论著·生物医学工程

机器学习驱动的个性化氨甲环酸用药推荐改善骨科手术患者围手术期预后:一项基于大样本数据库的研究

李 剑[1],周 秘[1],刘 祥[2],朱易紫婷[2],舒 欣[2],张序昊[2],贺文泉[2] [1]重庆市合川区人民医院麻醉科,重庆;[2]陆军军医大学(第三军医大学)第一附属医院麻醉科,重庆

[摘要] **目的** 基于机器学习构建骨科手术围手术期氨甲环酸使用的个性化推荐策略,以减少围手术期出血及相关并发症,改善患者预后。**方法** 从INSPIRE数据库中纳入了11 727例骨科手术患者,通过多重插补方法处理缺失数据,并采用Boruta分析筛选出相关特征变量。构建了多种机器学习模型,包括梯度提升机(gradient boosting machine, GBM)、广义线性模型(generalized linear model, GLM)、极端梯度提升机(eXtreme gradient boosting, XGBoost)、K最近邻(k-nearest neighbors, KNN)、神经网络(neural network, NNET)、朴素贝叶斯(naive bayes, NB)及随机森林(random forest, RF),评估各模型在预测术中出血及术后住院时间延长的效能。筛选最优模型后采用加权集成法进行整合,以达到最佳预后为目标对氨甲环酸使用策略进行推荐。通过测试集验证及与医生决策对比,完成模型性能评估。**结果** 针对术中出血,RF模型的受试者操作曲线下面积(area under the subject curve, AUC)达到0.73,显著优于其他模型;而针对术后住院时间延长,XGBoost模型表现得最佳(AUC=0.84)。基于上述最佳模型,实施了集成策略,结果显示符合推荐策略的患者术中出血量降低,术后住院时间缩短。**结论** 氨甲环酸的使用与术中出血和术后住院时间相关,应用构建的模型进行个性化决策推荐能有效改善患者预后。

[关键词] 骨科手术;围手术期;氨甲环酸;策略推荐

[中图法分类号] R319;R687.3;R973.1 [文献标志码] A

Machine learning-driven personalized tranexamic acid administration recommendations improve perioperative outcomes in orthopedic surgery patients: A large-scale database study

LI Jian[1], ZHOU Mi[1], LIU Xiang[2], ZHU Yiziting[2], SHU Xin[2], ZHANG Xuhao[2], HE Wenquan[2] [1]Department of Anesthesiology, Hechuan People's Hospital, Chongqing; [2]Department of Anesthesiology, First Affiliated Hospital, Army Medical University (Third Military Medical University), Chongqing, China

[**Abstract**] Objective To develop a personalized recommendation strategy for tranexamic acid administration during the perioperative period of orthopedic surgery based on machine learning, aiming to reduce perioperative bleeding and related complications and improving clinical outcomes. Methods A total of 11 727 patients undergoing orthopedic surgery from the INSPIRE database were subjected in this study. Missing data were handled using multiple imputation methods, and relevant feature variables were screened using Boruta analysis. We constructed various machine learning models, including Gradient Boosting

[通信作者] 贺文泉,E-mail: hewenquan7@tmmu.edu.cn

2097-0927/Copyright © 2025, 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/)

引用格式:李剑,周秘,刘祥,等. 机器学习驱动的个性化氨甲环酸用药推荐改善骨科手术患者围手术期预后:一项基于大样本数据库的研究[J]. 陆军军医大学学报,2025,47(22):2868-2880. DOI:10.16016/j.2097-0927.202507025.

Citation: LI J, ZHOU M, LIU X, *et al*. Machine learning-driven personalized tranexamic acid administration recommendations improve perioperative outcomes in orthopedic surgery patients: A large-scale database study[J]. J Army Med Univ, 2025, 47(22): 2868-2880. DOI: 10.16016/j.2097-0927.202507025.

