DOI:10.16016/j.2097-0927.202212045

基于机器学习建立腹部手术术后脓毒症患者死亡风险预测模型

舒 欣[1],李昊洋[2],李雨捷[1],宋艾璘[1],胡小艳[1],陈芋文[1,3],张 炬[3],易 斌[1] 400038 重庆,陆军军医大学(第三军医大学):第一附属医院麻醉科[1],基础医学院学员五大队[2];400714 重庆,中国科学院重庆绿色智能技术研究院[3]

[摘要] **目的** 探讨机器学习算法构建腹部手术术后脓毒症患者死亡风险预测模型的可行性。**方法** 采用病例-对照研究设计方案,从公共重症监护医学信息数据库(Medical Information Mart for Intensive Care Ⅳ,MIMIC-Ⅳ v1.0)中筛选出行腹部手术后发生脓毒症的患者,研究终点事件定义为患者入院后 90 d 内死亡。根据死亡与否将数据集随机拆分为训练数据集(70%)与测试数据集(30%),在训练数据集上基于 Logistic 回归(logistic regression,LR)、梯度提升树(gradient boosting decision tree,GBDT)、随机森林(random forest,RF)、支持向量机(support vector machine,SVM)和自适应提升(adaptive boosting,AdaBoost)算法构建预测死亡风险模型;在测试数据集上通过受试者工作曲线(receiver operating characteristic curve,ROC)和曲线下面积(area under the ROC curve,AUC)、敏感性、特异性、阳性预测值、阴性预测值、F1 分数和准确率来评估模型效能。**结果** 最终 986 例患者纳入本研究,其中 251 例(25.5%)患者入院后 90 d 内死亡,LR、GBDT、RF、SVM 及 AdaBoost 模型的 AUC 依次为 0.852、0.903、0.921、0.940 和 0.906,其中 SVM 的 AUC 最高,预测性能更好,而 LR 模型效能最差。**结论** 基于 GBDT、RF、SVM 及 AdaBoost 这 4 种算法建立的腹部手术术后脓毒症死亡率预测模型的效能优于传统的 LR 模型,可能有助于临床决策,改善不良结局。

[关键词] 术后脓毒症;机器学习;腹部手术;预测模型;死亡风险
[中图法分类号] R319;R619.3;R656.07 [文献标志码] A

Prediction of postoperative sepsis mortality risk based on machine learning in patients undergoing abdominal surgery

SHU Xin[1], LI Haoyang[2], LI Yujie[1], SONG Ailin[1], HU Xiaoyan[1], CHEN Yuwen[1,3], ZHANG Ju[3], YI Bin[1]
[1]Department of Anesthesiology, First Affiliated Hospital, [2]Regiment Five, Basical Medicine College, Army Medical University (Third Military Medical University), Chongqing, 400038; [3]Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, 400714, China

[**Abstract**] **Objective** To explore the feasibility of constructing prediction models of postoperative sepsis mortality risk based on machine learning in patients undergoing abdominal surgery. **Methods** A case-control trial was designed and conducted on the patients diagnosed with sepsis after abdominal surgery from Medical Information Mart for Intensive Care Ⅳ (MIMIC-Ⅳ) database, and 90-day mortality was defined as the primary endpoint event after hospitalization. The dataset was ramdomly split into training (70%) and test (30%) datasets according to wether diagnosed with postopertive sepsis or not. On the training dataset, logistic regression (LR), gradient boosting decision tree (GBDT), random forest (RF), support vector machine (SVM) and adaptive boosting (AdaBoost) were used to develop the prediction model for death. The area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value, negative predictive value, accuracy and F1 score were used for model evaluation on the test dataset.

[基金项目] 国家重点研发计划课题(2018YFC0116702);重庆英才计划"包干制"项目(CQYC202103080)
[通信作者] 易 斌,E-mail:yibin1974@163.com

2097-0927/Copyright © 2022,陆军军医大学(第三军医大学)。这篇开放获取文章遵循 CC BY 许可协议(https://creativecommons.org/licenses/by/4.0/)

Results A total of 986 patients were finally analyzed, of whom 251 patients (25.5%) died within 90 d after hospitalization. The AUC values of LR, GBDT, RF, SVM and AdaBoost prediction models were 0.852, 0.903, 0.921, 0.940 and 0.906, respectively. The model based on SVM yielded the best AUC value, higher differentiation and better prediction performance, while LR performed the worst among them. **Conclusion** The performances of the prediction model of postoperative sepsis mortality based on GBDTT, RF, SVM and AdaBoost are all better than that of traditional LR model, which may help to assist clinical decision making and improve adverse outcomes.

