

· 影像组学与人工智能专题 ·

钆塞酸二钠增强 MRI 肝胆期影像组学对乙肝患者肝储备功能的评估价值

胡景卉, 王芳, 侯承师, 黄京城, 杨鑫, 王文剑, 陈磊, 叶靖, 孙骏, 罗先富

【摘要】 目的:探讨钆塞酸二钠(Gd-EOB-DTPA)增强 MRI 肝胆期影像组学对乙型肝炎病毒感染患者肝储备功能的评估价值。**方法:**将 73 例慢性乙型肝炎肝功能正常者、136 例 Child-Pugh A 级和 70 例 Child-Pugh B/C 级乙型肝炎患者按照 8:2 随机分成训练组和测试组。从肝胆期 MRI 图像中勾画全肝实质进行影像组学特征提取。将得到的影像组学得分值(Rad-score)和多因素回归分析筛选出的临床变量联合,建立临床-影像组学联合模型。采用受试者工作特征(ROC)曲线分析评估效能。另一时间段 64 例(17 例肝功能正常、30 例 Child-Pugh A 级和 17 例 Child-Pugh B/C 级)乙肝患者用来验证。**结果:**多因素回归分析筛选出血小板、总胆红素和凝血酶原时间国际化比值用以建立鉴别肝功能正常与 Child-Pugh A 级的临床模型,利用 4 个最相关的影像组学特征构建影像组学模型,训练组、测试组和验证组中临床模型鉴别肝功能正常与 Child-Pugh A 级的曲线下面积(AUC)分别为 0.897、0.884、0.780,组学模型的 AUC 分别为 0.890、0.914、0.824,临床-影像组学联合模型的 AUC 分别为 0.951、0.978、0.886。联合模型与临床模型、组学模型相比,净重新分类改善指数(NRI)和综合判别改善指数(IDI)的检验值均 <0.05 。血小板和总胆红素被筛选出构建鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的临床模型,7 个最相关的影像组学特征构建影像组学模型,训练组、测试组和验证组中影像组学模型鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的 AUC 分别为 0.862、0.865、0.818,联合模型(AUC=0.940、0.934、0.951)与临床模型(AUC=0.916、0.893、0.914)的评估效能差异无统计学意义($P>0.05$)。**结论:**基于 Gd-EOB-DTPA 增强 MRI 肝胆期影像组学对乙肝患者的肝储备功能具有良好的评估价值,联合临床变量后可以提高 Child-Pugh A 级的早期检出效能,但对于 Child-Pugh A 级与 Child-Pugh B/C 级的评估,单独的临床模型已经表现出了良好的效能。

【关键词】 钆塞酸二钠; 乙型肝炎病毒; 肝储备功能; 磁共振成像; 影像组学; 列线图

【中图分类号】 R512.6; R445.2 **【文献标志码】** A **【文章编号】** 1000-0313(2024)01-0022-08

DOI:10.13609/j.cnki.1000-0313.2024.01.005

开放科学(资源服务)标识码(OSID):



Radiomics based on Gd-EOB-DTPA enhanced MR hepatobiliary phase images for evaluating liver reserve function of HBV patients HU Jing-hui, WANG Fang, HOU Cheng-shi, et al. The Clinical Medicine College of Yangzhou University, Jiangsu 225001, China

【Abstract】 Objective: To investigate the value of radiomics based on the hepatobiliary phase images of Gd-EOB-DTPA enhanced liver MR in evaluating liver reserve function of hepatitis B virus (HBV) patients. **Methods:** 73 chronic HBV patients with normal liver function, 136 HBV patients with Child-Pugh A liver function and 70 patients with Child-Pugh B/C liver function were randomly divided into the training group and the test group in the ratio of 8:2. Radiomics features were extracted from the hepatobiliary phase MR images via delineating the whole liver parenchyma. The clinic-radiomics model combining the radiomics score (Rad-score) and clinical variables selected by multivariate regression analysis was built. Evaluation performances of the model were analyzed using receiver operating characteristic (ROC) curve. 64 HBV patients were used for validation from another period, including 17 cases with normal liver function, 30 with Child-Pugh A and 17 with Child-Pugh B/C. **Results:**

作者单位: 225001 江苏,扬州大学医学院(胡景卉、王文剑); 225001 江苏,苏北人民医院医学影像科(胡景卉、侯承师、黄京城、杨鑫、王文剑、叶靖、孙骏、罗先富); 200232 上海,上海联影智能有限公司研发部(王芳、陈磊); 116031 辽宁,大连医科大学临床医学院(侯承师、黄京城、杨鑫)

作者简介: 胡景卉(1994—),女,安徽滁州人,硕士研究生,主要从事腹盆腔疾病影像诊断工作。

通讯作者: 罗先富, E-mail: xianfu-luo@hotmail.com

基金项目: 江苏省“333 高层次人才培养工程”(2022-3-6-139)

Platelet (PLT), total bilirubin (TBIL) and international normalized ratio of prothrombin time (INR) were selected by multivariate regression analysis to construct clinical model for distinguishing normal liver function from Child-Pugh A. The radiomics model was constructed using the four most relevant radiomics features. The areas under the curve (AUCs) of clinical model for distinguishing normal liver function from Child-Pugh A were 0.897, 0.884 and 0.780 in the training, test and validation cohorts, respectively. The AUCs of radiomics model were 0.890, 0.914 and 0.824 in the training, test and validation cohorts, respectively. The AUCs of clinic-radiomics combined model were 0.951, 0.978 and 0.886 in the training, test and validation cohorts, respectively, whose net reclassification index (NRI) and integrated discrimination improvement (IDI) were significantly different from the other two models ($P < 0.05$). PLT and TBIL were selected to construct clinical model for distinguishing Child-Pugh A from Child-Pugh B/C. The radiomics model was constructed using the seven most relevant radiomics features. The AUCs of radiomics model for distinguishing Child-Pugh A from Child-Pugh B/C were 0.862, 0.865 and 0.818 in the training, test and validation cohorts, respectively. There was no significant difference between the clinic-radiomics combined model (AUC=0.940, 0.934 and 0.951, respectively) and the clinical model (AUC=0.916, 0.893 and 0.914, respectively) ($P > 0.05$). **Conclusions:** Radiomics based on the hepatobiliary phase images of Gd-EOB-DTPA enhanced MR has a high evaluation value for liver reserve function of HBV patients, which can improve the early detection performance of Child-Pugh A when combined with clinical variables. However, the single clinical model has shown good evaluation efficacy for Child-Pugh A vs. Child-Pugh B/C.

