

• 中枢神经影像学 •

基于 MRI 影像组学特征构建脑胶质瘤 IDH1 突变预测模型

江少凡, 宋阳, 薛蕴菁, 蒋日烽

【摘要】 目的:探讨基于不同 MRI 序列的影像组学特征构建的机器学习(ML)模型预测胶质瘤 IDH1 突变的价值。**方法:**回顾性搜集经手术病理证实的 161 例胶质瘤患者(70 例 IDH1 突变型/91 例野生型)的临床资料,主要包括年龄、性别、Karnofsky 功能状态(KPS)评分和肿瘤的病理分级。所有患者术前行 MRI 检查获得 T_2 WI、 T_2 -FLAIR、ADC 图及对比增强 T_1 WI 图像,术后病理标本均行 IDH1 检测。将 161 例患者按照 7:3 的比例随机分配为训练集和测试集。由 2 位影像医师利用 Image J 软件共同对病灶在配准过的 T_2 -FLAIR 或对比增强 T_1 WI 序列上进行逐层 ROI 的勾画,最后形成感兴趣区容积(VOI),然后使用 FAE 软件在各序列图像上提取 VOI 的影像组学特征,基于训练集的数据,通过均值归一化、方差分析的特征选择方法、皮尔逊相关系数的特征降维方法、4 种 ML 分类器(线性判别分析、LASSO 回归、逻辑回归、支持向量机)以及十折交叉验证法构建 15 种 ML 模型,并采用 ROC 曲线和 Delong 检验在测试集中筛选出最佳序列或序列组合;再基于最佳序列或序列组合,使用均值归一化、2 种特征选择方式(方差分析、特征权重算法)、2 种降维方式(皮尔逊相关系数和主成分分析法)、4 种 ML 分类器构建了 16 种 ML 模型,通过 FAE 软件的 one-standard error 法、Delong 检验和 ROC 曲线筛选拟合度较好且 AUC 最高的 ML 模型,并将此模型联合临床参数构建联合模型,评价各联合模型的诊断效能。**结果:**基于 ADC 图+对比增强 T_1 WI 序列组合提取的组学特征构建的 4 种 ML 模型在测试集中的 AUC 分别为 0.888、0.872、0.896 和 0.877,均高于其它序列或序列组合构建的 ML 模型;利用此序列组合构建的 16 种 ML 模型中,以方差分析特征选择法、主成分分析方法的降维方式及分类器为 LASSO 回归时所构建的 ML 模型具有较好的拟合度且测试集中的 AUC 最高,为 0.829(95%CI: 0.658~0.966),此模型结合 KPS 评分和肿瘤病理分级所构建的联合模型在测试集中的 AUC 为 0.874(95%CI: 0.722~0.985)。**结论:**基于 ADC 图+对比增强 T_1 WI 序列组合提取的组学特征构建的 ML 模型对胶质瘤 IDH1 突变有具有较好的预测效能,联合 KPS 评分和病理分级可提高预测效能。

【关键词】 胶质瘤; 异柠檬酸脱氢酶突变; 影像组学; 磁共振成像; 预测模型

【中图分类号】 R445.2; R739.41 **【文献标识码】** A **【文章编号】** 1000-0313(2023)12-1493-07

DOI:10.13609/j.cnki.1000-0313.2023.12.001

开放科学(资源服务)标识码(OSID):



Prediction model construction based on MRI radiomics features for evaluating IDH1 mutation of glioma

JIANG Shao-fan, SONG Yang, XUE Yun-jing, et al. Departments of Radiology, Fujian Medical University Union Hospital, Fuzhou 350003, China

【Abstract】 Objective: To investigate the value of different machine learning models based on radiomics features of different MRI sequences in predicting IDH1 mutations in glioma. **Methods:** The age, sex, Karnofsky Performance Status (KPS) score and pathological grading data of 161 patients with glioma (70 IDH1 mutant and 91 wild-type) confirmed by operation and pathology were retrospectively collected. All patients underwent preoperative MRI to obtain T_2 WI, T_2 -FLAIR, ADC and contrast-enhanced T_1 WI images, and postoperative pathological specimens were detected by IDH1. 161 patients were randomly assigned to a training set and a test set in a 7:3 ratio. Two radiologists delineated the each lesion to obtain 2D-ROIs with Image J software on each slice of the lesion on registered T_2 -FLAIR or enhanced T_1 WI sequences and volume of interest (VOI) was formed. Then the imaging

作者单位: 350003 福建福州, 福建医科大学附属协和医院放射科(江少凡、薛蕴菁、蒋日烽); 200241 上海, 华东师范大学上海磁共振重点实验室(宋阳)

