

论 著

基于增强CT影像组学联合临床特征预测肝细胞癌病理分化程度的应用研究

吕 娜 马春雨 朱 林
郭 飞*

蚌埠医学院第一附属医院放射科
(安徽 蚌埠 233000)

【摘要】目的 探讨增强CT影像组学模型联合临床特征对术前预测肝细胞癌(HCC)的病理分化程度的应用价值。**方法** 回顾性分析经手术切除或病理穿刺证实为HCC的196例患者的病理学及术前增强CT影像学资料。按照WHO标准将患者分为高分化组及中-低分化组,按照7:3比例将患者随机分为训练组(137例)和验证组(59例),保存增强CT的动脉期(AP)、静脉期(VP)及延迟期(DP)影像学图像,在医准-达尔文科研平台中提取并筛选各期图像的影像组学特征,应用“最大值归一化法、最优特征筛选、最小绝对收缩和选择算子(LASSO)回归”进行降维、筛选,构建动脉期+静脉期、动脉期+延迟期、静脉期+延迟期及三期联合的影像组学模型。应用单因素及多因素分析方法从临床资料中筛选具有统计学差异的危险因素并建立临床模型。用逻辑回归(Logistic)方法分析影像组学模型、临床模型及影像组学联合临床模型,使用受试者工作特征曲线(ROC)评估各模型预测肝癌病理分化程度的效能,并计算曲线下面积(AUC)、准确度、特异度及敏感度等。**结果** 1.在影像组学模型中,三期联合模型预测肝癌病理分化程度效能最佳,训练组及验证组曲线下面积(AUC)分别为0.877及0.801。2.单因素及多因素分析结果显示AFP($P=0.010$)及ALT($P=0.024$)最终为预测HCC病理分化程度的危险因素。临床特征AFP及ALT构建临床模型,训练组及验证组曲线下面积(AUC)分别为0.695及0.816。3.联合模型的预测效能优于各期影像组学模型及临床模型。训练组及验证组曲线下面积(AUC)分别为0.899及0.890。**结论** 基于增强CT影像组学模型联合临床模型能够准确地预测肝细胞癌(HCC)的病理分化程度。

【关键词】 肝细胞癌; 影像组学; 病理分化程度; 体层摄影术; X线计算机

【中图分类号】 R735.7

【文献标识码】 A

DOI:10.3969/j.issn.1672-5131.2024.03.039

Application Study on Predicting the Degree of Pathological Differentiation of Hepatocellular Carcinoma Based on Enhanced CT Imaging Histology Combined with Clinical Features

LV Na, MA Chun-yu, ZHU Lin, GUO Fei*.

Department of Radiology, the First Affiliated Hospital of Bengbu Medical College, Bengbu 233000, Anhui Province, China

ABSTRACT

Objective To investigate the application value of enhanced CT imaging histological model combined with clinical features for preoperative prediction of pathological differentiation of hepatocellular carcinoma (HCC). **Methods** Pathological and preoperative enhanced CT imaging data of 196 patients with HCC confirmed by surgical resection or pathologic puncture were retrospectively analyzed. The patients were divided into high differentiation group and middle-low differentiation group according to WHO standards, and randomly divided into training group (137 cases) and validation group (59 cases) according to the ratio of 7:3. The arterial-phase (AP), venous-phase (VP), and delayed-phase (DP) imaging images of the enhanced CT were preserved, and the imaging histological features of the images of the various phases were extracted and screened in the MediAll-In-Darwin scientific research platform, and the "maximum normalization method, optimal feature screening, minimum absolute contraction, and delayed-phase imaging" were applied. "We extracted and screened the imaging features of each phase in the Medical-Quasi-Darwin Research Platform, and applied the maximum normalization method, optimal feature screening, minimum absolute shrinkage and selection operator (LASSO) regression for dimensionality reduction and screening, and constructed the imaging histological models of arterial + venous, arterial + delayed, venous + delayed, and the three-phase combination. Single-factor and multifactor analysis methods were applied to screen statistically different risk factors from clinical data and establish clinical models. Imaging histology models, clinical models and combined imaging histology clinical models were analyzed by logistic methods, and the efficacy of each model in predicting the degree of pathological differentiation of hepatocellular carcinoma was evaluated by using the subject operating characteristic curve (ROC), and the area under the curve (AUC), accuracy, specificity and sensitivity were calculated. **Results** 1. Among the imaging models, the three-phase combined model had the best efficacy in predicting the degree of pathological differentiation of hepatocellular carcinoma (HCC), with the area under the curve (AUC) of 0.877 and 0.801 in the training and validation groups, respectively. 2. The results of the unifactorial and multifactorial analyses showed that AFP ($P=0.010$) and ALT ($P=0.024$) were ultimately the risk factors for the prediction of the degree of pathologic differentiation of HCC. Clinical models were constructed with clinical features AFP and ALT, and the area under the curve (AUC) was 0.695 and 0.816 in the training and validation groups, respectively. 3. The predictive efficacy of the combined model was better than that of the imaging histology model and the clinical model in all phases. The area under the curve (AUC) was 0.899 and 0.890 in the training and validation groups, respectively. **Conclusion** Enhanced CT image-based histologic models combined with clinical models can accurately predict the degree of pathologic differentiation in hepatocellular carcinoma (HCC).