【Key words】 Gd-EOB-DTPA; Hepatitis B virus; Liver reserve function; Magnetic resonance imaging; Radiomics; Nomogram

慢性乙肝是亚洲地区导致肝硬化的最常见原因,在早期阶段,临床表现常常较为隐蔽,当症状出现时,往往伴有严重的并发症。全球每年约有 200 万人死于肝病,其中超过 100 万人死于肝硬化并发症^[1]。随着肝炎的进展,肝细胞出现变性和坏死、纤维结缔组织增生和假小叶再生,导致正常肝细胞数量减少,影响肝脏的合成、分泌及代谢功能,从而导致肝储备功能的下降,临床上表现为总胆红素和转氨酶的升高、凝血障碍、脾肿大等。对慢性肝病患者进行早期的评估、管理和治疗至关重要。

目前,临床上主要通过血液生化指标来评估肝储备功能,但这些指标不仅不能显示肝脏的解剖结构或提供血液灌注信息,并且容易受急性感染等因素的影响。由于肉眼可见的结构变化往往在慢性疾病的发生过程中出现较晚,其他影像学检查如超声、CT 和常规 MRI 对于早期慢性肝病的检测能力有限^[2]。钆塞酸二钠 (gadolinium-ethoxybenzyl-diethylenetriamine-pentaacetic acid, Gd-EOB-DTPA) 作为一种肝细胞特异性对比剂,可在肝胆期被正常肝细胞通过肝细胞膜表面的有机阴离子转运系统阴离子转运多肽 (organic anion transporting polypeptide, OATP) 特异性摄取,并且肝胆期图像的多个测量参数已被证实可以用来评估肝功能,如肝实质相对强化程度、肝胆期的肝门静脉比等^[3-5]。然而,通过信号强度测量计算得出的参数来

评估肝功能是有限的,因为它们仅仅通过勾画肝脏几个层面的局部感兴趣区,这种方法并不能评估完整的肝实质储备功能。

影像组学是近几年新兴的影像研究领域,它可以通过勾画整个病变区域的体积,再通过机器学习分析大量的高维数据,提取出重要特征并定量表示,从而提高预测、诊断和预后的准确性^[6],这对于早期评估肝功能具有重要价值。周玮等^[7]仅从肝胆期影像组学方面评估了 Child-Pugh A 与 B/C 级肝硬化患者的肝储备功能,未联合临床进行分析是否临床已经具有较高的评估价值。张智星等^[8]采用了 3 期增强 CT 图像对 Child-Pugh 进行分级评估。本研究旨在探讨是否能够单独基于 Gd-EOB-DTPA 增强 MRI 肝胆期图像来建立临床-影像组学联合模型,对慢性乙肝患者肝功能进行更好的分级评估。

材料与方法

1. 病例资料

回顾性搜集 2018 年 12 月至 2021 年 7 月就诊于苏北人民医院行 Gd-EOB-DTPA 增强 MRI 检查的 1205 例存在慢性乙型肝炎感染的患者。病例纳入标准:①实验室证据表明存在乙型肝炎病毒感染;②具有完整的临床数据资料,包括是否有腹水、肝性脑病、机体白蛋白水平以及 MRI 检查前后 1 周内的总胆红素

(total bilirubin, TBIL)、丙氨酸氨基转移酶(alanine aminotransferase, ALT)、天冬氨酸氨基转移酶(aspartate transaminase, AST)、血小板(platelet, PLT)、凝血酶原时间(prothrombin time, PT)、凝血酶原时间国际化比值(international normalized ratio, INR);③肾功能正常。病例排除标准:①肝脏存在较大的病灶(最大直径 >5 cm)或多个病灶(数量 >5),导致正常肝实质范围较小影响测量准确性;②有肝介入治疗史或肝切除手术史;③门静脉存在栓子,影响血流动力学者^[9];④合并有其他肝炎病毒感染;⑤肝胆期图像质量不佳。根据收集的临床及实验室指标评估出 Child-Pugh 分级,包括总胆红素、白蛋白水平、凝血酶原延长时间、腹水量级以及是否有肝性脑病^[10]。采用相同的纳入和排除标准,从 2022 年 7 月至 12 月就诊的慢性乙型肝炎感染患者中筛选病例。本研究为回顾性临床研究,经苏北人民医院伦理委员会批准(批准文号:2021ky219),免除受试者知情同意。

2. MRI 检查方法

MRI 检查均采用 3.0T MRI 扫描仪(GE Discovery 750, 750W, 美国),8 通道体部相控阵线圈。扫描前嘱患者禁食 6~8 h,采用呼吸门控技术,对患者进行平静呼吸及屏气训练,扫描范围从膈顶至双肾下缘。所有患者均进行 MRI 平扫及钆塞酸二钠增强扫描。 T_1 WI 平扫及增强采用肝脏容积加速采集(liver acquisition with volume acceleration, LAVA)序列,分别于注射对比剂后 20 s(动脉期)、60 s(门静脉期)、3 min(移行期)及 20 min(肝胆特异期)进行扫描得到 4 期图像。LAVA 序列扫描参数:重复时间 3.7 ms,回波时间 1.7 ms,层厚 5.0 mm,层间距 2.5 mm,矩阵 320×224 ,视野 $40 \text{ cm} \times 40 \text{ cm}$,翻转角 18° 。对比剂采用钆塞酸二钠(Gd-EOB-DTPA, Primovist, 德国拜耳医药保健有限公司),注射流率 2.0 mL/s(剂量 0.025 mmol/kg),注射对比剂后再以相同流率注射 20 mL 0.9%氯化钠溶液进行冲洗。

3. 研究方法

临床模型构建:将临床变量(包括年龄、性别、身体质量指数、TBIL、ALT、AST、PLT、PT、INR)进行单因素和多因素逻辑回归分析,筛选出与 Child-Pugh 分

级相关的临床变量,即独立危险因素,将其联合构建出临床评估模型。采用受试者工作特征(receiver operating characteristic, ROC)曲线的曲线下面积(area under the curve, AUC)分析模型的评估能力。

