



DOI:10.11817/j.issn.1672-7347.2019.03.004

<http://xbyxb.csu.edu.cn/xbwk/fileup/PDF/201903251.pdf>

## 基于CT影像组学对食管鳞状细胞癌病理分化程度的预测

程蕾舒<sup>1</sup>, 吴磊<sup>1</sup>, 陈舒婷<sup>2</sup>, 叶维韬<sup>2</sup>, 刘再毅<sup>1,2</sup>, 梁长虹<sup>1,2</sup>

(1. 华南理工大学医学院, 广州 510006; 2. 广东省人民医院放射科, 广州 510080)

**[摘要]** **目的:** 探讨CT影像组学在预测食管鳞状细胞癌分化程度中的价值。**方法:** 回顾性分析2008年1月到2016年8月经手术病理切片证实为食管鳞状细胞癌, 临床资料完善和术前行CT增强扫描并获得完整图像的160例, 随机分成训练集(103例)和验证集(57例), 运用Matlab软件对图像进行特征提取并进行筛选, 建立影像组学标签。结合获得的影像组学标签和临床资料建立多变量logistic回归分析, 建立影像组学模型(radiomics predictive mode)并进行验证。运用受试者操作特征(receiver operator characteristic, ROC)曲线评价模型术前预测食管鳞状细胞癌分化程度的效能。**结果:** 影像组学标签为食道癌病理分化程度显著的独立预测因素。影像组学预测模型在训练组中的曲线下面积(area under curve, AUC)值为0.791, 敏感度为81.6%, 特异度为72.3%; 在验证集中的AUC值为0.757, 敏感度为70.0%, 特异度为73.0%。**结论:** 基于CT影像提取的影像组学特征构建的模型, 对于术前预测食管癌分化程度具有一定的效能。

**[关键词]** 食管鳞状细胞癌; 分化程度; 影像组学

## CT-based radiomics analysis for evaluating the differentiation degree of esophageal squamous carcinoma

CHENG Leishu<sup>1</sup>, WU Lei<sup>1</sup>, CHEN Shuting<sup>2</sup>, YE Weitao<sup>2</sup>, LIU Zaiyi<sup>1,2</sup>, LIANG Changhong<sup>1,2</sup>

(1. School of Medicine, South China University of Technology, Guangzhou 510006;

2. Department of Radiology, Guangdong Provincial People's Hospital, Guangzhou 510080, China)

### ABSTRACT

**Objective:** To build a CT-based radiomics predictive mode to evaluate the differentiation degree of the esophageal squamous carcinoma.

**Methods:** A total of 160 patients with surgical pathology, complete clinical data and chest CT scanning before operation were retrospectively collected from January 2008 to August 2016. All patients were assigned randomly to a primary data set and an independent validation. Texture analysis was performed on CT images, while the carcinomas were performed by manual segmentation to extract the radiomics features. Radiomics features were extracted and 9 radiomics

收稿日期(Date of reception): 2018-08-07

第一作者(First author): 程蕾舒, Email: kilimanjory@qq.com, ORCID:0000-0003-2673-6078

通信作者(Corresponding author): 梁长虹, Email: cjr.lchh@vip.163.com, ORCID: 0000-0001-8267-150X

基金项目(Foundation item): 国家重点研发计划(2017YFC1309100); 国家自然科学基金(81771912); 广东省省级科技计划项目(2017B020227012)。This work was supported by the National Key R&D Program (2017YFC1309100), the National Natural Science Foundation (81771912), and Science and Technology Planning Project of Guangdong Province (2017B020227012), China.

signatures were finally selected after dimension reduction. Radiomics features were extracted and established via Matlab. Multivariable logistic regression analysis was performed to build the predictive model. A 10-fold cross-validation was used for selecting parameters in the least absolute shrinkage and selection operator (LASSO) model by minimum criteria. The receiver operating characteristic (ROC) curves and areas under ROC curve (AUC) were used to compare the model performance in the primary validation and the independent validation for evaluating the differentiation degree of esophageal squamous carcinoma.