Keywords: Hepatocellular Carcinoma; Radiomics; Pathological Differentiation Degree; Tomography; X-ray Computed

肝癌是全球范围内较为高发的恶性肿瘤,在全球恶性肿瘤发病率中位居第5位,死亡率位居第2位^[1]。在我国最为多见的病理类型是肝细胞癌,占我国肝肿瘤的85%-90%以上^[2]。预后决定了患者的生存期限,其中病理分化程度被认为影响肝癌术后复发率及生存预后的最重要因素。根据目前常用的WHO分类方法和Edmondson-Steiner分级标准将肝癌病理分化程度分为高分化、中分化、低分化、未分化^[3]。肿瘤组织病理分化程度和异型性的程度相关联,肝癌病理分化程度高,表示和正常肝脏组织结构差异性较小,肿瘤细胞异型性不显著,恶性程度较低;而肝癌病理分化程度低,表示和正常肝脏组织结构差异性大,肿瘤细胞异型性显著,恶性程度较高^[4]。临床医生在手术切除前明确肿瘤的病理组织学类型对于后续的治疗至关重要。多个研究结果显示肝癌的影像学表现与病理分化程度之间具有一定相关性。在2012年,荷兰Philippe Lambin教授提出影像组学这一概念,指从医学影像数据中大量提取描述肿瘤特性的组学特征,并对这些特征进行整合与分析,筛选出最有利于疾病诊断的特征,解析影像学、病理学及临床数据的内在联系,指导临床治疗和评估预后^[5]。动态增强CT是诊断肝癌及评估疗效的主要影像学方法,它能够清晰地显示出病灶的数量范围、形态特征、密度差异及血流供应等情况。

本研究使用增强CT影像组学方法预测肝癌病理分化程度,评估其统计结果的准确性及临床应用价值。

【第一作者】 吕 娜,女,住院医师,主要研究方向:影像诊断学。E-mail: 675297942@qq.com

【通讯作者】 郭 飞,男,副主任医师,主要研究方向:影像诊断学。E-mail: gf233003@126.com

1 资料与方法

1.1 一般资料 回顾性分析2018年6月至2022年1月在蚌埠医学院第一附属医院经手术切除或病理穿刺证实为HCC的196例患者。将样本按照7：3比例随机分为训练组(137例)和验证组(59例)。查阅并整理患者的临床信息，包括年龄、性别、有无肝炎病史、病灶最大直径及实验室相关检查结果包括甲胎蛋白(AFP)、天门冬氨酸转移酶(AST)、r-谷氨酰转移酶(GGT)、丙氨酸氨基转移酶(ALT)、乙肝表面抗原(HBsAg)、总胆红素(TBIL)。

纳入标准：经病理穿刺活检或手术切除证实为肝细胞癌，并且有明确的WHO分型或Edmondson-Steiner病理学分级；具有术前30天内的肝脏三期增强CT扫描影像；术前未经过其他HCC相关治疗。临床数据完整。排除标准：影像伪影较多，影响感兴趣区的勾画；临床数据缺失。术后病理结果未获得具体的肝癌病理分化程度。

