

DOI: 10.16781/j.CN31-2187/R.20250005

· 论 著 ·

上尿路结石术后尿源性脓毒血症人工智能预警模型：基于临床多模态数据

陈永文¹, 罗小燕¹, 梁燕秋¹, 王于路¹, 谭宝飞¹, 陈一峰¹, 梁 彬¹, 黄倍源¹, 韦嘉佳², 王祖恒², 王富博^{2*}, 庞桂建^{1*}

1. 广西医科大学第六附属医院泌尿外科, 玉林 537000

2. 广西医科大学数字医学与健康高校工程研究中心, 南宁 530000

[摘要] **目的** 利用多种机器学习算法构建并验证上尿路结石患者术后发生尿源性脓毒血症的预测模型。**方法** 纳入 2018 年 6 月至 2023 年 6 月在广西医科大学第六附属医院接受手术治疗的上尿路结石患者 7 464 例, 其中 622 例 (8.33%) 术后发生尿源性脓毒血症。将所有患者按 7 : 3 比例分为训练集 (5 224 例) 和验证集 (2 240 例)。采用极限梯度提升 (XGBoost)、logistic 回归、轻量级梯度提升机 (LightGBM)、随机森林 (RF)、自适应增强 (AdaBoost)、梯度提升决策树 (GBDT) 6 种机器学习算法构建术后发生尿源性脓毒血症预测模型, 并通过 ROC 曲线、沙普利加性解释 (SHAP) 分析、校准曲线和决策曲线分析 (DCA) 评估模型的预测能力和临床效益等。**结果** 纳入模型的临床参数包括 BMI、手术次数、心率、Barthel 指数、静脉血栓栓塞 (VTE) 风险评分、性别、美国麻醉医师协会 (ASA) 分级、尿亚硝酸盐及尿白细胞。在训练集中, XGBoost、LightGBM 和 RF 模型表现优异, ROC 曲线 AUC 值均达到 1.00。在验证集中, logistic 回归模型表现最佳, ROC 曲线 AUC 值为 0.76, 具有较好的预测稳定性和校准度; AdaBoost 和 GBDT 模型次之, AUC 值分别为 0.74 和 0.75, 而 LightGBM、XGBoost 和 RF 模型的 AUC 值分别为 0.71、0.70 和 0.68。在模型的可解释性方面, SHAP 分析显示, 变量贡献度由大到小依次为心率、尿白细胞、性别、BMI、Barthel 指数、VTE 风险评分、尿亚硝酸盐、手术次数和 ASA 分级。**结论** 成功构建上尿路结石术后发生尿源性脓毒血症的早期风险预测 logistic 回归模型, 该模型具有较好的预测性能和校准度, 可有效辅助临床诊断。

[关键词] 上尿路结石; 尿源性脓毒血症; 机器学习; 风险预测模型

[引用本文] 陈永文, 罗小燕, 梁燕秋, 等. 上尿路结石术后尿源性脓毒血症人工智能预警模型: 基于临床多模态数据 [J]. 海军军医大学学报, 2025, 46(7): 889-897. DOI: 10.16781/j.CN31-2187/R.20250005.

Artificial intelligence warning model for urosepsis after upper urinary tract stone surgery: based on clinical multimodal data

CHEN Yongwen¹, LUO Xiaoyan¹, LIANG Yanqiu¹, WANG Yulu¹, TAN Baofei¹, CHEN Yifeng¹, LIANG Bin¹, HUANG Beiyuan¹, WEI Jiajia², WANG Zuheng², WANG Fubo^{2*}, PANG Guijian^{1*}

1. Department of Urology, the Sixth Affiliated Hospital of Guangxi Medical University, Yulin 537000, Guangxi Zhuang Autonomous Region, China

2. University Engineering Research Center of Digital Medicine and Healthcare, Guangxi Medical University, Nanning 530000, Guangxi Zhuang Autonomous Region, China

[Abstract] **Objective** To construct and validate a prediction model for urosepsis in patients after upper urinary tract stone surgery using various machine learning algorithms. **Methods** A total of 7 464 upper urinary tract stone patients who underwent surgery at the Sixth Affiliated Hospital of Guangxi Medical University from Jun. 2018 to Jun. 2023 were enrolled and randomly assigned to training (5 224 cases) or validation sets (2 240 cases) at a ratio of 7 : 3. Among them, 622 (8.33%) cases developed urosepsis postoperatively. Six machine learning algorithms, including extreme gradient boosting (XGBoost), logistic regression, light gradient boosting machine (LightGBM), random forest (RF), adaptive boosting (AdaBoost), and gradient boosting decision tree (GBDT), were used to construct prediction models for postoperative urosepsis. The model's

[收稿日期] 2025-01-03 **[接受日期]** 2025-05-19

[基金项目] 广西科技重大专项(桂科AA22096030, 桂科AA22096032), 广西医科大学基因组重点实验室开放课题(GXGPMC202304). Supported by Science and Technology Major Project of Guangxi (AA22096030, AA22096032) and Open Project of Key Laboratory of Genomics of Guangxi Medical University (GXGPMC202304).