作者简介: 江少凡(1987—), 男, 福建福州人, 硕士研究生, 主治医师, 主要从事磁共振功能成像及影像组学研究工作。

通讯作者: 蒋日烽, E-mail: 26630706@qq.com

基金项目: 福建医科大学启航基金项目(2019QH1034)

radiomics features of VOI are extracted from each sequence of images using FAE software. Based on the data from the training set, Fifteen ML models were constructed by means normalization, feature selection method of ANOVA, feature dimensionality reduction method of Pearson correlation coefficient, four ML classifiers (linear discriminant analysis, LASSO regression, logistic regression, support vector machine) and ten-folds cross-validation method, and then the best sequence or sequences combination was selected out in test set by ROC curve analysis and Delong test. Based on the best sequence or sequence combination, 16 ML models were constructed through mean normalization, two feature selection method of ANOVA and feature weight algorithm, two feature dimensionality reduction Pearson correlation coefficient/principal component analysis method, and 4 ML classifiers. The ML models with good fit and highest AUC value were selected out by one standard error method, Delong test and ROC curve. Finally, clinical parameters were combined to evaluate the diagnostic efficacy of each combined model. **Results:** The AUC values of the four ML models constructed based on the radiomics features extracted from the ADC and contrast enhanced T₁WI in the test set were 0.888, 0.872, 0.896 and 0.877, respectively, which were higher than those of other ML models or sequence combinations. Among the 16 ML models constructed by this sequence combination, the ML model constructed by ANOVA feature selection method, principal component analysis method dimension reduction method and LASSO regression classifier showed a good fit and the highest AUC of 0.829 (95%CI: 0.658~0.966) in the test set, and the AUC value of the combined model built by the combination of KPS score and tumor pathological grade in the test set was 0.874 (95%CI: 0.722~0.985). **Conclusion:** The ML model constructed based on the combination of ADC map and contrast enhanced T₁WI sequence extraction has a good predictive efficacy for glioma IDH1 mutation, and the combined KPS score and pathological grade can improve the predictive efficacy.

【Key words】 Glioma; Isocitrate dehydrogenase mutation; Radiomics; Magnetic resonance imaging; Prediction model

胶质瘤是颅内最常见的恶性肿瘤,即使通过积极的手术、放化疗和靶向药物等治疗,预后仍然很差。近几年来有研究者逐渐开展了胶质瘤分子、基因层面的研究,所研究的基因主要有 IDH、EGFR、MGMT 及 1p/19q 等^[1-2]。IDH 基因中以 IDH1 突变最常见,其作用是降低肿瘤细胞增殖程度和减少血管生成,使胶质瘤侵袭性相对减低,故有 IDH1 突变的胶质瘤患者具有相对良好的预后^[3]。针对 IDH1 突变型胶质瘤的靶向药物治疗手段已逐步开展。因此,对于胶质瘤尤其是无法切除的胶质瘤患者而言,无创性预测 IDH1 突变有助于指导临床尽早制订合理准确的治疗策略。

影像组学技术可基于常规 MRI 图像提取的肿瘤病灶的高通量定量特征,在脑肿瘤的诊断方面具有高度特异性和无创性的综合优势。既往的研究结果表明单一序列对脑肿瘤的诊断效能有限,多个序列组合能获得更好的诊断效能^[4-5]。但既往对多种序列组合和多种机器学习(machine learning, ML)模型进行比较的相关研究报道较少或不全面。本研究中通过多种序列及组合、特征选择方式、降维方式及 ML 分类器的组合使用来构建多种 ML 影像组学模型,评价各模型对 IDH1 突变胶质瘤的预测效能。

材料与方法

1. 一般资料

回顾性分析 2014 年 1 月—2021 年 12 月在本院于术前接受 3.0T MR 检查并经病理证实的脑胶质瘤患者的病例资料。纳入标准:①病理诊断为脑胶质瘤;②首次接受胶质瘤手术治疗且术前未接受任何相关治疗;③术前行 3.0T MR 检查,检查序列需包括标准化采集的 T₂WI、T₂-FLAIR、DWI 及对比增强 T₁WI。排除标准:①图像质量较差,图像信噪比较低和/或有明显伪影等;②病灶过小(肿瘤直径<1.0 cm);③免疫组化检查未检测 IDH1。根据纳入和排除标准,最终纳入 161 例脑胶质瘤患者:低级别(WHO I~II 级)66 例,高级别(WHO III~IV 级)95 例;IDH1 突变型 70 例,IDH1 野生型 91 例。IDH1 突变型病例中,男 41 例,女 29 例;年龄 23~68 岁,平均(44.05±12.15)岁。IDH1 野生型病例中,男 54 例,女 37 例;年龄 3~77 岁,平均(51.36±16.77)岁。按照 7:3 的比例将 161 例患者随机分为训练集和测试集。训练集 113 例,IDH1 突变型与野生型例数比为 50/63;测试集共 48 例,IDH1 突变型与野生型例数比为 20/28。