1.2 仪器与方法 使用256排GE Revolution CT扫描仪进行检查，采用非离子型造影剂-碘佛醇(350mgI/mL)，用双筒高压注射器将碘佛醇推注至患者肘正中静脉，使用的剂量为成人1.5mL/kg，流速为4-5mL/s，剂量为50-60mL，对比剂注射后于30-35s(动脉期)、60-75s(门脉期)、100-120s(延迟期)进行CT扫描，使用数字成像和通信标准(DICOM)文件格式将扫描获得的三期CT增强图像存入图像存储和传输系统(PACS)中。

1.3 图像后处理 将符合上述研究标准的影像学图像从PACS系统中导出并保存至医准-达尔文智能科研平台中。由一名副主任医师在动脉期、静脉期及延迟期图像中手动沿着病灶的最大直径层面进行感兴趣区(region of interest, ROI)勾画，见图1A-1C。

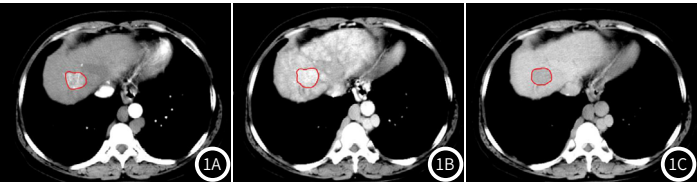


图1A-图1C 患者，男性，48岁，病理确诊为低分化肝癌，图1A、1B、1C分别为动脉期、静脉期、延迟期图像中肿瘤病灶区域的手动勾画，红色区域内为感兴趣区ROI。

1.4 特征筛选与模型构建 首先采取最大最小值归一化法对提取的影像组学特征进行预处理，把原始数据用线性化方法把区间拉伸至(0, 1)之间。再用特征选择组件-最优特征筛选器(百分比)，根据经验将K设置为15%，利用方差分析F检验(f_classif)统计量来选择按F值排序的前15%特征。最后使用L1范数正则化-最小绝对收缩和选择运算符(least absolute shrinkage and selection operator, LASSO)，把选择后的组学特征放入模型中进行拟合。采用逻辑回归模型分析降维筛选后的影像组学特征，建立动脉期+静脉期(AP+VP)模型、动脉期+延迟期(AP+DP)模型、静脉期+延迟期(VP+DP)模型及三期联合(AP+VP+DP)影像组学模型，再用筛选后的临床危险因素构建临床特征模型。最后将影像组学模型与临床模型相结合构建联合模型。

1.5 统计学方法 选用IBM SPSS Statistics 20.0统计软件对临床数据和影像组学特征进行分析。采用Kolmogorov-Smirnov检验分析定量资料是否服从正态分布，对于符合正态分布的定量资料，两组之间比较采用独立样本t检验，用均数±标准差表示；而对于不符合正态分布的定量资料，两组之间比较采用Mann-Whitney U检验，用中位数M(P₂₅,P₇₅)表示。采用皮尔逊卡方检验对定性资料进行分析，用频数(百分比)表示。采用逻辑回归分析绘制影像组学模型、临床特征模型及联合模型的训练组和验证组的受试者工作特征曲线(ROC)，计算出ROC曲线下面积(AUC)、敏感度、特异度、约登指数及95%CI (P<0.05具有统计学意义)，评价各模型预测肝癌病理分化程度的效能。

2 结果

2.1 影像组学模型的构建 筛选影像组学特征：共提取2817个影像组学特征。使用Select K Best(score_func=f_classif, k=15)