[作者简介] 陈永文. E-mail: mnwkeyw@163.com

* 通信作者(Corresponding authors). E-mail: wangfubo@gxmu.edu.cn, pangguijian@163.com

predictive ability and clinical benefits were evaluated using receiver operating characteristic (ROC) curves, Shapley additive explanation (SHAP) analysis, calibration curves, and decision curve analysis (DCA). **Results** The clinical features included body mass index (BMI), number of surgeries, heart rate, Barthel index, venous thrombo embolism (VTE) risk assessment, gender, American Society of Anesthesiologists (ASA) grade, urinary nitrite, and urinary leukocyte in the models. In the training set, the XGBoost, LightGBM, and RF models performed excellently, with area under curve (AUC) values of ROC curves reaching 1.00. In the validation set, the logistic regression model performed the best, with an AUC value of ROC curve of 0.76, showing good predictive stability and calibration. The AdaBoost and GBDT models followed with AUC values of 0.74 and 0.75, respectively, while the AUC values of the LightGBM, XGBoost, and RF models were 0.71, 0.70, and 0.68. In terms of model interpretability, SHAP analysis showed the contribution of variables in a descending order as: heart rate, urinary leukocytes, gender, BMI, Barthel index, VTE risk assessment, urinary nitrite, number of surgeries, and ASA grade. **Conclusion** A logistic regression model for early risk prediction of postoperative urosepsis in upper urinary tract stone patients has been successfully constructed. This model has good predictive performance and calibration, and can effectively assist clinical diagnosis.

[**Key words**] upper urinary tract calculus; urosepsis; machine learning; risk prediction model

[**Citation**] CHEN Y, LUO X, LIANG Y, et al. Artificial intelligence warning model for urosepsis after upper urinary tract stone surgery: based on clinical multimodal data[J]. Acad J Naval Med Univ, 2025, 46(7): 889-897. DOI: 10.16781/j.CN31-2187/R.20250005.

泌尿系统结石是一种常见的泌尿系统疾病,全球的终身患病率为 10%~20%, 每年发病率约为 1%^[1]。泌尿系统结石中,上尿路结石的占比较高,其中肾结石、输尿管结石的患病率分别为 8.1% 和 7.8%^[2]。尽管体外冲击波碎石术、经皮肾镜取石术、输尿管镜碎石术及输尿管软镜碎石术等外科技术具有较高的结石清除率和手术安全性,但术后仍存在感染风险^[3]。尿源性脓毒血症作为源自泌尿道感染的全身炎症反应综合征,其经皮肾镜取石术和输尿管软镜碎石术的术后总体发生率为 1%~6%^[4-5]。脓毒血症进展极快,一旦发展为重症脓毒血症、感染性休克甚至是多器官功能衰竭,救治难度将显著增加,不仅医疗成本巨大,病死率亦大幅升高^[6]。因此,早期预警和干预对改善结石术后脓毒血症患者的预后至关重要^[7]。目前,对于上尿路结石手术治疗后发生尿源性脓毒血症的危险因素与风险预测研究不足,尚无早期准确、临床可用性高的预测模型。

本研究基于临床数据,探讨上尿路结石患者术后发生尿源性脓毒血症的影响因素,并利用多种机器学习算法构建并验证尿源性脓毒血症风险预测模型,旨在为医护人员提供一种能够早期识别与预防术后尿源性脓毒血症的评估工具,进而减轻患者的负担与病痛。

1 资料和方法

1.1 研究对象 选择 2018 年 6 月至 2023 年 6 月于

广西医科大学第六附属医院接受手术治疗的 7 464 例上尿路结石患者为研究对象。纳入标准:(1)于广西医科大学第六附属医院就诊并确诊为上尿路结石;(2)患者的手术适应证严格遵循《2015 欧洲泌尿外科协会尿石症指南》的相关标准^[8];(3)一般资料和各项临床检查资料完整;(4)年龄≥18 岁。排除标准:(1)合并泌尿系统畸形的患者;(2)合并其他系统感染性疾病的患者;(3)合并恶性肿瘤的患者;(4)合并精神系统疾病或语言表达能力欠缺的患者。

尿源性脓毒血症诊断标准:血、尿培养分离出相同病原菌且快速序贯器官衰竭评估(quick sequential organ failure assessment, qSOFA)评分≥2 分。qSOFA 评分由呼吸频率≥20 min⁻¹、意识状态变化、收缩压≤100 mmHg(1 mmHg=0.133 kPa)3 项组成,每满足一项条件计 1 分^[9]。

本研究获得广西医科大学医学伦理委员会审批(KY20230266),并在中国临床试验注册中心注册(ChiCTR2400079409)。由于本研究采用回顾性研究设计,伦理委员会免除了患者知情同意的要求。

按 7:3 的比例将研究对象随机分为训练集和验证集,同时根据患者术后是否发生尿源性脓毒血症分为尿源性脓毒血症组($n=622$)和非尿源性脓毒血症组($n=6\ 842$)。研究流程如图 1 所示。

