 和 HER2 的表达情况,检测结果由病理科医师进行判读,当至少 1%的肿瘤细胞核呈现 ER 或 PR 染色时即判定为 ER 或 PR 阳性,HER2 表达结果为-或+者判定为 HER2 阴性,+++者判定为 HER2 阳性,++者进一步用荧光原位杂交法进行验证^[5]。根据

ER 和 PR 的表达情况将所有病例分为 Luminal 型[ER 阳性和(或)PR 阳性]和非 Luminal 型(ER 阴性和 PR 阴性)。

1.4 图像分割及影像组学特征提取

肿瘤的分割采用手动分割法,从 PACS 系统中调取所有病人的原始 DCE-MRI 图像,根据时间信号曲线,选取强化最明显的时相导入 RadCloud 软件(慧影医疗科技有限公司),由两名乳腺诊断放射医师(分别具有 3 年和 12 年工作经验)在不知道分子亚型的情况下逐层勾画病灶边界,高年资医师的勾画结果用来提取影像组学特征,低年资医师的勾画结果用来评估预测性能的可靠性^[6]。共计 1 029 个特征值被提取,可归纳分为 4 组:①15 个形态特征,描述了肿瘤的位置、大小和形状等几何学参数,如直径、表面积、体积等;②95 个直方图特征(一阶统计特征),反映了图像中像素的灰度分布情况,如病灶密度的平均值、偏差、峰度、偏度等;③295 个纹理特征,如能量、熵、自相关、逆差距、同质性和异质性等,这些特征通过灰度共生矩阵(gray-level co-occurrence matrix, GLCM)、灰度游程长度矩阵(gray-level run length matrix, GLRLM)及灰度区域大小矩阵(gray-level size zone matrix, GLSZM)获得,描述了病灶中体素灰度的空间关系,能较好地反映肿瘤的异质性;④624 个小波特征,小波特征是在直方图特征和纹理特征的基础上进一步使用滤波器,经过多尺度细化分析获得,可将图像中的不同时频特征成分进行分离,获得更加精准的影像特征信息^[7]。

数据降维及最优特征筛选:由于提取的影像特征数量较大,为避免出现过拟合,使用了方差选择和最小绝对收缩算子(least absolute shrinkage and selection operator,LASSO)进行降维并筛选最优特征,首先使用方差选择法,阈值设为 0.8,计算每个特征值的方差,保留方差大于 0.8 的所有特征值,然后使用单变量方差分析衡量特征和分类结果的关系,选择 $P < 0.05$ 的特征,最后用 LASSO 进行最优特征筛选,LASSO 使用 5 折交叉验证法,最终筛选出 4 个最优的影像组学特征(图 1)。

1.5 模型建立与性能预测

以 Luminal 型和非 Luminal 型为标签、高通量特征为输入、分类判别结果为输出进行机器学习。构建基于逻辑回归(logistic regression,LR)、随机森林(random forest,RF)、K 近邻(K nearest neighbor,KNN)和支持向量机(support vector machine,SVM)的分类器,在建立模型前先对输入的特征值进行了标准化处理,去掉均值并缩放到单位方差,然后采用 5 折交叉验证法对分类性能进行验证,计算 4 种分类器的敏感度、特异度、准确度和受试者操作特性(receiver operating characteristic,ROC)曲线下面积(area under curve,AUC)。本文采用的影像组学流程如图 2 所示。

1.6 统计学处理

采用 SPSS 17.0 进行统计分析,用组内相关系数(intra-class correlation coefficient,ICC)评价两名医师的一致性,取 $ICC > 0.75$ 为信度良好;采用 Spearman 相关分析评估特征值与 ER、PR 和 HER2 的相关性。检验水准 $\alpha = 0.05$ 。