**Results:** Radiomics signature showed great effect in discriminating primary data set and independent validation. The predictive model had a good performance in primary data set. The AUC was 0.791, the sensitivity was 81.6%, and specificity was 72.3%. In the independent validation, the AUC was 0.757, the sensitivity was 70.0%, and the specificity was 73.0%.

**Conclusion:** The predictive model can be used for evaluating the differentiation degree of esophageal squamous carcinoma efficiently, which can be helpful to clinicians in diagnosis and choice of treatment for esophageal squamous carcinoma.

## KEY WORDS

esophageal squamous carcinoma; differentiation degree; radiomics

食管癌在全世界癌症死因中排名第六, 每年死亡人数约40万<sup>[1-2]</sup>。我国食管癌发病和病死率一直居高不下, 约占全球总数的50%。食管癌的5年生存率极低, 发达国家的5年生存率在18%左右, 发展中国家低于5%<sup>[3]</sup>。我国食管癌患者多为鳞状细胞癌。鳞状细胞癌的分化程度与患者治疗及预后息息相关<sup>[4]</sup>。采用相同治疗手段的患者可因为分化程度的不同, 产生完全不同的效果<sup>[5]</sup>, 因此在术前判断食管癌分化程度对于食管癌患者的治疗及预后至关重要<sup>[6]</sup>。传统的CT检查方法更多地应用于食管癌的常规诊断、分期, 对食管鳞状细胞癌分化程度的研究较为匮乏<sup>[7]</sup>。目前新兴的影像组学方法是基于CT图像, 运用机器将图像转化为高维特征, 实现了图像数据化, 克服了传统影像医师自身因素对诊断的影响, 使结果更加客观, 亦能使影像医师更好地分析肿瘤内部环境, 有利于肿瘤生物学包括分化程度的研究<sup>[8-9]</sup>。目前影像组学并未开展对于食管癌分化程度的研究。本研究旨在基于CT图像, 结合临床信息、病理等, 运用最新的影像组学方法, 构建预测模型, 对食管鳞状细胞癌分化程度进行研究, 以期为临床决策提供更多的帮助<sup>[10-12]</sup>。

## 1 资料与方法

### 1.1 病例资料

回顾性分析在2008年1月到2016年8月广东省人民医院接受治疗的食管癌患者的资料。纳入标准: 1) 手术病理结果证实为食管鳞状细胞癌患者; 2) 无放射治疗(以下简称放疗)、化学药物治疗(以下简称化疗)

史的患者; 3) 无造影剂过敏史及其他相关检查禁忌证的患者; 4) 临床相关资料完善的患者。排除标准: 图像质量过差, 影响正常诊断的患者资料; 肿块位于食管、胃交界处患者的资料。收集病例数203例, 剔除43例, 入组160例。其中男117例, 女43例。

### 1.2 CT图像采集

对所有入组患者采用GE 64排螺旋电子计算机断层扫描仪(computed tomography, CT)行胸部增强扫描, 视野(field of view, FOV)为350 mm×350 mm, 管电压为120 kV, 管电流为160 mA。先行CT平扫, 然后经肘正中静脉注射对比剂碘普安, 流率为3 mL/s, 剂量为90~100 mL, 30 s延迟后行动脉期、60 s延迟后行实质期CT增强扫描。将扫描获得的CT图像上传至医学影像信息系统(Picture Archiving and Communication Systems, PACS)。

#### 1.2.1 肿瘤分割

在获得的图像中, 选取质量满足诊断要求的图像, 由2名具有2年以上诊断经验的放射科医师在互不影响的情况下, 利用3D Slicer软件(<https://www.slicer.org/>)手动勾画感兴趣区域(region of interest, ROI), 勾画时选取肿块最大层面的平扫、动脉期、实质期图像均一致的图像, 沿肿瘤边界尽可能地包含肿瘤区域, 并尽量避免坏死、出血区域, 以免影响测定结果的准确性。