[**Key words**] postoperative sepsis; machine learning; abdominal surgery; prediction model; mortality risk

Supported by the National Key R&D Program of China (2018YFC0116702) and the "Lump Sum Contract" Project of Chongqing Talent Program (CQYC202103080). Corresponding author: YI Bin, E-mail: yibin1974@163.com

脓毒症是由宿主对感染反应失调所致的危及生命的器官功能障碍,是一种致死率高且花费高昂的疾病[1-2]。术后脓毒症是脓毒症的主要表现形式之一,发生于手术后或术后住院期间[3]。腹部手术由于手术时间长,肠道内寄生着多种细菌,手术时易造成细菌移位,同时机体免疫力降低,术后肠道功能减退,术后脓毒症成为患者器官功能障碍和死亡的主要原因之一[4]。早期预测术后脓毒症患者的死亡风险,量化脓毒症患者的疾病严重程度,及早进行干预,可以改善患者结局。虽然目前已有多种评分应用于危重症患者病情的危险程度评估及不良预后预测,如简化急性生理功能评分Ⅱ(simplified acute physiology score Ⅱ, SAPS-Ⅱ)、急性生理学和慢性健康状况评分Ⅱ(acute physiology and chronic health evaluation Ⅱ,APACHE-Ⅱ)及序贯器官衰竭评分(sequential organ failure assessment, SOFA)等,但它们大多是基于重症监护室(intensive care unit, ICU)患者设计的,未能针对术后脓毒症患者的特点(如缺乏手术相关因素),对死亡风险预测缺乏疾病特异性。随着计算机性能的不断提升及算法的优化,机器学习已广泛应用于临床研究中,在过去的十年里,已有许多团队尝试创建不同的机器学习模型用于脓毒症的死亡风险预测[5-7]。虽然大多取得了良好的效果,但在腹部手术术后脓毒症患者的临床应用上仍受到了限制。本研究以腹部手术术后脓毒症患者为研究对象,拟采用目前应用最广泛的4种集成学习算法建立患者90 d内死亡风险的预测模型,并与传统Logistic回归模型进行比较,为临床医生早期识别高危患者和精准干预提供新思路。

1 资料与方法

1.1 一般资料

1.1.1 数据来源 本研究的医疗数据来自于公共重症监护医学信息数据库(Medical Information Mart for Intensive Care Ⅳ,MIMIC-Ⅳ v1.0)[8],因其为开放性数据库,仅需通过网上参加伦理考核并获得许可后即可使用其数据(许可认证号:46086293)。所有遵守数据库使用要求者均可使用该数据库(https://mimic.physionet.org)。由于该项目不影响临床护理,并且隐藏了所有受保护的信息,因此不需要患者签署同意书。

1.1.2 研究对象 MIMIC-Ⅳ数据库包含了2008年至2019年共383 220名贝斯以色列迪康医学中心ICU收治住院的患者临床数据。纳入标准:行腹部手术[参考第十版国际疾病分类(international classification of diseases-tenth revision, ICD-10)腹部手术包含的不同类型进行纳入:胃肠手术、肝胆胰手术、脾手术、泌尿系统手术、女性生殖系统手术];年龄>18岁;诊断包含脓毒症、严重脓毒症、脓毒性休克;符合脓毒症3.0诊断标准[1],即SOFA评分≥2。排除标准:手术操作部位表浅,未涉及腹腔;手术记录信息缺失>30%或同一操作重复记录者;术前患有脓毒症或可疑感染,或严重合并症者;确诊后无法再次行SOFA评分者。

1.1.3 数据收集 根据临床经验、既往研究及MIMIC-Ⅳ数据库数据记录情况,本研究共募集了4类共43个指标:①一般资料指标:性别、年龄、入院方式、体质指数(body mass index,BMI);②术前检验(选取距离手术开始时间最近的一次结果):碱性磷酸酶(alkaline phosphatase, ALP)、血钠、血钾、血氯、尿蛋白、血小板计数(platelet, PLT)、白细胞计数(white blood cell, WBC)、凝血酶原时间(prothrombin time, PT)、动脉血酸碱度(arteria blood pondus hydrogenii, PH)、动脉血CO_2分压(arterial partial pressure of carbon dioxide,$PaCO_2$)、动脉血氧分压(arterial partial pressure of oxygen,PO_2)、血乳酸(lactic acid,Lac)、中性粒细胞与淋巴细胞比(neutrophil-lymphocyte ratio,NLCR)、血红蛋白(hemoglobin,Hb)、国际标准化比值(international normalized ratio,INR)、红细胞比容(hematocrit,HCT)、血葡萄糖、血肌酐(creatinine,sCr)、血尿素氮(blood urea nitrogen,BUN)、总胆红素(total bilirubin,TBIL)、