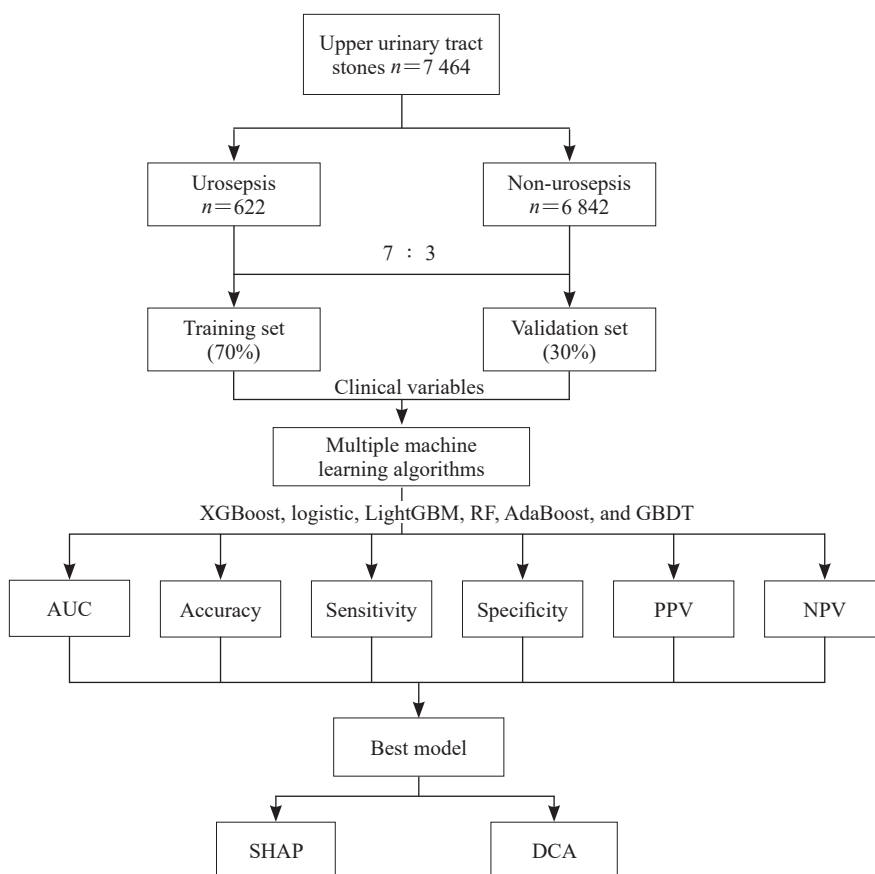


图 1 研究流程图

Fig 1 Flowchart of this study

XGBoost: Extreme gradient boosting; LightGBM: Light gradient boosting machine; RF: Random forest; AdaBoost: Adaptive boosting; GBDT: Gradient boosting decision tree; AUC: Area under curve; PPV: Positive predictive value; NPV: Negative predictive value; SHAP: Shapley additive explanation; DCA: Decision curve analysis.

1.2 一般资料及临床指标收集 纳入的多模态临床参数包括年龄、性别、BMI、手术次数、呼吸频率、体温、心率、收缩压、舒张压、日常生活活动能力量表（Barthel 指数）、静脉血栓栓塞（venous thrombus embolism, VTE）风险评分、Morse 跌倒风险评分（Morse 评分），是否患糖尿病、高血压，美国麻醉医师学会（American Society of Anesthesiologists, ASA）分级，尿培养结果，以及尿亚硝酸盐、尿白细胞检测结果。所有资料均通过医院系统检索录入。上述数据均来源于患者入院至手术前的阶段。

1.3 机器学习算法模型构建与验证 排除尿源性脓毒血症患者和无尿源性脓毒血症患者之间差异无统计学意义的临床变量。特征筛选后，采用 6 种机器学习算法基于训练集数据构建尿源性脓毒症风险预测模型，机器学习算法包括极限梯

度提升（extreme gradient boosting, XGBoost）、logistic 回归、轻量级梯度提升机（light gradient boosting machine, LightGBM）、随机森林（random forest, RF）、自适应增强（adaptive boosting, AdaBoost）、梯度提升决策树（gradient boosting decision tree, GBDT）。随后，利用验证集数据检验这些模型的有效性。

1.4 统计学处理 所有数据应用 SPSS 26.0 软件和 Python 3.4.4 软件进行统计分析。计量资料先进行正态性检验，符合正态分布的计量资料以 $\bar{x} \pm s$ 表示，两组间比较采用 Student's t 检验；不符合正态分布的计量资料以 $M(Q_1, Q_3)$ 表示，两组间比较采用 Mann-Whitney U 检验。计数资料、等级资料以例数和百分数的形式描述，两组间比较采用 χ^2 检验。特征重要性排名采用最小绝对收缩和选择算子（least absolute shrinkage and selection operator,

LASSO) 回归方法, 利用 ROC 曲线的 AUC 来量化 6 种机器学习算法模型的预测能力。使用沙普利加性解释 (Shapley additive explanation, SHAP) 工具对模型参数的贡献进行排名。采用决策曲线分析 (decision curve analysis, DCA) 评估最终模型的临床可用性。检验水准 (α) 为 0.05 (双侧检验)。

2 结 果

2.1 两组患者基线资料比较 7 464 例上尿路结石患者中术后发生尿源性脓毒血症者共 622 例

(8.33%), 未发生尿源性脓毒血症者 6 842 例 (91.67%)。622 例发生尿源性脓毒血症的患者经补液、抗感染、应用去甲肾上腺素等治疗后, 527 例 (84.73%) 患者情况转为稳定; 95 例 (15.27%) 患者转 ICU 进一步治疗, 其中 3 例 (0.48%) 患者死亡, 其余均好转后出院。与非尿源性脓毒血症组相比, 尿源性脓毒血症组患者 BMI 更低, 手术次数更多, 心率、VTE 风险评分、ASA 分级更高, Barthel 指数更低, 女性患者占比更高, 尿亚硝酸盐阳性率及尿白细胞阳性率更高, 差异均有统计学意义 (均 $P<0.05$)。见表 1。