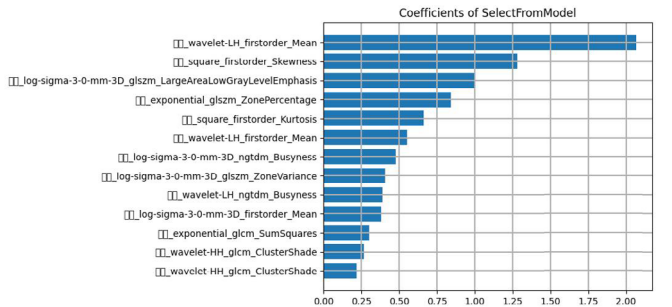
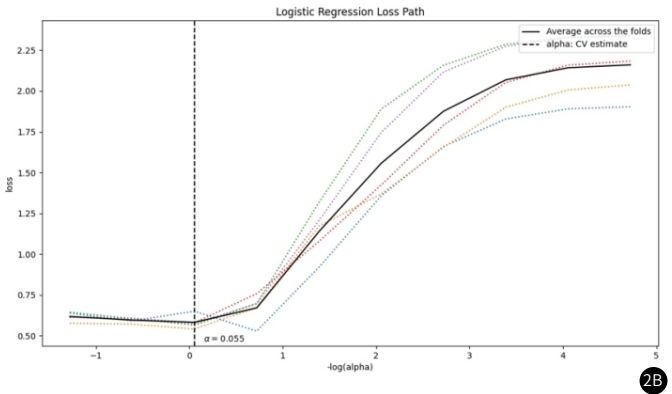
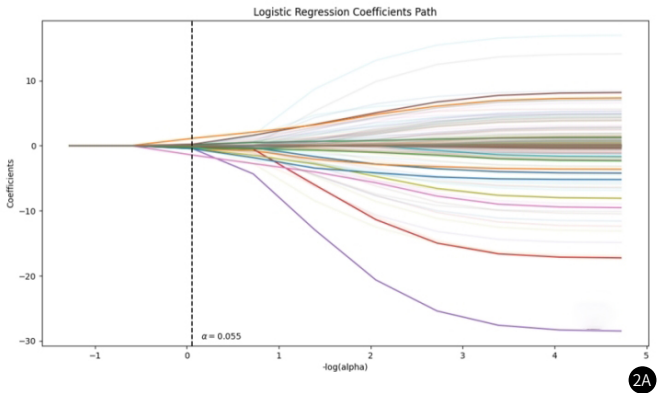
算法对2817个影像组学特征进行筛选，共得出 282个组学特征，再使用L1正则化算法降维得到13个最有代表性的组学特征，见图2A-2C。本研究建立AP+VP模型、AP+DP模型、VP+DP模型及AP+VP+DP模型。结果：AP+VP+DP模型预测肝癌的病理分化程度效能最佳。对AP+VP+DP模型进行训练与交叉验证后，得到其训练组及测试组的AUC值、特异度、灵敏度、95%CI及约登指数，见表1-表2、图3A-图3B。

表1 在训练组中各模型预测HCC病理分化程度的效能

模型	敏感度(%)	特异度(%)	约登指数	AUC值	95%CI	P值
AP+VP	80.15	78.01	0.5816	0.807	0.715-0.897	<0.01
AP+DP	82.37	85.65	0.6802	0.851	0.762-0.933	<0.01
VP+DP	81.74	83.27	0.6501	0.834	0.746-0.921	<0.01
AP+VP+DP	85.42	93.83	0.7925	0.877	0.809-0.945	<0.01

表2 在验证组中各模型预测HCC病理分化程度的效能

模型	敏感度(%)	特异度(%)	约登指数	AUC值	95%CI	P值
AP+VP	80.42	61.57	0.4199	0.754	0.587-0.921	0.05
AP+DP	73.93	69.26	0.4319	0.721	0.546-0.896	0.16
VP+DP	81.87	64.66	0.4653	0.709	0.555-0.863	0.22
AP+VP+DP	82.64	76.91	0.5955	0.801	0.639-0.963	0.01