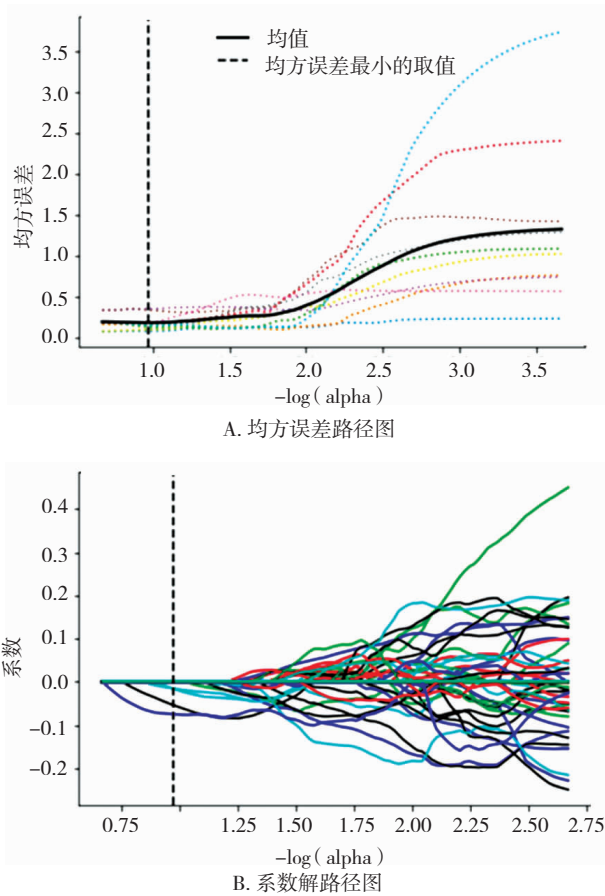


图 A 为利用 5 折交叉验证选择最优的 LASSO 模型参数,纵轴是均方误差,横坐标为 $-\log(\alpha)$,不同颜色的虚线表示交叉验证的每组样本对应不同的 $-\log(\alpha)$ 有不同的均方误差,左侧的垂线表示交叉验证中当均方误差最小时对应的参数取值,图 B 为影像组学特征在图 A 对应的参数下的特征系数曲线图,垂线对应于图 A 左侧的垂线,表示通过最优的参数值筛选出 4 个影像组学特征

图 1 LASSO 筛选最优特征

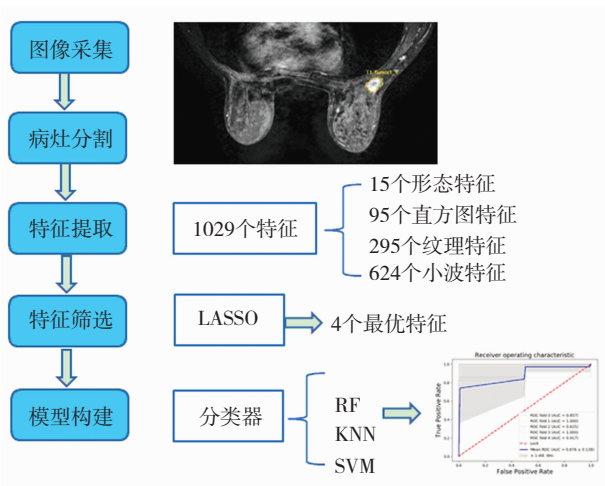


图 2 影像组学流程图

2 结 果

最终 42 例患者纳入研究,均为浸润性导管癌,年龄 33~77 岁 (50.95 ± 9.06) 岁,包括 32 例 Luminal 型和 10 例非 Luminal 型。2 名医师测量结果的 ICC 值为 0.928,2 名医师的一致性良好。经过方差选择和 LASSO 处理,最终筛选出 4 个最优特征,包括 1 个形状特征和 3 个小波特征(表 1),其中 Original_Shape_Maximum 2D Diameter Row、Wavelet-LLH_GLCM_Idn 和 Wavelet-HHH_GLCM_Correlation 三个特征值与 ER 和 PR 的表达情况均具有显著相关性,而 Wavelet-LHH_GLSZM_Large Area High Gray Level Emphasis 与 ER 和 PR 的表达情况不具有显著相关性,4 个特征值与 HER2 的表达情况均不具有显著相关性(表 2)。

