



DOI:10.11817/j.issn.1672-7347.2021.200074

http://xbyxb.csu.edu.cn/xbwk/fileup/PDF/202104385.pdf

基于影像组学的logistic回归模型预测胶质瘤分级

孙贤婷¹, 廖伟华¹, 曹冬¹, 赵跃龙², 周高峰¹, 王冬翠¹, 毛弈韬¹

(1. 中南大学湘雅医院放射科, 长沙 410008; 2. 华南理工大学计算机科学与工程学院, 广州 510640)

[摘要] **目的:** 胶质瘤是最常见的颅内原发中枢神经系统肿瘤, 胶质瘤的分级对临床治疗及随访方案的选择、预后的评估有重要指导意义。本研究目的在于探讨基于影像组学的logistic回归模型预测胶质瘤病理分级的可行性。**方法:** 回顾性收集2012年1月至2018年12月经手术病理切片证实为胶质瘤的146例患者。手动分割患者增强T₁加权成像(contrast-enhanced T₁-weighted imaging, T₁WI+C)图像中的胶质瘤区域, 形成3D感兴趣区(region of interest, ROI); 提取41个影像特征; 采用最小绝对收缩和选择运算(least absolute shrinkage and selection operator, LASSO)二元logistic回归法筛选与胶质瘤病理分级最相关的特征并计算影像组学得分(radiomics score, Rad-score); 采用单因素logistic回归建模方法建立预测模型; 用受试者操作特征(receiver operating characteristic, ROC)曲线评估模型的区分能力, 评估指标为曲线下面积(area under the curve, AUC)。利用Hosmer-Lemeshow检验衡量模型预测的准确性。**结果:** 筛选出5个与胶质瘤病理分级最相关的特征, 用这5个特征构建的预测胶质瘤病理分级的logistic回归模型的ROC曲线AUC为0.919, 具有很好的区分能力, 其校准曲线经Hosmer-Lemeshow检验, 与理想曲线的差异无统计学意义($P=0.808$), 预测准确性高。**结论:** 基于影像组学的logistic回归模型可以有效地对胶质瘤病理分级进行预测, 有望成为术前预测胶质瘤分级的辅助方法。

[关键词] 胶质瘤; 分级; 最小绝对收缩和选择运算; 影像组学; logistic回归

A logistic regression model for prediction of glioma grading based on radiomics

SUN Xianting¹, LIAO Weihua¹, CAO Dong¹, ZHAO Yuelong², ZHOU Gaofeng¹,
WANG Dongcui¹, MAO Yitao¹

(1. Department of Radiology, Xiangya Hospital, Central South University, Changsha 410008;

2. School of Computer Science and Engineering, South China University of Technology, Guangzhou 510640, China)

ABSTRACT

Objective: Glioma is the most common intracranial primary tumor in central nervous system. Glioma grading possesses important guiding significance for the selection of clinical treatment and follow-up plan, and the assessment of prognosis. This study aims to

收稿日期(Date of reception): 2020-02-06

第一作者(First author): 孙贤婷, Email: sunxianting2009@163.com, ORCID: 0000-0001-6013-5976

通信作者(Corresponding author): 毛弈韬, Email: maoyt@csu.edu.cn, ORCID: 0000-0001-5734-1744

基金项目(Foundation item): 国家自然科学基金(81701847); 湖南省自然科学基金(2017JJ3497, 2017JJ2387)。This work was supported by the National Natural Science Foundation (81701847) and the Natural Science Foundation of Hunan Province (2017JJ3497, 2017JJ2387), China.

explore the feasibility of logistic regression model based on radiomics to predict glioma grading.

Methods: Retrospective analysis was performed on 146 glioma patients with confirmed pathological diagnosis from January, 2012 to December, 2018. A total of 41 radiomics features were extracted from contrast-enhanced T₁-weighted imaging (T₁WI+C) lesion by manual segmentation. Least absolute shrinkage and selection operator (LASSO) was used to select the most-predictive radiomics features for pathological grading and to calculate radiomics score (Rad-score) of each patient. A logistic regression model was built to explore the correlation between glioma grading and Rad-score. Receiver operating characteristic (ROC) curve was performed to evaluate the model's predictive ability with area under the curve (AUC) for the evaluation index. Hosmer-Lemeshow test was used to measure the model's predictive accuracy.

Results: A total of 5 imaging features selected by LASSO were used to establish a logistic regression model for predicting glioma grading. The model showed good discrimination with AUC value of 0.919. Hosmer-Lemeshow test showed no significant difference between the calibration curve and the ideal curve ($P=0.808$), indicating high predictive accuracy of the model.