本研究经福建医科大学附属协和医院伦理委员会批准。为保护患者隐私,隐去患者姓名等识别信息,所有入组患者的影像资料以数字编号表示。

2. MRI 检查方法

使用 GE Discovery 750 3.0T、Siemens Trio Tim 3.0T 或 Siemens Prisma 3.0T 磁共振仪,收集的横轴面序列有 T_2 WI、 T_2 -FLAIR、DWI 及对比增强 T_1 WI,各序列标准化采集参数如下。 T_2 WI:TR 4000 ms,TE 90 ms,体素分辨率 $0.5\text{ mm} \times 0.5\text{ mm} \times 5.0\text{ mm}$; T_2 -FLAIR:TR 9000 ms,TE 145 ms,体素分辨率 $0.6\text{ mm} \times 0.6\text{ mm} \times 5.0\text{ mm}$;DWI 序列:TR 4000 ms,TE 60 ms,体素分辨率 $1.4\text{ mm} \times 1.4\text{ mm} \times 5.0\text{ mm}$,b 值选用 0 和 1000 s/mm^2 ,自动生成 ADC 图(本研究提取的是 ADC 图的组学特征);对比增强 T_1 WI:TR 1750 ms,TE 9.0 ms,体素分辨率 $0.6\text{ mm} \times 0.6\text{ mm} \times 5.0\text{ mm}$ 。将患者 4 个序列图像的 DICOM 格式转化为 Nifti 格式后,采用 SPM12 软件将所有序列的图像向 T_2 WI 图像进行配准。

3. ROI 勾画及特征提取

由 2 位放射科医师(分别具有 5 和 10 年工作经验)利用 ImageJ 软件(<https://imagej.nih.gov/ij>)共同进行 ROI 的勾画。勾画方法:对每例患者的前期已配准好的所有序列图像进行观察,对强化不明显或无强化的脑胶质瘤选择 T_2 -FLAIR 序列进行 ROI 的勾画(图 1);对于强化明显的脑胶质瘤则在对比增强 T_1 WI 图像上进行 ROI 的勾画(图 2);通过观察 T_2 WI 和 T_2 -FLAIR 图像,在肿瘤实体区域逐层勾画 ROI,注意避开肿瘤内的液化坏死区。因所有图像先期均进行了配准,故在任意一个序列上勾画的 ROI 均可同步到其它序列。最后生成感兴趣体积(volume of interest, VOI)。利用 FAE Pro V0.4.1 软件(<https://github.com/salan668/FAE>)中的 pyradiomics 模块提取各序列 VOI 的影像组学特征。

4. 最佳序列的筛选和模型的构建

使用 FAE 软件,将训练集和测试集中各序列所提取的组学特征按数据平衡化、数据归一化、特征选择、特征降维和分类建模的顺序进行。

首先,在测试集中筛选最佳的序列或组合,具体步骤如下。①数据平衡化:上采样;②数据归一化:均值归一化;③特征选择:采用方差分析法(analysis of variance, ANOVA)进行特征选择,特征数的范围设置为 1~20;④特征降维:对每两个特征使用皮尔逊相关系数(Pearson correlation coefficient, PCC)法进行分析,如果其 r 值大于 0.9,则在 FAE 软件中随机删除其中一个特征;⑤分类建模:分别采用四种算法即线性判别分析(linear discriminant analysis, LDA)、最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)回归(LASSO regression, LR)、逻辑回归(logistic regression, LG)和支持向量机(support vector machine, SVM)建立机器学习模型。在训练集中采用 10 折交叉验证对模型进行训练。序列或组合总计 15 种,通过比较其 ROC 曲线的 AUC 筛选出最佳序列或组合。以上所有流程都是在 Python 3.7.6 上使用 FAE Pro V0.4.1 软件来完成的^[6]。

然后,基于最佳序列组合,将其提取的组学特征采用多种算法按照上述步骤重新进行建模,主要步骤和方法如下。①数据平衡化:上采样;②数据归一化:均值归一化;③采用 2 种方式进行特征选择,分别是方差分析和特征权重算法(Relief),特征数的范围设置为 1~20;④选用 2 种方式进行特征降维,分别是 PCC 和主成分分析(principal component analysis, PCA);⑤然后,采用 LDA、LR、LG 和 SVM 四种算法分别建立机器学习模型。通过上述步骤和算法的组合共获得 16 种机器学习模型。