肝半自动分割:由一位具有 5 年工作经验的放射科医师将所有纳入研究患者的肝胆期图像导入“uAI 科研平台”(uRP, United Imaging Healthcare Co, 中国上海)。一个名为 V-Net 的器官分割深度学习模型用于自动分割肝脏的感兴趣体积(volume of interest, VOI)^[11],再进一步进行手动修正,以排除每个层面上的脂肪、空气、肝脏病变、大血管和胆管区域(图 1、2)。慢性肝炎肝功能正常、Child-Pugh A 级和 Child-Pugh B/C 级的患者分别标记为 0、1 和 2。然后,由另一位具有 10 年工作经验的放射科医生检查核对 VOI 以及标记的标签。

影像组学特征提取与选择:使用 Python 3.7 中的 Pyradiomics 工具箱,从 Gd-EOB-DTPA 肝脏增强 MRI 肝胆期图像中提取符合国际生物标记物标准化倡议的 2600 个肝实质区域的影像特征。这些特征包括了 4 个特征组:18 个一级特征、14 个体积和形状特征、72 个纹理特征和 2496 个基于滤波器的特征。采用随机抽样方法将数据集按 8:2 的比例分成训练组和测试组,另一时间段的数据集构建验证组。为了避免

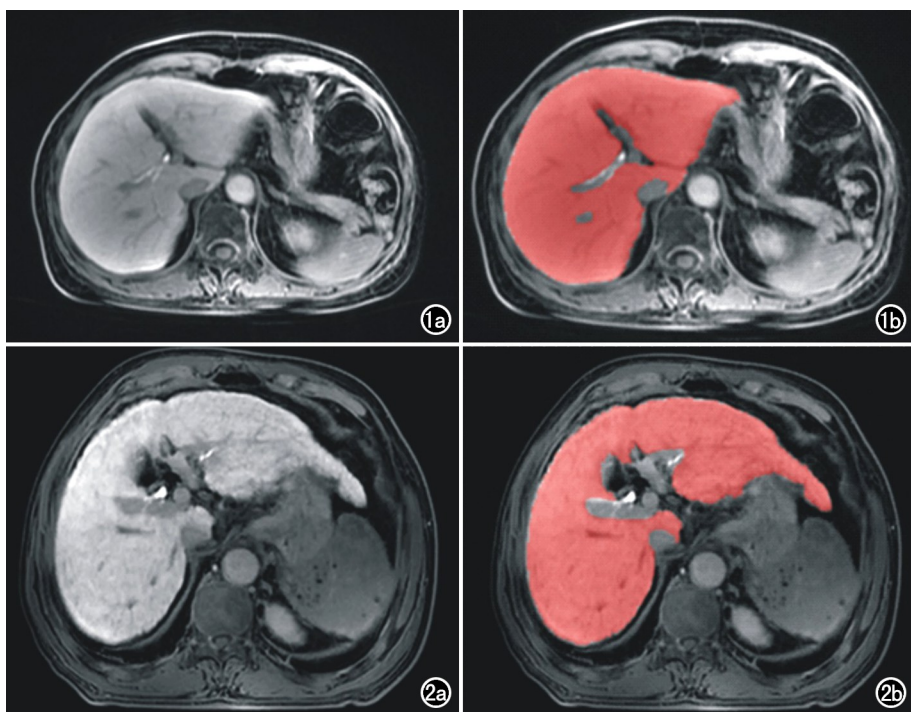


图 1 慢性乙肝肝功能正常者的 Gd-EOB-DTPA 增强 MRI 肝胆期图像, Rad-score 值为 0.565。a) 单个层面原图; b) 该层面勾画示例图。

图 2 Child-Pugh A 级患者的 Gd-EOB-DTPA 增强 MRI 肝胆期图像, Rad-score 值为 0.809。a) 单个层面原图,肝包膜稍欠光整; b) 该层面勾画示例图。

特征之间的维度影响,使用 Z 分数预处理方法进行特征归一化。应用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO) 算法进行降维处理,筛选出最相关的系数非 0 的影像组学特征。将得到的组学特征进行线性组合计算,得出每例患者对应的影像组学得分(Rad-score)。

影像组学模型及临床-组学联合模型构建:AUC 用于量化 Rad-score 的评估能力,在训练组、测试组和验证组中,分别建立 Rad-score 影像组学评估模型。将临床模型中的独立危险因素和 Rad-score 联合进行多因素逻辑回归分析,构建出临床-影像组学联合评估模型,用于对慢性肝炎患者进行肝储备功能评估。并根据训练组的约登指数确定临界值,以得到相应的敏感度、特异度、准确度以及阳性预测值、阴性预测值。净重新分类改善指数(net reclassification index, NRI)、综合判别改善指数(integrated discrimination improvement, IDI)用以评价联合模型较其他模型的检出效能改善情况。

列线图的建立和评估:利用训练组建立并绘制整合了临床变量和影像组学 Rad-score 的组学列线图。绘制列线图的校准曲线,通过 Hosmer-Lemeshow 检验来评估拟合优度。决策曲线分析(decision curve analysis, DCA)用于评估不同阈值概率下列线图的净收益。

4.统计学分析

采用 SPSS 24.0 软件和 R 语言(version 3.3.3, Vienna, Austria)进行统计学分析。根据正态性检验结果,将定量数据以中位数(四分位数间距,不符合正态

分布)或平均值±标准差(符合正态分布)表示,定性数据以百分数(%)表示。定量数据采用 Mann-Whitney U 检验或 Student's t 检验进行组间比较,定性数据采用卡方检验进行组间比较。以 $P<0.05$ 为差异有统计学意义。

结 果

1.病例资料

1205 例慢性乙肝患者经纳入、排除标准,排除临床数据资料不完整者 42 例,肝脏存在较大病灶(最大直径>5 cm)者 287 例,多个病灶(数量>5)者 336 例,有肝脏介入治疗史或肝切除手术史者 114 例,门静脉存在栓子者 56 例,合并丙肝者 55 例,戊肝者 17 例,肝胆期图像质量不佳者 19 例。最终 279 例患者纳入本研究,其中男 171 例,女 108 例,年龄中位数为 60 岁,按照临床 Child-Pugh 分级标准将 279 例患者分为 73 例慢性乙型肝炎肝功能正常者、136 例 Child-Pugh A 级患者和 70 例 Child-Pugh B/C 级患者。肝功能正常和 Child-Pugh A 级的 209 例患者被随机分配到训练组($n=167$)和测试组($n=42$)中,Child-Pugh A 和 Child-Pugh B/C 级的 206 例被随机分配到训练组($n=164$)和测试组($n=42$)中。患者的临床资料见表 1。另一时间段筛选出的 64 例患者用于验证,包括 17 例慢性乙肝肝功能正常者、30 例 Child-Pugh A 级和 17 例 Child-Pugh B/C 级乙肝患者。

