

机器学习优化能谱 CT 预测胃腺癌的浸润性

葛慧婷, 王莉莉, 刘元芬, 邹添秀, 林维文, 薛蕴菁

摘要: **目的** 探讨机器学习(ML)算法结合能谱 CT 定量参数和临床模型在预测胃腺癌(GAC)患者淋巴血管浸润(LVI)和神经周围浸润(PNI)的潜在价值。**方法** 收集 2017 年 12 月—2022 年 5 月经病理证实的 GAC 患者 114 例。研究参数涉及血清肿瘤标志物、CT-TN 分期、CT 评估壁外血管浸润(CT-EMVI)以及能谱 CT 定量参数。通过 WEKA 软件的 Best-First 算法进行特征筛选,并运用贝叶斯网络(BN)及支持向量机(SVM)算法建立模型。**结果** 相较于 LVI/PNI 阴性组,LVI/PNI 阳性组中 CT-T₃₋₄ 期、CT-N 阳性、CT-EMVI 阳性、血清肿瘤标志物[糖类抗原(CA)72-4 和 CA19-9]更为常见,能谱 CT 参数也更高,差别均有统计学意义($P<0.05$)。经特征选择,关键变量包括 CT-T 分期、CT-EMVI、VP-NIC 和 EP-70 keV CT 值。基于这些变量,分别使用 BN 和 SVM 构建临床参数模型、能谱 CT 参数模型和混合模型,共 6 个模型。6 个模型均展现了高预测性能,无过拟合现象。BN 的混合模型预测性能最佳,AUC 值为 0.890~0.933,Delong 检验显示其在统计学上具有显著优势($P<0.05$);而 SVM 的混合模型与另外 2 种模型间的差别无统计学意义($P>0.05$)。**结论** 结合临床和能谱 CT 参数的 ML 模型能够高效能评估 GAC 患者的 LVI 和 PNI 状态,其中 BN 混合模型的准确性最高。

关键词: 胃腺癌;能谱 CT;淋巴血管浸润;神经周围浸润

文献标志码: A **文章编号:** 1672-4194(2024)03-0193-07

胃癌是全球范围内常见的癌症之一,也是导致癌症相关死亡的第三大原因^[1],而 95% 以上的胃癌为胃腺癌(gastric adenocarcinoma, GAC)。术前评估肿瘤的深入侵犯性特征,如淋巴血管浸润(lymphovascular invasion, LVI)和神经周围浸润(perineural invasion, PNI),有利于制定精细化的治疗策略。LVI 和 PNI 与胃癌的分期、浸润深度、淋巴结转移及远处转移紧密相关,被认为是预后不良的生物标志,并在多项研究中被证实为新辅助和辅助治疗的重要考量因素^[2-6]。目前,临床上主要通过术后组织病理学确认 LVI 和 PNI,大大限制了术前评估的可行性。能谱 CT 引入了多参数的定量分析,将 CT 成像的能力扩展到微观层面。本研究旨在通过机器学习(machine learning, ML)技术整合临床参数和能谱 CT 参数,并于术前预测 GAC 组织学的 LVI 和 PNI 状态,报道如下。

1 对象与方法

1.1 对象 收集 2017 年 12 月—2022 年 5 月收治的 GAC 患者 154 例。纳入标准:(1)经根治性手术切除且病理诊断为 GAC,具有完整的 LVI 和 PNI 病理信息;(2)术前 2 周内进行腹部能谱 CT 三期增强扫描;(3)CT 检查前及术前均未进行局部或全身

治疗。排除标准:(1)CT 图像上病灶不清晰;(2)胃充盈不良;(3)图像质量差。共排除 40 例,余 114 例纳入研究。114 例中,LVI 阳性 51 例,阴性 63 例;PNI 阳性 59 例,阴性 55 例。将 LVI 和/或 PNI 阳性的患者作为阳性组(LVI/PNI 阳性组, $n=69$),LVI 及 PNI 均阴性的患者作为阴性组(LVI/PNI 阴性组, $n=45$)。2 组患者的一般资料见表 1。本研究经福建医科大学附属协和医院伦理委员会批准(2022KJT016)。

1.2 方法

1.2.1 收集患者的临床资料 由 1 位有 5~10 a 工作经验的主治医师负责收集整理患者的临床资料,包括性别、年龄、cTNM 分期、肿瘤部位、血清肿瘤标志物、组织学类型、肿瘤分化程度、pTNM 分期、完整的 LVI 和 PNI 信息。血清肿瘤标志物包括癌胚抗原(carcinoembryonic antigen, CEA)、糖类抗原(carbohydrate antigen, CA)19-9 (CA19-9)、CA125、CA72-4 和甲胎蛋白(alpha-fetoprotein, AFP)。