Conclusion: The logistic regression model using radiomics exhibits a relatively high accuracy for predicting glioma grading, which may serve as a complementary tool for preoperative prediction of glioma grading.

KEY WORDS

glioma; grading; least absolute shrinkage and selection operator; radiomics; logistic regression

胶质瘤是最常见的颅内原发中枢神经系统肿瘤, 发病率(4.67~5.73)/100 000^[1]。根据2016版WHO中枢神经系统肿瘤分类^[2], 胶质瘤包括星形细胞瘤、少突胶质细胞瘤、室管膜瘤、混合性神经元-神经胶质肿瘤及其他来源胶质瘤。根据病理组织学, 胶质瘤分为I~IV级, 其中I~II级为低级别, III~IV级为高级别, 胶质瘤的分级对临床治疗及随访方案的选择、患者的预后有重要指导意义^[2-4]。目前胶质瘤的分级主要还是依靠活检或手术后的病理诊断, 病理诊断虽是分级的金标准, 但有创、耗时长等缺点限制了其应用。在常规影像学图像中有很多潜在的信息无法被人眼发现, 导致其在胶质瘤精准分级方面的应用仍面临巨大挑战。随着人工智能及精准医疗的飞速发展, 影像组学已成为临床研究的热点。影像组学通过挖掘数据提取疾病影像数据中所包含的定量特征, 进而找出疾病的影像标志物或预测疾病分类、分级, 具有无创、经济、高效和可重复的优点, 有助于临床决策。基于此, 近年来研究者们利用影像

组学数据对胶质瘤进行分级取得了较好的效果, 但使用何种影像序列及何种模型效果最好尚有待探索^[5-11]。本研究采用手动方式分割患者增强T₁加权成像(contrast-enhanced T₁-weighted imaging, T₁WI+C)图像中的胶质瘤区域, 并进一步提取影像特征, 通过筛选特征并构建logistic回归模型在术前对高、低级别胶质瘤进行分级, 以期满足临床上胶质瘤的术前精准预测及个性化治疗的需求。

1 对象与方法

1.1 对象

回顾性收集中南大学湘雅医院2012年1月至2018年12月经手术病理证实的胶质瘤患者。纳入标准: 1)术前在同一台机器采集MRI数据; 2)扫描MRI前未接受穿刺、放射治疗和化学治疗; 3)术后病理学诊断具有精确分级。排除标准: 1)MRI数据伪影大; 2)合并其他严重的神经系统疾病。

1.2 MRI 扫描方法

采用 GE Signa HDx 3.0T 磁共振成像仪及 8 通道头部相控阵线圈采集患者的颅脑 MRI 平扫+增强数据, 并在线圈的内部填充适量的 MRI 兼容的泡沫以固定被试头部, 减少扫描过程中的头动伪影。1) 轴位 T_2 加权成像 (T_2 -weighted imaging, T_2 WI) 序列参数: 重复时间 (time of repetition, TR)=4 480 ms, 回波时间 (time of echo, TE)=120 ms, 层厚=5 mm, 层间距=1.5 mm, 层数=20, 视野 (field of view, FOV)=240 mm×240 mm, 矩阵=320×288。2) 轴位 T_2 液体衰减反转恢复成像 (T_2 -fluid attenuated inversion recovery, T_2 FLAIR) 序列参数: TR=8 500 ms, TE=168 ms, 矩阵=320×256, 层厚、层间距、层数及 FOV 与 T_2 WI 一致。3) 轴位 T_1 加权成像 (T_1 -weighted imaging, T_1 WI) 序列参数: TR=2 000 ms, TE=24 ms, 矩阵=320×224, FOV、层厚、层间距、层数与 T_2 WI 一致。4) 轴位增强 T_1 加权成像 (contrast enhancement T_1 -weighted imaging, T_1 WI+C) 序列参数: TR=1 725 ms, TE=24 ms, 矩阵=320×256, 层厚、层间距、层数及 FOV 与 T_2 WI 一致。

1.3 图像分析

1.3.1 图像和病灶感兴趣区的获取

图像来自中南大学湘雅医院放射科影像归档和通信系统 (picture archiving and communication systems, PACS)。基于 MATLAB 的图形用户界面 (graphic user interface, GUI) 获取病灶感兴趣区 (region of interest, ROI)。首先由 2 名具有 3 年以上工作经验的影像医师独立手动分割每例患者 T_1 WI+C 图像中的胶质瘤区域, 形成 3D ROI 图像; 然后交由 1 位高级职称影像医师进行审核, 当审核者认为 ROI 绘制错误或者差异过大时, 则返回上一步骤进行修改; 最终使用的 ROI 为 2 名医师各自勾画 ROI 的重叠区域。胶质瘤区域的分割标准基于轴位 T_1 WI+C 图像中的肿瘤实性区域, 包括肿瘤异质性导致的无强化区, 排除内部明显的坏死、囊变区域。对于部分无对比增强或强化不明显的低级别胶质瘤, 肿瘤分割需要同时参考轴位 T_2 WI 及 T_2 FLAIR 序列。