2.临床模型

单因素及多因素逻辑回归分析结果显示,PLT、TBIL 和 INR 是乙肝 Child-Pugh A 级的独立危险因

表 1 患者基线资料 (n,%)

指标	肝功能正常和 Child-Pugh A 级患者				Child-Pugh A 和 Child-Pugh B/C 级患者			
	训练组 (n=167)	测试组 (n=42)	统计量	P 值	训练组 (n=164)	测试组 (n=42)	统计量	P 值
年龄(岁) ^a	60.00(18.00)	59.50(19.00)	-0.221 ^d	0.825	57.50(18.50)	62.50(14.75)	-1.669 ^d	0.095
性别(%) ^b			0.097 ^e	0.755			0.104 ^e	0.748
男性	103(61.68)	27(64.29)			101(61.59)	27(64.29)		
女性	64(38.32)	15(35.71)			63(38.41)	15(35.71)		
身高(cm) ^a	165.00(10.00)	165.00(12.00)	-1.069 ^d	0.285	165.00(10.00)	168.00(10.00)	-0.687 ^d	0.492
体重(kg) ^a	65.00(12.00)	65.00(10.50)	-1.767 ^d	0.077	65.00(11.75)	65.00(10.25)	-0.553 ^d	0.580
BMI(kg/m ²) ^c	24.11±3.25	23.61± 3.10	0.897 ^f	0.371	24.22±2.99	23.67±3.81	1.027 ^f	0.306
TBIL(μmol/L) ^a	13.80(9.10)	13.50(8.38)	-0.508 ^d	0.611	20.15(17.88)	17.60(11.73)	-1.559 ^d	0.119
ALT(U/L) ^a	24.00(19.00)	27.50(25.78)	-0.545 ^d	0.586	33.00(31.60)	28.45(24.25)	-2.131 ^d	0.033
AST(U/L) ^a	27.00(17.00)	28.00(19.00)	-0.258 ^d	0.796	38.00(37.75)	35.50(30.75)	-0.775 ^d	0.438
PLT(10 ⁹ /L) ^a	141.00(102.00)	113.50(55.00)	-2.956 ^d	0.003	91.00(86.25)	102.50(54.25)	-1.371 ^d	0.170
PT(s) ^c	13.20(1.70)	12.95(1.85)	-0.251 ^d	0.802	14.15(2.85)	14.10(2.43)	-0.680 ^d	0.496
INR ^a	1.05(0.14)	1.05(0.12)	-0.247 ^d	0.805	1.13(0.25)	1.16(0.14)	-0.791 ^d	0.429
Cr(μmol/L) ^a	74.20(14.00)	76.00(16.50)	-1.402 ^d	0.161	75.00(14.00)	76.50(22.25)	-0.892 ^d	0.372
Child-Pugh 分级 ^b			0.014 ^e	0.905			0.010 ^e	0.921
肝功能正常	58(34.73)	15(35.71)			—	—		
Child-Pugh A	109(65.27)	27(64.29)			108(65.85)	28(66.67)		
Child-Pugh B/C	—	—			56(34.15)	14(33.33)		

注: BMI 身体质量指数, TBIL 总胆红素, ALT 丙氨酸氨基转移酶, AST 天冬氨酸氨基转移酶, PLT 血小板, PT 凝血酶原时间, INR 凝血酶原时间国际标准化比值, Cr 肌酐。 ^a中位数(四分位间距); ^b百分数(%); ^c平均值±标准差; ^dZ 值; ^e卡方值; ^ft 值。

素,PLT、TBIL 是乙肝 Child-Pugh B/C 级的独立危险因素(表 2、3),分别进行联合,构建出临床评估模型,训练组、测试组和验证组中临床模型鉴别肝功能正常与 Child-Pugh A 级的 AUC 分别为 0.897、0.884、0.780,训练组、测试组和验证组中临床模型鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的 AUC 分别为 0.916、0.893、0.914。

3.影像组学特征及影像组学模型的评估效能

在训练组中,慢性肝炎肝功能正常者和 Child-Pugh A 级患者的数量分别为 58 例和 109 例。通过 LASSO 回归分析后,选择出 4 个相关特征,包括 2 个灰度依赖性矩阵(gray-level dependence matrix, GLDM)、1 个灰度大小区域矩阵(gray-level size zone matrix, GLSZM)和 1 个灰度行程矩阵(gray-level run length matrix, GLRLM)。所有特征的定量值在两组中均具有统计学差异($P<0.05$)。根据以上 4 个特征及其系数、截距值计算得到影像组学标签 Rad-score (图 3):

$$\text{Rad-score}_{\text{肝功能正常 vs. Child-Pugh A}} = 0.038 \times \text{normalize_glrlm_ShortRunHighGrayLevelEmphasis} + 0.026 \times \text{normalize_gldm_SmallDependenceEmphasis} - 0.005 \times \log_gldm_log\text{-}\sigma\text{-}1\text{-}5\text{mm}\text{-}3\text{D}\text{-}LargeDependenceHighGrayLevelEmphasis - 0.083 \times \text{wavelet_glszm_wavelet\text{-}HHH\text{-}ZoneEntropy} + 0.651$$

在训练组、测试组和验证组中,Rad-score 鉴别肝

功能正常与 Child-Pugh A 级的 AUC 分别为 0.890、0.914、0.824(表 4)。

Child-Pugh A 级与 Child-Pugh B/C 级患者通过 LASSO 回归分析后,选择出 7 个相关特征,包括 4 个一阶特征(first order)、1 个灰度共生矩阵(gray-level cooccurrence matrix, GLCM)和 2 个 GLRLM。所有特征的定量值在两组中均具有统计学差异($P<0.05$)。根据以上 7 个特征及其系数、截距值计算得到影像组学标签 Rad-score:

$$\text{Radscore}_{\text{Child-Pugh A vs. B/C}} = 0.067 \times \text{normalize_glrlm_ShortRunHighGrayLevelEmphasis} + 0.028 \times \text{wavelet_firstorder_wavelet\text{-}LLH\text{-}Kurtosis} + 0.003 \times \text{normalize_glrlm_RunLengthNonUniformityNormalized} - 0.006 \times \text{wavelet_glcm_wavelet\text{-}HHH\text{-}InverseVariance} - 0.010 \times \text{normalize_firstorder_Maximum} - 0.011 \times \text{normalize_firstorder_TotalEnergy} - 0.074 \times \text{normalize_firstorder_RootMeanSquared} + 0.340$$

