

• 骨骼肌肉影像学 •

基于常规 MRI 影像组学模型预测软组织肉瘤复发

周慧, 陈基明, 吴莉莉, 邵颖, 范海云, 陈亮亮

【摘要】 目的:探讨基于常规 MRI 的影像组学模型对预测软组织肉瘤(STS)复发的价值。**方法:**回顾性分析 2012 年 1 月—2021 年 6 月在本院经手术病理证实的 92 例 STS 患者的临床和影像资料。术后每 3 个月进行一次影像学检查,随访时间至少 12 个月以上,根据随访结果有无复发或远处转移分为复发组(27 例),无复发组(65 例)。采用完全随机方法将所有患者按 7:3 的比例分为训练集($n=65$)和验证集($n=27$)。使用 ITK-SNAP 软件,分别在 T_1 WI 和压脂 T_2 WI 上逐层沿肿瘤边缘手动勾画 ROI 并进行三维融合(VOI),然后使用 AK 软件提取纹理特征,使用最小冗余最大相关(mRMR)和最小绝对值收敛和选择算子(LASSO)回归分析方法分别对 T_1 WI 序列、压脂 T_2 WI 序列和联合序列的纹理特征进行降维和筛选,并建立影像组学模型,根据各个组学特征的权重系数计算影像组学评分(Radscore),运用 100 次留组交叉验证(LGOCV)方法来评估模型的可靠性。将临床病理、常规 MRI 特征与预测效能最高的影像组学模型的 Radscore 相结合,采用多因素 logistic 回归(LR)、随机森林(RF)和支持向量机(SVM)三种机器学习算法分别建立机器学习模型。采用受试者工作特征(ROC)曲线评价各模型的预测效能,应用决策曲线分析(DCA)评估模型的临床应用价值。**结果:**临床模型在训练集和验证集中预测 STS 复发的 ROC 曲线下面积(AUC)分别为 0.71(95%CI:0.58~0.85)和 0.74(95%CI:0.52~0.97)。基于 T_1 WI、压脂 T_2 WI 和联合序列的影像组学模型在训练集中预测 STS 复发的 AUC 分别为 0.81(95%CI:0.70~0.93)、0.92(95%CI:0.86~0.99)和 0.91(95%CI:0.84~0.99),在验证集中分别为 0.84(95%CI:0.63~1.00)、0.92(95%CI:0.81~1.00)和 0.86(95%CI:0.72~1.00)。采用机器学习算法构建的 LR、RF 和 SVM 模型在训练集中预测 STS 复发的 AUC 分别为 0.93(95%CI:0.87~0.99)、0.91(95%CI:0.84~0.99)和 0.77(95%CI:0.63~0.91),在验证集中分别为 0.93(95%CI:0.83~1.00)、0.86(95%CI:0.71~1.00)和 0.83(95%CI:0.66~1.00)。DCA 分析结果表明,压脂 T_2 WI 和联合序列的影像组学模型、以及 LR 和 RF 模型的临床受益均较好。**结论:**基于常规 MRI 序列中的压脂 T_2 WI 和联合序列构建的影像组学模型对预测 STS 复发具有较高的预测效能和较好的临床受益,基于不同机器学习算法构建的预测模型的预测效能并无明显提高。

【关键词】 软组织肉瘤; 肿瘤复发; 磁共振成像; 影像组学; 预测模型

【中图分类号】 R738.6; R445.2 **【文献标志码】** A **【文章编号】** 1000-0313(2022)12-1561-07

DOI:10.13609/j.cnki.1000-0313.2022.12.017

开放科学(资源服务)标识码(OSID):



The value of conventional MRI imaging diagnostic model in predicting the recurrence of soft tissue sarcoma
ZHOU Hui, CHEN Ji-ming, WU Li-li, et al. Imaging Center, Yijishan Hospital, Wannan Medical College, Anhui 241001, China

【Abstract】 Objective: To investigate the value of conventional MRI imaging diagnostic model in predicting the recurrence of soft tissue sarcoma (STS). **Methods:** The clinical and imaging data of 92 patients with STS confirmed by surgery and pathology in Yijishan Hospital of Wannan Medical University from January 2012 to June 2021 were retrospectively analyzed. Imaging examination was performed every 3 months after operation, and each of the follow-up interval was at least 12 months. According to the follow-up results, the patients were divided into recurrence group (27 cases) and non-recurrence group (65 cases). Patients were divided in a ratio of 7:3 into the training set ($n=65$) and the validation set ($n=27$) using a completely randomized approach. Using ITK-SNAP software, ROIs were manually delineated layer by layer along the tumor edge on T_1 WI and fat suppressed (FS)