1.3.2 影像特征的提取

根据既往研究^[10]提取 41 个影像特征。1) 3 个全局特征: 峰度、偏度、方差。2) 7 个灰度共生矩阵 (gray-level co-occurrence matrix, GLCM) 特征: 对比度、相关性、能量、熵、均数和、同质性、方差。3) 13 个灰度游程矩阵 (gray-level run-length matrix, GLRLM) 特征: 长程强调度、短程强调度、游程非一致性、游程方差、游程百分比、灰度非一致性、灰度方差、

高灰度游程强调度、低灰度游程强调度、长程低灰度强调度、短程低灰度强调度、长程高灰度强调度、短程高灰度强调度。4) 13 个灰度区域大小矩阵 (gray-level size zone matrix, GLSZM) 特征: 高灰度区域强调度、低灰度区域强调度、区域百分比、区域大小非一致性、区域方差、灰度非一致性、灰度方差、大区域高灰度强调度、小区域高灰度强调度、大区域低灰度强调度、小区域低灰度强调度、大区域强调度、小区域强调度。5) 5 个邻域灰度差矩阵 (neighborhood gray-tone difference matrix, NGTDM) 特征: 繁度、糙度、复杂度、对比度、劲度。与既往研究^[10]不同的是, 这 41 个影像特征均基于 3D 图像进行计算。本研究建立 logistic 回归模型的最终目的是预测胶质瘤样本的病理分级, 而单一的样本数据无法进行类似群体数据的归一化处理, 故在本研究中, 影像特征均使用其原始值, 不进行归一化处理。

1.3.3 影像特征的筛选

上述 41 个影像特征对于胶质瘤的病理分级并非都有用, 其中包含一些冗余的特征, 而选取特征过多会增加过拟合的风险, 因此降低影像特征的维度以选择与肿瘤级别高度相关的指标是非常必要的^[12]。本研究采用最小绝对收缩和选择运算 (least absolute shrinkage and selection operator, LASSO) 二元 logistic 回归法筛选变量, 该方法通过在回归方程中引入 1 个惩罚系数 λ , 对冗余变量进行“惩罚”, 即将其系数逐步压缩至 0, 可以有效地筛选出与结局最相关的变量, 降低自变量维度, 从而避免过拟合的风险^[13]。

1.4 模型的建立

采用单因素 logistic 回归建模方法建立预测模型。首先, 将筛选得到的影像特征的权重加和, 得到总体的影像组学得分 (radiomics score, Rad-score); 然后, 用 Rad-score 与胶质瘤的病理分级进行单因素 logistic 回归分析, 得到拟合后的 logistic 回归方程, 建立模型。该模型计算出样本为高级别胶质瘤的概率, 并根据约登指数最大原则划定预测概率阈值, 大于阈值的为高级别胶质瘤, 小于阈值的为低级别胶质瘤^[14]。

1.5 统计学处理

采用 MATLAB 中的统计工具箱进行统计学分析。分别使用两样本 t 检验、卡方检验分析高级别胶质瘤组与低级别胶质瘤组年龄、性别构成差异是否有统计学意义。采用组间相关系数 (inter-class correlation coefficient, ICC) 评估 2 名医师独立获取的肿瘤 ROI 影像特征的一致性, ICC>0.75 被认为一致性良好^[15]。

用受试者操作特征(receiver operating characteristic, ROC)曲线评估模型的区分能力, 评估指标为曲线下面积(area under the curve, AUC)。利用 Hosmer-Lemeshow 检验来判断校准曲线与理想曲线是否存在差异, 继而衡量模型预测的准确性^[16-17]。P<0.05 为差异有统计学意义。

2 结果

2.1 一般资料

共纳入胶质瘤患者 146 例, 其中高级别胶质瘤患者(高级别组)70 例, 低级别胶质瘤患者(低级别组)76 例。高级别组与低级别组在年龄、性别构成上的差异无统计学意义(均 P>0.05, 表 1)。