#### 1.2.2 影像组学特征提取以及特征筛选

对于分割后的图像, 运用Matlab软件(<https://cn.mathworks.com/>)对图像进行影像组学特征提取。将纳入的160例患者随机分成训练集(103例)和验证

集(57例), 每例提取9 790个影像组学特征, 提取的影像组学特征包括以下4类。1)一阶灰度统计量: 峰度、偏度、能量等。2)形态学特征: 肿瘤直径、体积、表面积等。3)纹理特征: 基于灰度共生矩阵(GLCM)、灰度区域矩阵(GLSZM)、灰度游程矩阵(GLRLM)等的特征。4)小波变换等特征。从提取的上述特征中, 采用组内相关系数(intraclass correlation coefficient, ICC), 选取ICC>0.9的特征, 其中6 140个特征通过ICC检验; 随后, 在训练集上计算特征与特征间的皮尔逊相关系数( $r$ ), 其 $|r|>0.9$ 的5 922个特征被舍弃, 再对剩余的218个特征进行多重共线性检测, 剔除其中5个具有多重共线的特征; 再对剩余的213个特征进行病理分化程度的效能评价, 利用受试者操作特征(receiver operator characteristic, ROC)曲线下面积(are a under the curve, AUC)作为评价指标, 从213(0.441<AUC<0.616)个特征中选出46个(0.550<AUC<0.616)特征。最后, 运用最小绝对收缩和选择算子(the least absolute shrinkage and selection operator, LASSO)回归模型对剩下的特征进行分析并从中选取9个特征用于构建影像组学模型。LASSO是一种压缩估计, 通过构造一个惩罚函数, 压缩部分变量的系数使之为零, 从而得到一个比较精炼的模型。验证集用于验证该模型的效能。

### 1.3 病理分级

依据WHO病理分级标准<sup>[13]</sup>, 请有多年从业经验

的病理科医生将采集的病理样本分为高中分化和低分化两组(由于收集的病例中高分化病例较少, 故将高、中分化病例并为一组)。

### 1.4 统计学处理

运用R软件(<https://www.r-project.org/>, 3.4.0版)进行所有统计学分析。对于连续变量如年龄、影像组学评分采用非参数秩和检验; 分类变量如性别、肿瘤部位, 采用皮尔逊 $\chi^2$ 检验;  $P<0.05$ 为差异具有统计学意义。

利用AUC对提取的影像组学特征进行评估筛选; 运用AUC对构建的相关模型评价食管鳞状细胞癌分化程度的效能进行评估。AUC值在0.7~0.8之间代表具有一定的诊断效能, AUC值>0.8代表具有良好的诊断效能<sup>[14]</sup>。

## 2 结果

### 2.1 患者基线资料、肿块位置、影像组学模型

收集的基线资料包括患者的年龄[(58.11±7.85)岁]、性别、食管鳞状细胞癌生长部位, 各基线指标在训练集和验证集间差异均不具有统计学意义( $P>0.05$ )。构建的影像组学评分在训练集和验证集中差异均有统计学意义( $P=0.001$ , 表1)。

表1 食管癌患者训练集和验证集基线资料结果比较

Table 1 Comparisons of baseline data in primary and independent validation

组别	年龄/岁	性别分布/例		肿块位置分布/例			影像组学评分
		男	女	食管上段	食管中段	食管下段	
训练集							
高中分化组	58.38±8.08	48	17	6	36	23	0.768(0.456~0.979)**
低分化组	57.42±7.97	26	12	5	19	14	0.298(0.057~0.534)
验证集							
高中分化组	58.08±7.87	26	11	1	8	10	0.674(0.541~0.987)††
低分化组	58.60±7.33	17	3	5	19	14	0.473(0.064~0.594)