作者单位: 241001 安徽芜湖, 皖南医学院弋矶山医院影像中心

作者简介: 周慧(1995—), 女, 安徽铜陵人, 硕士研究生, 住院医师, 主要从事骨骼肌肉系统影像诊断工作。

通讯作者: 陈基明, E-mail: yjsycjm@126.com

T₂WI, and then 3D fusion (VOI) was performed. Then the AK software was imported to extract the texture features. The minimum redundancy maximum correlation (mRMR) and least absolute shrinkage and selection operator (LASSO) regression analysis methods were used to reduce the dimension of texture features of T₁WI sequence, fat suppressed T₂WI sequence and combined sequence, respectively, to select out the best feature subset, and establish the radiomics model. Then the radiomics score (Radscore) was calculated according to the feature weights, and the reliability of the model was evaluated by 100 leave-group-out cross validation (LGOCV). The clinicopathological indexes and conventional MRI features with the Radscore based on T₂WI sequence were combined to establish machine learning model using three machine learning algorithms, including multivariate logistic regression (LR), random forest (RF) and support vector machine (SVM). The predictive efficacy of different model was evaluated by receiver operating characteristic (ROC) curve, and the clinical value of the model was evaluated by decision curve analysis (DCA). **Results:** The area under the ROC curve (AUC) of the clinical model for predicting STS recurrence was 0.71 (95%CI: 0.58~0.85) in the training set and 0.74 (95%CI: 0.52~0.97) in the validation set. The AUC of the radiomics model of T₁WI sequence, FS-T₂WI sequence and combined sequence in predicting STS recurrence in the training set were 0.81 (95%CI: 0.70~0.93), 0.92 (95%CI: 0.86~0.99) and 0.91 (95%CI: 0.84~0.99), respectively. In the validation set, they were 0.84 (95%CI: 0.63~1.00), 0.92 (95%CI: 0.81~1.00) and 0.86 (95%CI: 0.72~1.00), respectively. The AUC of LR, RF and SVM model constructed by machine learning algorithm for predicting STS recurrence in training set were 0.93 (95%CI: 0.87~0.99), 0.91 (95%CI: 0.84~0.99) and 0.77 (95%CI: 0.63~0.91), respectively. In the validation set, they were 0.93 (95%CI: 0.83~1.00), 0.86 (95%CI: 0.71~1.00) and 0.83 (95%CI: 0.66~1.00), respectively. The results of DCA analysis showed that the radiomics model of FS-T₂WI sequence and combined sequence, LR and RF models had better clinical benefits. **Conclusion:** Radiomics models based on fat suppression T₂WI and combined sequences of conventional MRI sequences have high predictive efficiency and good clinical benefit for predicting STS recurrence, while prediction models based on different machine learning algorithms have no significant improvement in the predictive efficiency.

【Key words】 Soft tissue sarcoma; Tumor recurrence; Magnetic resonance imaging; Radiomics; Prediction model

软组织肉瘤 (soft tissue sarcoma, STS) 是一组高度异质性的恶性肿瘤, 其特点为具有局部侵袭性、呈浸润性或破坏性生长。目前, 手术切除是治疗软组织肉瘤的主要方式, 尽管对于高度恶性的Ⅲ期肿瘤根治性或广泛手术切除能尽量达到 R0 切除, 获得安全的手术切除边界, 但仍有 15%~40% 的 STS 患者出现肿瘤复发, 严重影响预后, 患者 5 年生存率仅为 55%~70%^[1-2]。STS 生物学行为存在多样性, STS 患者术后复发受多种因素影响, 不仅与切缘状态密切相关, 还与患者的年龄、病灶大小、肿瘤位置及其与周围结构关系等因素有关。MRI 是 STS 诊断与鉴别诊断、术前分期和手术治疗方案制订以及术后随访的首选影像检查方法。常规 MRI 检查最为常用, 但采用常规 MRI 预测 STS 复发的价值有限^[3-4]。影像组学通过自动化数据算法将医学图像转化为深层次特征空间数据, 可量化分析病灶内部潜在的病理生理学异质性, 是当前的研究热点^[5]。目前已有基于机器学习影像组学分析

应用于 STS 的鉴别诊断、术前分级等方面的研究^[6-8], 但应用于预测 STS 复发却鲜有报道。笔者应用不同机器学习方法建立基于常规 MRI 的影像组学模型, 旨在探讨其在预测软组织肉瘤复发中的价值。

材料与方法

1. 临床资料

回顾性搜集本院 2012 年 1 月—2021 年 6 月经手术病理证实 STS 患者的完整临床及影像资料。纳入标准: ①经手术病理证实为软组织肉瘤, 病理诊断依据为 2020 年第五版 WHO 骨肿瘤及软组织肿瘤分类标准; ②术前 2 周内进行常规 MRI 扫描; ③术后随访时间至少 12 个月。排除标准: ①肿瘤边界显示不清晰或 MRI 质量不佳, 难以准确勾画 ROI 及后处理; ②术后失访。最终纳入 92 例 STS 患者, 男 49 例、女 43 例, 年龄 5~87, 平均 (53±19) 岁。STS 病理类型包括纤维肉瘤 (28 例)、脂肪肉瘤 (15 例)、滑膜肉瘤 (15 例)、