表 1 低级别胶质瘤与高级别胶质瘤患者的基本情况比较
Table 1 Comparison of basic conditions between high-grade glioma and low-grade glioma

组别	n	男/女	年龄/岁
高级别组	70	29/41	20.5~70.2(50.3±12.7)
低级别组	76	30/46	19.7~73.3(44.2±11.3)
总体	146	59/87	19.7~73.3(47.6±11.9)

2.2 一致性检验结果

2 名影像医师提取的 ROI 影像特征 ICC 为 0.893~0.967, 提示 2 名医师提取 ROI 具有较高的组间一致性。

2.3 特征筛选结果

在图 1A 中, 横坐标表示惩罚系数 λ 的对数 (logλ), 纵坐标表示模型的 AUC, 曲线上的数字表示模型中剩余的特征数量, 可见随着惩罚系数 λ 逐渐变大, 特征数量逐步减少。左侧虚线表示 AUC 最大时对应的 λ 值和筛选到的特征数量, 右侧虚线表示在最大 AUC 的基础上增加 1 倍标准误时模型中剩余特征的数量进一步减少。选用 1 倍标准误时对应的参数可以在简化模型的同时保证一定的准确率, 提高模型的效能^[16]。

在图 1B 中, 横坐标表示惩罚系数 λ 的对数 (logλ), 纵坐标表示模型中特征的系数, 曲线上的数字表示模型中系数不为 0 的特征的数量。随着 λ 逐渐变大, 越来越多的特征系数被压缩为 0, 在 1 倍标准误时(右侧虚线处)只剩余 5 个非 0 特征。因此, 通过 LASSO 二元 logistic 回归法共筛选得到 5 个最相关的影像特征(表 2)。最终 Rad-score 的计算便是将这 5 个影像特征的值与其系数(权重)的乘积相加。

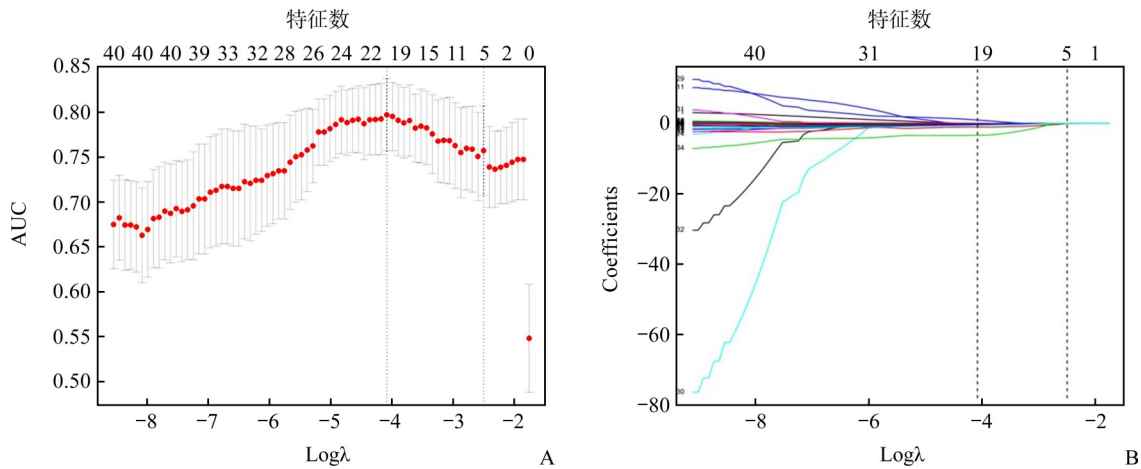


图 1 LASSO 回归变量筛选图

Figure 1 Feature selection using LASSO

A: Variation of AUC. The horizontal axis shows logλ and the vertical axis shows the AUC. The numbers above the curve represent the number of feature of nonzero coefficient. The left dotted line represents the feature number corresponding to the maximum AUC, while the right dotted line represents that corresponding to 1 standard error of AUC. B: The shrinkage plot of coefficients. The horizontal axis shows logλ and the vertical axis shows coefficient. The numbers above the curve represent the number of features with nonzero coefficient.