天门冬氨酸氨基转移酶(aspartate aminotransferase,AST)、白蛋白(albumin,ALB);③术后评分:确诊时SOFA评分、SAPS-Ⅱ、格拉斯哥昏迷评分(Glasgow coma scale,GCS)、牛津急性疾病严重度评分(Oxford acute severity of illness score,OASIS);④围术期相关事件及合并症:诊断数、弥散性血管内凝血(disseminated intravascular coagulation,DIC)、心律失常、贫血、呼吸衰竭、心力衰竭、糖尿病、慢性阻塞性肺疾病(chronic obstructive pulmonary diseases,COPD)、肾脏疾病、肝脏疾病、恶性肿瘤、输注血制品、手术部位。

1.2　方法

1.2.1　模型构建　　为增加模型泛化能力和减少过拟合,我们采用LASSO回归分析筛选有意义的变量纳入模型构建[9]。同时,为使数据符合机器学习的输入规则,我们对离散型变量采用one-hot编码处理,连续型变量则采用Z-score归一化处理。根据入院后90 d死亡与否,将数据集按7∶3的比例,随机拆分为训练数据集和测试数据集。训练数据集采用Logistic回归(logistic regression,LR)、梯度提升树(gradient boosting decision tree,GBDT)、随机森林(random forest,RF)、支持向量机(support vector machine,SVM)和自适应提升(adaptive boosting,AdaBoost)这5种算法分别构建死亡风险预测模型。其中LR为传统线性回归分析模型,后4种为目前应用最为广泛的集成学习模型,它们可结合多个学习器进而获得比单一学习器更优越的泛化性能。此外,测试数据集用于模型的验证与评估。本研究的技术路线流程图见图1。

1.2.2　模型评估　　采用受试者工作曲线(receiver operating characteristic curve,ROC)和曲线下面积(area under the ROC curve,AUC)、敏感性、特异性、阳性预测值(positive predictive value,PPV)、阴性预测值(negative predictive value,NPV)、准确率和F1分数(精确率和召回率的一种加权平均)来评估各模型的预测效能。

1.3　统计学分析

采用SPSS软件(V.26.0,IBM)和R软件(V.4.2.2)进行数据分析。偏态分布的计量资料以$M(P_{25},P_{75})$表示,组间比较采用Mann-Whitney U检验。计数资料采用频数(%)表示,组间比较采用χ^2检验或*Fisher*确切概率法。所有的统计检验均为双侧,$P<0.05$为差异有统计学意义。

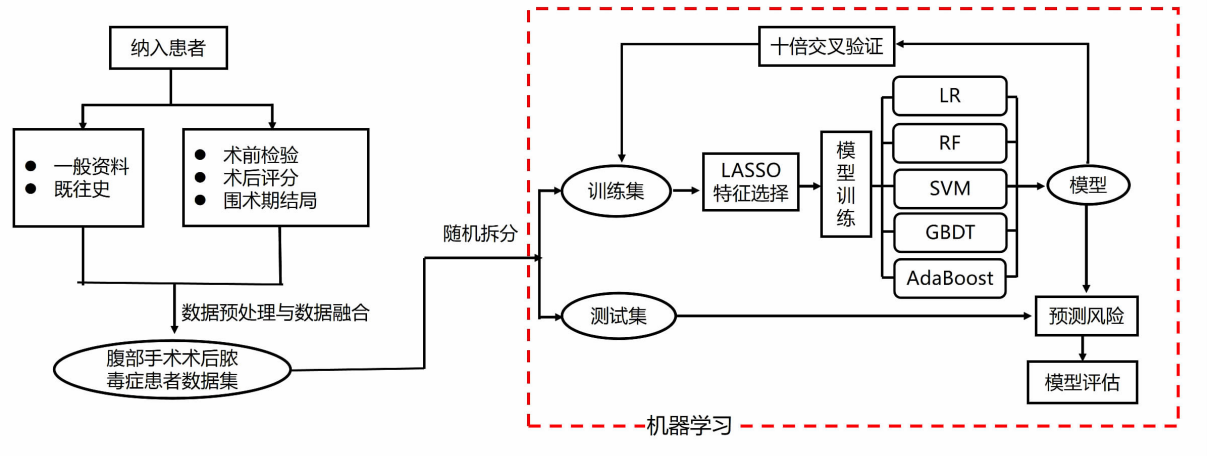
2　结果

2.1　患者基线资料

根据纳入排除原则,MIMIC-IV数据库中共986例患者行腹部手术发生了术后脓毒症,其中有251例患者于入院后90天内死亡,死亡率为25.5%(图2)。死亡组与存活组的一般资料比较见表1。与存活组患者相比,死亡组患者诊断数、确诊时SOFA评分、SAPS-Ⅱ评分、OASIS评分和血乳酸水平更高,患有DIC、呼吸衰竭、心力衰竭、肾脏疾病、肝脏疾病和输注血制品的更多(均$P<0.001$)。