表 1 LASSO 筛选出的最优影像特征

特征名称	特征描述
Original_Shape_Maximum 2D Diameter Row	原始图像上基于形态学的最大行 2D 直径
Wavelet-LHH_GLSZM_Large Area High Gray Level Emphasis	基于小波变换的灰度区域大小矩阵中高灰度值的较大区域联合分布在图像中的比例
Wavelet-LLH_GLCM_Idn	基于小波变换的灰度共生矩阵中的反向差分归一化
Wavelet-HHH_GLCM_Correlation	基于小波变换的灰度共生矩阵中灰度值与各自像素之间的线性依赖关系

表 2 最优特征与 ER、PR 及 HER2 的相关性

影像特征	ER		PR		HER2	
	r 值	P 值	r 值	P 值	r 值	P 值
Original_Shape_Maximum 2D Diameter Row	-0.422	0.005 ^a	-0.383	0.012 ^a	0.128	0.418
Wavelet-LHH_GLSZM_Large Area High Gray Level Emphasis	0.074	0.643	-0.006	0.969	0.045	0.775
Wavelet-LLH_GLCM_Idn	-0.520	0.000 ^a	-0.355	0.021 ^a	-0.045	0.775
Wavelet-HHH_GLCM_Correlation	-0.538	0.000 ^a	-0.416	0.006 ^a	-0.045	0.775

注:ER,雌激素受体;PR,孕激素受体;HER2:人表皮生长因子受体 2;a:结果有统计学意义($P<0.05$)

4 种分类器预测乳腺癌 Luminal 型和非 Luminal 型的敏感度、特异度、准确度及 AUC 见表 3, 可见基于 RF 的分类器的预测性能最佳, 其敏感度、特异度、准确度和 AUC 为 0.838、0.900、0.853 和 0.876。

表 3 基于 4 种模型的预测性能

分类器	敏感度	特异度	准确度	AUC
LR	0.805	0.800	0.803	0.868
RF	0.838	0.900	0.853	0.876
KNN	0.776	0.900	0.831	0.851
SVM	0.933	0.100	0.736	0.854

注: LR, 逻辑回归; RF: 随机森林; KNN: K 近邻; SVM: 支持向量机; AUC: 曲线下面积

3 讨 论

乳腺癌在分子水平上具有高度异质性, 病理类型和临床分期相同的患者治疗反应和预后可呈现明显差异。近年研究发现, ER 和 PR 是乳腺癌分子分型的重要参考依据, 对判断乳腺肿瘤预后、指导内分泌治疗都具有及其重要的价值^[7]。根据 ER 和 PR 的表达情况, 可将乳腺癌分为 Luminal 型和非 Luminal 型; Luminal 型包括传统意义的 Luminal A 型和 Luminal B 型, 对内分泌治疗敏感, 远处转移率低, 预后较好^[8], 非 Luminal 型包括 HER2 过表达型和三阴性 (triple negative, TN), 对内分泌治疗不敏感, 且更易发生肿瘤的复发和转移, 预后较差^[9-10]。左婷婷等^[11]对北京 4 所医院的 4 531 例乳腺癌患者进行随访, 发现 Luminal 型乳腺癌患者的 5 年生存率明显高于非 Luminal 型, 早期判断乳腺癌的分子分型对选择治疗方案和预测预后有积极意义。临床上, 免疫组化法是确定乳腺癌分子分型的常用方法, 但该方法存在有创性且耗时较长的缺点。

影像组学是影像医学的一个新兴研究方法, 具有易获取、无创性和可量化等优点, 通过提取与肿瘤组织的显微结构、生物学特性等相关的信息特征, 从而提高疾病诊断和预测的准确性^[12]。影像组学对肿瘤的鉴别诊断、淋巴结转移、术前分级和预后评估等方面具有较好的应用价值^[13-14]。DCE-MRI 是一种通过评估肿瘤形态和血流动力学来诊断乳腺癌的有效方法, 可以提供高时间分辨率、高空间分