表2 LASSO 二元 logistic 回归法筛选得到的影像特征

Table 2 Radiomics features selected by LASSO binary logistic regression

类别	中文名称	英文名称	英文缩写	系数
灰度共生矩阵	均数和	sum average	SA	0.048
	方差	variance		0.009
灰度游程矩阵	灰度非一致性	gray-level non-uniformity	GLN	-0.015
灰度区域大小矩阵	区域大小非一致性	zone-size non-uniformity	ZSN	-0.092
邻域灰度差矩阵	对比度	contrast		0.116

2.4 Logistic 回归模型

146 例胶质瘤患者的 Rad-score 瀑布图见图 2A: 红色为高级别胶质瘤, 绿色为低级别胶质瘤, 每根柱子所对应高度的纵坐标值即为相应患者的 Rad-score。根据 ROC 曲线(图 2B), 算得截断值为 2.914, AUC 值为 0.919, 敏感度为 0.886, 特异度为 0.921。用 Rad-score 进行单因素 logistic 回归建模, 以胶质瘤

级别为结局变量(高级别为 1, 低级别为 0), 算得 Rad-score 回归系数为 1.584($P < 0.0001$), 表明 Rad-score 与胶质瘤级别显著相关。其最终的 logistic 回归方程为:

$$P = \frac{1}{1 + \exp(4.780 - 1.584 \times \text{Rad-score})}$$

其中 P 表示样本为高级别胶质瘤的概率。

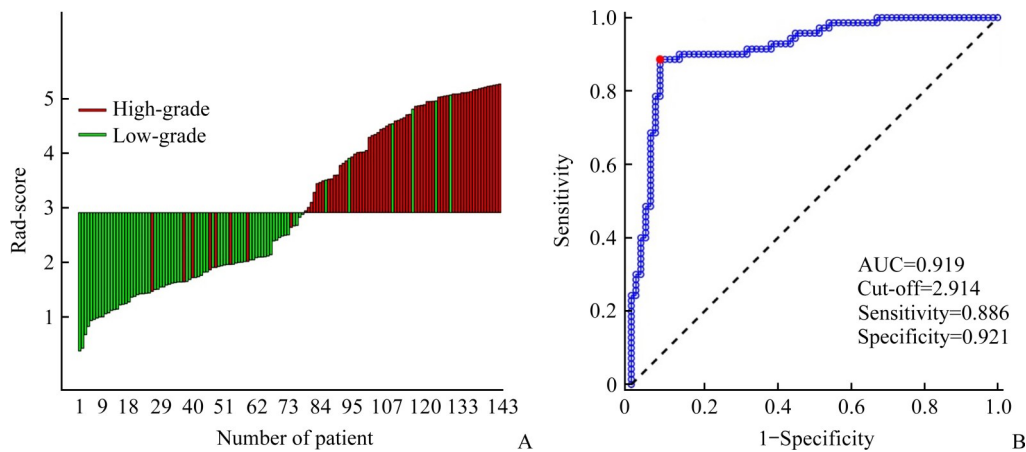


图2 Rad-score瀑布图及ROC曲线

Figure 2 Waterfall plot and ROC curve

A: Waterfall plot of Rad-score. The vertical axis shows the value of Rad-score; red bars represent patients with actual high-grade glioma; green bars represent patients with actual low-grade glioma; the transverse line represents the cut-off point in ROC curve. Bars under the transverse line represent the predicted patients with low-grade glioma, bars above the transverse line represent the predicted patients with high-grade glioma. B: ROC curve of Rad-score. The red point represents the sensitivity and specificity at the cut-off point.

2.5 模型校准度评估

通过建立的模型可计算得到各样本为高级别胶质瘤的概率(预测概率)。为比较预测概率与实际概率是否相符, 进一步行预测模型的校准曲线分析^[18]。图

3 显示该模型的校准曲线, 虚线所示的校准线与理想模型的对角线稍有偏差, 但接近吻合, Hosmer-Lemeshow 检验结果显示差异无统计学意义($P = 0.808$)。

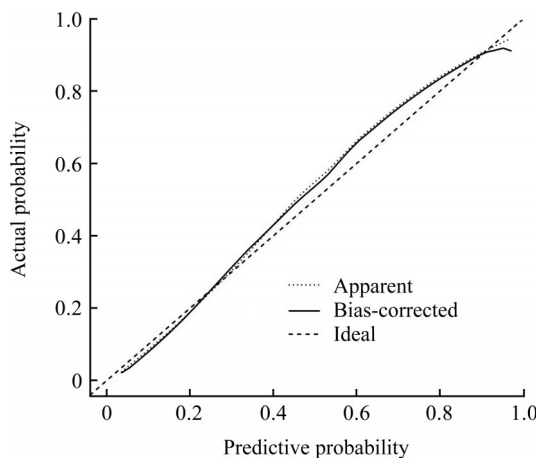


图3 预测模型的校准曲线

Figure 3 Calibration curve for the predictive model

The P value of Hosmer-Lemeshow test for calibration is 0.808, showing no significant difference. This result indicates good agreement between the predictive probability and the actual probability.