与训练集低分化组比较, \*\* $P<0.001$ ; 与验证集低分化组比较, †† $P<0.001$

### 2.2 影像组学模型的构建

对9 790个影像组学特征降维, 运用LASSO-logistic回归对剩余的46个影像组学特征进行进一步筛选。在本研究中, 使用10倍交叉验证, 选取交叉验证中误差最小、AUC较大处的Lambda值用于

特征选取(图1)。选出的9个特征用于构建影像组学模型, 9个特征分别为LoG\_1.5\_RLV, LoG\_2.0\_NGTDI\_Contrast, bior3.7\_1\_SRLGLE, bior5.5\_4\_MP, rbio1.1\_1\_HGZE, db10\_5\_IDN, db10\_5\_LGZE, db8\_4\_SRLGLE, rbio3.1\_1\_SZN。构建模

型如下: 影像组学标签评分(radiomics-score, Rad-score) $=0.591-0.214\times\text{LoG}_{1.5\_RLV}+0.075\times\text{LoG}_{2.0\_NGTDM\_Contrast}+0.242\times\text{bior}_{3.7\_1\_SRLGLE}+0.34\times\text{rbio}_{1.1\_1\_HGZE}-0.263\times\text{rbio}_{3.1\_1\_SZN}+0.211\times\text{db}_{8\_4\_SRLGLE}-0.04\times\text{bior}_{5.5\_4\_MP}-0.372\times\text{db}_{10\_5\_IDN}-0.271\times\text{db}_{10\_5\_LGZE}$ 。1)RLV(run length variance, 游程方差): 用于度量运行长度的运行差异; 2) contrast(对比度): 用于衡量空间强度的变化; 3) SRLGLE(short run low gray level emphasis, 短游程低灰度游程强度): 衡量较低灰度值的短游程的分布情况; 4)MP(maximum probability, 最大概率): 反映矩阵中出现最多的共生对的概率; 5)HGZE(high gray-level zone emphasis, 高灰度值空间强度): 衡量高灰度

值的连通区域的分布; 6)IDN(inverse difference moment normalized, 归一化逆差): 衡量图像局部均匀性; 7)LGZE(low gray-level zone emphasis, 低灰度值空间强度): 衡量低灰度值的连通区域的分布; 8)SZN(size zone non-uniformity, 空间不均匀性): 衡量整个图像中连通区域的可变性。

### 2.3 影像组学模型的效能评价

运用影像组学构建的模型进行效能评价, 在训练集中, AUC值为0.791, 敏感度为81.6%, 特异度为72.3%; 在验证集中, AUC值为0.757, 敏感度为70.0%, 特异度为73.0%(图2A和2B)。