辨率和高信噪比的图像。基于 DCE-MRI 的影像组学在乳腺癌的分子分型中的预测价值是近年来的研究热点, Elizabeth 等^[15]使用肿瘤中心层面的第一期增强图像鉴别 Luminal、HER-2 和 TN 的总体准确率为 0.712; 吴佩琪等^[16]对 79 例患者选用肿瘤最大层面第 4 期增强图像进行影像组学分析, 鉴别 Luminal A 与非 Luminal A、Luminal B 与非 Luminal B、TN 与非 TN 准确率分别为 0.696、0.709、0.772; 而 Wang 等^[17]逐层勾画了整个肿瘤区域最接近强化峰值的图像, 鉴别 TN 和非 TN 的准确率为 0.900。以上文献得到的结果并不完全一致, 可能是因为增强时相选择、层面选择和分子分型方法不同。

本研究中根据时间信号曲线选择强化最明显的时相逐层勾画了整个肿瘤层面, 获得了较高的准确性, 与 Wang 等^[17]的研究结果一致。原因可能有: ①强化最明显的时相更能反映肿瘤的异质性及肿瘤的侵袭性^[18], 而且该时相病灶边界更为清晰, 可减少勾画病灶的误差。②本研究对病灶进行了逐层勾画提取特征值, 比仅使用肿瘤最大层面更好地反映了肿瘤的整体情况及异质性。③不同的分型方法也可能影响预测结果, ER 和 PR 对肿瘤影像特征和异质性的影响比较相似^[18], 非 Luminal 型肿瘤细胞具有较高的细胞增殖能力、易发生 TP53 突变^[19], Luminal A 型和 Luminal B 型之间以及 HER2 过表达型和 TN 之间的差异相对较小。

在本研究结果中, 最终筛选出 4 个特征值与乳腺癌的分子分型有关, 其中 Original_Shape_Maximum 2D Diameter Row、Wavelet-LLH_GLCM_Idn 和 Wavelet-HHH_GLCM_Correlation 3 个特征值与 ER 和 PR 的表达情况均具有显著相关性, 研究认为 ER 和 PR 的表达情况能较好地反映乳腺癌的生物学特性, 是乳腺癌分子分型的重要参考依据^[20]。Original_Shape_Maximum 2D Diameter Row 是基于形状和大小的特征, 提示肿瘤的大小与乳腺癌的分子分型相关, 该结论与之前文献报道一致: 不同激素受体表达的病灶形状和大小存在差异, 激素受体阴性的肿瘤比激素受体阳性的肿瘤往往体积更大^[21]。Wavelet-LLH_GLCM_Idn 和 Wavelet-HHH_GLCM_Correlation 是通过小波变换得到的高阶统计特征, 能较好地显

示肿瘤内部的异质性以及灰度和纹理特征的细微差异^[2], Correlation 反映了各体素灰度值的相关性, Idn 反映了图像局部的异质性。本课题组通过 42 例乳腺癌患者的 DCE-MRI 分析,发现影像组学在乳腺癌 Luminal 和非 Luminal 分子亚型预测方面具有很好的预测性能,对比 4 种分类器,基于 RF 的分类器预测性能最佳,鉴别乳腺癌 Luminal 型和非 Luminal 型的敏感度、特异度、准确度及 AUC 分别为 0.838、0.900、0.853 和 0.876。

本研究尚存在一些不足之处:由于纳入病例数不多,未对乳腺癌 Luminal 和非 Luminal 的分子亚型,即 Luminal A 型、Luminal B 型、HER2 过表达型和 TN 型进行进一步组间比较分析,以后的研究将纳入更多乳腺癌免疫组织化学分子分型,并进一步探讨影像组学特征在乳腺癌分子表达中的预测价值。