未分化肉瘤(11 例)、横纹肌肉瘤(7 例)、软骨肉瘤(5 例)、骨外骨肉瘤(4 例)、血管肉瘤(3 例)、上皮样肉瘤(3 例)和腺泡样软组织肉瘤(1 例)。

术后每 3 个月进行一次胸部 X 线平片或 CT 检查及病灶部位的 MRI 检查,根据随访结果(其中二次手术 23 例、电话或门诊随访 7 例),将 STS 术后随访期间发生局部复发(26 例)或远处转移者(1 例肺转移)纳入复发组(27 例),未发生局部复发或远处转移者纳入无复发组(65 例)。采用计算机完全随机方法将患者按 7:3 的比例分为训练集 65 例(复发组 19 例,无复发组 46 例)和验证集 27 例(复发组 8 例,无复发组 19 例)。本研究经本院临床伦理委员会审批通过。

2. MRI 检查方法

使用 Siemens Avanto 1.5T 磁共振扫描仪,将靠近边缘区的病灶尽可能置于扫描中心。扫描序列及参数如下。横轴面和冠状面 T_1 WI:TR 460~580 ms,TE 8~11 ms;横轴面、冠状面和矢状面压脂 T_2 WI:TR 3000~5900 ms,TE 49~100 ms,矩阵 $205 \times 256 \sim 384 \times 288$,视野 $18 \text{ cm} \times 18 \text{ cm} \sim 40 \text{ cm} \times 40 \text{ cm}$,层厚 $4.0 \sim 6.0 \text{ mm}$,层间距 2.0 mm。52 例进行了 MR 增强扫描。

3. 临床病理和常规 MRI 征象的评价

搜集每例患者的临床病理资料,包括性别、年龄、病理分级和临床分期等。病理分级参考法国国家癌症研究中心(FNCLCC)的分级标准,将肿瘤的病理分级分为 I~III 级, I 级为低级别, II、III 级为高级别^[9];临床分期参考美国癌症联合会(AJCC)的分期标准,分为 I~IV 期^[10]。由 2 位分别具有 5 年和 15 年工作经验的放射科医师共同评估病灶的常规 MRI 特征,意见不一致时经讨论达成共识。常规 MRI 特征:①肿瘤最长径;②生长部位;③位置深浅:以浅筋膜为界,处于浅筋膜以外为浅部,处于浅筋膜以内为深部;④形态,即肿瘤边缘是否呈明显的不规则分叶状改变;⑤边界,包括肿瘤是否具有包膜、与邻近肌肉等组织分界是否清晰;

⑥肿瘤周围水肿,定义为肿瘤周围异常信号区,在 T_1 WI 上呈斑片状低信号、在压脂 T_2 WI 上呈高信号;⑦筋膜、血管和神经侵犯,表现为在增强 T_1 WI 或压脂 T_2 WI 上可见肿瘤与邻近筋膜、血管和神经分界不清或呈包绕改变;⑧骨和骨膜受累,骨受累表现为肿瘤包绕受累骨或侵及骨髓腔,骨膜受累表现为肿瘤与累及骨之间无薄层脂肪组织。在横轴面压脂 T_2 WI 上测量肿瘤最长径;分别在冠状面或矢状面 T_1 WI 和压脂 T_2 WI 上在病灶内和邻近的正常肌肉组织内勾画 ROI,测量病灶和正常肌肉组织的信号强度,所有 ROI 的形状、大小相同,放置 ROI 时应尽量避开肿瘤出血、坏死及囊变区。每个部位均测量 3 次,取其平均值。分别计算 T_1 WI 和压脂 T_2 WI 上病灶与正常肌肉信号的比值(Rt1 和 Rfst2)。

4. 图像预处理、纹理特征提取、筛选及影像组学模型的建立

由于不同部位病灶的扫描参数存在差异,故在提取纹理特征之前,对图像进行预处理以提高纹理分辨力。首先,使用 $\mu \pm 3\delta$ 方法标准化图像的灰阶水平,提高图像对比度^[11];然后,采用灰度均衡化,来减少计算时间,提高纹理结果的信噪比^[12];最后,应用 3D 插补技术进行重采样,确保病灶空间特征的一致性^[13]。

将所有患者的 T_1 WI 和压脂 T_2 WI 图像从 PACS 系统以 DICOM 格式导入 ITK-SNP 3.6.0 软件。由上述 2 位医师采用人工手动方法分别在横轴面 T_1 WI 和 T_2 WI 图像上(T_1 WI 上病灶 ROI 的勾画参考压脂 T_2 WI),沿病灶边缘逐层勾画 ROI,随后软件自动融合生成病灶的 3D-ROI(图 1)。然后将 3D-ROI 导入 AK 3.2.0 软件进行纹理特征的提取,软件经过自动计算得到基于 T_1 WI 和压脂 T_2 WI 的影像组学特征各 1316 个,特征类型包括一阶特征、形态特征、灰度共生矩阵特征、灰度区域大小矩阵特征、灰度游程矩阵特征、邻域灰度差矩阵特征、灰度依赖矩阵特征以及拉普拉斯-高斯滤波、小波变换和局部二值模式变换,其中,小波