2.2　特征筛选结果

本研究选取LASSO回归分析特征选择系数在一个方差范围内仍不为0的变量,最终有意义的变量如下:①一般信息:性别、BMI、入院方式;②术前检验:ALB、尿蛋白、PLT、WBC、PT、$PaCO_2$、Lac、NLCR、Hb、血葡萄糖;③术后评分:确诊时SOFA评分、SAPS-Ⅱ评分、OASIS评分;④围术期相关事件及合并症:输注血制品、手术部位、诊断数、心力衰竭、呼吸衰竭、DIC、COPD、肝脏疾病、恶性肿瘤。相关特征系数见表2。



LR:Logistic回归;GBDT:梯度提升树;RF:随机森林;SVM:支持向量机;AdaBoost:自适应提升

图1　研究技术路线流程图

表 1 腹部手术术后脓毒症死亡组与存活组的一般资料比较[$M(P_{25},P_{75})$,例(%)]

变量	死亡组(n=251)	存活组(n=735)	P
年龄/岁	69.33(60.23,79.73)	69.56(58.77,80.41)	0.818
BMI/kg·m^{-2}	27.61(23.83,32.78)	28.16(24.67,33.33)	0.263
女性	120(47.8)	353(48.0)	0.952
诊断数	26.0(21.0,33.0)	22.0(17.0,28.0)	<0.001
急诊入院	161(64.1)	501(68.2)	0.242
合并症			
DIC	31(12.4)	16(2.2)	<0.001
心律失常	122(48.6)	277(37.7)	0.002
贫血	77(30.7)	200(27.2)	0.291
呼吸衰竭	161(64.1)	258(35.1)	<0.001
心力衰竭	87(34.7)	160(21.8)	<0.001
糖尿病	79(31.5)	230(31.3)	0.957
COPD	42(16.7)	78(10.6)	0.010
肾脏疾病	197(78.5)	457(62.2)	<0.001
肝脏疾病	98(39.0)	125(17.0)	<0.001
恶性肿瘤	79(31.5)	162(22.0)	0.003
输注血制品	176(70.1)	284(38.6)	<0.001
手术部位			<0.001
肝胆胰手术	67(26.7)	260(35.4)	
女性生殖系统手术	2(0.8)	6(0.8)	
脾手术	1(0.4)	3(0.4)	
泌尿系统手术	16(6.4)	160(21.8)	
胃肠手术	165(65.7)	306(41.6)	
术前检验			
ALP/U·L^{-1}	128.0(73.0,249.0)	120.0(74.0,234.0)	0.756
血钠/mmol·L^{-1}	137.0(134.0,140.0)	137.0(135.0,140.0)	0.360
血钾/mmol·L^{-1}	4.20(3.70,4.70)	4.0(3.60,4.40)	0.006
血氯/mmol·L^{-1}	104.0(99.0,108.0)	104.0(100.0,108.0)	0.947
尿蛋白/mg·dL^{-1}	30.0(0,30.0)	30.0(0,30.0)	0.274
PLT/×10^{12}·L^{-1}	185.0(110.0,274.0)	196.0(139.0,284.0)	0.013
WBC/×10^{9}·L^{-1}	10.10(5.80,17.50)	12.70(7.90,18.80)	<0.001
pT/s	15.90(13.20,21.30)	15.10(13.20,17.50)	0.004
pH	7.32(7.21,7.39)	7.34(7.28,7.40)	<0.001
$PaCO_2$/mmHg	41.0(34.0,47.0)	40.0(35.0,45.0)	0.265
PaO_2/mmHg	94.0(67.0,186.0)	93.0(51.0,179.0)	0.346
Lac/mmol·L^{-1}	3.60(1.80,9.0)	1.40(1.0,1.80)	<0.001
NLCR	0.83(0.58,0.93)	0.84(0.66,0.96)	0.200
Hb/g·dL^{-1}	10.30(8.70,11.70)	10.80(9.10,12.30)	0.017
INR	1.50(1.20,1.90)	1.30(1.20,1.60)	<0.001
HCT/%	32.0(27.10,35.90)	32.90(28.20,37.20)	0.031
血葡萄糖/mg·dL^{-1}	116.0(90.0,155.0)	122.0(100.0,158.0)	0.034
sCr/mg·dL^{-1}	1.20(0.80,2.10)	1.10(0.80,1.80)	0.071
BUN/mg·dL^{-1}	27.0(18.0,45.0)	22.0(16.0,38.0)	0.002
TBIL/mg·dL^{-1}	1.60(0.60,4.50)	1.60(0.60,4.0)	0.421
AST/IU·L^{-1}	59.0(28.0,135.0)	54.0(25.0,133.0)	0.402
ALB/g·dL^{-1}	2.60(2.20,3.10)	2.80(2.40,3.30)	0.003
术后评分			
确诊时 SOFA/分	5.0(3.0,7.0)	3.0(2.0,5.0)	<0.001
SAPS-Ⅱ/分	57.0(47.0,69.0)	42.0(34.0,52.0)	<0.001
GCS/分	15.0(14.0,15.0)	15.0(15.0,15.0)	0.155
OASIS/分	45.0(37.0,51.0)	36.0(29.0,43.0)	<0.001