1.2.2 CT 检查 检查前禁食 8~12 h,饮水 800~1 000 mL(无显影剂),使胃腔充盈。仰卧位,扫描范围包括中上腹。右臂肘正中静脉留置规格为 20 G 的留置针。所有患者均在 CT(Revolution 型,美国 GE 集团)上行能谱扫描。管电压:80、140 kV 高低电压瞬时切换;管电流:200 mA;探测器宽度:80 mm;螺距:0.992:1;转速:每转 0.6~0.8 s;矩阵:512×512。重建 70 keV 单能量图像,厚度为

收稿日期:2024-02-06

资助项目:福建省科技厅社会发展引导性(重点)项目(2022Y0025)

作者单位:福建医科大学 附属协和医院影像科,福州 350001

作者简介:葛慧婷,女,主治医师,医学硕士

通信作者:薛蕴菁. Email: xyj10022023@163.com

表 1 2 组患者的一般资料
Tab. 1 General data of two groups

一般资料	LVI/PNI 阴性组 (n=45)	LVI/PNI 阳性组 (n=69)
性别		
男	34	56
女	11	13
年龄/岁		
≥60	32	50
<60	13	19
c-T		
T ₁	9	1
T ₂	16	7
T ₃	12	15
T ₄	8	46
c-N		
N ₀	33	14
N _{1~3}	12	55
p-T		
T ₁	31	4
T ₂	8	4
T ₃	4	31
T ₄	2	30
p-N		
N ₀	33	16
N _{1~3}	12	53
部位		
胃上部	12	18
胃中部	7	23
胃下部	25	23
全胃、多发	1	5
分化程度		
低分化	11	26
中低分化	10	16
中分化	23	27
中高分化	1	0
高分化	0	0

LVI:淋巴血管浸润;PNI:神经周围浸润。

1.25 mm。采用高压注射器注入非离子型水溶性对比剂(典迈伦,400 mg/mL,上海博莱科信谊药业有限责任公司),剂量 1.5 mL/kg,注射速率 2.5~3.0 mL/s。分别于注射造影剂开始后 20~25、60、180 s 行动脉期(arterial phase,AP)、门脉期(portal venous phase,VP)及平衡期(equilibrium phase,EP)进行扫描。

1.2.3 图像分析 首先对 2 位有 5~10 a 工作经验的影像科主治医师进行统一培训和测试,分别盲法阅片,再由 1 位有 15 a 以上工作经验的放射科主任医师核实并确认。所有 TN 分期、EMVI 评分均在工作站(AW4.6,美国 GE 公司)多重重建平面进行。根据第八版美国癌症联合委员会(American Joint Committee on Cancer,AJCC)癌症分期手册^[7],在 CT 图像上评估胃壁浸润深度(CT-T)和淋

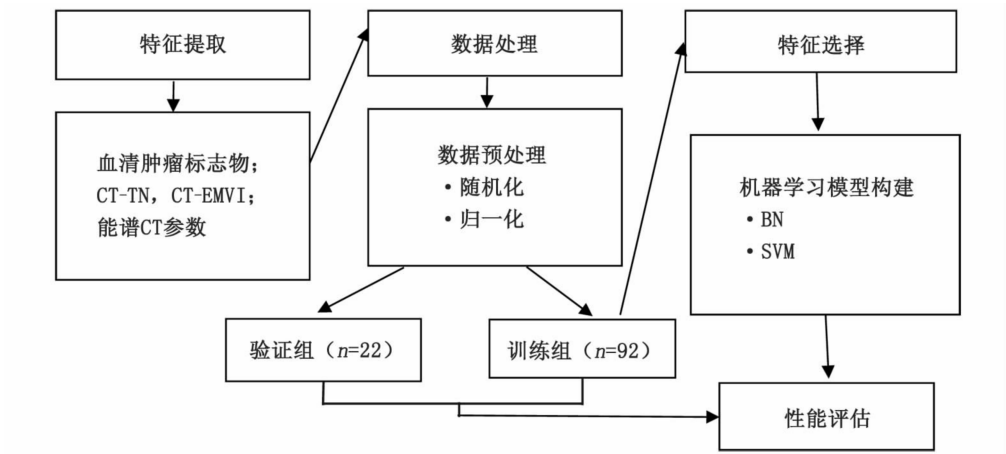
巴结状态(CT-N)。根据 YANG 等^[8]的报道,在 CT 图像上对胃癌壁外血管浸润(CT-detected extramural venous invasion,CT-EMVI)进行定义及评分:0 分,肿瘤未浸透胃壁,病灶旁无壁外血管;1 分,肿瘤浸透胃壁,病灶旁无壁外血管;2 分,病灶浸透胃壁,见迂曲血管与病灶相连,但血管腔内未见肿瘤密度影;3 分,肿块浸透胃壁,向壁外浸润,受累血管轻度迂曲扩张,内见肿瘤密度影;4 分,肿瘤浸透胃壁,壁外血管腔明显扩张,内见稍低密度充盈缺陷。0~2 分为 CT-EMVI 阴性,3~4 分为 CT-EMVI 阳性。