3 讨论

目前, 影像组学已成为影像学领域的研究热点之一。相比于医学影像的传统视觉阅片方式, 影像组学可通过机器学习的方法深入挖掘医学影像的生物学本质, 提取影像图像潜在的定量特征。影像组学是在大数据背景下医学影像、计算机、人工智能等多学科交叉的产物, 可为精准医疗、个体化诊疗的实施提供支持。近年来, 已尝试将影像组学应用于肿瘤分级、鉴别诊断、预后、疗效检测甚至分子水平诊断^[9, 19-21]。本研究采用基于影像组学的 logistic 回归模型作为分类器, 可以有效地对胶质瘤病理分级进行预测, 并发掘出基于 5 个影像组学指标的 Rad-score。Rad-score 可作为影像学上的标志物, 基于影像组学的 logistic 回归模型预测胶质瘤分级有望成为术前预测胶质瘤分级的辅助方法。

本研究通过 Rad-score 建立 logistic 回归模型并将其应用于胶质瘤分级的预测, 结果表明该模型可以准确地对胶质瘤的高、低级别进行二分类判别。关于影像组学研究采用哪些序列图像建模更有效, 目前尚未达成一致意见。Tian 等^[5]采集 153 例胶质瘤患者(低级别胶质瘤 42 例, 高级别胶质瘤 111 例)的术前多模态 MRI(包括 T_1 WI、 T_2 WI、 T_1 WI+C、多 b 值弥散加权成像、三维动脉自旋标记灌注加权像)进行影像组学分析, 预测胶质瘤病理分级, 结果表明: 单序列中 T_1 WI+C 序列纹理特征效能最高, 全部模态组合的预测效能高于单序列, 结构模态组合(T_1 WI、

T_2 WI、 T_1 WI+C)、扩散模态(多 b 值弥散加权成像)的预测效能均高于灌注模态(三位动脉自旋标记灌注加权像), 但是该研究存在样本量不均衡的局限性。邓慧媛等^[8]利用 logistic 多因素回归分析发现: 从弥散加权成像(diffusion-weighted imaging, DWI)序列提取的特征 Percentile-75 对高、低级别胶质瘤的诊断效能最高(AUC=0.7930), 而该研究在 T_2 WI 序列及 T_1 WI+C 序列中未筛选出与肿瘤分级相关的特征。本研究采用 T_1 WI+C 序列中影像组学信息, 通过 LASSO 二元 logistic 回归法筛选与胶质瘤病理分级最相关的特征, 对高、低级别胶质瘤进行预测, 取得较满意的效果(AUC=0.919), 说明常规 MRI 图像中也包含有意义的影像组学特征。

ROI 分割包括手动分割、半自动分割、自动化分割 3 种方式。尽管有研究^[22]提出用软件自动分割 ROI 的算法可减轻人力负担及避免主观因素的影响; 但是在病理状态下, 胶质瘤的影像学表现差异较大, 由高水平的影像医师进行手动分割仍是目前研究的主流^[5]。本研究采用 2 名影像医师各自独立勾画病灶的 ROI, 再由另一名高年资上级医师审核, 这种肿瘤 ROI 分割方法耗时相对较长, 但是结果更准确, 也便于理解。

与既往研究仅关注一个或一类纹理特征分析不同, 本研究基于胶质瘤患者术前 T_1 WI+C 影像图像, 通过手工分割病灶, 提取 146 例患者的 41 个影像特征, 通过 LASSO 二元 logistic 回归法筛选特征, 旨在构建更简洁且更具诊断价值的综合影像特征。LASSO 二元 logistic 回归法是目前在数据挖掘与建模分析中应用较广的一种变量降维方法^[23-24]。本研究采用 LASSO 二元 logistic 回归法筛选出 5 个特征, 这 5 个特征与我们之前研究^[10]中人工神经网络的最终输入特征一致。在之前的人工神经网络研究中, 我们用两样本 t 检验筛选出 7 个特征, 将这 7 个特征逐一加入模型, 发现仅使用前 5 个特征便能使模型的准确率达到平台水平。显然本研究用 LASSO 回归筛选变量效率更高, 且能计算出每个变量的系数(即相对权重), 从而进一步构建出一个综合的 Rad-score, 定量计算样本为高级别胶质瘤的概率。

本研究采用 ROC 曲线的 AUC 及校准曲线的 P 值共同评估模型对高、低级别胶质瘤的分类效能。AUC 越接近 1, 表明分类效能越好^[25-26]。本研究构建的模型区分高、低级别胶质瘤的 AUC 为 0.919, 表明模型对高、低级别胶质瘤具有很好的区分能力; 且校准曲线与理想曲线吻合度较好, 偏差很小($P=0.808$), 说明模型预测的概率与实际情况十分接近。