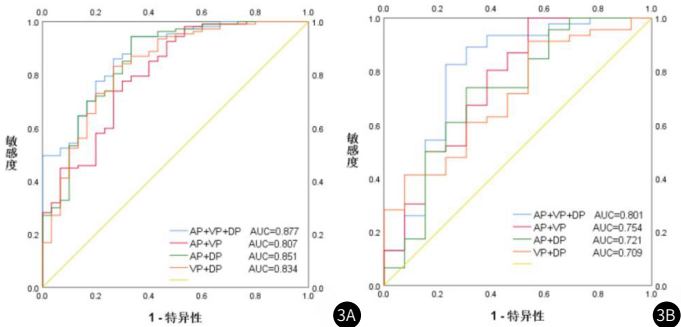


图2A-图2C LASSO-Logistic回归模型筛选出13个影像学特征

2.2 临床模型的构建 本研究将高、高-中分化并为高分化组，中、中-低、低分化并为中-低分化组。统计结果表明各临床特征在训练组及验证组间对比均无统计学意义($P>0.05$)，见表3。单因素分析结果显示，ALT($P=0.004$)、AST($P=0.001$)及AFP($P=0.008$)在高分化组和中-低分化组之间差异具有统计学意义，它们为预测HCC病理分化程度的临床危险因素，见表4。最后采用二元逻辑回归方法将单因素分析结果中 $P<0.05$ 的变量，即ALT($P=0.001$)、AST($P=0.001$)及AFP($P=0.004$)进行多因素分析，结果显示AFP($P=0.010$)及ALT($P=0.024$)为影响肝癌病理分化程度的临床危险因素，见表5。

表3 训练组及验证组中病例的临床特征

	训练组(n=137)	验证组(n=59)	P值
年龄	58.0(51.5,67.0)	56.0(49.0,66.0)	0.321
性别	男 111(81.0)	49(83.1)	0.737
	女 26(19.0)	10(16.9)	
ALT	40.0(20.5,60.0)	50.0(32.0,63.0)	0.114
AST	43.0(27.50,77.50)	46.0(29.0,68.0)	0.371
GGT	54.0(26.0,102.0)	66.0(33.0,139.0)	0.271
TBIL	14.5(10.45,20.45)	15.3(10.4,24.9)	0.346
AFP	≥ 20 85(62.0)	38(64.4)	0.754
	< 20 52(38.0)	21(35.6)	
病灶最大直径	4.2(2.8,6.4)	5.0(3.0,7.7)	0.336
肝炎病史	有 108(78.8)	52(81.6)	0.680
	无 29(21.2)	7(18.4)	

表4 预测 HCC 病理分化程度的单因素分析

	高分化	中-低分化	P
年龄	61.0(51.0,68.0)	57.0(50.5,65.5)	0.213
性别	男 33(76.7)	127(83.0)	0.349
	女 10(23.3)	26(17.0)	
ALT	30.0(16.0,44.0)	48.0(30.0,70.0)	0.001
AST	32.0(26.0,45.0)	49.0(31.50,90.5)	0.001
GGT	49.0(25.0,85.0)	61.0(27.5,117.5)	0.256
TBIL	15.4(11.0,20.4)	14.50(10.35,21.25)	0.561
AFP	≥ 20 19(44.2)	104(68.0)	0.004
	< 20 24(55.8)	49(32.0)	
病灶最大直径	4.6(3.0,6.5)	4.3(2.8,7.15)	0.700
肝炎病史	有 32(74.4)	128(83.7)	0.207
	无 11(25.6)	25(16.3)	

表5 预测HCC病理分化程度的多因素分析

变量及常数	B	标准误差	Wald	P	OR	OR的95%CI	
						下限	上限
ALT	0.024	0.011	5.125	0.024	1.025	1.003	1.047
AST	0.002	0.007	0.058	0.809	1.002	0.989	1.015
AFP	0.967	0.375	6.662	0.010	2.631	1.262	5.483
变量	-0.411	0.400	1.055	0.304	0.663	-	-

2.3 三种模型的诊断效能 单因素和多因素分析结果显示AFP及ALT为预测肝癌的病理分化程度的临床危险因素，使用Logistic回归分析构建临床特征模型(APF+ALT)。从上述结论得出最佳影像学模型为AP+VP+DP。再将多因素分析结果中的AFP+ALT与AP+VP+DP组合成联合模型。利用受试者工作曲线(ROC)评估它们对肝细胞癌病理分化程度的预测效能，见表6。结果表明在训练组及验证组中，联合模型的AUC均大于其他模型，见图3C-图3D。

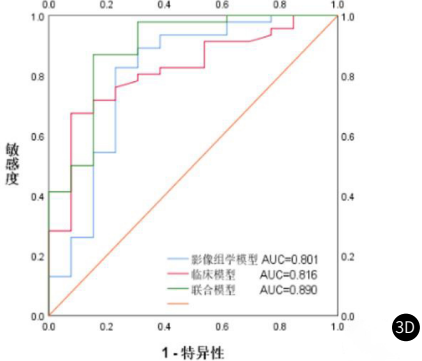
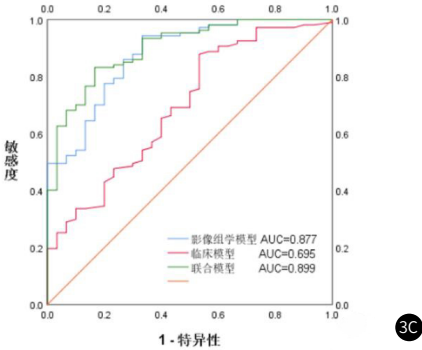