训练集通过 10 折交叉验证对模型进行训练。在

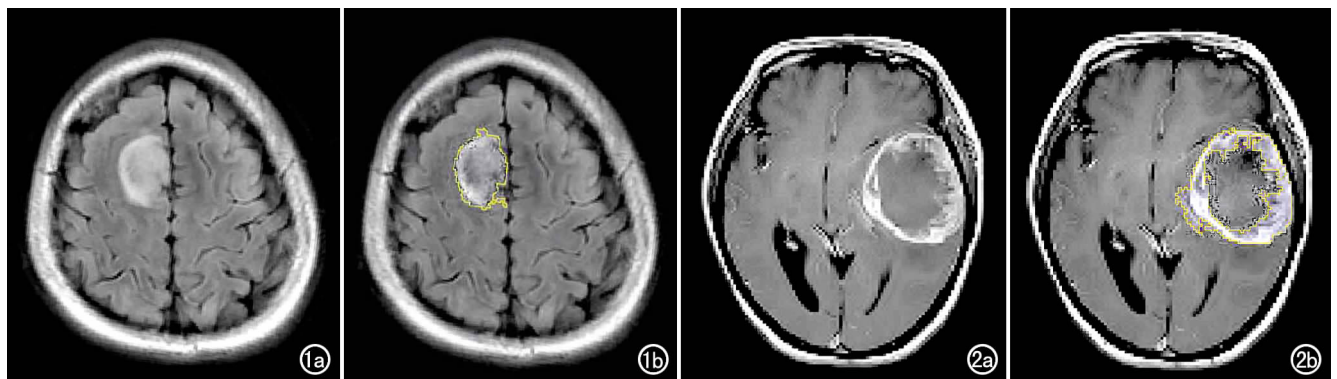


图 1 右额叶脑胶质瘤 ROI 勾画示意图。a) T_2 -FLAIR 序列显示肿瘤呈不均匀高信号;b) T_2 -FLAIR 序列,沿肿瘤边缘勾画 ROI。图 2 左侧额叶胶质瘤 ROI 勾画示意图。a) 对比增强 T_1 WI 序列,显示肿瘤周边呈明显强化,中心区域无明显强化;b) 沿肿瘤强化明显的周边实质区域勾画 ROI。

测试集中对 16 种机器学习模型进行验证,通过 FAE 软件的 one-standard error 法、ROC 曲线和 Delong 检验筛选出具有较好拟合度和 AUC 最大的机器学习模型,记录其相应的诊断敏感度、特异度和符合率等效能指标值,然后,将此最佳机器学习模型联合临床指标(年龄、性别、病理分级及 KPS 评分)构建联合模型,分析联合模型的诊断效能。

所有 AUC 值的 95%CI 采用 bootstrape 法进行 1000 次重采样得到。

5. 统计学方法

临床资料的组间比较使用 SPSS 23.0 统计软件。计数资料用频数表示,符合正态分布的计量资料采用均数±标准差表示。采用独立样本 *t* 检验比较训练集和测试集之间患者年龄和 KPS 评分的差异,采用卡方检验比较训练集和测试集之间患者性别构成、肿瘤病理级别和 IDH1 突变率的差异。采用 Delong 检验比较不同机器学习模型 AUC 的差异。以 *P*<0.05 为差异有统计学意义。

结 果

1. 临床资料

训练集和测试集中患者年龄、性别、KPS 评分、肿瘤分级和 IDH1 突变率的差异均无统计学意义(*P*>0.05),详见表 1。

表 1 训练集和测试集临床资料的比较

指标	训练集 (n=113)	测试集 (n=48)	χ^2/t 值	<i>P</i> 值
年龄/岁	48.00±15.15	47.33±18.23	-0.215	0.830
KPS 评分	75.73±11.27	79.74±9.86	1.923	0.057
性别比(男/女)	66/47	29/19	0.056	0.813
病理级别/例			2.684	0.101
低	51	15		
高	62	33		
IDH1/例			0.091	0.762
突变型	50	20		
野生型	63	28		

2. 特征提取结果

T₂WI、T₂-FLAIR、ADC 图及对比增强 T₁WI 四个序列均各自提取了 98 个影像组学特征,包括灰度共生矩阵(gray level co-occurence matrix, GLCM)特征 24 个、灰度行程长度矩阵(gray level run-length matrix, GLRLM)特征 16 个、灰度区域大小矩阵(gray level size zone matrix, GLSZM)特征 16 个、邻域灰度差矩阵(neighbouring gray tone difference matrix, NGTDM)特征 5 个、一阶特征 18 个和形态特征 19 个。

3. 筛选最佳序列或组合

基于 15 个序列或序列组合提取的组学特征所构建的预测 IDH1 突变型胶质瘤的四种机器学习模型在训练集和测试集中的 AUC 值详见表 2。基于 ADC 图 and 对比增强 T₁WI 序列提取的组学特征所构建的 4 种机器学习模型在测试集中的 AUC 分别为 0.888、0.872、0.896 和 0.877,均高于其它序列或组合的 AUC 值,故此序列组合为最佳。

4. 基于最佳序列构建 ML 模型及其效能分析

基于 ADC 图+增强 T₁WI 序列提取的组学特征,采用 2 种降维方式,2 种特征选择方式以及 4 种分类器共构建了 16 种 ML 模型,通过 FAE 软件的 one-standard error 法排除过拟合或明显欠拟合数据后剩下 5 种 ML 模型,以其构建的算法组合命名,分别为 PCA_ANOVA_LDA、PCA_ANOVA_LR、PCA_ANOVA_LG、PCA_ANOVA_SVM 和 PCC_Relief_LDA。