表 1 上尿路结石患者的基线资料

Tab 1 Baseline data of patients with upper urinary tract stones

Variable	Non-urosepsis group $N=6\ 842$	Urosepsis group $N=622$	Statistic	P value
Age/year, $M(Q_1, Q_3)$	54.00 (46.00, 63.00)	55.00 (48.00, 63.75)	$Z=1.590$	0.112
Gender, n (%)			$\chi^2=85.048$	<0.001
Female	2 622 (38.32)	356 (57.23)		
Male	4 220 (61.68)	266 (42.77)		
BMI/($\text{kg}\cdot\text{m}^{-2}$), $M(Q_1, Q_3)$	23.80 (21.48, 25.97)	23.26 (20.83, 25.52)	$Z=3.761$	<0.001
Diabetes mellitus ^a , n (%)			$\chi^2=2.433$	0.119
No	6 351 (93.00)	568 (91.32)		
Yes	478 (7.00)	54 (8.68)		
Hypertension ^a , n (%)			$\chi^2=0.133$	0.716
No	5 356 (78.44)	484 (77.81)		
Yes	1 472 (21.56)	138 (22.19)		
Number of surgeries, $M(Q_1, Q_3)$	1 (1, 2)	1 (1, 2)	$Z=3.955$	<0.001
Respiratory rate/ min^{-1} , $M(Q_1, Q_3)$	20 (20, 20)	20 (20, 20)	$Z=0.928$	0.200
Temperature/ $^{\circ}\text{C}$, $M(Q_1, Q_3)$	36.6 (36.5, 36.7)	36.6 (36.5, 36.8)	$Z=1.179$	0.233
Heart rate/ min^{-1} , $M(Q_1, Q_3)$	70 (60, 81)	82 (70, 98)	$Z=15.722$	<0.001
SBP/mmHg, $M(Q_1, Q_3)$	130 (120, 142)	130 (120, 143)	$Z=0.060$	0.948
DBP/mmHg, $M(Q_1, Q_3)$	80 (74, 89)	82 (75, 89)	$Z=1.550$	0.121
Barthel index, $M(Q_1, Q_3)$	40 (30, 85)	35 (25, 80)	$Z=2.417$	0.015
VTE risk score, $M(Q_1, Q_3)$	2 (1, 3)	2 (1, 3)	$Z=2.640$	0.007
Morse score, $M(Q_1, Q_3)$	4 (4, 5)	4 (4, 5)	$Z=0.274$	0.748
ASA grade ^a , n (%)			$\chi^2=12.000$	0.007
1	1 037 (15.28)	66 (10.65)		
2	5 605 (82.60)	539 (86.94)		
3	142 (2.09)	14 (2.26)		
4	2 (0.03)	1 (0.16)		
Urinary culture ^a , n (%)			$\chi^2=1.726$	0.189
Negative	5 925 (98.80)	528 (98.14)		
Positive	72 (1.20)	10 (1.86)		
Urine nitrite, n (%)			$\chi^2=71.849$	<0.001
Negative	6 557 (95.83)	549 (88.26)		
Positive	285 (4.17)	73 (11.74)		
Urine leukocyte, n (%)			$\chi^2=173.123$	<0.001
Negative	2 637 (38.54)	121 (19.45)		
+	2 007 (29.33)	148 (23.79)		
++	349 (5.10)	47 (7.56)		
+++	388 (5.67)	48 (7.72)		
++++	1 461 (21.35)	258 (41.48)		

^a: There were missing data. The mean rank of the number of surgeries in the urosepsis group (4 052.57) was higher than that in the non-urosepsis group (3 696.26); the mean rank of the VTE risk score in the urosepsis group (1 796.35) was higher than that in the non-urosepsis group (1 632.83). 1 mmHg=0.133 kPa. BMI: Body mass index; SBP: Systolic blood pressure; DBP: Diastolic blood pressure; VTE: Venous thrombus embolism; ASA: American Society of Anesthesiologists.

2.2 上尿路结石术后发生尿源性脓毒血症预测模型的构建与验证 所有患者按 7 : 3 的比例随机分为训练集 ($n=5\,224$) 和验证集 ($n=2\,240$) , 训练集