Machine (GBM), Generalized Linear Model (GLM), eXtreme Gradient Boosting (XGBoost), K-Nearest Neighbors (KNN), Neural Network (NNET), Naive Bayes (NB), and Random Forest (RF), to evaluate their performance in predicting intraoperative bleeding and prolonged postoperative length of hospital stay. The optimal model was then selected and further integrated using a weighted ensemble, aiming to achieve the best prognosis by recommending usage strategies for tranexamic acid. The predictive performance of the constructed model was then verified against the testing set, and compared with the physician decision-making to complete the evaluation. **Results** In predicting intraoperative bleeding, the RF model achieved an area under the receiver operating characteristic curve (AUC) of 0.73, which was significantly better than other models. In predicting the prolonged postoperative length of hospital stay, the XGBoost model performed the best, with an AUC value of 0.84. Based on the above best-performing models, an ensemble strategy was implemented. The patients who followed the recommended strategy had reduced intraoperative bleeding and shorter postoperative length of hospital stay. **Conclusion** The use of tranexamic acid is associated with intraoperative bleeding and postoperative length of hospital stay. Personalized decision-making recommendation based on our constructed model can effectively improve the outcomes of the patients undergoing orthopedic surgery.

[**Key words**] orthopedic surgery; perioperative period; tranexamic acid; strategy recommendation

Corresponding author: HE Wenquan, E-mail: hewenquan7@tmmu.edu.cn

骨科手术是治疗各种骨骼和关节疾病的重要医疗干预,广泛应用于创伤、退行性疾病及先天性畸形等情况。这类手术通常涉及较大的软组织剥离和骨骼操作,因此出血风险相对较高。术中出血不仅会影响手术视野,还可能导致患者需要输血,从而增加并发症的风险[1-2]。因此,控制出血是骨科手术成功的重要因素之一。

氨甲环酸(tranexamic acid, TXA)是一种抗纤溶药物,广泛应用于各种外科手术中以减少出血。在骨科手术中,TXA的使用被认为可以有效降低术中和术后的出血量[3-6]。然而,TXA的应用也有导致癫痫发作、血栓栓塞事件和急性肾损伤等不良反应,且在不同类型骨科手术中的具体应用时机、途径,以及应用效果仍存在争议[7-9]。目前,外科手术中TXA的推荐应用方案虽基于手术类型(如关节置换术、脊柱手术、创伤骨科手术)及患者静脉血栓栓塞症风险分层进行制定,但其本质还是依据医生主观经验判断,难以推广。此外,TXA决策推荐涉及多种相关因素,包括患者基线信息、手术细节和术前检查等大量指标,难以权衡相关指标以及平衡多种结局。因此,如何在骨科手术中实现TXA的科学、合理应用,已成为一个亟待解决的临床决策难题。

机器学习算法凭借其强大的数据处理与模式识别能力,能够高效整合大规模、多维度临床数据,并自动挖掘其中隐含的深层次规律,从而为制定个性化的TXA用药策略提供数据驱动的指导。本文旨在基于机器学习技术,构建一套用于骨科手术TXA用药辅助决策的模型,以协助临床医生优化TXA给药方案,以期在降低相关并发症风险的同时,最终提升骨科手术的整体疗效与安全性。

1 方法

1.1 数据采集

本研究数据来源于INSPIRE数据库[10]。该数据库是一个公开的围手术期医学研究数据集,其包含了2011—2020年在韩国一家学术机构接受手术的约130 000例病例的麻醉相关信息(约占该机构同期总手术量的50%)。数据从PhysioNet平台下载获取,鉴于所用数据为公开的匿名数据,本研究豁免患者知情同意。本研究旨在对接受骨科手术患者的TXA使用策略进行分析。研究对象的纳入标准包括:①接受骨科手术;②年龄≥18岁。排除标准包括:①同一次住院期间接受多次手术者;②接受体外循环手术的患者;③关键指标——血细胞比容(hematocrit, Hct)数据缺失的患者;④手术时间或麻醉时间记录有误的患者;⑤ASA分级>Ⅴ级的患者。患者纳入流程如图1所示。骨科手术病例的识别根据ICD-10-PCS编码进行,具体纳入了前两位为0L、0M、0P、0Q、0R、0S的编码。

为促进临床推广应用,本研究纳入了与围手术期出血相关、能有效反映患者病情变化且临床常用易测的指标。其中,TXA作为研究关注的干预措施,贯穿整个围手术期,并根据用药时机将使用策略定义为四类:术前术中均不使用、仅术前使用、仅术中使用、术