在训练组、测试组和验证组中,Rad-score 鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的的 AUC 分别为 0.862、0.865、0.818。

4.临床-影像组学联合模型的评估效能

对于肝功能正常与 Child-Pugh A 级患者,将 PLT、TBIL、INR 与 Rad-score 进行多因素回归分析后,所有参数均纳入临床-影像组学联合模型(表 3)。训练组中,联合模型鉴别肝功能正常与 Child-Pugh A

表 2 临床参数的单因素和多因素回归分析结果

变量	肝功能正常和 Child-Pugh A 级				Child-Pugh A 和 Child-Pugh B/C 级患者			
	单因素回归分析		多因素回归分析		单因素回归分析		多因素回归分析	
	OR (95%CI)	P 值	OR (95%CI)	P 值	OR (95%CI)	P 值	OR (95%CI)	P 值
年龄	0.987(0.948~1.026)	0.503			1.015(0.966~1.067)	0.558		
性别	0.943(0.298~2.989)	0.921			1.459(0.451~4.715)	0.528		
BMI	1.084(0.906~1.296)	0.379			0.837(0.693~1.010)	0.063		
TBIL	1.119(1.019~1.229)	0.019	1.149(1.051~1.256)	0.002	1.116(1.043~1.194)	0.001	1.140(1.082~1.200)	<0.001
ALT	1.059(0.998~1.124)	0.060	1.001(0.983~1.019)	0.918				
AST	1.041(0.979~1.108)	0.200	1.004(0.984~1.024)	0.678				
PLT	0.990(0.982~0.998)	0.018	0.992(0.986~0.998)	0.012	0.986(0.975~0.999)	0.028	0.983(0.973~0.994)	0.002
PT	0.976(0.550~1.731)	0.934			1.695(0.993~2.895)	0.053		
INR	6.361(2.474~16.352)	<0.001	5.418(2.718~10.802)	<0.001	1.622(0.913~2.880)	0.099		

表 3 临床和影像组学参数的多因素回归分析结果

变量	肝功能正常与 Child-Pugh A 级		Child-Pugh A 与 Child-Pugh B/C 级患者	
	OR (95%CI)	P 值	OR (95%CI)	P 值
TBIL	1.140(1.039~1.250)	0.005	1.104(1.043~1.169)	0.001
PLT	0.990(0.982~0.998)	0.012	0.982(0.971~0.993)	0.001
INR	4.290(1.904~9.666)	<0.001	—	—
Rad-score	6.101(2.915~12.768)	<0.001	2.188(1.420~3.373)	<0.001

表 4 肝功能正常与 Child-Pugh A 级组的临床模型、影像组学模型和临床-组学联合模型评估效能

指标	临床模型			影像组学模型			联合模型		
	训练组	测试组	验证组	训练组	测试组	验证组	训练组	测试组	验证组
AUC (95%CI)	0.897 (0.849~0.945)	0.884 (0.787~0.981)	0.780 (0.649~0.911)	0.890 (0.838~0.942)	0.914 (0.824~1.000)	0.824 (0.707~0.940)	0.951 (0.920~0.982)	0.978 (0.944~1.000)	0.886 (0.789~0.984)
敏感度(%)	89.0	77.8	86.7	76.1	81.5	63.3	91.7	88.9	90.0
特异度(%)	79.3	73.3	47.1	87.9	80.0	76.5	89.7	86.7	76.5
准确度(%)	85.6	76.2	72.3	80.2	81.0	68.1	91.0	88.1	85.1
阳性预测值(%)	89.0	84.0	74.3	92.2	88.0	82.6	94.3	92.3	87.1
阴性预测值(%)	79.3	64.7	66.7	66.2	70.6	54.2	85.2	81.3	81.3

级的 AUC 为 0.951, 较临床模型 ($\text{NRI}=1.022, P<0.001$; $\text{IDI}=0.172, P<0.001$) 和影像组学模型 ($\text{NRI}=1.217, P<0.001$; $\text{IDI}=0.197, P<0.001$), 显著提高了评估效能。测试组中, 联合模型鉴别肝功能正常与 Child-Pugh A 级的 AUC 为 0.978, 同样显著提高了评估效能 (与临床模型比较: $\text{NRI}=1.452, P<0.001$; $\text{IDI}=0.331, P=0.001$ 。与组学模型比较: $\text{NRI}=1.644, P<0.001$; $\text{IDI}=0.224, P=0.002$)。验证组中, 联合模型鉴别肝功能正常与 Child-Pugh A 级的 AUC 为 0.886, 较临床模型 ($\text{NRI}=0.302, P=0.007$; $\text{IDI}=0.170, P=0.002$) 和组学模型 ($\text{NRI}=0.639, P<0.001$; $\text{IDI}=0.311, P<0.001$) 也显著提高了评估效能 (表 4、图 4)。

对于 Child-Pugh A 与 Child-Pugh B/C 级患者, 将 PLT、TBIL 与 Rad-score 进行多因素回归分析后, 所有参数均纳入临床-影像组学联合模型 (表 3)。在训练组、测试组和验证组中, 联合模型鉴别 Child-Pugh A 级与 Child-Pugh B/C 级的 AUC 分别为 0.940、0.934、0.951, 相较于临床模型 ($\text{AUC}=0.916$ 、0.893、0.914) 并未有明显提高, 差异无统计学意义 (Delong 检验, P 均 >0.05)。

5. 列线图的建立和评估

对于肝功能正常与 Child-Pugh A 组, 临床-影像组学联合模型的列线图如图 5 所示, Hosmer-Lemeshow 拟合优度检验的 P 值无统计学意义 (训练组 $P=0.443$, 测试组 $P=1.000$, 验证组 $P=0.343$, 图