和验证集之间除糖尿病 ($P=0.032$) 和体温 ($P=0.026$) 外, 其余各基线变量的差异均无有统计学意义 (均 $P>0.05$) 。见表 2。

表 2 训练集和验证集中上尿路结石患者的基线数据

Tab 2 Baseline data of patients with upper urinary tract stones in training and validation sets				
Variable	Training set $N=5\,224$	Validation set $N=2\,240$	Statistic	P value
Age/year, $M(Q_1, Q_3)$	54.00 (46.00, 63.00)	54.00 (46.00, 62.00)	$Z=1.039$	0.299
Gender, n (%)			$\chi^2=0.981$	0.322
Female	2 104 (40.28)	874 (39.02)		
Male	3 120 (59.72)	1 366 (60.98)		
BMI/($\text{kg}\cdot\text{m}^{-2}$), $M(Q_1, Q_3)$	23.73 (21.48, 25.91)	23.83 (21.48, 26.02)	$Z=1.013$	0.311
Diabetes mellitus ^a , n (%)			$\chi^2=4.575$	0.032
No	4 819 (92.44)	2 100 (93.83)		
Yes	394 (7.56)	138 (6.17)		
Hypertension ^a , n (%)			$\chi^2=0.282$	0.595
No	4 077 (78.22)	1 763 (78.78)		
Yes	1 135 (21.78)	475 (21.22)		
Number of surgeries, $M(Q_1, Q_3)$	1 (1, 2)	1 (1, 2)	$Z=0.312$	0.755
Respiratory rate/ min^{-1} , $M(Q_1, Q_3)$	20 (20, 20)	20 (20, 20)	$Z=1.188$	0.235
Temperature/ $^{\circ}\text{C}$, $M(Q_1, Q_3)$	36.60 (36.50, 36.70)	36.60 (36.40, 36.70)	$Z=2.226$	0.026
Heart rate/ min^{-1} , $M(Q_1, Q_3)$	71 (60, 82)	72 (62, 83)	$Z=1.287$	0.198
SBP/mmHg, $M(Q_1, Q_3)$	130.00 (120.00, 142.00)	130.00 (120.75, 143.00)	$Z=0.667$	0.505
DBP/mmHg, $M(Q_1, Q_3)$	80 (74, 89)	80 (74, 89)	$Z=0.803$	0.422
Barthel index, $M(Q_1, Q_3)$	40 (30, 85)	40 (30, 85)	$Z=0.679$	0.497
VTE risk score, $M(Q_1, Q_3)$	2 (1, 3)	2 (1, 3)	$Z=0.673$	0.501
Morse score, $M(Q_1, Q_3)$	4 (4, 5)	4 (4, 5)	$Z=0.055$	0.956
ASA grade ^a , n (%)			$\chi^2=0.265$	0.607
1	787 (15.18)	316 (14.23)		
2	4 286 (82.66)	1 858 (83.66)		
3	109 (2.10)	47 (2.12)		
4	3 (0.06)	0		
Urinary culture ^a , n (%)			$\chi^2=0.418$	0.518
Negative	4 509 (98.64)	1 944 (98.98)		
Positive	62 (1.36)	20 (1.02)		
Urine nitrite, n (%)			$\chi^2=1.477$	0.226
Negative	4 988 (95.48)	2 118 (94.55)		
Positive	236 (4.52)	122 (5.45)		
Urine leukocyte, n (%)			$\chi^2=5.810$	0.214
Negative	1 921 (36.77)	837 (37.37)		
+	1 503 (28.77)	652 (29.11)		
++	287 (5.49)	109 (4.87)		
+++	324 (6.20)	112 (5.00)		
++++	1 189 (22.76)	530 (23.66)		

^a: There were missing data. 1 mmHg=0.133 kPa. BMI: Body mass index; SBP: Systolic blood pressure; DBP: Diastolic blood pressure; VTE: Venous thrombus embolism; ASA: American Society of Anesthesiologists.

在训练集中, XGBoost模型表现最佳, AUC 值为 1.00 (95% CI 1.00~1.00) , 准确度为 1.00, 灵敏度为 1.00, 特异度为 1.00, 阳性预测值为 0.98, 阴性预测值为 1.00, F1 分数为 0.99, $Kappa$ 值为 0.99 (95% CI 0.98~1.00) 。此外, 在训练集中, LightGBM 和 RF 模型的 AUC 值均也达到 1.00, GBDT 模型次之 (AUC 值为 0.88) , 而 AdaBoost 和 logistic 回归模型的 AUC 值较低, 分别为 0.83 和 0.77。见表 3。

在验证集中, logistic 回归模型表现最佳, AUC 值为 0.76 (95% CI 0.66~0.86) , 准确度为 0.69, 灵敏度为 0.66, 特异度为 0.70, 阳性预测值为 0.16, 阴性预测值为 0.96, F1 分数为 0.25, $Kappa$ 值为 0.15 (95% CI 0.11~0.19) ; AdaBoost 和 GBDT 模型次之, AUC 值分别为 0.74 和 0.75; 而 LightGBM、XGBoost 和 RF 模型的 AUC 值分别为 0.71、0.70 和 0.68。见表 4。

表 3 训练集中 6 个预测模型对上尿路结石术后发生尿源性脓毒血症的预测能力

Tab 3 Predictive ability of 6 prediction models for urosepsis after upper urinary tract stone surgery in training set

Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 score	Kappa (95% CI)
XGBoost	1.00 (1.00, 1.00)	1.00	1.00	1.00	0.98	1.00	0.99	0.99 (0.98, 1.00)
Logistic regression	0.77 (0.74, 0.80)	0.70	0.71	0.07	0.17	0.97	0.27	0.17 (0.15, 0.19)
LightGBM	1.00 (1.00, 1.00)	0.99	1.00	0.99	0.92	1.00	0.95	0.95 (0.94, 0.96)
RF	1.00 (NaN, NaN)	1.00	1.00	1.00	0.99	1.00	1.00	1.00 (0.99, 1.00)
AdaBoost	0.83 (0.80, 0.85)	0.75	0.75	0.75	0.20	0.97	0.32	0.22 (0.21, 0.24)
GBDT	0.88 (0.85, 0.90)	0.84	0.76	0.85	0.30	0.98	0.43	0.36 (0.34, 0.38)

XGBoost: Extreme gradient boosting; LightGBM: Light gradient boosting machine; RF: Random forest; AdaBoost: Adaptive boosting; GBDT: Gradient boosting decision tree; AUC: Area under curve; CI: Confidence interval; PPV: Positive predictive value; NPV: Negative predictive value; NaN: Not a number.

表 4 验证集中 6 个预测模型对上尿路结石术后发生尿源性脓毒血症的预测能力

Tab 4 Predictive ability of 6 prediction models for urosepsis after upper urinary tract stone surgery in validation set

Model	AUC (95% CI)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 score	Kappa (95% CI)
XGBoost	0.70 (0.59, 0.81)	0.89	0.19	0.95	0.25	0.93	0.21	0.15 (0.11, 0.19)
Logistic regression	0.76 (0.66, 0.86)	0.69	0.66	0.70	0.16	0.96	0.25	0.15 (0.11, 0.19)
LightGBM	0.71 (0.60, 0.81)	0.89	0.23	0.95	0.27	0.94	0.24	0.19 (0.13, 0.24)
RF	0.68 (0.56, 0.79)	0.91	0.15	0.97	0.30	0.93	0.20	0.15 (0.11, 0.19)
AdaBoost	0.74 (0.63, 0.84)	0.73	0.63	0.74	0.17	0.96	0.26	0.16 (0.13, 0.20)
GBDT	0.75 (0.64, 0.85)	0.80	0.56	0.82	0.21	0.96	0.30	0.22 (0.17, 0.26)

XGBoost: Extreme gradient boosting; LightGBM: Light gradient boosting machine; RF: Random forest; AdaBoost: Adaptive boosting; GBDT: Gradient boosting decision tree; AUC: Area under curve; CI: Confidence interval; PPV: Positive predictive value; NPV: Negative predictive value.