虽然本研究构建模型的预测效能稍低于人工神

神经网络模型(AUC=0.9486)^[10], 但是本研究的模型具有以下优势: 1)模型基于肿瘤的3D ROI数据更能反映肿瘤的异质性。针对本研究中的数据, 3D方式建模的logistic回归模型的AUC明显高于3D方式建模(之前的研究^[10]为2D方式建模)的神经网络模型(0.919 vs 0.6873, $P<0.001$)。2)Logistic回归模型可以精确地量化预测概率, 且结果显示预测概率与实际概率一致性高, 这使得logistic回归模型的临床可解释性更强, 这一点是神经网络模型无法做到的。3)对于一个模型的评价, 光有区分度是不够的, 还需要考察其准确度(也就是校准度)。Logistic回归模型很容易做到这一点, 但是神经网络模型的结果并非概率值, 也无法进行校准曲线分析。因此, 我们认为logistic回归模型优于人工神经网络模型。

本研究的局限性在于: 1)样本量相对不足, 因此无法将病例再分为训练集和验证集来进行外部验证。根据目前较为公认的样本量估计方法^[27], 二分类结局中数量较少组的样本量至少应为进入模型中变量数的10倍, 达到15倍以上较为适宜。尽管最终我们的建模变量仅为1个合成后的变量(Rad-score), 但我们在LASSO二分类逻辑回归中筛选到5个变量, 因此样本数至少应为100例以上, 达到150例可认为LASSO回归结果相对稳定。因此我们现阶段的样本量刚好仅适用于建模, 而无法再分出验证集。2)仅依靠影像组学指标建模, 未将临床资料纳入建模分析。我们将会深化目前的工作, 在后续研究中进一步完善上述不足。

利益冲突声明: 作者声称无任何利益冲突。

参考文献

- [1] Ostrom QT, Bauchet L, Davis FG, et al. The epidemiology of glioma in adults: A "state of the science" review[J]. *Neuro-oncology*, 2014, 16(7): 896-913.
- [2] Louis DN, Perry A, Reifenberger G, et al. The 2016 world health organization classification of tumors of the central nervous system: A summary[J]. *Acta Neuropathol*, 2016, 131(6): 803-820.
- [3] Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system[J]. *Acta Neuropathol*, 2007, 114(2): 97-109.
- [4] Wu W, Lamborn KR, Buckner JC, et al. Joint NCCTG and NABTC prognostic factors analysis for high-grade recurrent glioma[J]. *Neuro-oncology*, 2010, 12(2): 164-172.
- [5] Tian Q, Yan LF, Zhang X, et al. Radiomics strategy for glioma grading using texture features from multiparametric mri[J]. *J Magn Reson Imaging*, 2018, 48(6): 1518-1528.
- [6] Cho HH, Lee SH, Kim J, et al. Classification of the glioma grading using radiomics analysis[J]. *Peer J*, 2018, 6: e5982.
- [7] 杨志炜, 郭天, 谢海滨, 等. 影像组学在脑胶质瘤分级中的应用研究[J]. *磁共振成像*, 2018, 9(6): 439-445. YANG Zhiwei, GUO Tian, XIE Haibin, et al. Application of radiomics in the grading of brain tumor[J]. *Chinese Journal of Magnetic Resonance Imaging*, 2018, 9(6): 439-445.
- [8] 邓慧媛, 杨军, 柯腾飞, 等. 颅内胶质瘤MRI影像组学评估颅内高级别与低级别胶质瘤[J]. *放射学实践*, 2017, 32(12): 1229-1233. DENG Huiyuan, YANG Jun, KE Tengfei, et al. Diagnosis efficacy of MRI radiomics in grading of cerebral glioma before operation[J]. *Radiologic Practice*, 2017, 32(12): 1229-1233.
- [9] 李惠明, 张军. 基于影像组学的人工智能在脑胶质瘤MRI诊断中的应用[J]. *国际医学放射学杂志*, 2019, 42(5): 531-534, 546. LI Huiming, ZHANG Jun. Application of artificial intelligence based on radiomics in glioma MR imaging[J]. *International Journal of Medical Radiology*, 2019, 42(5): 531-534, 546.
- [10] 毛弈韬, 廖伟华, 曹冬, 等. 基于影像信息的人工神经网络的脑胶质瘤分级[J]. *中南大学学报(医学版)*, 2018, 43(12): 1315-1322. MAO Yitao, LIAO Weihua, CAO Dong, et al. An artificial neural network model for glioma grading using image information[J]. *Journal of Central South University. Medical Science*, 2018, 43(12): 1315-1322.
- [11] Raja R, Sinha N, Saini J, et al. Assessment of tissue heterogeneity using diffusion tensor and diffusion kurtosis imaging for grading gliomas[J]. *Neuroradiology*, 2016, 58(12): 1217-1231.
- [12] Teschendorff AE. Avoiding common pitfalls in machine learning omic data science[J]. *Nat Mater*, 2019, 18(5): 422-427.
- [13] Tibshirani RJ. Regression shrinkage and selection via the LASSO[J]. *J Roy Stat Soc B*, 1996, 73(1): 273-282.
- [14] Li C, Chen J, Qin G. Partial Youden index and its inferences[J]. *J Biopharm Stat*, 2019, 29(2): 385-399.
- [15] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research[J]. *J Chiropr Med*, 2016, 15(2): 155-163.
- [16] 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.
- [17] Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The hosmer-lemeshow test revisited[J]. *Crit Care Med*, 2007, 35(9): 2052-2056.
- [18] Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: Users' guides to the medical literature[J]. *JAMA*, 2017, 318(14): 1377-1384.
- [19] 冷月爽, 王小宜, 廖伟华, 等. 影像组学在胶质瘤临床研究中的应用进展[J]. *中南大学学报(医学版)*, 2018, 43(4): 354-359. LENG Yueshuang, WANG Xiaoyi, LIAO Weihua, et al. Radiomics in gliomas: A promising assistance for glioma clinical research[J]. *Journal of Central South University*.