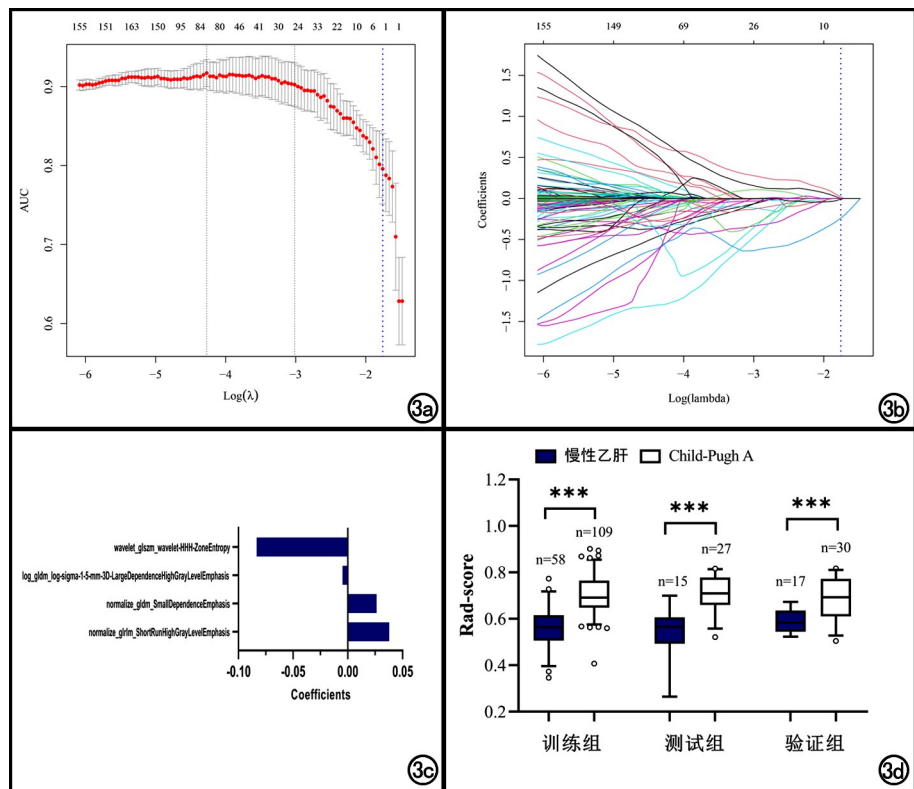


图 3 肝功能正常与 Child-Pugh A 级组采用 LASSO 算法进行影像组学特征筛选。a) 调整参数 λ 使拟合损失值二项偏差最小, 以筛选出最优的影像组学特征; b) 筛选出最优影像组学特征的系数收敛图, 在最佳 λ 值处画一条垂直线, 得到 4 个系数非零的特征; c) 选定的 4 个最优特征及其相关系数; d) 训练组、测试组和验证组中每例患者的影像组学得分, 在鉴别慢性乙肝与 Child-Pugh A 级患者中差异有统计学意义 ($P<0.05$)。

6), 表明列线图具有良好的校准性能。在 DCA 中的所有阈值概率下, 临床-影像组学联合模型的列线图表现出优于其他模型的临床净收益 (图 7)。

讨论

本研究从肝脏 Gd-EOB-DTPA 增强 MRI 肝胆期图像中提取出与慢性乙型肝炎肝储备功能最相关的影像组学特征, 并通过计算得到每例患者对应的影像组学标签即 Rad-score, 结果显示影像组学模型具有良好

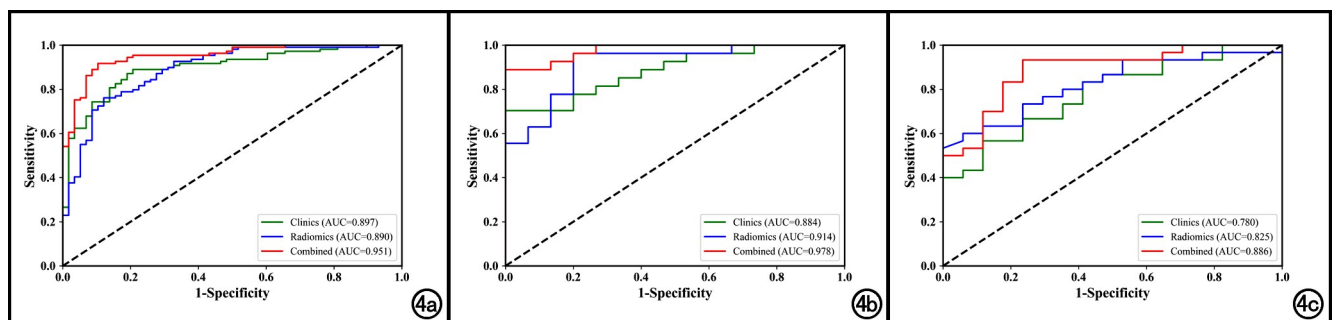


图 4 肝功能正常与 Child-Pugh A 级组中临床模型、影像组学模型和临床-组学联合模型的 ROC 曲线。a) 训练组; b) 测试组; c) 验证组。

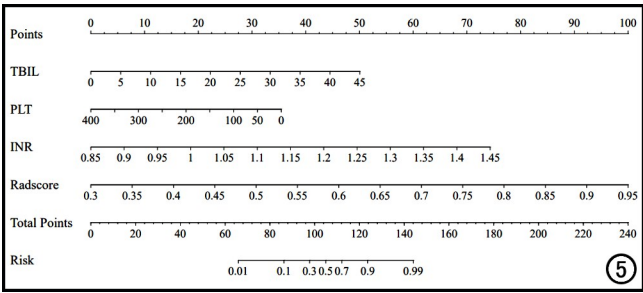


图 5 肝功能正常与 Child-Pugh A 级组联合模型的列线图。

的评估效能。肝功能正常与 Child-Pugh A 级组在联合了影像组学和临床模型后,联合模型对于评估效能有所提高,优于单纯的临床模型和影像组学模型。但在 Child-Pugh A 级与 Child-Pugh B/C 级组中,临床模型已经表现出了较好的评估价值。

经过单因素和多因素回归分析后,肝功能正常与 Child-Pugh A 级组筛选出了 3 个临床参数纳入临床模型,分别是 TBIL、PLT 和 INR。Child-Pugh A 级与 Child-Pugh B/C 级组则筛选出 TBIL 和 PLT 建立临床模型。TBIL 是肝功能中的一个重要指标,肝炎进展过程中,肝功能下降,机体对于胆红素的代谢能力也下降,导致其在血液中积聚、增高。相关研究表明 PLT 与肝脏的再生功能密切相关,肝功能差的患者往往由于正常肝细胞减少,PLT 数量也相应下降,导致脾肿大、脾功能亢进,PLT 单因素对于肝功能的预测

就具有一定意义^[12,13]。同样,肝功能的下降,凝血酶原相对也会缺乏,凝血因子的合成发生障碍,INR 也是一种反映肝脏合成能力的重要指标^[14,15]。两组的临床模型测试组 AUC 分别达到了 0.884、0.893,验证组为 0.780、0.914,具有良好的评估效能。