综合分析上述 6 个模型数据,训练集和验证集中按 AUC 值排序最优模型分别为 XGBoost 和 logistic 回归模型,两者不一致。XGBoost 模型在训练阶段表现优异,但在验证阶段性能下降明显,提示其可能存在过拟合现象;而 logistic 回归模型在训练和验证阶段的性能较为一致,表明其具有较好的稳定性。因此,本研究最终选择 logistic 回归模

型作为最终预测模型。

2.3 模型可解释性 采用 SHAP 方法,通过计算每个变量对模型预测能力的贡献度来解释最终模型的输出。结果显示,变量贡献度由大到小依次为心率、尿白细胞、性别、BMI、Barthel 指数、VTE 风险评分、尿亚硝酸盐、手术次数和 ASA 分级。见图 2。

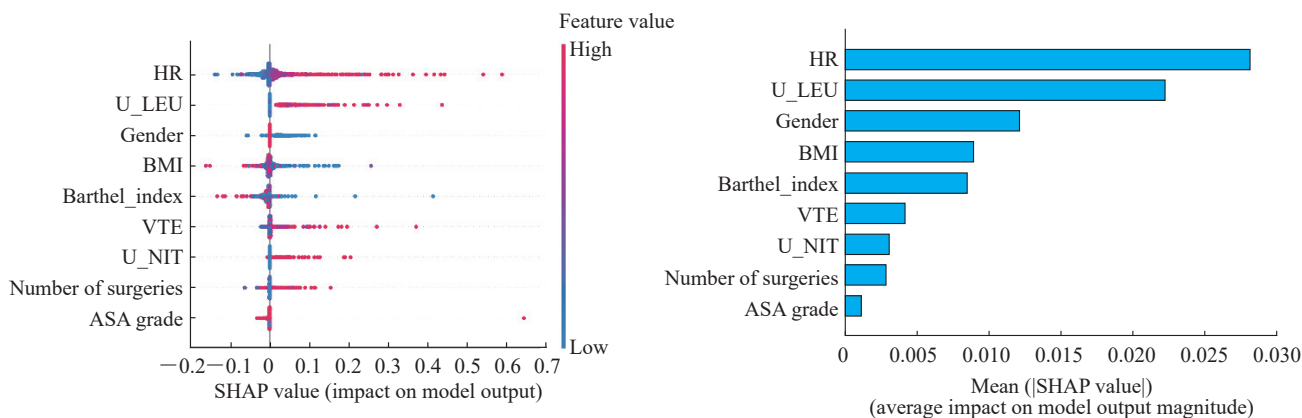


图 2 SHAP 方法分析上尿路结石术后发生尿源性脓毒血症预测模型参数的重要性排序

Fig 2 Importance of parameters of prediction model for urosepsis after upper urinary tract stone surgery by SHAP method

SHAP: Shapley additive explanation; HR: Heart rate; U_LEU: Urine leukocyte; BMI: Body mass index; VTE: Venous thrombus embolism; U_NIT: Urine nitrite; ASA: American Society of Anesthesiologists.

2.4 预测模型的临床可用性评价 通过绘制DCA曲线评估模型的临床效益。其中,“treat all”曲线表示假设所有上尿路结石患者术后均发生尿源性脓毒血症并对其进行干预所获得的效益,“treat none”曲线表示假设所有上尿路结石患者术后均未发生尿源性脓毒血症且未进行任何干预时的效益(即为0),“logistic”曲线则反映了基于logistic回归模型将术后患者分为尿源性脓毒血症组和非尿源性脓毒血症组后的临床效益。DCA曲线越靠近图右上方(即AUC越大),表明预测模型的临床效益越高。根据DCA曲线结果,logistic回归预测模型在阈值概率为10%~100%范围内表现出正向临床收益(图3),即在这些概率区间内logistic回归模型具有较高的临床应用价值。

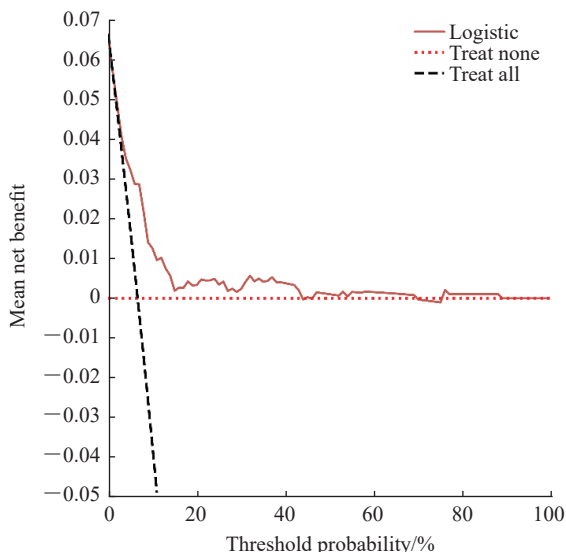


图3 上尿路结石术后发生尿源性脓毒血症 logistic 回归预测模型的DCA

Fig 3 Decision curve analysis (DCA) of logistic regression prediction model for urosepsis after upper urinary tract stone surgery