表6 影像学模型、临床模型和联合模型分别预测训练组与验证组HCC病理分化程度的ROC曲线结果

模型	训练组(n=137)						验证组(n=59)					
	约登指数	特异度(%)	敏感度(%)	95%CI	AUC值	P值	约登指数	特异度(%)	敏感度(%)	95%CI	AUC值	P值
影像学模型	0.611	66.70	94.40	0.809-0.945	0.877	<0.01	0.595	76.90	82.60	0.639-0.963	0.801	<0.01
临床模型	0.346	46.70	87.90	0.588-0.802	0.695	<0.01	0.520	84.60	67.40	0.693-0.939	0.816	<0.01
联合模型	0.665	83.30	83.20	0.841-0.958	0.899	<0.01	0.716	84.60	87.00	0.779-0.981	0.890	<0.01

3 讨 论

我国是世界肝癌的高发区域，其发病率和死亡率逐年递增。所以肝癌的早期筛查及诊断至关重要。肝癌的病理分化程度作为患者预后的重要因素，病理的分化程度越低，意味着病灶更易侵犯邻近的正常肝组织，所以低分化肝癌患者术后复发率高，预后差^[6]。近几年随着技术的不断发展，多种医学影像检查可用于肝癌的诊断及疗效评估。但肿瘤的发生及发展是复杂多变的，一般常规影像学检查无法探勘出肿瘤病理分化程度方面的深层次信息。大量研究证明^[7-8]，肿瘤常表现出异质性(heterogeneity of tumors)，同样的恶性肿瘤处在同一病人体内的不同位置或是在不同病人个体间，以及不同生长阶段的肿瘤的免疫特性、生长方式及侵袭程度都有所不同。影像组学是一种新兴技术，它能够使用算法或统计分析工具将医学图像转换为具有高预测性和高代表性的影像数据，寻找出影像图像信息和肿瘤细胞表型特征这之间的潜在关系^[9]，从分子水平上反映人体内解剖组织结构和基因代谢水平的动态变化。影像组学在食管癌^[10]、直肠癌^[11]、宫颈癌^[12]、肺癌^[13]等肿瘤诊断、组织分子分型、预测肿瘤基因表型及生存期评估等方面已经显示出重要的临床价值。多项研究证明在肝癌的诊断、评估病理分化程度及术后预后等方面，影像组学具有较好的应用价值^[14-16]。

与既往研究不同，本研究使用医准一达尔文智能平台提取组学特征，选用LASSO-L1正则化回归分析筛选出最有价值的组学数据，建立AP+VP模型、AP+DP模型、VP+DP模型及AP+VP+DP的影像组学模型，其中AP+VP+DP模型训练组AUC为0.877，验证组AUC为0.801；AP+VP+DP的预测效能均优于其他影像组学模型。经单因素及多因素分析结果显示，肝癌的病理分化程度与AFP和ALT具有一定的关联性(P<0.05)，其中AFP与肝癌的相关性已经被Bai^[17]、曾^[18]等研究证明。但与ALT的相关性还有待考究，这可能与收集的病例较少有关，后续将进一步扩大病例数进行研究。最后把临床特征加入影像组学模型，构建联合模型进行效能评估。联合模型训练组AUC为0.899，验证组AUC为0.890。结果显示：影像组学通过联合临床特征，预测效能更高。

本研究存在的一些局限性：(1)本研究是回顾性研究，收集的病例有限，结果易受偏移的影响，后续将继续收集病例来进行研究。(2)本研究肝癌的病理分化程度仅分成两组，这样的分组方式可能会影响预测效能，后续将进一步细化分组。(3)影像组学研究需要庞大的数据支撑，不同机器的图像采集方式和不同的扫描参数可能对提取的特征产生一些影响，目前尚未建立统一的标准^[19]。

综上所述，基于增强CT影像组学模型联合临床模型有助于术前预测肝细胞癌的病理分化程度，为肝癌的临床诊断及制定诊疗方案提供一定的参考价值。

参考文献

[1]Cotrim HP,Oliveira CP,Coelho HS,et al.Nonalcoholic steatohepatitis and hepatocellular carcinoma: Brazilian survey[J].Clinics (Sao Paulo, Brazil), 2016, 71 (5): 281-284.
[2]中华人民共和国国家卫生健康委员会医政医管局.原发性肝癌诊疗规范(2019年版)[J].临床肝胆病杂志, 2020, 36 (2): 277-292.