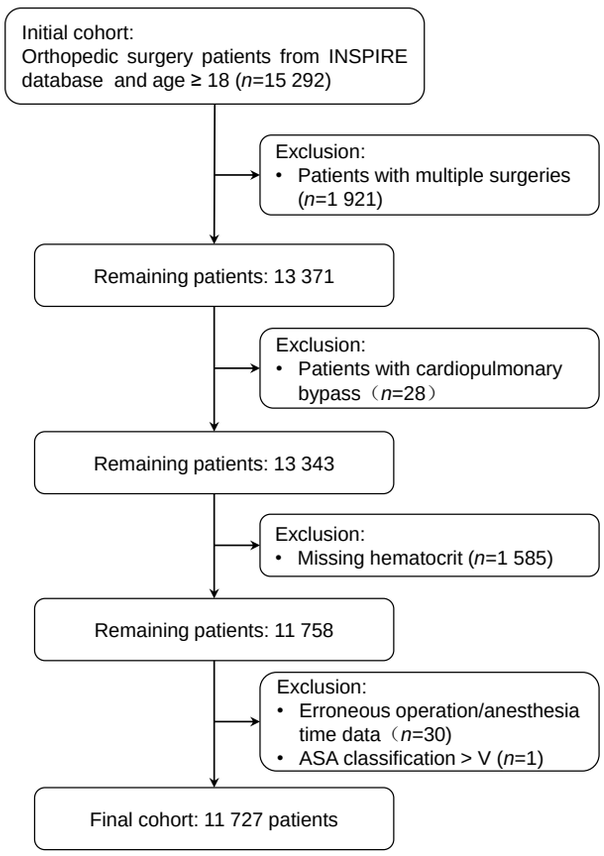


图1 患者纳入与排除流程图

Figure 1 Patients inclusion and exclusion flowchart

前术中均使用,其余分析指标详见表1。

术前使用抗凝药物定义为手术前7 d内使用肝素、华法林、低分子量肝素、阿哌沙班、利伐沙班、依度沙班、达比加群、磺达肝癸钠;术前使用抗血小板药物定义为手术前7 d内使用阿司匹林、氯吡格雷、替罗非班、替格瑞洛等;术中出血定义为术中出血≥200 mL,术后住院时间延长定义为≥10 979 min(训练集均值);术中失血量通过Gross[11]方程算,总失血量(total blood loss, TBL)=患者的总血容量(patient's total blood volume, PBV)×(pre-Hct − post-Hct)/ave-Hct。Pre-Hct代表患者手术前的Hct,Post-Hct是指手术后的Hct,ave-Hct表示手术前后的平均Hct。PBV使用Nadler公式[12]计算:PBV=K1×[身高(m)]^3+K2×体质量(kg)+K3。对于男性,K1=0.366 9,K2=0.0321 9,K3=0.604 1,而对于女性,K1=0.356 1,K2=0.033 08,K3=0.183 3。

1.2 数据处理

本研究的整个过程包括数据预处理、特征选择、模型拟合和评估,以及TXA的用药推荐(图2)。

在数据预处理阶段,对于缺失值>40%的指标进行删除,分类变量采用众数插补,其余缺失值采用R语言MICE包进行多重插补,以确保数据的完整性和准确性。随后,为了消除不同量纲特征对模型训练的影响,对连续变量进行了归一化处理和标准化,使得这些特征值被缩放到一个统一的范围内。

1.3 特征筛选以及模型构建

数据集随机分为训练集和测试集,分别为80%和20%,使用训练数据集进行特征选择和模型拟合,在测试数据集进行模型评估进行内部验证。

使用Boruta进行特征选择,有助于降低模型的复杂度,减少过拟合的风险。针对每个目标(术中出血、术后住院时间延长),分别训练模型。模型构建选择7种

表1 最终分析指标

Table 1 Final indicators for analysis

	Category	Indicator
	General information	Age, gender, BMI
	Comorbidities	Diabetes, hypertension, coronary heart disease
Preoperative	Medication	Anticoagulant drugs, antiplatelet drugs, tranexamic acid
	Laboratory test	Creatinine, BUN, phosphorus, chlorine, potassium, sodium, ALT, AST, ALP, TBIL, albumin, STP, calcium, fibrinogen, APTT, INR, glucose, LYMPH%, NEUT%, PLT, hematocrit, hemoglobin, WBC, CRP
	Anesthesia	ASA classification, anesthesia method, anesthesia time
Intraoperative	Surgery	Surgical site, surgical time, emergency surgery
	Medication	Tranexamic acid
Postoperative	outcome	HCT, length of hospital stay, extended length of hospital stay

BMI: Body mass index; BUN: Blood urea nitrogen; APTT: Activated partial thromboplastin time; INR: International normalized ratio; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; ALP: Alkaline phosphatase; TBIL: Total bilirubin; STP: Serum total protein; LYMPH%: Lymphocyte percentage; NEUT%: Neutrophil percentage; PLT: Platelet count; WBC: White blood cell count; CRP: C-reactive protein