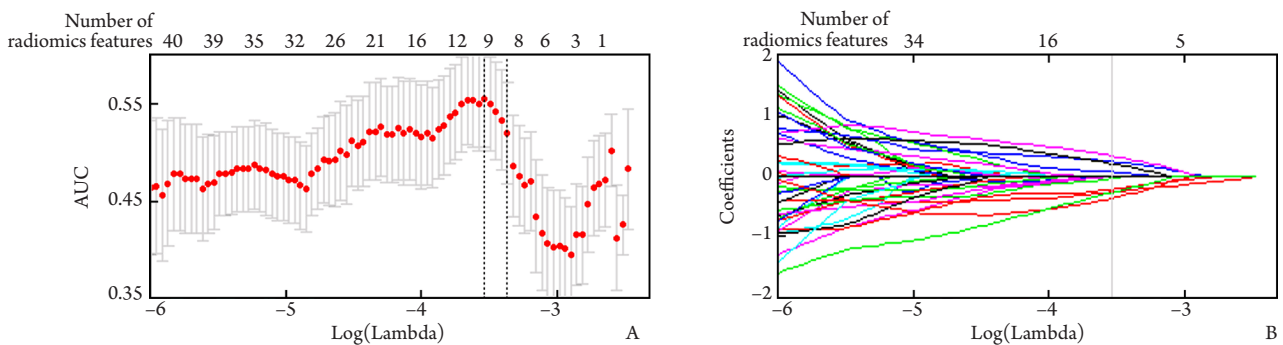


图1 Logistic回归模型构建的纹理特征

Figure 1 Texture feature construction using the binary logistic regression model

A: Ten-fold cross-validation was used for selecting parameters in the least absolute shrinkage and selection operator (LASSO) model by minimum criteria. At the optimal values  $\log(\lambda)$ , dotted vertical lines were set via the minimum criteria, where 9 features were selected. B: Coefficient of convergence in LASSO via 10-fold cross-validation. Each curve represents the change of its independent variable coefficient. Vertical line was set at the non-zero coefficients selected via 10-fold cross-validation, where 9 non-zero coefficients were included

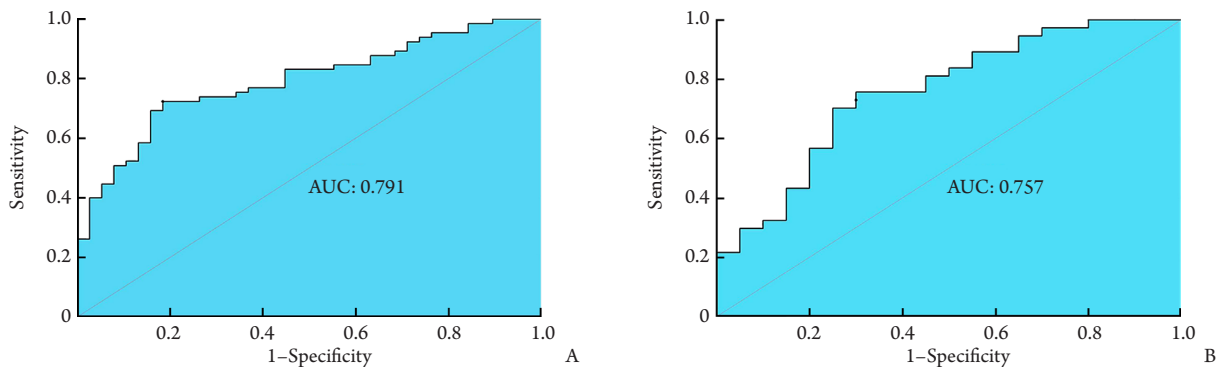


图2 ROC曲线

Figure 2 ROC curve

A: In validation cohort, AUC is 0.791, sensitivity is 81.6%, and specificity is 72.3%. B: In validation cohort, AUC is 0.757, sensitivity is 70.0%, and specificity is 73.0%

### 3 讨论

食管鳞状细胞癌与患者自身分化缺陷有关, 分化程度与患者的治疗效果及预后有着很大的关联, 因此术前对食管鳞状细胞癌分化程度进行预测, 能使患者受益。新兴的影像组学能更好地描述肿瘤本身异质性及生物学特性<sup>[14]</sup>, 在很大程度上弥补了传统诊断方法的不足。本研究将影像组学方法创新性地应用于食管鳞状细胞癌分化程度的判断, 通过从CT影像中提取肿瘤的影像特征, 结合相关临床信息, 构建出具有筛选效能的预测模型。对9 790个影像组学特征降维后, 运用LASSO-logistic回归对剩余的46个影像组学特征进一步筛选, 提取出9个有意义的组学特征, 在此基础上构建了多变量logistic模型<sup>[15]</sup>。运用构建的模型对训练集和验证集的食管鳞状细胞癌的病理分化程度进行预测。在训练集中, AUC值为0.791, 在验证集中, AUC值为0.757, 两组AUC值均在0.7~0.8之间, 表明该模型对食管鳞状细胞癌的分化程度具有一定的评价效能。研究还对收集的临床信息进行了研究, 结果表明: 对于食管鳞状细胞癌来说, 年龄、性别、肿块的生长部位都不能作为预测病理分化程度的独立预测因素, 无法用于模型的构建。但由于本研究收集的病例数不足, 且未进行多中心研究, 临床资料是否对预测病理分化程度构成影响还有待商榷。不论是在训练集还是验证集中, 构建的影像组学模型均对食管鳞状细胞癌病理分化程度具有一定的预测效能, 可以对术前食管鳞状细胞癌的病理分化程度预测提供帮助。