BMI:体质指数,DIC:弥散性血管内凝血,COPD:慢性阻塞性肺疾病,ALP:碱性磷酸酶,PLT:血小板计数,WBC:白细胞计数,PT:凝血酶原时间,pH:动脉血酸碱度,$PaCO_2$:动脉血二氧化碳分压,PaO_2:动脉血氧分压,Lac:血乳酸,NLCR:中性粒细胞与淋巴细胞比,Hb:血红蛋白,INR:国际标准化比值,HCT:红细胞比容,sCr:血肌酐,BUN:血尿素氮,TBIL:总胆红素,AST:谷草转氨酶,ALB:白蛋白,SOFA:序贯器官衰竭评分,SAPS-Ⅱ:简化急性生理功能评分Ⅱ,GCS:格拉斯哥昏迷评分,OASIS:牛津急性疾病严重度评分

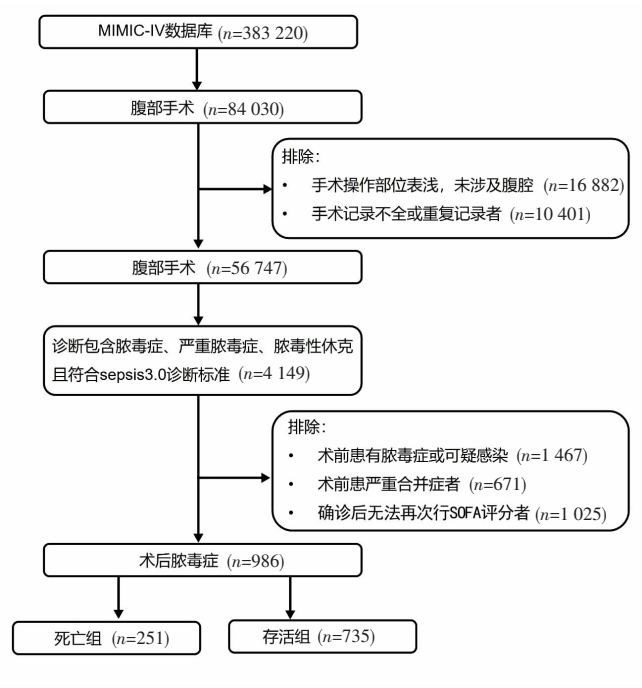


图 2 病例筛选流程图

2.3 模型性能

在测试数据集验证并评估 5 种算法对腹部手术术后脓毒症患者 90 d 死亡率的预测能力。如表 3 所示，LR、GBDT、RF、SVM 及 AdaBoost 模型的 AUC 依次为 0.852、0.903、0.921、0.940 和 0.906，其中，SVM 的 AUC 最高，相较于其他 4 种机器学习模型的区分度更高，预测性能更好(图 3)。同时，SVM 的准确率、PPV 及 F1 分数均高于其他 4 种模型，其对所有样本预测正确的比例较高，且对阳性样本的识别能力强。而 GBDT 的灵敏度更高，能更灵敏地发现死亡风险高的患者。整体上，这 4 种集成学习算法模型的预测效能均比传统的 LR 模型更高，且 SVM 的预测性能更令人满意。