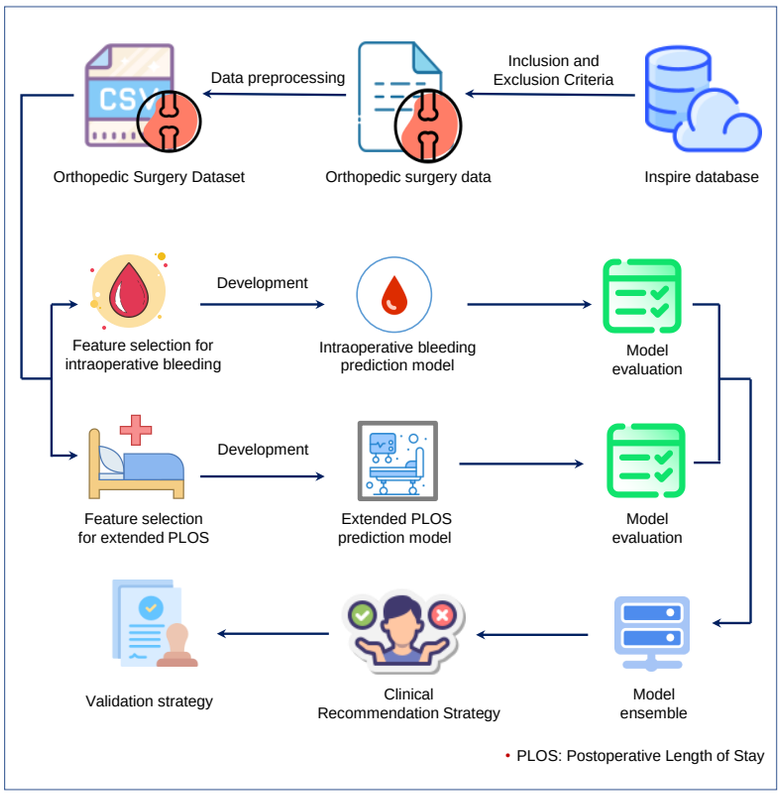


图2 研究流程示意图

Figure 2 Schematic diagram of research process

传统的机器学习算法,包括梯度提升机(gradient boosting machine, GBM)、广义线性模型(generalized linear model, GLM)、极端梯度提升(eXtreme gradient boosting, XGBoost)、神经网络(neural network, NNET)、K最近邻(k-nearest neighbors, KNN)以及随机森林(random forest, RF)。构建完成后,通过受试者操作曲线下面积(area under the subject curve, AUC)、灵敏度、特异性、阳性预测值(positive predictive value, PPV)、阴性预测值(negative predictive value, NPV)、准确性(accuracy, ACC)、F1评分以及Kappa值来评价模型的性能。最后,通过决策曲线分析(decision curve analysis, DCA)对模型进行评价,评估模型的临床效益和效用。同时,还利用校准曲线评价模型的校准度。

由于推荐决策需要考虑多目标的结局,针对术中出血和术后住院时间延长这两个结局分别构建最佳模型后,进行模型集成。集成策略选择加权集成[13],根据治疗目的,希望术中出血风险最小,也希望术后住院时间不会延长,给予这两个结局同等权重,即给予每个模型的预测结果相同的权重(权重均为0.5)。对于每个患者,计算两个模型的预测结果的加权平均值,作为最终的综合预测结果。具体公式为:综合预测结果=0.5×术中出血风险预测结果+0.5×术后住院时间延长预测结果。最后根据综合预测结果找到风险最小的治疗策略,进行决策推荐。

1.4 特征重要性分析

为分别探究两个结局的最优模型,本研究分析了影响其预测的特征重要性,采用SHAP方法以分析特征对模型输出的影响。SHAP法基于博弈论中的Shapley值概念,该概念旨在公平地分配合作博弈中各方的贡献度。在机器学习背景下,Shapley值被用以量化每个特征对模型最终预测的贡献程度。

1.5 决策效果验证

根据构建的模型,基于更好的预后选择治疗方案,将模型推荐决策和医生决策对比,利用测试集进行决策推荐,符合推荐策略的患者和不符合推荐策略的术中平均出血量和术后平均住院时间进行对比,以判断模型对患者预后的改善。