综上所述,基于 DCE-MRI 的影像组学在预测乳腺癌 Luminal 型和非 Luminal 型中具有较好的应用价值,基于 RF 的分类器预测性能最佳。

参 考 文 献

- [1] Goss PE, Strasser-Weippl K, Lee-Bychkovsky BL, et al. Challenges to effective cancer control in China, India, and Russia[J]. *Lancet Oncol*, 2014, 15(5):489-538.
- [2] Chen W, Sun K, Zheng R, et al. Cancer incidence and mortality in China, 2014[J]. *Chin J Cancer Res*, 2018, 30(1):1-12.
- [3] Anderson WF, Rosenberg PS, Prat A, et al. How many etiological subtypes of breast cancer: two, three, four, or more?[J]. *J Natl Cancer Inst*, 2014, 106(8):165.
- [4] Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach[J]. *Nat Commun*, 2014, 5:4006.
- [5] Hammond ME, Hayes DF, Dowsett M, et al. American society of clinical oncology/college of American pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer[J]. *J Clin Oncol*, 2010, 28(16):2784-2795.
- [6] Liu C, Ding J, Spuhler K, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer by radiomic signatures from dynamic contrast-enhanced MRI[J]. *J Magn Reson Imaging*, 2019, 49(1):131-140.
- [7] 张作文, 刘胜春, 吴凯南, 等. 雌、孕激素受体状况对青年乳腺癌的远期预后价值[J]. *重庆医科大学学报*, 2013, 38(1):94-97.
- [8] Schnitt SJ. Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy[J]. *Mod Pathol*, 2010, 23(S2):60-64.
- [9] Su Y, Zheng Y, Zheng W, et al. Distinct distribution and prognostic significance of molecular subtypes of breast cancer in Chinese women: a population-based cohort study[J]. *BMC Cancer*, 2011, 11:292.
- [10] Engstrom MJ, Opdahl S, Hagen AI, et al. Molecular subtypes, histopathological grade and survival in a historic cohort of breast cancer patients[J]. *Breast Cancer Res Treat*, 2013, 140(3):463-473.
- [11] 左婷婷, 曾红梅, 郑荣寿, 等. 北京妇女绝经状态与不同分子亚型对乳腺癌患者生存的影响[J]. *中华预防医学杂志*, 2017, 51(5):409-414.
- [12] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data[J]. *Radiology*, 2016, 278(2):563-577.
- [13] Lee SE, Han K, Kwak JY, et al. Radiomics of US texture features in differential diagnosis between triple-negative breast cancer and fibroadenoma[J]. *Sci Rep*, 2018, 8(1):13546.
- [14] Huang SY, Franc BL, Harnish RJ, et al. Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis[J]. *NPJ Breast Cancer*, 2018, 4:24.
- [15] Sutton EJ, Dashevsky BZ, Oh JH, et al. Breast cancer molecular subtype classifier that incorporates MRI features[J]. *J Magn Reson Imaging*, 2016, 44(1):122-129.
- [16] 吴佩琪, 赵可, 吴磊, 等. 基于扩散加权成像和动态增强 MRI 的影像组学特征与乳腺癌分子分型的关系初探[J]. *中华放射学杂志*, 2018, 52(5):338-343.
- [17] Wang J, Kato F, Oyama-Manabe N, et al. Identifying triple-negative breast cancer using background parenchymal enhancement heterogeneity on dynamic contrast-enhanced mri: a pilot radiomics study[J]. *PLoS One*, 2015, 10(11):e143308.
- [18] Chen JH, Baek HM, Nalcioğlu O, et al. Estrogen receptor and breast MR imaging features: a correlation study[J]. *J Magn Reson Imaging*, 2008, 27(4):825-833.
- [19] Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study[J]. *JAMA*, 2006, 295(21):2492-2502.
- [20] Györfy B, Hatzis C, Sanft T, et al. Multigene prognostic tests in breast cancer: past, present, future[J]. *Breast Cancer Res*, 2015, 17:11.
- [21] Li H, Zhu Y, Burnside ES, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set[J]. *NPJ Breast Cancer*, 2016, 2:16012.
- [22] Huynh E, Coroller TP, Narayan V, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer[J]. *Radiother Oncol*, 2016, 120(2):258-266.

(责任编辑:唐秋姗)