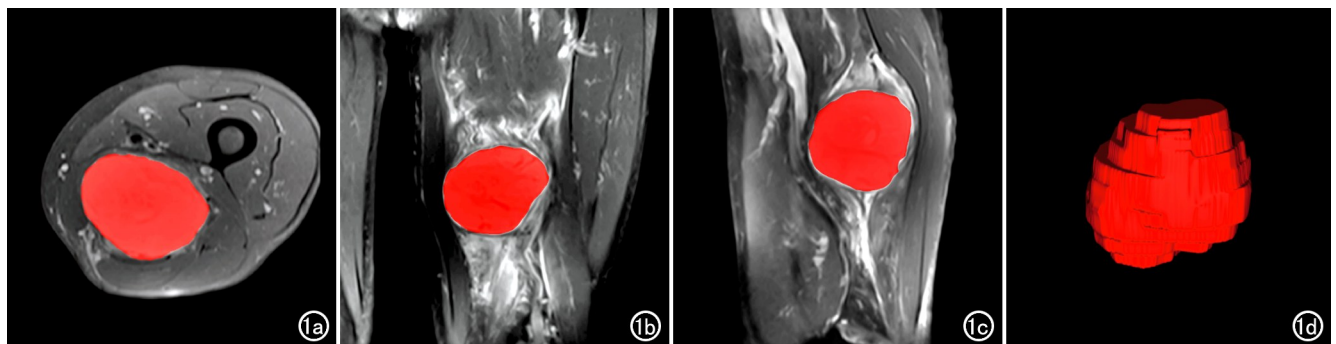


图 1 压脂 T_2 WI 图像上肿瘤 ROI 的勾画。a) 横轴面图像,沿肿瘤边缘勾勒 2D-ROI(红色区域); b~c)在横轴面 ROI 勾勒完成后,经 ITK-SNP 3.6.0 软件重组形成冠状面和矢状面图像上的 2D-ROI(红色区域); d) 所有层面的 2D-ROI 经软件融合生成病灶的 3D-ROI。

变换又包含 HLL、HHL、HLH、LHH、LLH、LHL、LLL 和 HHH 八种子带, H 和 L 分别代表病灶区域在 x、y 和 z 轴方向上的低频分量(low-frequency component)和高频分量(high-frequency component)。采用组内相关系数(ICC)评价 2 位医师提取特征的一致性, ICC>0.80 为一致性较好, 保留一致性较好的特征进行后续分析。采用最小冗余最大相关性(minimum redundancy maximum correlation, mRMR)和最小绝对收缩与选择算子(least absolute shrinkage and selection operator, LASSO)回归分析对纹理特征进行降维, 筛选出最佳特征子集, 构建基于 T₁WI 和压脂 T₂WI 和联合序列的 3 个影像组学模型, 采用 logistic 回归(logistic regression, LR)、随机森林(random forest, RF)及支持向量机(support vector machine, SVM)三种机器学习算法分别建立机器学习模型, 再

根据各个特征的权重计算影像组学评分(Radscore)。

5. 统计分析

采用 SPSS 24.0、R 语言 4.0.4 和 MedCalc 19.0.2 软件进行统计分析。计量资料采用 Kolmogorov-Smirnov 检验进行正态性检验, 符合正态分布的计量资料以均数±标准差表示, 并采用独立样本 *t* 检验进行组间比较; 计数资料以频数表示, 采用 χ^2 检验或 Fisher 精确检验进行组间比较。采用 ICC 评价两位医师提取纹理特征的一致性。绘制校正曲线并使用 Hosmer-Lemeshow 检验分析预测模型的拟合优度, 计算受试者工作特征(ROC)曲线下面积(AUC), 通过 DeLong 检验比较多个模型间 AUC 的差异。采用决策曲线分析(decision curve analysis, DCA)比较不同模型的临床净收益。以 *P*<0.05 为差异具有统计学意义。

表 1 无复发组和复发组 STS 患者在训练集和验证集中的临床资料和常规 MRI 特征

指标或特征	训练集(n=65)		<i>t</i> / χ^2 值	<i>P</i> 值	验证集(n=27)		<i>t</i> / χ^2 值	<i>P</i> 值
	无复发组 (n=46)	复发组 (n=19)			无复发组 (n=19)	复发组 (n=8)		
男/女	29/17	12/7	0.000	0.993	5/14	3/5	0.338	0.658
年龄/岁	53±17	49±18	−0.853	0.397	57±21	44±21	−1.419	0.168
肿瘤长径/cm	10.0±6.9	7.2±4.3	−1.670	0.100	8.5±6.0	5.8±5.8	−1.084	0.289
生长部位/例			1.488	0.475			3.398	0.183
四肢	32	11			12	7		
躯干	13	8			6	0		
其它(颈部)	1	0			1	1		
位置/例			7.490	0.006			7.467	0.017
深部	7	9			1	4		
表浅	39	10			18	4		
形态/例			3.721	0.054			3.431	0.145
规则	7	7			3	4		
不规则	39	12			16	4		
边界/例			0.591	0.442			0.016	1.000
清晰	32	15			9	4		
不清晰	14	4			10	4		
瘤周水肿/例			3.582	0.058			3.283	0.103
有	15	11			7	6		
无	31	8			12	2		
筋膜侵犯/例			1.685	0.194			7.467	0.017
有	6	5			1	4		
无	40	14			18	4		
血管、神经侵犯/例			0.530	0.466			8.882	0.006
有	15	8			3	6		
无	31	11			16	2		
骨、骨膜受累/例			4.208	0.040			1.067	0.633
有	33	18			13	7		
无	13	1			6	1		
病理分级/例			2.464	0.116			2.220	0.201
低级别	5	5			1	2		
高级别	41	14			18	6		
临床分期/例			7.027	0.071			3.732	0.155
I	4	2			1	2		
II	7	7			4	3		
III	35	9			14	3		
IV	0	1			0	0		
Rt1	1.18±0.58	1.35±0.69	1.038	0.303	1.22±0.54	1.01±0.19	−1.072	0.294
Rfst2	2.49±1.78	3.54±3.35	1.639	0.106	2.44±1.13	3.37±1.61	1.713	0.099