本研究中肝功能正常与 Child-Pugh A 级组提取出了 4 个最相关的肝胆期影像组学纹理特征,包括 2 个 GLDM 特征、1 个 GLRLM 特征和 1 个 GLSZM 特征。GLSZM 纹理特征可以用来描述灰度值的分布,来自于 GLSZM 类的区域熵特征具有最高的绝对值相关性系数,可以度量灰度不均匀性或随机性,值越高表示肝实质组织的异质性越高^[16]。其他包括 GLDM 纹理特征可以量化肝胆期 MRI 图像的灰度依赖性,GLRLM 纹理特征则提供了具有相同灰度的连续像素运行的空间分布信息,它们都可以反映出一阶纹理特征无法显示的空间异质性改变^[17]。Child-Pugh A 级与 Child-Pugh B/C 级组提取出 7 个最相关的肝胆期影像组学纹理特征,包括 4 个一阶特征、1 个 GLCM 和 2 个 GLRLM 特征。一阶直方图特征描述了图像区域内体素强度的分布,晚期肝硬化有大小不一的结节和纤维间隔,导致体素强度分布不均;GLCM 则可以捕获具有预定义灰度强度的像素对或体素之间的空间关系。病理上,随着肝炎的进展,肝细胞变性坏死导致正常的肝小叶被破坏,假小叶内再生的肝细胞不具备正常肝细胞的功能,从而对于 Gd-EOB-DTPA 的摄取减少,肝胆期图像信号强度降低,肝脏组织的异质性也随

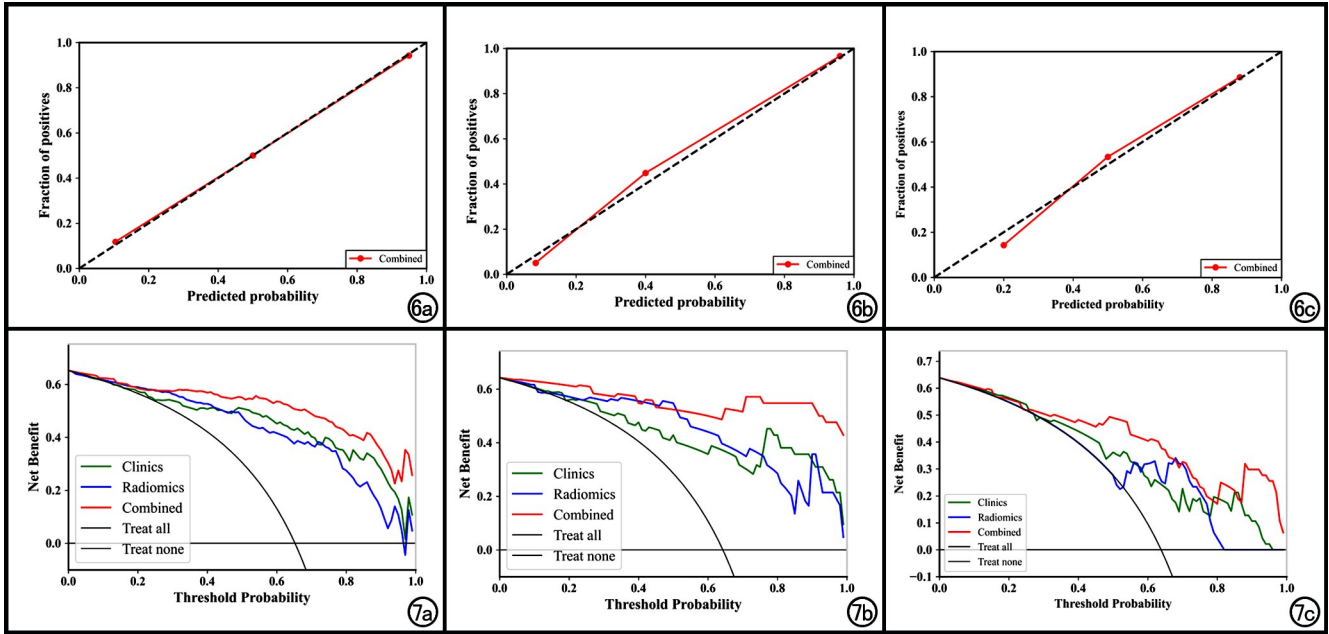


图 6 校准曲线表明肝功能正常与 Child-Pugh A 级组中训练组、测试组和验证组的列线图校准良好,Hosmer-Lemeshow 拟合优度检验的 P 值均 >0.05 。a)训练组;b)测试组;c)验证组。图 7 决策曲线分析结果表明,与临床模型和影像组学模型相比,肝功能正常与 Child-Pugh A 级组中训练组、测试组和验证组中联合模型列线图在各概率阈值下均具有更高的临床净效益。a)训练组;b)测试组;c)验证组。

着肝脏硬度的增加而增加,逐渐演变形成肝硬化^[18-20]。影像组学在肝胆期图像上进行全肝组织的体积分割,准确捕捉细微的组织变化,提取出高维的成像特征,通过筛选出的这些纹理特征计算得到 Rad-score 组学模型,也具有较高的评估效能,测试组分别为 0.914、0.865,验证组分别为 0.824、0.818。在联合了影像组学和临床模型后,诊断效能得到了显著提高,并且在两组的多因素回归分析中,Rad-score 都具有最高的权重系数(OR=6.101、2.188),这充分说明了影像组学在肝储备功能评估中的重要作用。肝功能正常与 Child-Pugh A 级组的联合模型 AUC 值,训练组为 0.951,测试组为 0.978,验证组为 0.886,并且 NRI 和 IDI 表明均改善了模型的评估能力,但对于 Child-Pugh A 级与 Child-Pugh B/C 级组,联合模型相对于临床模型并没有明显优势。

既往多项研究表明,影像组学可以更好地评估肝脏功能储备。相关研究表明通过评估术前肝功能,可以预测肝癌患者术后肝功能衰竭^[21,22];还可以用来预测、识别慢性肝病患者是否发生晚期肝纤维化以及准确分期患者的显著纤维化($\geq F2$)水平^[23,24]。然而,肝纤维化程度并不能直接反映整体的肝功能。Nitsch 等^[25]曾利用基于肝脏和脾脏的 MRI 影像特征开发了一种肝硬化疾病严重程度的预测模型,但它们仅限于临床失代偿的患者。本研究更全面地评估了乙肝患者的肝储备功能,为临床迅速采取干预措施提供了有价值的辅助信息。