1.2.4 数据测量 在 GE AW4.6 工作站、GSI Viewer (gemstone spectral imaging viewer,GSI)软件上进行图像分析。由 2 位有 5~10 a 工作经验的影像科主治医师在肿瘤最大层面选取感兴趣区域(region of interest,ROI),再由 1 位有 15 a 以上工作经验的影像科主任医师核实并确认。ROI 约占病变面积的 2/3,避开坏死、钙化和血管区域。每个病灶测量 3 次,计算平均值。得到 AP、VP、EP 的能谱 CT 参数,包括 40~140 keV 的单能量 CT 值(间隔 10 keV,3 期共 33 个数据)、能谱曲线、碘浓度(iodine concentration,IC)和有效原子序数(effective-Z,Zeff)。计算能谱曲线斜率($K_{40\sim70}$)和标准化碘浓度比(the ratio of normative iodine concentration,NIC):

$$K_{40\sim70}=(CT_{40\text{ keV}}-CT_{70\text{ keV}})/30 \tag{1}$$

$$NIC=IC_{\text{病变区域}}/IC_{\text{同平面主动脉}} \tag{2}$$

1.2.5 个性化预测模型构建 按照 8:2 的比例,将 114 例 GAC 患者随机分为 2 组,即训练组 92 例(LVI/PNI 阳性 54 例、阴性 38 例),验证组 22 例(LVI/PNI 阳性 15 例、阴性 7 例)。对训练数据集执行特征降维处理,包括 Z 分数标准化、相关性特征选择(correlation-based feature selection,CFS)子集评估器和最佳优先(Best-First)启发式搜索方法;该组合在十倍交叉验证的框架下进行,剔除 49 个重复特征,最终选取 4 个被 Best-First 方法频繁选中且选择概率≥90%的特征,即 CT-T 分期、CT-EMVI、VP-NIC 和 EP-70 keV CT 值,构建基于贝叶斯网络(Bayesian network,BN)^[9]和支持向量机(support vector machine,SVM)的学习模型各 3 个,包括能谱 CT 参数模型(VP-NIC 和 EP-70 keV CT 值)、临床参数模型(CT-T 分期和 CT-EMVI)及混合模型。采用 WEKA 数据挖掘软件包(版本 3.8.3,新西兰怀卡托大学)及 R 软件(版本 3.4.3)进行预测模型的构建和绘图,并在验证组中进行评估。影像特征的筛选过程以及 ML 模型的构建流程见图 1。



CT-TN:CT 评估胃壁浸润深度和淋巴结状态;CT-EMVI:CT 评估壁外血管浸润;BN:贝叶斯网络;SVM:支持向量机。

图 1 影像特征筛选和机器学习模型构建工作流程图

Fig. 1 Workflow diagram for image feature screening and machine learning model construction

1.3 统计学处理 采用 SPSS 26.0 软件进行统计分析。计数资料采用 χ^2 检验进行分析;组间能谱 CT 定量参数采用 Kolmogorov-Smirnov 检验进行正态性检验,符合正态分布则采用独立样本 t 检验,非正态分布数据采用 Mann-Whitney U 检验。双侧检验, $P<0.05$ 为差别有统计学意义。