结 果

1. 临床资料和常规 MRI 特征的比较

在训练集和验证集中,无复发组与复发组之间临床病理特征和常规 MRI 征象的比较见表 1。在 2 个数据集中,病灶位置(深、浅)在两组之间的差异有统计学意义($P < 0.05$);骨及骨膜受累在训练集中的两组之间的差异有统计学意义($P < 0.05$),在验证集中的两组之间的差异无统计学意义($P > 0.05$);筋膜侵犯和神经血管侵犯两个征象在验证集中的两组之间的差异有统计学意义($P < 0.05$);其它变量在 2 个数据集中的两组之间的差异均无统计学意义($P > 0.05$)。

2. 影像组学特征的一致性分析

两位医师测量的各项影像组学特征之间的一致性均较好(P 均 < 0.05),ICC 的平均值为 0.821,95%CI 为 0.632~0.916。因此,本研究中选取高年资医师提取的组学特征进行后续分析。

3. 纹理特征筛选及影像组学模型的建立

AK 软件自 T_1 WI 和压脂 T_2 WI 上各提取了 1316 个纹理特征,2 个序列联合则共有 2632 个纹理特征,对这 3 个特征集分别采用 mRMR 方法去除冗余和不相关特征后各保留 30 个特征,然后采用 LASSO 回归进一步降维(图 2),最终基于 T_1 WI、压脂 T_2 WI 和联合序列各得到 9、13 和 12 个最具有预测价值的纹理特征(图 3)。采用 logistic 回归分析方法分别建立影像组学模型,将模型中各个特征与所对应的加权系数进行线性组合得到每个患者的 Radscore 值(图 4)。100 次留组交叉验证(100 leave-group-out cross validation, LGOCV)显示 2 个单独序列及联合序列的组学模型均具有良好的稳定性。

4. 临床和影像组学模型的效能评估

将组间差异有统计学意义的临床病理指标和常规

MRI 征象(包括病灶位置、骨及骨膜受累)纳入多因素 logistic 回归分析,构建临床模型,结果显示病灶位置为独立危险因素,其优势比(odds ratio, OR)为 5.014 ($P = 0.009$)。

采用 ROC 曲线分析临床模型和 3 个序列影像组学模型的诊断效能,结果见表 2 和图 5。在训练集和验证集中临床模型预测 STS 复发的 AUC 分别为 0.71 和 0.74。在 3 个序列影像组学模型中,压脂 T_2 WI 的影像组学模型在训练集和验证集中的诊断效能均较好(AUC 均为 0.92);但 Delong 检验结果显示,3 个序列影像组学模型之间 AUC 的差异均无统计学意义(训练集中 Z 值分别为 1.669、1.520 和 0.159,验证集中 Z 值分别为 1.065、0.193 和 0.592; P 均 > 0.05)。三个序列的影像组学模型的 AUC 均大于临床模型,但仅在训练集中压脂 T_2 WI 和联合序列影像组学模型的 AUC 与临床模型之间的差异有统计学意义($Z = 3.106, P = 0.002; Z = 2.545, P = 0.011$);在验证集中各序列影像组学模型与临床模型间 AUC 的差异均无统计学意义(Z 值分别为 0.700、1.535 和 0.935, P 均 > 0.05)。

5. 机器学习模型的建立及效能评估

因在训练集和验证集中压脂 T_2 WI 影像组学模型的诊断效能均较好,故将临床病理指标和常规 MRI 特征与压脂 T_2 WI 序列的 Radscore 相结合,分别采用多因素 LR、RF 及 SVM 三种机器学习方法建立机器学习模型。采用 ROC 曲线分析各模型诊断效能,结果见表 2 和图 5。在 LR、RF 和 SVM 三个机器学习模型中,LR 模型在训练集和验证集中的预测效能均相对最高(AUC 分别为 0.93 和 0.93)。在训练集中,LR 模型的 AUC 与 RF 模型之间的差异无统计学意义($Z = 0.362, P = 0.718$),与 SVM 模型的差异有统计学意义($Z = 2.444, P = 0.015$);在验证集中,LR 模型与 RF、