5 种 ML 模型在训练集和测试集中预测 IDH1 突变型的效能分析结果详见表 3。5 种模型在测试集中的 AUC 为 0.656、0.766、0.854、0.814 和 0.810,在测试集中的 AUC 分别为 0.596、0.829、0.808、0.821 和 0.818。Delong 检验结果(图 3):在训练集中,PCA_

表 2 基于不同序列及组合的四种机器学习模型预测胶质瘤 IDH1 突变的 AUC 值

序列/组合	训练集				测试集			
	LDA	LR	LG	SVM	LDA	LR	LG	SVM
ADC 图	0.887	0.738	0.738	0.737	0.739	0.705	0.703	0.707
T ₂ WI	0.796	0.780	0.791	0.777	0.662	0.639	0.639	0.624
增强 T ₁ WI	0.710	0.693	0.699	0.690	0.781	0.760	0.771	0.756
T ₂ -FLAIR	0.745	0.710	0.808	0.744	0.684	0.641	0.735	0.663
ADC 图+T ₂ WI	0.792	0.748	0.766	0.756	0.822	0.719	0.787	0.745
ADC 图+T ₂ -FLAIR	0.802	0.794	0.838	0.797	0.636	0.615	0.628	0.668
ADC 图+增强 T ₁ WI	0.728	0.699	0.725	0.710	0.888	0.872	0.896	0.877
T ₂ -FLAIR+T ₂ WI	0.737	0.732	0.733	0.724	0.718	0.735	0.729	0.719
T ₂ -FLAIR+增强	0.791	0.806	0.825	0.800	0.621	0.628	0.625	0.647
T ₂ WI+增强 T ₁ WI	0.843	0.835	0.824	0.847	0.760	0.736	0.742	0.747
增强 T ₁ WI+ADC 图+T ₂ WI	0.812	0.860	0.806	0.856	0.663	0.567	0.583	0.588
增强 T ₁ WI+T ₂ -FLAIR+T ₂ WI	0.824	0.817	0.787	0.819	0.784	0.778	0.771	0.766
T ₂ WI+T ₂ -FLAIR+ADC 图	0.833	0.795	0.832	0.805	0.727	0.615	0.693	0.666
增强 T ₁ WI+ADC 图+T ₂ -FLAIR	0.736	0.783	0.737	0.736	0.789	0.786	0.800	0.824
4 个序列联合	0.865	0.804	0.791	0.781	0.623	0.591	0.620	0.604

ANOVA_LDA 的 AUC 低于另外 4 种 ML 模型 ($P=0.047, 0.006, 0.041, 0.040$), 另外 4 种 ML 模型之间 AUC 的差异均无统计学意义 ($P>0.05$); 在测试集中, PCA_ANOVA_LDA 的 AUC 低于另外 4 种 ML 模型 ($P=0.011, 0.025, 0.013, 0.017$), 另外 4 种 ML 模型之间 AUC 的差异无统计学意义 ($P>0.05$), 其中以 PCA_ANOVA_LR 的 AUC 最大, 为 0.829(95%CI:0.658~0.966), 特异度为 0.864, 敏感度为 0.765, 符合率为 0.821, 构建该 ML 模型的组学特征及权重系数见表 4。

表 4 PCA_ANOVA_LR 模型中的组学特征及其权重系数

组学特征	贡献权重
ADC-glszm_LargeAreaHighGrayLevelEmphasis	0.787
ADC-glcmm_MCC	3.544
ADC-glcmm_Correlation	-1.428
T1C-glrlm_LongRunEmphasis	3.437
T1C-glszm_LargeAreaEmphasis	1.461
T1C-glszm_LargeAreaHighGrayLevelEmphasis	0.787
T1C-glszm_LargeAreaLowGrayLevelEmphasis	1.591
T1C-glcmm_ClusterProminence	-15.982
T1C-glcmm_ClusterShade	-1.232
T1C-ngtdm_Strength	-0.897
T1C-ngtdm_Complexity	14.812

注: T1C 代表基于对比增强 T₁WI。

5. 临床资料结合最佳影像组学模型的效能分析

将最佳影像组学模型 PCA_ANOVA_LR 进一步与四项临床指标相结合构建组合模型, 通过 FAE 软件的一标准误差法排除了过拟合和欠拟合的组合模型后, 拟合度较好的各组合模型在训练集和测试集中的效能分析结果详见表 5。在测试集中, 以结合了 KPS 评分、病理分级和影像组学模型的联合模型的 AUC 最大, 为 0.874(95%CI:0.722~0.985)。