2 结果

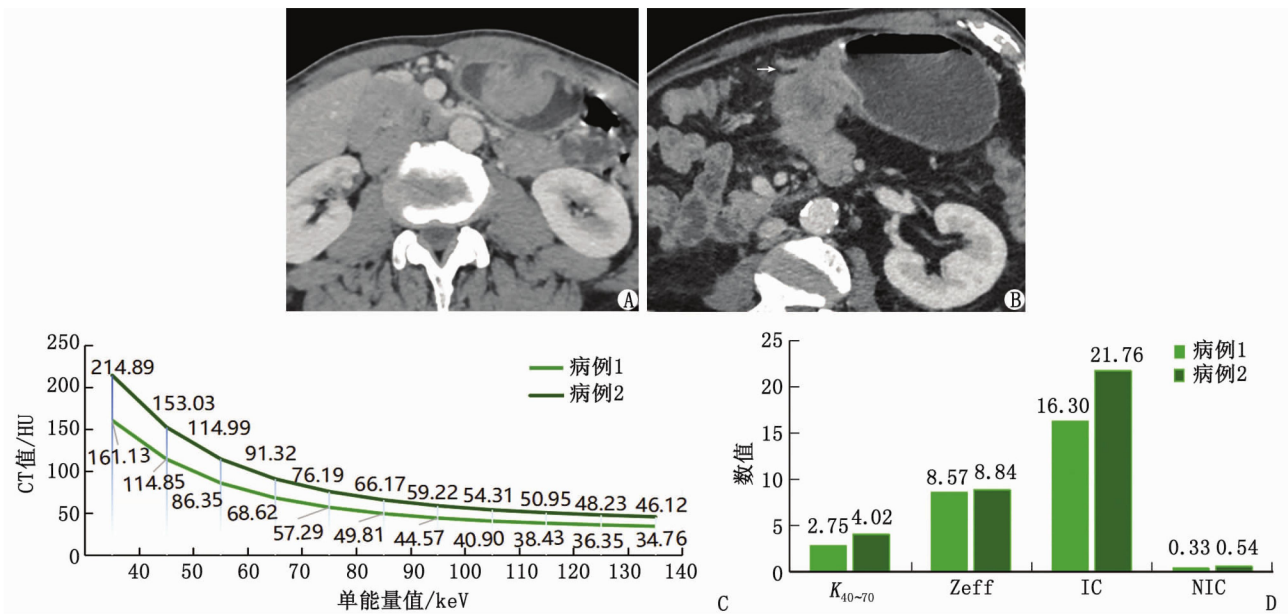
2.1 组间临床特征、能谱参数比较分析 与 LVI/PNI 阴性组比较,LVI/PNI 阳性组 CT- $T_{3\sim4}$ 期、CT-N 阳性、CT-EMVI 阳性更为常见($P<0.05$),CA72-4 和 CA19-9 阳性率更高($P<0.05$),AP、VP、EP 的 40~140 keV CT 值、 $K_{40\sim70}$ 、Zeff、IC、NIC 也更高($P<0.05$,表 2)。2 组间的 AFP、CEA 和 CA125 阳性率比较,差别均无统计学意义($P>0.05$)。

2.2 典型病例 病例 1,女性,55 岁,病理诊断为 GAC,LVI、PNI 均为阴性。VP-CT 横断位显示,胃癌病变浸润黏膜下层,浸润深度超过胃壁的 50%且外侧低密度区仍可见,胃壁外膜完整,局部周边未见血管影,评估为 CT- T_2N_0 ,CT-EMVI 0 分,CT-EMVI 阴性(图 2A)。病例 2,男性,47 岁,病理诊断为 GAC,LVI、PNI 均为阳性。VP-CT 横断位显示,胃癌病变浸润胃壁全层,病变周围脂肪间隙可见束状增厚血管影,血管内见低密度充盈缺损,CT 图像评估为 CT- $T_{4a}N_0$,CT-EMVI 4 分,为 CT-EMVI 阳性(图 2B)。能谱曲线及能谱参数直方图见图 2C 和图 2D。

表 2 2 组间临床参数、能谱 CT 参数比较分析
Tab. 2 Comparative analysis of clinical indicators and energy spectrum CT parameters between two groups

临床参数	LVI/PNI 阴性组 (n=45)	LVI/PNI 阳性组 (n=69)	统计值	P
CT-T			40.990	<0.001
T _{1~2}	35(77.8)	12(17.4)		
T _{3~4}	10(22.2)	57(82.6)		
CT-N			13.597	<0.001
N ₀	37(82.2)	33(47.8)		
N _{1~3}	8(17.8)	36(52.2)		
CT-EMVI			32.858	<0.001
阴性	39(86.7)	22(31.9)		
阳性	6(13.3)	47(68.1)		
CA72-4			10.292	0.001
阴性	44(97.8)	52(75.4)		
阳性	1(2.2)	17(24.6)		
AFP			—	1.000☆
阴性	44(97.8)	67(97.1)		
阳性	1(2.2)	2(2.9)		
CEA			1.224	0.269
阴性	40(88.9)	56(81.2)		
阳性	5(11.1)	13(18.8)		
CA19-9			10.292	0.026
阴性	43(95.6)	56(81.2)		
阳性	2(4.4)	13(18.8)		
CA125			—	0.647☆
阴性	44(97.8)	65(94.2)		
阳性	1(2.2)	4(5.8)		
能谱 CT 参数				
VP-NIC	0.36(0.33,0.49)	0.52(0.45,0.63)	-5.043	<0.001
EP-70 keV CT/HU	68.36±16.17	91.44±17.74	7.027	<0.001