传统影像学主要采用<sup>18</sup>F-FDG正电子发射计算机断层显像(positron emission tomography, PET)/CT融合图像(FDG PET/CT)以及宝石CT能谱的方法对食管鳞状细胞癌分化程度进行预测研究。冯瑞等<sup>[16]</sup>的研究表明标准摄取值(standard uptake value, SUV)值与食管鳞状细胞癌分化程度呈正相关, 恶性肿瘤的分化程度一般与其代谢水平相关, SUV值则为反映肿瘤葡萄糖代谢情况的半定量指标, 因此认为SUV值可以反映肿瘤的分化程度。但FDG PET/CT检查费用高昂, 很难在临床中进行推广应用。刘月华等<sup>[17]</sup>则利用宝石CT能谱对食管鳞状细胞癌分化程度进行研究, 结果表明: 不同分化程度的食管鳞状细胞癌的能谱特征参数和能谱曲线具有差异性。因此他们认为通过分析能谱特征参数和能谱曲线可以对食管鳞状细胞癌分化程度的判断提供帮助。但此研究纳入的病例数仅有9例, 研究结果具有一定局限性, 有待验证。相比传统的研究方法, 影像组学在经济、研究样本量以及可靠性方面均具有一定优越性。

影像组学虽然近年来才兴起, 但对于肿瘤分

期、分级方面的预测等研究已经取得了一定的进展。Huang等<sup>[9]</sup>创新性地应用影像组学对早期非小细胞肺癌分期进行研究, 从大量图像中提取筛选出了有意义的影像组学特征, 运用LASSO方法构建了有效的影像组学模型, 在非小细胞肺癌术前预测方面具有一定效能, 为临床决策提供了帮助。此外, 影像组学在乳腺癌、结直肠癌等多种癌症中的预测作用的研究均为临床提供了诊断新思路。影像组学方兴未艾, 食管癌相关研究仍需进一步开展。

本研究的创新点在于运用影像组学对食管鳞状细胞癌的术前病理分化程度进行了预测。但本研究的不足在于纳入的病例数不足, 收集的临床资料不足; 没有把相关基因纳入研究, 构建的模型不够完善; 验证集数量不足, 模型未能得到充分验证; 收集的病例仅来自一家医院, 并未开展多中心研究。今后可补充收集更多的病例, 并完善临床资料, 以期得到更具有预测效能的影像组学模型。