表 3 基于 ADC 和增强 T₁WI 序列的组学特征构建的不同机器学习模型的诊断效能

模型序号	降维方式	特征选择	分类器	训练集				测试集			
				AUC(95%CI)	特异度	敏感度	符合率	AUC(95%CI)	特异度	敏感度	符合率
1	PCA	ANOVA	LDA	0.656(0.532~0.773)	0.860	0.410	0.663	0.596(0.427~0.751)	0.546	0.647	0.590
2	PCA	ANOVA	LR	0.766(0.667~0.859)	0.740	0.692	0.719	0.829(0.658~0.966)	0.864	0.765	0.821
3	PCA	ANOVA	LG	0.854(0.765~0.923)	0.840	0.769	0.809	0.808(0.622~0.946)	0.864	0.765	0.821
4	PCA	ANOVA	SVM	0.814(0.724~0.894)	0.920	0.590	0.775	0.821(0.647~0.958)	0.818	0.824	0.821
5	PCC	Relief	LDA	0.810(0.726~0.894)	0.740	0.718	0.730	0.818(0.642~0.968)	0.818	0.882	0.846

表 5 临床资料联合影像组学模型构建的联合模型的诊断效能

联合模型的类型	训练集				测试集			
	AUC(95%CI)	特异度	敏感度	符合率	AUC(95%CI)	特异度	敏感度	符合率
年龄+组学模型	0.819(0.723~0.903)	0.800	0.769	0.787	0.791(0.603~0.940)	0.773	0.824	0.795
病理分级+组学模型	0.815(0.715~0.901)	0.740	0.795	0.764	0.837(0.653~0.976)	0.857	0.824	0.821
性别+组学模型	0.847(0.768~0.920)	0.660	0.872	0.753	0.786(0.589~0.932)	0.863	0.765	0.821
KPS 评分+组学模型	0.868(0.790~0.941)	0.960	0.615	0.809	0.864(0.715~0.981)	0.910	0.766	0.846
KPS 评分+病理分级+组学模型	0.867(0.788~0.939)	0.940	0.667	0.820	0.874(0.722~0.985)	0.954	0.764	0.872

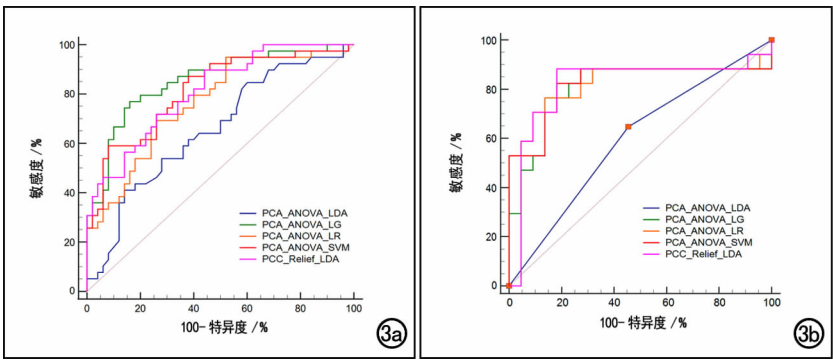


图 3 数据拟合较好的 5 种 ML 模型在训练集和测试集中的 ROC 曲线, 显示 PCA_ANOVA_LDA 模型在训练集和测试集中的 AUC 均低于另外 4 种 ML 模型。a) 训练集; b) 测试集。

讨论

本研究结果显示基于 ADC 图+对比增强 T₁WI 序列组合提取的影像组学特征所构建的 4 种机器学习模型在测试集中预测胶质瘤 IDH1 突变型的 AUC 值高于其它序列及组合, 提示这个序列组合具有较好且稳定的诊断效能。基于该序列组合进一步构建的 16 种 ML 影像组学模型中, PCA_ANOVA_LR、PCA_ANOVA_LG、PCA_ANOVA_SVM 和 PCC_Relief_LDA 这 4 种模型均具有较好的诊断效能, 其中 PCA_ANOVA_LR 在测试集中的 AUC 最大, 联合 KPS 评分、病理分级后的联合模型的 AUC 进一步提高。决定胶质瘤患者预后的最主要因素是细胞增殖程度和血供程度, 两者也是反映胶质瘤侵袭性强弱的指标, 而与之相对应的 MRI 征象是 DWI 上肿瘤的扩散受限程度和对比增强 T₁WI 上肿瘤的强化程度, 胶质瘤 IDH1 的异质性能很好地在 DWI 的 ADC 图和对比增强 T₁WI 序列上得以体现, 这与既往相关研究中序列的选择结果相一致^[7-9]。本研究结果亦显示出基于