- Medical Sciences, 2018, 43(4): 354-359.
- [20] 王佳, 胡粟, 胡春洪. 脑胶质瘤 MRI 影像基因组学研究进展[J]. 磁共振成像, 2018, 9(10): 725-730.
WANG Jia, HU Su, HU Chunhong. An update of radiogenomics based on MRI features in glioma[J]. Chinese Journal of Magnetic Resonance Imaging, 2018, 9(10): 725-730.
- [21] Liu ZY, Wang S, Dong D, et al. The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges[J]. Theranostics, 2019, 9(5): 1303-1322.
- [22] Shi Z, Xu B, Zhao M, et al. A joint ROI extraction filter for computer aided lung nodule detection[J]. Biomed Mater Eng, 2015, 26(Suppl 1): S1491-S1499.
- [23] 王进, 李智慧, 沈浮, 等. 基于高分辨 T₂WI 的影像组学对直肠癌术前分期的应用价值[J]. 放射学实践, 2019, 34(11): 1251-1254.
WANG Jin, LI Zhihui, SHEN Fu, et al. The value of high resolution T₂WI-based radiomics in the preoperative staging of rectal cancer[J]. Radiologic Practice, 2019, 34(11): 1251-1254.
- [24] 崔达华, 赵莹, 刘爱连, 等. 基于动态增强 MRI 的影像组学列线图预测肝细胞癌切除术后 3 年复发的价值[J]. 中国临床医学影像杂志, 2019, 30(12): 863-868.
CUI Dahua, ZHAO Ying, LIU Ailian, et al. The value of radiomics nomogram based on dynamic enhanced MRI for predicting the recurrence of HCC in three years after hepatectomy[J]. Journal of China Clinic Medical Imaging, 2019, 30(12): 863-868.
- [25] Eng J. Receiver operating characteristic analysis: A primer[J]. Acad Radiol, 2005, 12(7): 909-916.
- [26] Ma H, Bandos AI, Gur D. On the use of partial area under the roc curve for comparison of two diagnostic tests[J]. Biom J, 2015, 57(2): 304-320.
- [27] Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model[J]. BMJ, 2020, 368: m441.

(本文编辑 郭征)

本文引用: 孙贤婷, 廖伟华, 曹冬, 赵跃龙, 周高峰, 王冬翠, 毛弈韬. 基于影像组学的 logistic 回归模型预测胶质瘤分级[J]. 中南大学学报(医学版), 2021, 46(4): 385-392. DOI: 10.11817/j.issn.1672-7347.2021.200074

Cite this article as: SUN Xianting, LIAO Weihua, CAO Dong, ZHAO Yuelong, ZHOU Gaofeng, WANG Dongcui, MAO Yitao. A logistic regression model for prediction of glioma grading based on radiomics[J]. Journal of Central South University. Medical Science, 2021, 46(4): 385-392. DOI:10.11817/j.issn.1672-7347.2021.200074