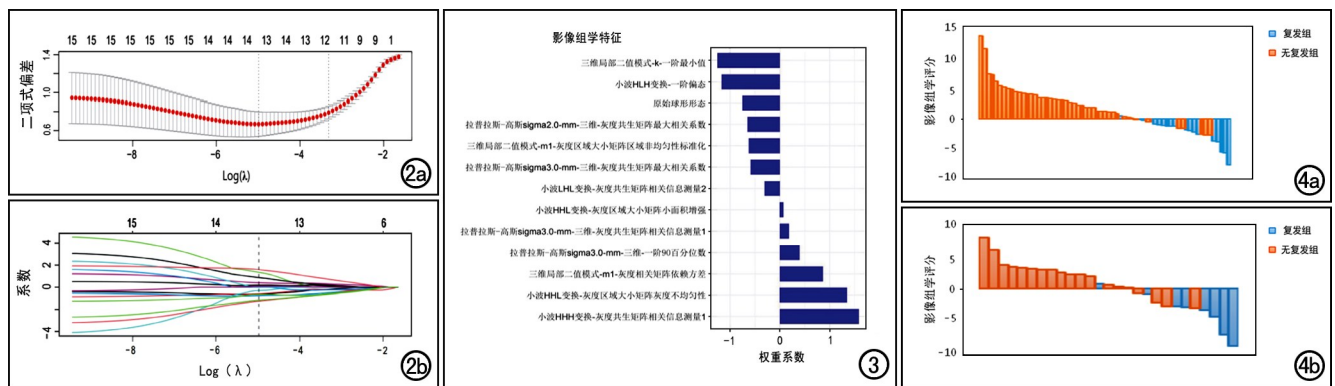


图 2 LASSO 回归对压脂 T_2 WI 序列提取的影像组学特征进行降维。a)使用 LASSO 方法降维后绘制二项式偏差和最优超参数 $\text{Log}(\lambda)$;b)纹理特征的 LASSO 收敛系数图,显示 13 个非零系数的纹理特征。

图 3 压脂 T_2 WI 影像组学模型中各项纹理特征及其权重系数值。图 4 训练集和与验证集中复发组与无复发组 Radscore 值柱状分布图,显示复发组的 Radscore 值显著低于无复发组。a)训练集;b)验证集。

表 2 临床模型、影像组学模型及不同机器学习模型 ROC 曲线分析结果

模型类型	训练集 (n=65)						验证集 (n=27)					
	AUC	Acc	Spe	Sen	PPV	NPV	AUC	Acc	Spe	Sen	PPV	NPV
临床模型	0.71(0.58~0.85)	0.58	0.47	0.85	0.40	0.88	0.74(0.52~0.97)	0.63	0.50	0.95	0.44	0.91
组学模型												
T ₁ WI	0.81(0.70~0.93)	0.82	0.84	0.72	0.66	0.89	0.84(0.63~1.00)	0.89	0.88	0.84	0.78	0.94
压脂 T ₂ WI	0.92(0.86~0.99)	0.92	0.95	0.87	0.89	0.94	0.92(0.82~1.00)	0.89	0.88	0.84	0.78	0.94
联合序列	0.91(0.84~0.99)	0.88	0.89	0.83	0.76	0.93	0.86(0.72~1.00)	0.85	0.88	0.74	0.75	0.89
机器学习模型												
LR	0.93(0.87~0.99)	0.94	0.95	0.87	0.89	0.96	0.93(0.83~1.00)	0.93	1.00	0.74	1.00	0.90
RF	0.91(0.84~0.99)	0.91	0.89	0.94	0.78	0.98	0.86(0.71~1.00)	0.78	0.75	0.90	0.58	0.93
SVM	0.77(0.63~0.91)	0.71	0.63	0.91	0.50	0.94	0.83(0.66~1.00)	0.85	0.88	0.79	0.75	0.89

注:Spe 为特异度,Sen 为敏感度,Acc 为符合率。括号内数据为 AUC 的 95%CI。

SVM 模型的 AUC 的差异均无统计学意义(Z 值分别为 0.921 和 1.263, P 均 >0.05)。

LR 模型的 AUC 与各序列影像组学模型之间的差异均无统计学意义(在训练集中 Z 值分别为 1.808、0.816 和 0.401,在验证集中 Z 值分别为 1.062、0.212、0.716, P 均 >0.05)。Hosmer-Lemeshow 检验分析表明 LR 模型在训练集和验证集中拟合度均较好,差异均无统计学意义(训练集: $\chi^2=6.593$, $P=0.781$;验证集: $\chi^2=4.564$, $P=0.803$)。

DCA 显示压脂 T₂WI 影像组学模型、RF 及 LR 模型均获得较好的净收益(图 6)。

讨 论

软组织肉瘤具有复发率高的特点,明确软组织肉瘤术后复发的影响因素,指导临床提早干预,对改善预后具有重要意义。本研究中纳入多种影响 STS 复发的因素,构建了临床模型、影像组学模型及不同机器学习模型,探索能有效预测软组织肉瘤复发的因素。本研究结果显示压脂 T₂WI、联合序列影像组学模型以及 LR、RF 模型均具有较高的预测效能(AUC 分别为 0.92、0.91、0.93 和 0.91),且临床受益均较好。