1.6 统计学分析

本研究所涉及的全部数据分析工作,均在R 4.2.2语言环境中完成。连续变量用$\bar{x}\pm s$或$M(P_{25},P_{75})$表示,使用t检验或Mann-Whitney U检验进行组间比较。分类变量用频率和百分比表示,采用卡方检验或Fisher精确检验进行分析。所有统计分析遵循双尾检验的原则,$P<0.05$时认为差异具有统计学意义。

2 结果

2.1 数据基线特征

本研究从INSPIRE数据库纳入了11 727例患者,其中训练集9 383例患者,其中5 161例出现了术中出血>200 mL,2 476例出现了术后住院时间延长;测试集2 344例患者,其中1 315例出现了术中出血>200 mL,616例出现了术后住院时间延长,见表2。

表2 原始数据基线资料[$M(P_{25},P_{75})$,例(%)]

Table 2 Baseline data of original information [$M(P_{25},P_{75})$, n (%)]

Indicators	Training dataset(n=9 383)	Test dataset(n=2 344)	P
Age /years	60.00 (45.00, 70.00)	60.00 (45.00, 70.00)	0.513
BMI/(kg/m^2)	24.24 (21.80, 26.78)	24.24 (21.78, 26.81)	0.849
Surgical time /min	95.00 (60.00, 150.00)	95.00 (65.00, 155.00)	0.327
Anesthesia time /min	140.00 (100.00, 205.00)	140.00 (105.00, 205.00)	0.456
Creatinine/(mg/dL)	0.80 (0.68, 0.92)	0.80 (0.68, 0.92)	0.909
BUN/(mg/dL)	15.00 (12.00, 18.00)	15.00 (12.00, 18.00)	0.413
Phosphorus/(mg/dL)	3.60 (3.20, 3.90)	3.60 (3.20, 3.90)	0.831
Chlorine/(mmol/L)	103.00 (102.00, 106.00)	103.00 (102.00, 106.00)	0.932
Potassium/(mmol/L)	4.00 (3.60, 4.40)	4.00 (3.60, 4.40)	0.325
Sodium/(mmol/L)	139.00 (137.00, 141.00)	139.00 (137.00, 141.00)	0.519
ALT/(UI/L)	20.00 (13.00, 30.00)	20.00 (13.00, 30.00)	0.768
AST/(UI/L)	22.00 (18.00, 28.00)	22.00 (18.00, 26.00)	0.043
ALP/(UI/L)	68.00 (55.00, 86.00)	68.00 (55.00, 86.00)	0.367
TBIL/(mg/dL)	0.50 (0.50, 0.80)	0.50 (0.50, 0.80)	0.827
Albumin/(g/dL)	4.20 (4.00, 4.40)	4.20 (4.00, 4.40)	0.811
STP/(g/dL)	7.00 (6.80, 7.50)	7.00 (6.80, 7.50)	0.524
Calcium/(mg/dL)	9.20 (8.90, 9.60)	9.20 (8.90, 9.60)	0.363
Fibrinogen/(mg/dL)	313.00 (276.00, 376.00)	313.00 (276.00, 376.00)	0.429
APTT/s	30.60 (29.10, 33.00)	31.20 (29.10, 33.00)	0.593
INR	0.98 (0.93, 1.03)	0.98 (0.93, 1.03)	0.154
Glucose/(mg/dL)	105.00 (96.00, 126.00)	105.00 (96.00, 120.00)	0.031
LYMPH/%	31.30 (22.10, 37.80)	28.70 (22.10, 37.80)	0.438
NEUT/%	61.00 (53.10, 69.60)	61.00 (53.10, 69.60)	0.661
PLT/(×10^9)	245.00 (200.00, 291.00)	245.00 (200.00, 291.00)	0.270
Hematocrit/%	39.10 (35.20, 42.00)	39.10 (36.10, 42.00)	0.286
Hemoglobin/(g/dL)	13.20 (12.20, 15.00)	13.20 (12.20, 15.00)	0.682
WBC (×10^9)	6.67 (5.58, 8.38)	6.67 (5.58, 8.38)	0.024
CRP/(mg/dL)	0.11 (0.02, 0.50)	0.11 (0.02, 0.50)	0.791
Gender			0.536
Female	5 225 (55.69)	1 288 (54.95)	
Male	4 158 (44.31)	1 056 (45.05)	
ASA classification			0.662
Ⅰ	3 136 (33.30)	800 (34.93)	
Ⅱ	5 191 (56.78)	1 283 (56.04)	
Ⅲ	783 (8.56)	200 (8.73)	
Ⅳ	32 (0.35)	6 (0.26)	
Ⅴ	1 (0.01)	1 (0.04)	