通过绘制训练集和验证集的校准曲线评估模型的校准度,其中X轴表示预测上尿路结石患者术后发生尿源性脓毒血症的概率,Y轴表示实际上尿路结石患者术后发生尿源性脓毒血症的概率;“perfectly calibrated”线表示预测概率等于实际概率的情况,偏离该曲线越远说明预测结果误差越大;“logistic”曲线表示logistic回归模型实际预测上尿路结石患者术后发生尿源性脓毒血症的概率。如图4所示,该预测模型的校准度较好,预测概率与实际概率具有较为满意的一致性。

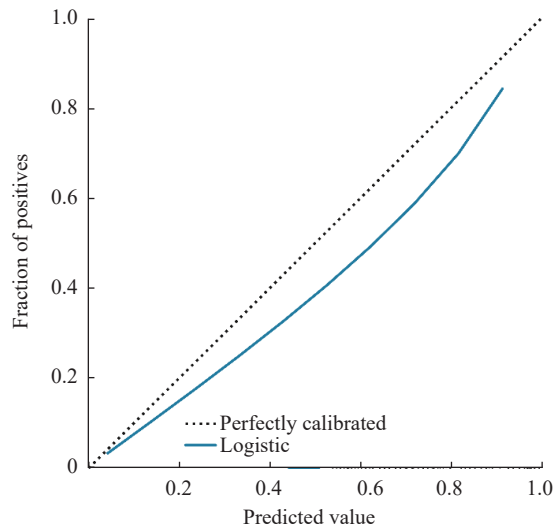


图4 上尿路结石术后发生尿源性脓毒血症 logistic 回归预测模型的校准曲线

Fig 4 Calibration curve of logistic regression prediction model for urosepsis after upper urinary tract stone surgery

3 讨论

脓毒血症是机体感染过程中的异常全身炎症反应,死亡率较高,其中尿源性脓毒血症占比为9%~30%^[10]。有研究表明,在发生尿路感染性休克的患者中,78%的男性和54%的女性脓毒血症的发生与尿路梗阻有关^[10]。尿源性脓毒血症可引发全身炎症反应、感染性休克,甚至死亡等不良后果,还可能遗留神经和认知后遗症,并且较高的住院费给患者带来较大负担。本研究纳入的7 464例上尿路结石手术治疗患者中,622例术后发生尿源性脓毒血症,发生率为8.33%。因此研究尿源性脓毒血症的影响因素、构建早期诊断系统十分必要,这不仅可以提高上尿路结石患者术后发生尿源性脓毒血症的早期诊断率,还可减轻患者病痛。

本研究结果显示,BMI、性别、手术次数、心率、Barthel指数、VTE风险评分、ASA分级、尿亚硝酸盐及尿白细胞为上尿路结石患者术后发生尿源性脓毒血症的影响因素。杨金儿等^[11]研究表明,BMI越高,患者发生尿源性脓毒血症的概率越大。然而,在本研究中,发生尿源性脓毒血症患者的BMI低于未发生尿源性脓毒血症者。这可能与广西本地的饮食习惯有关。低BMI的患者,尤其是有营养不良或慢性疾病的患者,体力较弱,免疫系统功能较差,术后的恢复能力较差、感染风险增加。女性尿道较短,细菌易经尿道口逆行,从而发生尿路感染,导致尿源性脓毒血症的发生风险上

升。因此,需要更加关注患上尿路结石的肥胖或女性患者,以减少手术治疗后尿源性脓毒血症的发生风险。文献报道,手术次数与脓毒血症发生显著相关^[12],本研究结果与之一致。

尽管机器学习算法因模型复杂而功能强大,但它仍存在难以直接解释的局限,即所谓的“黑匣子”问题^[13]。本研究基于SHAP分析发现,心率、Barthel指数、VTE风险分级及ASA评分也是上尿路结石患者术后发生尿源性脓毒血症的预警因素,但是关于这几个指标在尿源性脓毒血症中的研究较少,需要进一步研究其临床诊断价值。心动过速是严重脓毒症和脓毒症休克期间死亡的独立危险因素,心率升高可能会损害左心室舒张充盈,导致每搏输出量减少,同时损害冠状动脉血流并增加心肌需氧量^[14]。相关研究表明,将心率降低至 95 min^{-1} 可显著降低脓毒血症休克患者的死亡率^[15]。Deulofeu等^[16]研究发现,低Barthel指数评分是预测菌血症的危险因素。Hsieh等^[17]研究显示,低Barthel指数评分与急诊科入院患者7 d死亡率独立相关。ASA分级是术后肺部感染的7个术前预测因素之一,还是儿童非心脏手术后发生急性肾损伤的独立危险因素^[18]。

本研究构建的上尿路结石患者术后发生尿源性脓毒血症预测模型有一定的预测价值,在XGBoost、logistic回归、LightGBM、RF、AdaBoost、GBDT等6种模型中,最终选出logistic回归模型为最佳预测模型,AUC值为0.76(95% CI 0.66~0.86),准确度为0.69、灵敏度和特异度分别为0.66和0.70,且logistic回归模型的校准曲线也提示该模型具有较好的校准度,可为临床上减少上尿路结石患者术后尿源性脓毒血症的发生提供指导。本研究构建的logistic回归模型优点在于所有数据基于结构化临床数据,能够直接嵌入病历系统,对高危患者进行预警。然而,本研究亦存在不足之处,如研究样本均来自同一医院,代表度和可信度较为局限,在未来的研究中将考虑加入多中心数据进行验证,以增强研究结果的外推性和可靠性。此外,本预测模型的DCA结果表明模型的临床可用性仍待提高,模型构建仍需完善,未来将开展进一步研究,提高模型的精准性和可用性,更好地满足临床需求。同时,本研究将严重心、脑、肺器质性疾病的患者主要定义为不适合手术者,未纳入研究,这也导致模