STS 复发与性别、年龄、部位、位置及边界、肿瘤组织类型、病理分级及临床分期等多种因素有关。本

研究纳入多种反映肿瘤异质性的临床病理、常规 MRI 特征进行研究,结果显示仅肿瘤位置等少数指标有预测价值,由此构建的临床模型预测 STS 复发的效能较低。影像组学可以发现肿瘤肉眼无法识别的微小差异。我们提取 T₁WI 和压脂 T₂WI 上病灶的纹理特征并进行拉普拉斯-高斯滤波、小波变换和局部二值模式变换,筛选出最佳特征子集并构建影像组学模型,结果显示压脂 T₂WI 和联合序列的影像组学模型的预测效能明显优于临床模型;但是,将临床病理、常规 MRI 特征与 AUC 较高的压脂 T₂WI 序列的 Radscore 相结合构建不同机器学习模型,预测效能均没有明显提高,说明临床病理、常规 MRI 特征贡献较小。

由于压脂 T₂WI 影像组学模型在训练集和验证集中的诊断效能均较好,本研究中仅讨论压脂 T₂WI 序列。自该序列筛选出的 13 个最佳纹理特征中,12 个为拉普拉斯-高斯滤波、小波变换和局部二值模式变换特征,其中 5 个是经小波滤波分解获得,且权重较大。原始图像上的纹理特征经过滤波变换后可抑制图像噪声,能获得更真实的影像特征^[14-15]。近年来,小波变换被广泛应用于多个领域,并在医学图像处理和信号分析中取得了很好的成效^[15-16],小波变换利用对图像的多分辨率和时间频率的细节分析,通过在高频处时间细分和低频处频率细分,解决时间和频率之间的冲

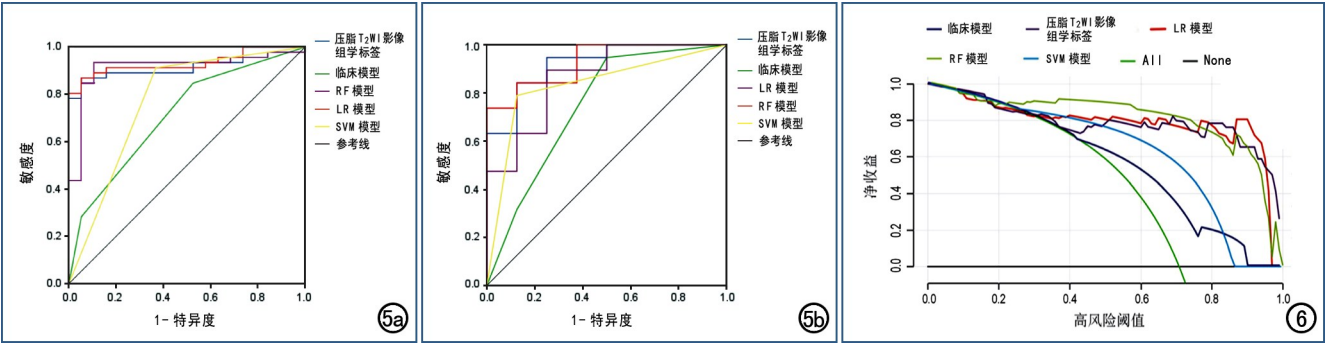


图 5 不同预测模型的 ROC 曲线,在训练集和验证集中均以 LR 模型的 AUC 最高,临床模型的 AUC 最低。
a)训练集;b)验证集。 图 6 临床模型、压脂 T₂WI 影像组学模型及 LR、RF 和 SVM 模型预测 STS 复发的决策曲线。All 表示所有 STS 复发患者, None 表示所有 STS 无复发患者。

突,从而实现对信号细节的任意聚焦,经小波变换后的纹理特征也更能反映肿瘤的异质性。

由于不同的机器学习算法各自具有不同的优点和不足,不同算法获得的模型表现各异,难以确定何种分类模型更适合预测 STS 复发。本研究探索性采用了多种机器学习算法构建预测模型,其中 LR 是最常用的分类方法,模型可解释性非常好;SVM 对于小样本、高维度、非线性数据有优势,泛化性能比较好^[18];而基于决策树方法的 RF 是具有监督集成功能的学习算法,能够处理不平衡数据集,在没有增加计算量的情况下能改进预测的准确性,抗过拟合能力强^[19-20]。本研究结果显示 LR 和 RF 模型均有较好的预测价值,明显优于 SVM 模型,LR 模型无论在训练集还是验证集中均具有很高的预测效能,具有较好的鲁棒性。

本研究的局限性:①为回顾性单中心研究,且 STS 的发病率低,故样本量相对较少,模型的泛化性和稳定性差,复发组与无复发组两组数据不均衡,需要继续扩大样本量、进行多中心研究,以改进模型的诊断效能;②筋膜、神经血管侵犯、骨和骨膜受累及位置深浅等特征为影像诊断结果,未经病理证实,存在一定的主观性;③本研究仅基于 T₁WI 和压脂 T₂WI 两个序列提取影像组学特征,缺少增强扫描序列,在反映肿瘤内部的特征信息方面可能存在缺失^[21]。