[3]Granata V,Fusco R,Filice S,et al.The current role and future perspectives of functional parameters by diffusion weighted imaging in the assessment of histologic grade of HCC[J].Infect Agent Cancer.2018 Jul 3;13: 23.
[4]David Liu,Shaohua Kevin Zhou.Histological differentiation grade prediction of hepatocellular carcinoma in computed tomography images [P].US9585627,2017-03-07.
[5]Davenport MS,Khalatbari S,Liu PS,et al.Repeatability of diagnostic features and scoring systems for hepatocellular carcinoma by using MR imaging[J].Radiology, 2014, 272 (1): 132-142.
[5] Welzel TM,Graubard BI,Quraishi S,et al.Population-attributable fractions of risk factors for hepatocellular carcinoma in the United States [J].Cancer, 2016, 122 (11): 1757-1765.
[7]Tellez-Gabriel M,Heymann MF,Heymann D.Circulating tumor cells as a tool for assessing tumor heterogeneity[J].Theranostics, 2019, 9 (16): 4580-4594.
[8]Patel AP,Tirosh I,Trombetta JJ,et al.Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma [J].Science, 2014, 344 (6190): 1396-1401.
[9]Parmar C,Grossmann P,Bussink J,et al.Machine learning methods for quantitative radiomic biomarkers[J].Sci Rep, 2015, 5: 13087.
[10]孔洁,祝淑钗,刘志坤,等.基于CT平扫影像组学模型预测食管癌淋巴结转移[J].中国肿瘤临床, 2022, 49 (19): 1333-1337.
[11]周彦汝,张岚,韩鼎盛,等.基于MRI影像组学构建新辅助化疗后局部进展期直肠癌的预测模型[J].中国医学影像学杂志, 2022, 30 (9): 881-888.
[12]王海波,崔薇,杨玮丽.基于多序列MRI影像组学预测早期宫颈癌淋巴血管侵犯的研究[J].中华全科医学, 2021, 19 (12): 2088-2092.
[13]陈巧玲,冯峰,李曼曼,等.联合含瘤周肿瘤全体积CT影像组学特征及临床相关独立预测因子列线图预测肺腺癌淋巴管侵犯[J].中国医学影像技术, 2022, 38 (8): 1182-1186.
[14]许露露,舒健,杨春梅.MRI影像组学评估肝细胞肝癌的病理分级[J].临床放射学杂志, 2020, 39 (8): 1520-1523.
[15]张加辉,徐丽萍,赵志新,等.基于MR影像组学及临床特征预测混合型肝癌和肿块型肝内胆管细胞癌术后生存[J].临床放射学杂志, 2022, 41 (2): 279-283.
[16]王晴,盛晔,刘海峰,等.基于术前增强MRI影像组学分析的列线图模型预测肝细胞癌切除术后复发风险的价值[J].磁共振成像, 2022, 13 (12): 93-99.
[17]Bai DS,Zhang C,Chen P,et al.The prognostic correlation of AFP level at diagnosis with pathological grade,progression,and survival of patients with hepatocellular carcinoma [J].Sci Rep, 2017, 7 (1): 12870.
[18]曾灏,翟嫒,张林芝.血清AFP、AFP-L3%、CA199、PIVKA-II联合GP73检测对肝细胞癌的诊断价值[J]. 2023, 13 (1): 149-152.
[19]Yang L,Gu D,Wei J,et al.A radiomics nomogram for preoperative prediction of microvascular invasion in hepatocellular carcinoma [J].Liver Cancer, 2019, 8 (5): 373-386.

(收稿日期: 2023-01-02)
(校对编辑: 姚丽娜)