型对高危人群的预测能力不足,未来将开展前瞻性研究以开发适用于高危人群的预测模型。最后,本研究纳入的临床参数有待扩充,如结石手术史、手术方式、临床检验数据、影像学检查结果等。未来研究应基于更多的临床多模态参数,提高模型的预测准确度。

[参 考 文 献]

- [1] MOE O W. Kidney stones: pathophysiology and medical management[J]. *Lancet*, 2006, 367(9507): 333-344. DOI: 10.1016/S0140-6736(06)68071-9.
- [2] TAN S, YUAN D, SU H, et al. Prevalence of urolithiasis in China: a systematic review and meta-analysis[J]. *BJU Int*, 2024, 133(1): 34-43. DOI: 10.1111/bju.16179.
- [3] 麦海星, 乔庐东, 吴文起, 等. 上尿路感染性结石诊断与治疗中国专家共识(2024版)[J]. *微创泌尿外科杂志*, 2024, 13(4): 217-223. DOI: 10.19558/j.cnki.10-1020/r.2024.04.001.
- [4] BONKAT G, CAI T, VEERATTERAPILLAY R, et al. Management of urosepsis in 2018[J]. *Eur Urol Focus*, 2019, 5(1): 5-9. DOI: 10.1016/j.euf.2018.11.003.
- [5] GROSSO A A, SESSA F, CAMPI R, et al. Intraoperative and postoperative surgical complications after ureteroscopy, retrograde intrarenal surgery, and percutaneous nephrolithotomy: a systematic review[J]. *Minerva Urol Nephrol*, 2021, 73(3): 309-332. DOI: 10.23736/S2724-6051.21.04294-4.
- [6] SEYMOUR C W, LIU V X, IWASHYNA T J, et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3)[J]. *JAMA*, 2016, 315(8): 762-774. DOI: 10.1001/jama.2016.0288.
- [7] WEISS S L, FITZGERALD J C, BALAMUTH F, et al. Delayed antimicrobial therapy increases mortality and organ dysfunction duration in pediatric sepsis[J]. *Crit Care Med*, 2014, 42(11): 2409-2417. DOI: 10.1097/CCM.0000000000000509.
- [8] BULTITUDE M, SMITH D, THOMAS K. Contemporary management of stone disease: the new EAU urolithiasis guidelines for 2015[J]. *Eur Urol*, 2016, 69(3): 483-484. DOI: 10.1016/j.eururo.2015.08.010.
- [9] SEYMOUR C W, LIU V X, IWASHYNA T J, et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3)[J]. *JAMA*, 2016, 315(8): 762-774. DOI: 10.1001/jama.2016.0288.
- [10] GULICIUC M, MAIER A C, MAIER I M, et al. The urosepsis—a literature review[J]. *Medicina (Kaunas)*, 2021, 57(9): 872. DOI: 10.3390/medicina57090872.

- [11] 杨金儿,苏飞月,蒋莉,等.上尿路结石患者术后并发尿源性脓毒血症的影响因素分析及护理对策[J].全科医学临床与教育,2023,21(10):952-954,957. DOI: 10.13558/j.cnki.issn1672-3686.2023.010.028.
- [12] SHEN R, MING S, QIAN W, et al. A novel post-percutaneous nephrolithotomy sepsis prediction model using machine learning[J]. BMC Urol, 2024, 24(1): 27. DOI: 10.1186/s12894-024-01414-x.
- [13] HU J, XU J, LI M, et al. Identification and validation of an explainable prediction model of acute kidney injury with prognostic implications in critically ill children: a prospective multicenter cohort study[J]. EClinicalMedicine, 2024, 68: 102409. DOI: 10.1016/j.eclinm.2023.102409.
- [14] BEDET A, VOIRIOT G, TERNACLE J, et al. Heart rate control during experimental sepsis in mice: comparison of ivabradine and β -blockers[J]. Anesthesiology, 2020, 132(2): 321-329. DOI: 10.1097/ALN.0000000000003045.
- [15] SCHORTGEN F, CHARLES-NELSON A, BOUADMAL, et al. Respective impact of lowering body temperature and heart rate on mortality in septic shock: mediation analysis of a randomized trial[J]. Intensive Care Med, 2015, 41(10): 1800-1808. DOI: 10.1007/s00134-015-3987-7.
- [16] DEULOFEU F, CERVELLÓ B, CAPELL S, et al. Predictors of mortality in patients with bacteremia: the importance of functional status[J]. J Am Geriatr Soc, 1998, 46(1): 14-18. DOI: 10.1111/j.1532-5415.1998.tb01007.x.
- [17] HSIEH M J, HSU N C, LIN Y F, et al. Developing and validating a model for predicting 7-day mortality of patients admitted from the emergency department: an initial alarm score by a prospective prediction model study[J]. BMJ Open, 2021, 11(1): e040837. DOI: 10.1136/bmjopen-2020-040837.
- [18] HAWKINS J, MPODY C, CORRIDORE M, et al. Risk factors and consequences of acute kidney injury after noncardiac surgery in children[J]. Anesth Analg, 2022, 135(3): 625-632. DOI: 10.1213/ANE.0000000000005901.

[本文编辑] 杨亚红