表 2 LASSO 特征选择结果及系数

变量	特征系数	变量	特征系数
年龄	0.000 000 00	PT	0.011 876 89
BMI	−0.001 201 15	pH	0.000 000 00
性别	0.185 971 33	$PaCO_2$	0.001 417 75
诊断数	0.018 685 82	PaO_2	0.000 000 00
入院方式	−0.039 617 03	Lac	0.469 754 18
DIC	0.348 786 15	NLCR	−0.000 638 85
心律失常	0.000 000 00	Hb	−0.009 728 46
贫血	0.000 000 00	INR	0.000 000 00
呼吸衰竭	0.577 193 47	HCT	0.000 000 00
心力衰竭	0.441 228 39	血葡萄糖	−0.002 076 03
糖尿病	0.000 000 00	sCr	0.000 000 00
COPD	0.497 430 35	BUN	0.000 000 00
肾脏疾病	0.000 000 00	TBIL	0.000 000 00
肝脏疾病	0.359 474 29	AST	0.000 000 00
恶性肿瘤	0.110 268 12	ALB	−0.064 269 11
ALP	0.000 000 00	输注血制品	0.394 188 04
血钠	0.000 000 00	手术部位	0.153 065 53
血钾	0.000 000 00	确诊时 SOFA	0.057 620 16
血氯	0.000 000 00	SAPS-Ⅱ评分	0.037 894 51
尿蛋白	−0.001 465 97	GCS 评分	0.000 000 00
PLT	−0.000 206 84	OASIS 评分	0.006 350 74
WBC	−0.016 378 27		

特征系数不为零的变量为重要变量，被纳入模型构建

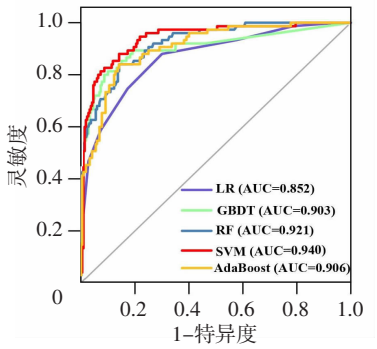


图 3 5 种模型预测腹部手术术后脓毒症 90 d 死亡的 ROC 曲线

表 3 5 种算法模型的预测性能比较

模型	AUC,95% *CI*	Sens(%)	Spec(%)	PPV(%)	NPV(%)	F1 分数	ACC (%)
LR	0.852,0.799~0.905	66.2	92.0	74.6	88.5	0.702	79.1
GBDT	0.903,0.856~0.950	78.7	91.4	75.6	92.6	0.771	85.0
RF	0.921,0.888~0.955	62.7	94.6	79.7	88.1	0.702	78.6
SVM	0.940,0.909~0.972	76.0	95.5	85.1	92.1	0.803	85.7
AdaBoost	0.906,0.868~0.944	66.7	90.9	71.4	88.9	0.690	78.8

3 讨论

脓毒症患者是外科手术进入ICU甚至死亡的主要原因,且腹部手术发生率最高[10]。1项Meta分析显示脓毒症患者住院期间死亡率为26.7%~41.9%[11],本研究腹部手术术后脓毒症患者入院后90 d死亡率约为25.5%。早期预测死亡风险,精准治疗可防止病情恶化并挽救患者。然而目前临床常用的评分工具依赖人工操作,费时费力且效率低下。准确、方便、快捷的病情评估能协助临床决策,及时采取救治措施,对脓毒症患者预后意义重大。

随着算法算力、计算机硬件水平的提升和大数据时代的来临,机器学习在挖掘、处理医疗数据中展现出极大的优势,目前已广泛应用于临床疾病的发生及预后的预测[12],如脓毒症及术后脓毒症的预测[13-14]、急性肾损伤的预测[15]、脓毒症患者心脏骤停的预测[16]、急性冠脉综合征后不良结局的预测[17]、ICU患者死亡风险的预测等[18-19]。

在本研究中,我们基于LR、GBDT、RF、SVM及AdaBoost算法构建了腹部手术术后脓毒症患者死亡风险预测模型,并比较了这5种模型的预测效能。GBDT、RF、SVM及AdaBoost是目前应用最广泛的4种集成学习算法,其中RF主要采用Bagging方法,GBDT和AdaBoost主要采用Boosting方法,SVM基于非线性映射理论寻找划分特征空间的最优平面,充分考虑属性之间的依赖关系,对存在依赖关系的属性添加关联弧,扩展了朴素贝叶斯模型的结构,可显著提高分类效果[20]。LR是一种经典的回归分析方法,被广泛应用于探讨疾病的危险因素或根据危险因素预测疾病发生的概率等。但XIAO等[21]发现与一些机器学习算法相比,传统的LR预测误差更大、预测性能更差。这一结论在本研究结果中得到了证实:LR模型AUC为0.852,敏感度为66.2%,特异度为92.0%,PPV为74.6%,NPV为88.5%,F1分数为0.702,准确度为79.1%,在5种模型中预测效能最差;而表现最好的SVM模型AUC为0.940,敏感度为76.0%,特异度为95.5%,PPV为85.1%,NPV为92.1%,F1分数为0.803,准确度为85.7%,该模型区分度好,对所有样本预测正确的比例高,且对识别死亡风险高的患者能力强。