综上所述, 基于CT图像提取的影像组学特征和临床资料联合构建的模型, 对于术前预测食管癌分化程度具有一定的效能, 能为临床医师的决策提供辅助作用。

**利益冲突声明:** 作者声称无任何利益冲突。

### 参考文献

- [1] 乔友林. 食管癌流行病学研究的重要里程碑[J]. 中国肿瘤临床, 2016, 43(12): 500-501.  
QIAO Youlin. An important milestone in the research on epidemiology of esophageal cancer[J]. Chinese Journal of Clinical Oncology, 2016, 43(12): 500-501.
- [2] Ferlay J, Soerjomatamm I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012[J]. Int J Cance, 2015, 136(5): E359-E386.
- [3] Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012[J]. CA Cancer J Clin, 2015, 65(2): 87-108.
- [4] 陆再英, 钟南山, 谢毅, 等. 内科学[M]. 7版. 北京: 人民卫生出版社, 2008: 376-377.  
LU Zaiying, ZHONG Nanshan, XIE Yi, et al. Internal Medicine[M]. 7th ed. Beijing: People's Medical Publishing House, 2008: 376-377.
- [5] Zhan C, Shi Y, Lu C, et al. Pyruvate kinase M2 is highly correlated with the differentiation and the prognosis of esophageal squamous cell cancer[J]. Dis Esophagus, 2013, 26(7): 746-753.
- [6] Ohashi S, Miyamoto S, Kikuchi O, et al. Recent advances from basic and clinical studies of esophageal squamous cell carcinoma[J]. Gastroenterology, 2015, 149(7): 1700-1715.

- [7] Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges[J]. *Magn Reson Imaging*, 2012, 30(9): 1234-1248.
- [8] 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.
- [9] Huang Y, Liu Z, He L, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer[J]. *Radiology*, 2016, 281(3): 947-957.
- [10] Lambin P, Rios-Vela, Quez E, et al. Radiomics: extracting more information from medical images using advanced feature analysis[J]. *Eur J cancer*, 2012, 48(4): 441-446.
- [11] Li Q, Kim J, Balagurunathan Y, et al. Imaging features from pretreatment CT scans are associated with clinical outcomes in nonsmall-cell lung cancer patients treated with stereotactic body radiotherapy[J]. *Med Phys*, 2017, 44(8): 4341-4349.
- [12] Hamihon SR, Aaltonen LA. World Health Organization Classification of Tumours[M]// Pathology and Genetics of Tumours of the Digestive System. IARC Press: Lyon, 2000: 16-17.
- [13] Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases[J]. *Radiology*, 1983, 148(3): 839-843.
- [14] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data[J]. *Radiology*, 2016, 278(2): 563-577.
- [15] Hang S, Hamit M, Yan C, et al. Texture analysis and classification on Xinjiang kazakh esophageal cancer images[C]. International Conference on Biotechnology and Medical Science, Nanjing, China, 2016: 297-306.
- [16] 冯瑞, 李明焕, 孔莉, 等. 食管癌原发灶PET-CT氟代脱氧葡萄糖摄取及其与临床病理参数的相关性[J]. *中华肿瘤杂志*, 2009, 31(6): 452-454.
- FENG Rui, LI Minghuan, LONG Li, et al. Correlation between PET-CT <sup>18</sup>F-FDG uptake in primary lesions and clinical pathological parameters in esophageal carcinoma patients[J]. *Chinese Journal of Oncology*, 2009, 31(6): 452-454.
- [17] 刘月华, 朱绍成, 史大鹏, 等. CT能谱成像术前评估食管鳞状细胞癌病理分级的临床价值[J]. *中华医学杂志*, 2017, 97(43): 3406-3411.
- LIU Yuehua, ZHU Shaocheng, SHI Dapeng, et al. Clinical value of spectral CT imaging in preoperative evaluation of pathological grading of esophageal squamous cell carcinoma[J]. *National Medical Journal of China*, 2017, 97(43): 3406-3411.

(本文编辑 陈丽文)

**本文引用:** 程蕾舒, 吴磊, 陈舒婷, 叶维韬, 刘再毅, 梁长虹. 基于CT影像组学对食管鳞状细胞癌病理分化程度的预测[J]. *中南大学学报(医学版)*, 2019, 44(3): 251-256. DOI:10.11817/j.issn.1672-7347.2019.03.004

**Cite this article as:** CHENG Leishu, WU Lei, CHEN Shuting, YE Weitao, LIU Zaiyi, LIANG Changhong. CT-based radiomics analysis for evaluating the differentiation degree of esophageal squamous carcinoma[J]. *Journal of Central South University. Medical Science*, 2019, 44(3): 251-256. DOI:10.11817/j.issn.1672-7347.2019.03.004