表中数据除能谱 CT 参数,余为 $n(\%)$ 。LVI:淋巴血管浸润;PNI:神经周围浸润;CT-T:CT 图像上的 T 分期;CT-N:CT 图像上 N 分期;CT-EMVI:CT 图像上壁外静脉浸润;CA72-4:糖类抗原 72-4;AFP:甲胎蛋白;CEA:癌胚抗原;CA19-9:糖类抗原 19-9;CA125:糖类抗原 125;VP-NIC:静脉期标准化碘浓度比值;EP-70 keV CT:平衡期 70 keV 单能量图像上 CT 值。☆:采用 Fisher 确切概率法。



A: 病例 1; B: 病例 2, 箭头示血管内低密度充盈缺损; C: 能谱曲线, 病例 2 门脉期 40~140 keV 单能量 CT 值均高于病例 1; D: 能谱参数直方图, 病例 2 门脉期 $K_{40\sim70}$ 、Zeff、IC、NIC 均高于病例 1。 $K_{40\sim70}$: 能谱曲线斜率; Zeff: 有效原子序数; IC: 碘物质浓度, 单位为 $100\text{ }\mu\text{g}/\text{cm}^3$; NIC: 标准化碘浓度比。

图 2 2 组典型病例示例
Fig. 2 Examples of typical case in two sets

2.3 个体化预测模型 ROC 分析和 Delong 检验提示,所有模型在分类 LVI/PNI 状态上均显示出显著性能(AUC:0.786~0.933, $P<0.001$),并未出现过拟合现象($P=0.206\ 2\sim0.898\ 6$, $P>0.05$)。临床参数模型和能谱 CT 模型在 2 组分类器中均表现出相似的分类准确度(AUC:0.790~0.905 vs 0.786~0.839),差别无统计学意义($P>0.05$,图 3A、B)。然而,基于 BN 的混合模型展现出更高

的 AUC 值(0.890~0.933),与同网络的能谱 CT 模型、临床参数模型和基于 SVM 的混合模型比较,差别均有统计学意义($P=0.029\ 2$ 、 $0.017\ 5$ 、 $0.036\ 6$,图 3A、C)。而基于 SVM 的混合模型与同一分类器的能谱 CT 模型和临床参数模型比较,差别无统计学意义($P>0.05$,图 3B)。BN 和 SVM 分类器在原始训练队列和独立内部验证队列上的预测性能(AUC 和 Kappa 系数)见表 3 和图 3。

表 3 机器学习算法驱动模型预测 LVI/PNI 的性能比较

Tab. 3 Performance comparison of machine learning algorithm-driven models for predicting LVI/PNI status