多序列组合所构建的影像组学模型较单序列模型更加具有优势,在建模时应重点关注 ADC 图和对比增强 T₁WI 序列所提取的组学特征。

本研究的样本量并不多,未纳入随机森林和决策树这 2 种分类器,原因是这两者都属于同一类型的分类器,比较适合对大样本数据进行评价,在小样本数据的分析中出现错误的概率较高,容易出现数据的过拟合。如郁义星等^[10]基于 132 例肝癌患者的肝胆期 MRI 图像提取的组学特征构建了 6 种机器学习模型来预测肝细胞癌微血管侵犯,结果显示随机森林和决策树模型在训练集中的 AUC 为 1 或接近 1,提示数据可能存在过拟合现象。而 LDA、LR、LG 和 SVM 这 4 种分类器对训练样本量的依赖较小,适合小样本高维数据的分类^[11-13]。Peng 等^[14]对 105 例胶质瘤患者基于其 T₂WI+增强 T₁WI 序列使用 SVM 分类器建立影像组学模型,结果显示模型在测试集中预测胶质瘤 IDH1 类型的 AUC 为 0.799。Chen 等^[15]基于 175 例低级别星形细胞瘤患者的增强 T₁WI 序列提取组学特征,并利用 LDA 分类器建立组学模型,模型在训练集和测试集中的 AUC 分别为 0.801 和 0.799。上述研究中的病例数与本研究相近且均未出现数据过拟合的情况,故本研究选择了对小样本二分类问题有较好训练效果的 LDA、LR、LG 和 SVM 这 4 种分类器进行建模。

本研究中还发现基于 ADC 图+增强 T₁WI 序列提取的组学特征所构建的 16 个机器学习模型中,有 11 个模型出现了较明显的数据欠拟合。数据欠拟合即训练集与测试集之间 AUC 的差异较大,且训练集的 AUC 值低于测试集,出现欠拟合现象的原因可能是数据分布不平衡。笔者发现本研究中无明显数据欠拟合的 5 种机器学习模型中有 4 种是由 PCA_ANOVA 方法构建的,笔者认为可能有以下两方面原因:一、要改善训练集中数据欠拟合的问题即提高模型的诊断效能,主要解决方法有增加特征数、减少正则化参数和使用非线性模型(如 SVM 等),而 PCA 作为一种高维数据的降维方式,可以使得降维造成的损失最小,主要作用就是保留训练集的特征信息,相当于增加了训练集的特征数,而 PCC 的降维方式是如果 2 个特征的相关系数值大于 0.9,则随机删除其中一个特征,因此相较于 PCC,PCA 能保留更多的训练集特征数,从而可提高训练集的诊断效能^[16];二、ANOVA 特征选择方式是通过选择与目标变量相关性最强的指标,可以减少过拟合风险和提高训练集的诊断效能,而 Relief 特征选择方法是将权重小于某个阈值的特征移除,减少了训练集的特征数,从而可导致模型在训练集中的诊断效能降低。同时本研究结果显示在训练集和

测试集中, LDA 分类器所构建的 PCA_ANOVA_LDA 与 PCC_Relief_LDA 两个模型之间 AUC 均存在差异,笔者认为 LDA 作为一种线性分类器,其稳定性可能不及 LR、LG 和 SVM。

患者的性别、年龄、肿瘤的病理分级及 KPS 评分均是评价胶质瘤的主要临床指标,既往研究中也有将临床资料与影像组学模型结合的报道^[4,17]。陈洋等^[18]通过 Log-rank 检验以及多因素逻辑回归分析,发现 KPS 评分、病理级别及 IDH 基因型均与胶质瘤预后密切相关,本研究结果亦显示影像组学模型结合 KPS 评分和病理分级后得到的联合模型的预测效能进一步提高。

此外,与以往的研究不同,本研究将高、低级别胶质瘤一起纳入研究,使组学模型不受病理分级的影响,更加适合在术前预测胶质瘤 IDH1 突变,在近期的研究中有类似报道^[17]。2021 年发布的第 5 版中枢神经系统脑肿瘤分类标准进一步弱化了肿瘤分级对胶质瘤分型的作用,而 IDH 基因则是首个得到公认的与胶质瘤分型相关的关键基因。

本研究具有一定的局限性:第一,为单中心研究,没有外部数据的验证,这是由于外部机构的 MRI 扫描参数无法做到与本研究完全一致;第二,由于本研究为回顾性分析,且入组条件较为严格,尽管纳入的样本是目前本中心能收集到的全部病例,但样本量还是相对较小,今后还将继续搜集相关病例以扩大样本量来进一步验证本研究的结果;第三,本研究中仅选用了常规序列进行研究和分析,没有将采用了新技术(如磁共振三维动静脉自旋标记、扩散峰度成像等)的相关数据纳入研究,原因是目前相应样本量还较少。

综上所述,基于 ADC+T₁WI 增强序列组合提取的影像组学特征,采用方差分析的特征选择方法、主成分分析的降维方法以及 LASSO 回归的分类方式所构建的机器学习模型对胶质瘤 IDH1 突变具有较好的预测效能,结合临床 KPS 评分和肿瘤病理分级所构建的联合模型可进一步提高预测效能。

参考文献:

- [1] Chaddad A, Sabri S, Niazi T. Prediction of survival with multi-scale radiomic analysis in glioblastoma patients[J]. Med Biol Eng Comput, 2018, 56(12): 2287-2300.
- [2] Bakas S, Akbari H, Sotiras A, et al. Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features[J/OL]. Sci Data, 2017, 5(4): e170117[2017 Sep 5]. DOI:10.1038/sdata.2017.117.
- [3] Suh CH, Kim HS, Jung SC, et al. Perfusion MRI as a diagnostic biomarker for differentiating glioma from brain metastasis: a systematic review and meta-analysis[J]. Eur Radiol, 2018, 28(9): 3819-3831.