综上所述,基于常规 MRI 的压脂 T₂WI 序列和联合序列的影像组学模型具有较高的预测效能,能获得较好的临床受益,可以无创、简便、有效地预测 STS 复发;不同机器学习算法构建的预测模型与常规影像组学模型相比,预测效能并无明显提高。

参考文献:

- [1] Kraybill WG, Harris J, Spiro IJ, et al. Long-term results of a phase 2 study of neoadjuvant chemotherapy and radiotherapy in the management of high-risk, high-grade, soft tissue sarcomas of the extremities and body wall: radiation therapy oncology group trial 9514[J]. *Cancer*, 2010, 116(19):4613-4621.
- [2] Chowdhary M, Chowdhary A, Sen N, et al. Does the addition of chemotherapy to neoadjuvant radiotherapy impact survival in high-risk extremity/trunk soft-tissue sarcoma[J]. *Cancer*, 2019, 125(21):3801-3809.
- [3] Costa FM, Martins PH, Canella C, et al. Multiparametric MR imaging of soft tissue tumors and pseudotumors[J]. *Magn Reson Imaging Clin N Am*, 2018, 26(4):543-558.
- [4] de Angelis F, Guy F, Bertaut A, et al. Limbs and trunk soft tissue sarcoma systematic local and remote monitoring by MRI and thoraco-abdomino-pelvic scanner: a single-centre retrospective study[J]. *Eur J Surg Oncol*, 2019, 45(7):1274-1280.
- [5] 张利文, 方梦捷, 臧亚丽, 等. 影像组学的发展与应用[J]. *中华放射*

- 学杂志, 2017, 51(1):75-77.
- [6] 朱逸峰, 王绍武. 影像组学方法在软组织肿瘤中的应用研究进展[J]. *放射学实践*, 2020, 35(11):1479-1484.
- [7] Wang H, Zhang J, Bao S, et al. Preoperative MRI-based radiomic machine-learning nomogram may accurately distinguish between benign and malignant soft-tissue lesions: a two-center study[J]. *J Magn Reson Imaging*, 2020, 52(3):873-882.
- [8] Wang H, Chen H, Duan S, et al. Radiomics and machine learning with multiparametric preoperative MRI may accurately predict the histopathological grades of soft tissue sarcomas[J]. *J Magn Reson Imaging*, 2020, 51(3):791-797.
- [9] Coindre JM. Grading of soft tissue sarcomas: review and update[J]. *Arch Pathol Lab Med*, 2006, 130(10):1448-1453.
- [10] Amin MB, Edge SB, Greene FL, et al. *AJCC cancer staging manual*, 8th ed[M]. New York:Springer, 2017.
- [11] Collewet G, Strzelecki M, Mariette F, et al. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification[J]. *Magn Reson Imaging*, 2004, 22(1):81-91.
- [12] Gibbs P, Turnbull LW. Textural analysis of contrast-enhanced MR images of the breast[J]. *Magn Reson Med*, 2003, 50(1):92-98.
- [13] Depeursinge A, Foncubierta-RA, van De Ville D, et al. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities[J]. *Med Image Anal*, 2014, 18(1):176-196.
- [14] Peuna A, Thevenot J, Saarakkala S, et al. Machine learning classification on texture analyzed T₂ maps of osteoarthritic cartilage: oulu knee osteoarthritis study[J]. *Osteoarthr Cartilage*, 2021, 29(6):859-869.
- [15] Trivizakis E, Ioannidis GS, Souglakos I, et al. A neural pathomics framework for classifying colorectal cancer histopathology images based on wavelet multi-scale texture analysis[J/OL]. *Sci Rep*, 2021, 11(1):e15546.
- [16] 沈莎莎, 韩丹, 赵伦, 等. 基于甲状腺乳头状癌结节 CT 静脉期图像小波纹理分析预测中央组淋巴结转移的可行性研究[J]. *中华放射学杂志*, 2019, 53(11):946-951.
- [17] 蓝文婷, 冯湛, 张艳, 等. 多模态 MRI 影像组学随机森林模型预测术前大脑胶质瘤 IDH1 基因表达类型效能的初步探讨[J]. *中华放射学杂志*, 2019, 53(10):864-870.
- [18] 王小莉, 瞿航, 成维艳, 等. X 线影像组学在鉴别乳腺良恶性病变中的应用价值及 3 种模型效能比较[J]. *实用临床医药杂志*, 2021, 25(8):21-24.
- [19] Zhang B, He X, Ouyang F, et al. Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma[J]. *Cancer Lett*, 2017, 403(9):21-27.
- [20] 刘明璐, 沈浮, 陆建平. 基于 MR T₂WI 的影像组学对直肠癌新辅助治疗疗效的评估[J]. *放射学实践*, 2021, 36(3):371-376.
- [21] Liu Z, Li Z, Qu J, et al. Radiomics of multiparametric MRI for pretreatment prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study[J/OL]. *Clin Cancer Res*, 2019, 25(12):e35383547.

(收稿日期:2021-12-10 修回日期:2022-04-17)