此外,本研究的单因素分析结果发现,高诊断数、高确诊时SOFA评分、高SAPS-Ⅱ评分、高OASIS评分、高BUN、高血乳酸和低白蛋白水平以及多合并症和输注血制品,可能会增加腹部手术术后脓毒症死亡风险。这一结论与既往研究及临床经验相一致,同时在LASSO回归分析中也得到了证实。其中,血乳酸水平(3.60 *vs* 1.40,死亡组 *vs* 存活组,$P<0.001$)是脓毒症休克的诊断指标[1],也是脓毒症休克患者死亡的独立预测因子,患者死亡率可随乳酸水平的增高而增高[22],此外血乳酸还用于指导脓毒症治疗[23]。

本研究存在一定的局限性。首先,本研究是基于公共数据库MIMIC-Ⅳ进行的单中心回顾性研究,尽管已经尽可能收集腹部手术术后脓毒症病例,但样本量依旧较一些大样本研究小,且未进行外部数据库的验证,模型的性能还需要多中心、大样本、前瞻性的研究进一步验证。其次,由于MIMIC-Ⅳ数据库时间跨度较长,近几年新发现的脓毒症特异性指标数据缺失率高,未能纳入构建模型,如C反应蛋白、降钙素原、IL-6、肝素结合蛋白等。最后,MIMIC-Ⅳ数据库来源于贝斯以色列女狄肯斯医疗中心,亚裔人群所占比例低,可能存在人群异质性。我们将在后期的工作中纳入本地患者数据,将现有模型与本地数据库相结合,验证模型性能,同时调整参数,提高模型性能,使之更符合本地人群特点。

综上所述,基于GBDT、RF、SVM及AdaBoost这4种集成学习算法建立的腹部手术术后脓毒症死亡率预测模型的研究可行,且这4种模型的预测效能均优于传统的LR模型,其中SVM模型的预测性能最令人满意。该模型可能提前识别出术后脓毒症死亡风险大的患者,并辅助临床医生进行决策,调整治疗方案,减少不良结局的发生。

参考文献:

[1] SINGER M, DEUTSCHMAN C S, SEYMOUR C W, *et al*. The third international consensus definitions for sepsis and septic shock (Sepsis-3)[J]. JAMA, 2016, 315(8): 801-810. DOI: 10.1001/jama.2016.0287.

[2] LIU V, ESCOBAR G J, GREENE J D, *et al*. Hospital deaths in patients with sepsis from 2 independent cohorts[J]. JAMA, 2014, 312(1): 90-92.

[3] PLAEKE P, DE MAN J G, COENEN S, *et al*. Clinical- and surgery-specific risk factors for post-operative sepsis: a systematic review and meta-analysis of over 30 million patients[J]. Surg Today, 2020, 50(5): 427-439. DOI: 10.1007/s00595-019-01827-4.

[4] WELEDJI E P, NGOWE M N. The challenge of intra-abdominal sepsis[J]. Int J Surg, 2013, 11(4): 290-295. DOI: 10.1016/j.ijsu.2013.02.021.

[5] NEMATI S, HOLDER A, RAZMI F, *et al*. An interpretable machine learning model for accurate prediction of sepsis in the ICU[J]. Crit Care Med, 2018, 46(4): 547-553. DOI: 10.1097/CCM.0000000000002936.

[6] LAURITSEN S M, THIESSON B, JØRGENSEN M J, *et al*. The Framing of machine learning risk prediction models illustrated by evaluation of sepsis in general wards[J]. NPJ Digit Med, 2021, 4(1): 158. DOI: 10.1038/s41746-021-00529-x.

[7] CILLONIZ C, WARD L, MOGENSEN M L, *et al*. Machine-learning model for mortality prediction in patients with community-acquired pneumonia: development and validation study[J]. Chest, 2022: S0012-S3692(22)01243-0. DOI: 10.1016/j.chest.2022.07.005.

[8] JOHNSON A B L, POLLARD T, HORNG S, *et al*. MIMIC-IV (version 1.0). PhysioNet2021. Available from: https://doi.org/10.13026/7vcr-e114.

[9] WANG J, ZHANG H Q, WANG J Z, *et al*. Feature selection using a neural network with group lasso regularization and controlled redundancy[J]. IEEE Trans Neural Netw Learn Syst, 2021, 32(3): 1110-1123. DOI: 10.1109/TNNLS.2020.2980383.

[10] BATEMAN B T, SCHMIDT U, BERMAN M F, *et al*. Temporal trends in the epidemiology of severe postoperative sepsis after elective surgery: a large, nationwide sample[J]. Anesthesiology, 2010, 112(4): 917-925. DOI: 10.1097/ALN.0b013e3181cea3d0.