- [4] Tan Y, Zhang ST, Wei JW, et al. A radiomics nomogram may improve the prediction of IDH genotype for astrocytoma before surgery[J]. Eur Radiol, 2019, 29(7):3325-3337.
- [5] Kim M, Jung SY, Park JE, et al. Diffusion- and perfusion-weighted MRI radiomics model may predict isocitrate dehydrogenase (IDH) mutation and tumor aggressiveness in diffuse lower grade glioma[J]. Eur Radiol, 2020, 30(4):2142-2151.
- [6] Song Y, Zhang J, Zhang YD, et al. Feature Explorer (FAE): a tool for developing and comparing radiomics models[J/OL]. PLoS One, 2020, 15(8):e0237587[2020 Aug 17]. DOI:10.1371/journal.pone.0237587.
- [7] Yan J, Zhang B, Zhang S, et al. Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients[J/OL]. NPJ Precis Oncol, 2021, 5(1):e72[2021 Jul 26]. DOI:10.1038/s41698-021-00205-z.
- [8] 廖天双, 陈东, 李操, 等. 基于多序列 MRI 影像组学模型预测脑膜瘤病理分级的价值[J]. 放射学实践, 2021, 36(12):1462-1466.
- [9] 卢俊, 李祥, 黎海亮. 基于 ADC 和增强 MRI 的影像组学模型预测低级别胶质瘤 TERT 启动子突变状态[J]. 放射学实践, 2022, 37(5):538-542.
- [10] 郁义星, 王希明, 胡春洪, 等. 钆塞酸二钠增强 MRI 影像组学和机器学习术前预测肝细胞癌微血管侵犯的价值[J]. 中华放射学杂志, 2021, 55(8):853-858.
- [11] 张会杰, 李恒芬, 李湘露. 支持向量机在临床医学中的应用研究进展[J]. 神经疾病与精神卫生, 2017, 17(11):812-815.
- [12] 韩蓓丽. SVM 和 RLR 算法的对比分析[J]. 应用数学进展, 2021, 10(12):4292-4303.
- [13] 曾维佳, 张日权. Lasso 变量选择的分布式算法[J]. 应用概率统计, 2022, 38(1):99-110.
- [14] Peng H, Huo J, Li B, et al. Predicting isocitrate dehydrogenase (IDH) mutation status in gliomas using multiparameter MRI radiomics features[J]. J Magn Reson Imaging, 2021, 53(5):1399-1407.
- [15] Chen B, Chen C, Wang J, et al. Differentiation of low-grade astrocytoma from anaplastic astrocytoma using radiomics-based machine learning techniques[J/OL]. Front Oncol, 2021, 11(11):e521313[2021 Jun 1]. DOI:10.3389/fonc.2021.521313. eCollection 2021.
- [16] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016:229-232.
- [17] 唐薇, 段俊艳, 余子意, 等. 增强 MRI 影像组学预测脑胶质瘤 IDH-1 基因突变的价值分析[J]. 磁共振成像, 2022, 13(5):111-114.
- [18] 陈洋, 阴鲁鑫, 李琳琳, 等. TOP II、IDH1 表达与胶质瘤预后的相关性及其临床价值[J]. 临床与实验病理学杂志, 2020, 36(10):1212-1214.

(收稿日期:2022-12-24 修回日期:2023-05-01)

《放射学实践》杂志微信公众平台开通啦!

遵照同行评议、价值导向、等效应用原则,国内各大学会、协会、组织机构通过科技工作者推荐、专家评议、结果公示等规定程序,《放射学实践》杂志入选中国科协发布 10 大领域《我国高质量科技期刊分级目录》业内认可的较高水平期刊。《放射学实践》杂志入选 2020 年版北京大学和北京高校图书馆期刊工作研究会共同主持的国家社会科学基金项目“学术期刊评价及文献计量学研究”研究成果——《中国核心期刊要目总览》。

《放射学实践》杂志微信公众平台立足于准确地传递医学影像领域的最新信息,致力于为关注医学影像领域的广大人士服务。欢迎大家通过微信平台,以文字、图片、音频和视频等形式与我们互动,分享交流最新的医学影像资讯。您还可以通过微信平台免费阅读及搜索本刊所有发表过的论文,投稿作者可以查询稿件状态等。

您可以通过以下方式关注《放射学实践》杂志微信公众平台:

1. 打开微信,通过“添加朋友”,在搜索栏里直接输入“放射学实践”进行搜索。
2. 在“查找微信公众号”栏里输入“放射学实践”即可找到微信公众号,点击“关注”,添加到通讯录。
3. 打开微信,点击“扫一扫”,手机镜头对准下面的二维码,扫出后点击关注即可。



注有惊喜!