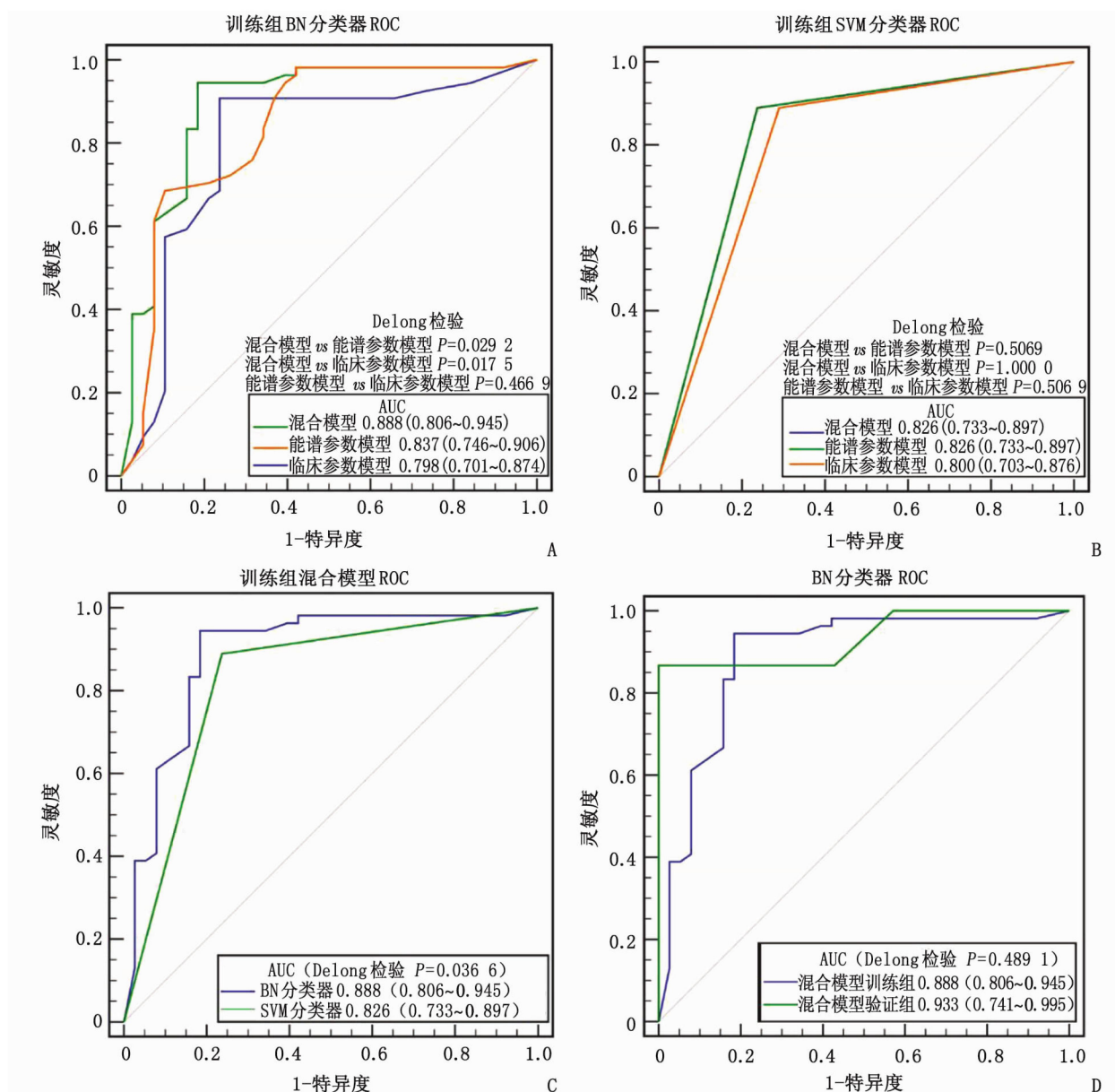
分 组	能谱 CT 参数模型		临床参数模型		混合模型	
	Kappa	AUC	Kappa	AUC	Kappa	AUC
训练组($n=92$)						
BN	0.441 7	0.839	0.659 8	0.798	0.772 3	0.890
SVM	0.611 3	0.800	0.659 8	0.826	0.659 8	0.826
验证组($n=22$)						
BN	0.505 6	0.810	0.455 4	0.905	0.581 0	0.933
SVM	0.645 2	0.786	0.581 0	0.790	0.581 0	0.790

LVI: 淋巴血管浸润; PNI: 神经周围浸润; BN: 贝叶斯网络; SVM: 支持向量机。

3 讨 论

肿瘤细胞可侵入淋巴管、血管和淋巴鞘、神经纤维,并向全身扩散,神经、血管和淋巴连接构成肿瘤微环境,在癌症的发展中起关键作用。本研究为术前预测 GAC 组织学的 LVI 和 PNI 状态,构建基于

BN 和 SVM 的学习模型来区分 LVI/PNI 状态。BN 的主要优势在于其能够在不确定的环境中应用概率推理处理临床实践中的因果关系,而 SVM 模型在处理非线性问题和高维数据集方面具有卓越性能。2 种方法均为临床参数分析中常用且有效的工具,可用于评估各类参数在实际临床场景中的应用



LVI:淋巴血管浸润;PNI:神经周围浸润;ROC:受试者工作特征;AUC:曲线下面积;BN:贝叶斯网络;SVM:支持向量机。A:基于BN分类器的3种模型;B:基于SVM分类器的3种模型;C:比较基于BN和SVM的混合模型的2条ROC曲线,差别有统计学意义($P<0.05$);D:基于BN的混合模型在训练集和验证集上的ROC曲线比较,差别无统计学意义($P>0.05$)。

图3 预测胃腺癌LVI/PNI的ROC曲线比较

Fig. 3 Comparison of ROC curves for predicting LVI/PNI status in gastric adenocarcinoma

价值。本研究构建的6个模型均展现出满意的预测性能,且在训练集和验证集中保持了一致性。

既往研究表明,临床T分期和N分期是预测LVI和PNI的重要因素^[9]。本研究进一步强调,CT-T_{3~4}期病例血管神经侵犯(LVI/PNI)的可能性较高,而CT-N分期的差别虽有统计学意义,但在特征选择阶段并未被纳入最终的ML模型。这可能是因为与T分期相比,术前CT对N分期的评估准确性较低^[10]。