[11] FLEISCHMANN-STRUZEK C, MELLHAMMAR L, ROSE N, *et al*. Incidence and mortality of hospital- and ICU-treated sepsis: results from an updated and expanded systematic review and meta-analysis[J]. Intensive Care Med, 2020, 46(8): 1552-1562. DOI: 10.1007/s00134-020-06151-x.

[12] 刘蓬然, 霍彤彤, 陆林, 等. 人工智能在医学中的应用现状与展望[J]. 中华医学杂志, 2021, 101(44): 3677-3683. DOI: 10.3760/cma.j.cn112137-20210313-00628
LIU P R, HUO T T, LU L, *et al*. Application status and prospect of artificial intelligence in medicine[J]. Natl Med J China, 2021, 101(44): 3677-3683. DOI: 10.3760/cma.j.cn112137-20210313-00628

[13] LAURITSEN S M, KALØR M E, KONGSGAARD E L, *et al*. Early detection of sepsis utilizing deep learning on electronic health record event sequences[J]. Artif Intell Med, 2020, 104: 101820. DOI: 10.1016/j.artmed.2020.101820.

[14] BUNN C, KULSHRESTHA S, BOYDA J, *et al*. Application of machine learning to the prediction of postoperative sepsis after appendectomy[J]. Surgery, 2021, 169(3): 671-677. DOI: 10.1016/j.surg.2020.07.045.

[15] DONG J Z, FENG T, THAPA-CHHETRY B, *et al*. Machine learning model for early prediction of acute kidney injury (AKI) in pediatric critical care[J]. Crit Care, 2021, 25(1): 288. DOI: 10.1186/s13054-021-03724-0.

[16] UENO R, XU L Y, UEGAMI W, *et al*. Value of laboratory results in addition to vital signs in a machine learning algorithm to predict in-hospital cardiac arrest: a single-center retrospective cohort study[J]. PLoS One, 2020, 15(7): e0235835. DOI: 10.1371/journal.pone.0235835.

[17] D'ASCENZO F, DE FILIPPO O, GALLONE G, *et al*. Machine learning-based prediction of adverse events following an acute coronary syndrome (PRAISE): a modelling study of pooled datasets[J]. Lancet, 2021, 397(10270): 199-207. DOI: 10.1016/S0140-6736(20)32519-8.

[18] THORSEN-MEYER H C, NIELSEN A B, NIELSEN A P, *et al*. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records[J]. Lancet Digit Health, 2020, 2(4): e179-e191. DOI: 10.1016/S2589-7500(20)30018-2.

[19] 齐霜, 徐浩然, 胡婕, 等. 基于机器学习的重症监护病房脓毒症患者早期死亡风险预测模型[J]. 解放军医学院学报, 2021, 42(2): 150-155, 181. DOI: 10.3969/j.issn.2095-5227.2021.02.006.
QI S, XU H R, HU J, *et al*. Early mortality risk prediction model for sepsis patients in intensive care unit based on machine learning[J]. Acad J Chin Pla Med Sch, 2021, 42(2): 150-155, 181. DOI: 10.3969/j.issn.2095-5227.2021.02.006.

[20] 张颖莹, 刘怡果, 赵丹, 等. 基于机器学习建立脓毒症心肾综合征患者早期死亡风险预测模型[J]. 中华肾脏病杂志, 2022, 38(9)785-793. DOI: 10.3760/cma.j.cn441217-20211126-00113
ZHANG Y Y, LIU Y G, ZHAO D, *et al*. Early mortality risk prediction models for patients with sepsis-induced cardiorenal syndrome based on machine learning[J]. Chin J Nephrol, 2022, 38(9): 785-793. DOI: 10.3760/cma.j.cn441217-20211126-00113

[21] XIAO J, DING R F, XU X L, *et al*. Comparison and development of machine learning tools in the prediction of chronic kidney disease progression[J]. J Transl Med, 2019, 17(1): 119. DOI: 10.1186/s12967-019-1860-0.

[22] RYOO S M, LEE J, LEE Y S, *et al*. Lactate level versus lactate clearance for predicting mortality in patients with septic shock defined by Sepsis-3[J]. Crit Care Med, 2018, 46(6): e489-e495. DOI: 10.1097/CCM.0000000000003030.

[23] JONES A E, SHAPIRO N I, TRZECIAK S, *et al*. Lactate clearance vs central venous oxygen saturation as goals of early sepsis therapy: a randomized clinical trial[J]. JAMA, 2010, 303(8): 739-746. DOI: 10.1001/jama.2010.158.

(收稿:2022-12-06;修回:2023-01-09)

(编辑 邓强庭)