本研究存在以下局限性:首先,本研究仅针对乙肝患者,然而在西方国家,丙肝和酒精性脂肪性肝炎也是肝病的常见原因,我们将在未来扩大纳入标准;其次,可能是由于选择偏倚或样本量较小,临床基线资料中肝功能正常与 Child-Pugh A 级组中的 PLT 和 Child-Pugh A 级与 Child-Pugh B/C 级组中的 ALT 在训练组和测试组中差异无统计学意义,并且由于 Child-Pugh C 级的病例数较少而和 B 级患者一起进行分析,未来将扩大样本量进行更全面研究。

综上所述,Gd-EOB-DTPA 增强 MRI 肝胆期影像组学模型可以对慢性乙肝患者进行较好的肝功能分级,临床-影像组学联合模型在乙肝患者肝功能损伤早期 Child-Pugh A 级的检出方面表现出了更好的性能。未来,采用基于影像组学的模型来评估肝功能,可能可以优化对于乙肝患者的管理及早期发现,更好地服务于精准医疗。

参考文献:

- [1] Asrani SK,Devarbhavi H,Eaton J,et al.Burden of liver diseases in the world[J].J Hepatol,2019,70(1):151-171.
- [2] Van Beers BE,Daire JL,Garteiser P.New imaging techniques for

- liver diseases[J].J Hepatol,2015,62(3):690-700.
- [3] Poetter-Lang S,Bastati N,Messner A,et al.Quantification of liver function using gadoxetic acid-enhanced MRI[J].Abdom Radiol (NY),2020,45(11):3532-3544.
- [4] 毕新军,张学琴,陆健,等.钆塞酸二钠增强 MRI 肝细胞摄取定量参数评估肝功能[J].放射学实践,2021,36(4):514-519.
- [5] Yang M,Zhang Y,Zhao W,et al.Evaluation of liver function using liver parenchyma,spleen and portal vein signal intensities during the hepatobiliary phase in Gd-EOB-D TPA-enhanced MRI [J].BMC Med Imaging,2020,20(1):119-128.
- [6] 徐向东,罗诗维,韦瑞丽,等.CT 影像组学预测肾透明细胞癌核分级:扫描期相 ROI 勾画策略[J].放射学实践,2022,37(12):1542-1547.
- [7] 周玮,胡红杰,沈博,等.基于钆塞酸二钠增强磁共振成像影像组学定量评估肝硬化患者肝脏储备功能的应用价值[J].中国医学科学院学报,2020,42(4):459-467.
- [8] 张智星,黄忠江,何生,等.基于增强 CT 影像组学评估肝硬化患者肝储备功能的应用[J].放射学实践,2022,37(6):676-682.
- [9] 陆力坚,黄仲奎,龙莉玲.磁共振门静脉血流动力学改变与肝功能相关性[J].放射学实践,2010,25(10):1125-1127.
- [10] Oliveira MS,Silva RPM,Valle SDCND,et al.Chronic hepatitis B and D:prognosis according to Child-Pugh score[J].Rev Bras Enferm,2017,70(5):1048-1053.
- [11] Zhang L,Zhang J,Shen P,et al.Block level skip connections across cascaded V-net for multi-organ segmentation [J].IEEE Trans Med Imaging,2020,39(9):2782-2793.
- [12] Takamoto T,Nara S,Ban D,et al.Application of albumin-bilirubin grade and platelet count to indocyanine green-based criteria for hepatectomy:predicting impaired liver function and postoperative outcomes of hepatocellular carcinoma[J].J Surg Oncol,2022,126(4):680-688.
- [13] Amygdalos I,Czigany Z,Bednarsch J,et al.Low postoperative platelet counts are associated with major morbidity and inferior survival in adult recipients of orthotopic liver transplantation[J].J Gastrointest Surg,2020,24(9):1996-2007.
- [14] Kaiser T,Kinny-Köster B,Bartels M,et al.Cholesterol esterification in plasma as a biomarker for liver function and prediction of mortality[J].BMC Gastroenterol,2017,17(1):57-63.
- [15] Gao F,Cai MX,Lin MT,et al.Prognostic value of international normalized ratio to albumin ratio among critically ill patients with cirrhosis[J].Eur J Gastroenterol Hepatol,2019,31(7):824-831.
- [16] Jajodia A,Gupta A,Prosch H,et al.Combination of radiomics and machine learning with diffusion-weighted MR imaging for clinical outcome prognostication in cervical cancer [J].Tomography,2021,7(3):344-357.
- [17] Mayerhoefer ME,Materka A,Langs G,et al.Introduction to radiomics[J].J Nucl Med,2020,61(4):488-495.
- [18] 吴娇艳,刘影.探讨肝脏特异性对比剂评价肝功能的价值[J].临床放射学杂志,2018,37(6):962-966.
- [19] Roehlen N,Crouch E,Baumert TF.Liver fibrosis:mechanistic concepts and therapeutic perspectives[J].Cells,2020,9(4):875-917.
- [20] Tamada T,Ito K,Higaki A,et al.Gd-EOB-DTPA-enhanced MR imaging:evaluation of hepatic enhancement effects in normal and

- cirrhotic livers[J].Eur J Radiol,2011,80(3):e311-316.
- [21] Zhu WS,Shi SY,Yang ZH,et al.Radiomics model based on pre-operative gadoxetic acid-enhanced MRI for predicting liver failure[J].World J Gastroenterol,2020,26(11):1208-1220.
- [22] Cai W,He B,Hu M,et al.A radiomics-based nomogram for the preoperative prediction of posthepatectomy liver failure in patients with hepatocellular carcinoma[J].Surg Oncol,2019,28(3):78-85.
- [23] Lu X,Zhou H,Wang K,et al.Comparing radiomics models with different inputs for accurate diagnosis of significant fibrosis in chronic liver disease[J].Eur Radiol,2021,31(11):8743-8754.
- [24] Zheng W,Guo W,Xiong M,et al.Clinic-radiological features and radiomics signatures based on Gd-BOPTA-enhanced MRI for predicting advanced liver fibrosis[J].Eur Radiol,2023,33(1):633-644.
- [25] Nitsch J,Sack J,Halle MW,et al.MRI-based radiomic feature analysis of end-stage liver disease for severity stratification[J].Int J Comput Assist Radiol Surg,2021,16(3):457-466.
- (收稿日期:2023-01-07 修回日期:2023-03-20)

《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。



关注有惊喜!