CT-EMVI是另一个与LVI/PNI密切相关的独立风险因素。EMVI通常与胃周神经、血管、淋巴

管的浸润共存,是肿瘤沿神经血管束扩散的一种方式^[11-12]。本研究结果也证实了CT-EMVI与LVI/PNI状态之间存在关联,表明CT-EMVI阳性的患者LVI/PNI阳性的风险也更高。

在肿瘤的诊断和治疗中,血清肿瘤标志物可为临床提供关键信息^[13]。本研究显示,LVI/PNI阳性的GAC患者,其CA72-4和CA19-9的阳性率均高于阴性患者,但在ML模型的特征选择中,CA72-4和CA19-9并未入选最终模型的预测因子,原因是其统计学意义低于其他特征。

能谱CT通过单能量CT值、IC和NIC提供了

关于肿瘤血供状态的定量信息^[14-15]。LVI/PNI 阳性患者的微血管密度更高,血管生成更加活跃,故病变强化程度增高。与阴性组比较,LVI/PNI 阳性组具有更高的碘吸收率(IC、NIC)和单能量 CT 值、能谱曲线斜率、Zeff,表明 2 组患者在能谱衰减特性上存在差异。Zeff 可能与肿瘤内血管和神经的病理变化相关,如瘤栓形成。因此,阳性组的肿瘤质量和密度更高,且组织成分也更加复杂。

选择能谱 CT 参数作为 LVI/PNI 预测因子时,VP 和 EP 参数比 AP 参数更为敏感。这可能是因为 VP 和 EP 更能反映肿瘤内细胞间隙的血供状态^[16]。经特征选择,VP-NIC 和 EP-70 keV 单能量 CT 值可作为预测因子。可见这些阶段参数的重要性。

本研究表明,利用能谱 CT 预测 VP 和 EP 的定量参数,其效能与传统的临床模型相当,两者的诊断性能均可靠。结合临床风险因素与能谱 CT 参数的混合模型,在 BN 框架下表现出更高的诊断效能,表明 BN 混合模型在预测 LVI/PNI 状态方面具有潜在的临床应用优势。此外,当模型整合了临床与能谱 CT 参数后,无论是训练组还是验证组,预测性能均显得更加稳定,且组间差别无统计学意义,表明这些参数具有较高的灵敏度和特异度,对 LVI/PNI 的术前评估具有实际价值,可为临床医师提供调整治疗方案的依据。

该研究尚存在局限性:样本量较小且分组不均衡,可能影响结果的普适性;研究范围仅限于 GAC,未涉及其他组织学类型;模型特征筛选时未包含一般资料等传统的临床病理因素;由于研究基于单一中心,其结果可能受中心效应的影响,未来需通过多中心研究进一步验证预测模型的临床可行性。

尽管本研究存在上述限制,但研究结果显示了其潜在的临床价值。本研究分析支持了能谱 CT 定量参数结合临床参数在预测 GAC 患者术前 LVI 和 PNI 方面的有效性和准确性。

参考文献:

- [1] BRAY F, FERLAY J, SOERJOMMATATAM I, et al. Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2018, 68(6):394-424.
- [2] 程非池, 邱佳辉, 郑扬, 等. 神经周围浸润、淋巴脉管浸润联合肿瘤间质比对胃癌预后的预测价值[J/OL]. 中华普通外科学文献(电子版), 2023, 17(3):186-192[2024-02-06]. <https://kns.cnki.net/kcms2/article/abstract?v=9CXCstbk-ttqX82XWef78430-REpPsVA7sks00neFTOzUyr07GDW04DWAbsSj>
- [3] 刘思佳. 光谱 CT 多参数成像预测结肠癌淋巴结转移及神经脉管侵犯的研究[D]. 昆明:昆明医科大学, 2023.
- [4] TAKANO K, ASHIKARI K, TAMURA S, et al. Clinicopathological features of endoscopically treated early gastric cancer with lymphovascular infiltration[J]. J Cancer Res Clin Oncol, 2023, 149(9):5781-5790.
- [5] 郭振江, 赵光远, 杜立强, 等. 进展期胃癌脉管侵犯术前列线图预测模型的建立和验证[J]. 天津医药, 2023, 51(12):1382-1386.
- [6] FUJIKAWA H, KOUMORI K, WATANABE H, et al. The clinical significance of lymphovascular invasion in gastric cancer[J]. In Vivo, 2020, 34(3):1533-1539.
- [7] AMIN M B, EDGE S B, GREENE F L, et al. AJCC cancer staging manual[M]. 8th ed. New York:Springer, 2017.
- [8] YANG Y T, DONG S Y, ZHAO J, et al. CT-detected extramural venous invasion is correlated with presence of lymph node metastasis and progression-free survival in gastric cancer[J]. Br J Radiol, 2020, 93(1116):20200673.
- [9] ZHANG F X, CHEN H X, LUO D D, et al. Lymphovascular or perineural invasion is associated with lymph node metastasis and survival outcomes in patients with gastric cancer[J]. Cancer Med, 2023, 12(8):9401-9408.
- [10] WANG Y, LIU W, YU Y, et al. CT radiomics nomogram for the preoperative prediction of lymph node metastasis in gastric cancer[J]. Eur Radiol, 2020, 30(2):976-986.
- [11] SALA L P, ETXEBERRIA L M, IGUÍÑIZ I E, et al. Gastric adenocarcinoma: A review of the TNM classification system and ways of spreading[J]. Radiologia (Engl Ed), 2023, 65(1):66-80.
- [12] GAO B, FENG C Z, CHAI F, et al. CT-detected extramural venous invasion-related gene signature for the overall survival prediction in patients with gastric cancer[J]. Cancer Med, 2021, 10(21):7816-7830.
- [13] FENG F, TIAN Y Z, XU G H, et al. Diagnostic and prognostic value of CEA, CA19-9, AFP and CA125 for early gastric cancer[J]. BMC Cancer, 2017, 17(1):737.
- [14] LIU G F, LI M Y, LI G S, et al. Assessing the blood supply status of the focal ground-glass opacity in lungs using spectral computed tomography[J]. Korean J Radiol, 2018, 19(1):130-138.
- [15] YANG C B, ZHANG S, JIA Y J, et al. Dual energy spectral CT imaging for the evaluation of small hepatocellular carcinoma microvascular invasion[J]. Eur J Radiol, 2017, 95:222-227.
- [16] REN H Y, SHEN J X, MAO X M, et al. Correlation between tumor vasculogenic mimicry and poor prognosis of human digestive cancer patients: A systematic review and meta-analysis[J]. Pathol Oncol Res, 2019, 25(3):849-858.

Machine Learning Optimization of Spectral CT for Predicting Invasiveness in Gastric Adenocarcinoma

GE Huiting, WANG Lili, LIU Yuanfen, ZOU Tianxiu, LIN Weiwen, XUE Yunjing

Department of Radiology, Fujian Medical University Union Hospital, Fuzhou 350001, China

ABSTRACT: **Objective** It is to investigate the potential value of machine learning (ML) algorithms combined with quantitative parameters of spectral CT and clinical models in predicting lymphovascular invasion (LVI) and perineural invasion (PNI) in patients with gastric adenocarcinoma (GAC). **Methods**

A total of 114 patients with GAC confirmed by pathology from December 2017 to May 2022 were collected.

The study parameters involved serum tumor markers, CT-TN staging, extramural venous invasion assessed by CT (CT-EMVI), and quantitative parameters of spectral CT. Feature screening was conducted using the Best-First algorithm in WEKA software, and models were built using Bayesian network (BN) and support vector machine (SVM) algorithms. **Results** Compared with the LVI/PNI(-) group, CT-T₃₋₄ stage, CT-N positive, CT-EMVI positive, and serum tumor markers (CA72-4, CA19-9) were more common in LVI/PNI(+) group, and the spectral CT parameters were higher in LVI/PNI(+) group, with statistically significant differences observed ($P < 0.05$). After feature screening, key variables included CT-T staging, CT-EMVI, VP-NIC, and EP-70 keV CT values. Based on these variables, six models were constructed using BN and SVM, including clinical parameter models, spectral CT parameter models, and combined models. All the six models demonstrated high predictive performance without overfitting. The combined model using BN showed the best predictive performance with an AUC range of 0.890 to 0.933, and Delong test showed that it was statistically significant ($P < 0.05$). In contrast, the SVM combined model did not show a statistical difference between the model and the other two models ($P > 0.05$). **Conclusion** Machine learning models combining clinical and spectral CT parameters can efficiently evaluate the LVI and PNI status of GAC patients, with the BN combined model achieving the highest accuracy.

KEY WORDS: gastric adenocarcinoma; spectral CT; lymphovascular invasion; perineural invasion

(编辑:何佳凤)

《福建医科大学学报》

双月刊 CN 35-1192/R 邮发代号 34-66

欢迎投稿

欢迎订阅