表2续下页

续表2 原始数据基线资料[$M(P_{25},P_{75})$, 例(%)]

Table 2(continued) Baseline data of original information [$M(P_{25},P_{75})$, n(%)]

Indicators	Training dataset(n=9 383)	Test dataset(n=2 344)	P
Emergency surgery			0.908
Non-emergency	8 376 (89.27)	2 095 (89.38)	
Emergency	1 007 (10.73)	249 (10.62)	
Anesthesia method			0.451
MAC	24 (0.26)	6 (0.26)	
Neuraxial	3 846 (40.99)	1 004 (42.83)	
Regional	42 (0.45)	10 (0.43)	
General	5 471 (58.30)	1 324 (56.48)	
Surgical Site			0.884
Tendon	652 (6.95)	159 (6.78)	
Bursa and ligament	341 (3.63)	75 (3.20)	
Upper limb bone	2 389 (25.46)	601 (25.64)	
Lower limb bone	707 (7.53)	178 (7.59)	
Upper limb joint	1 809 (19.28)	440 (18.77)	
Lower limb joint	3 485 (37.14)	891 (38.01)	
Tranexamic acid			0.447
Neither before nor during the operation	8 422 (89.76)	2 104 (89.76)	
Only before the operation	540 (5.76)	147 (6.27)	
Only during the operation	371 (3.95)	79 (3.37)	
Both before and during the operation	50 (0.53)	14 (0.60)	
Preoperative use of anticoagulants drugs			0.551
No	9 090 (96.88)	2 277 (97.14)	
Yes	293 (3.12)	67 (2.86)	
Preoperative use of antiplatelet drugs			0.862
No	9 072 (96.69)	2 264 (96.59)	
Yes	311 (3.31)	80 (3.41)	
Hypertension			0.649
No	8 491 (90.49)	2 129 (90.83)	
Yes	892 (9.51)	215 (9.17)	
Diabetes			0.910
No	8 943 (95.31)	2 236 (95.39)	
Yes	440 (4.69)	108 (4.61)	
Coronary heart disease			0.579
No	8 892 (94.77)	2 214 (94.45)	
Yes	491 (5.23)	130 (5.55)	
Intraoperative bleeding			0.351
Intraoperative bleeding ≤200 mL	4 222 (45.00)	1 029 (43.90)	
Intraoperative bleeding >200 mL	5 161 (55.00)	1 315 (56.10)	
Extended PLOS			0.936
No	6 907 (73.61)	1 728 (73.72)	
Yes	2 476 (26.39)	616 (26.28)	

BMI: Body mass index; BUN: Blood urea nitrogen; APTT: Activated partial thromboplastin time; INR: International normalized ratio; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; ALP: Alkaline phosphatase; TBIL: Total bilirubin; STP: Serum total protein; LYMPH%: Lymphocyte percentage; NEUT%: Neutrophil percentage; PLT: Platelet count; WBC: White blood cell count; CRP: C-reactive protein; MAC: Monitored anesthesia care; PLOS: Postoperative length of stay

剔除数据缺失率超过40%的患者后,有25个变量存在缺失值。使用MICE包进行多重插补后,插补后的数据与原始数据之间的分布差异无统计学意义,见表3。

2.2 特征筛选结果

经过Boruta分析,针对术中出血风险,筛选出36个变量作为模型特征输入变量,包括红细胞比容、钾离子、血清总蛋白、麻醉方式、手术部位、血红蛋白、手术时间、麻醉时间、白蛋白、钙离子、钠离子、抗血小板药、纤维蛋白原、诊断年龄、淋巴细胞百分比、血葡萄糖、中性粒细胞百分比、C反应蛋白、TXA、氯离子、碱性磷酸酶、抗凝血药、体质指数、总胆红素、冠心病、凝血酶原时间国际标准化比值、血小板、白细胞、丙氨酸氨基转移酶、活化部分凝血活酶时间、天门冬氨酸氨基转移酶、磷离子、肌酐、ASA分级、是否急诊、尿素氮,见图3。针对术后住院时间延长,筛选出39个变量作为模型特征输入变量,包括抗凝血药、麻醉时间、抗血小板药、手术时间、C反应蛋白、白蛋白、诊断年龄、

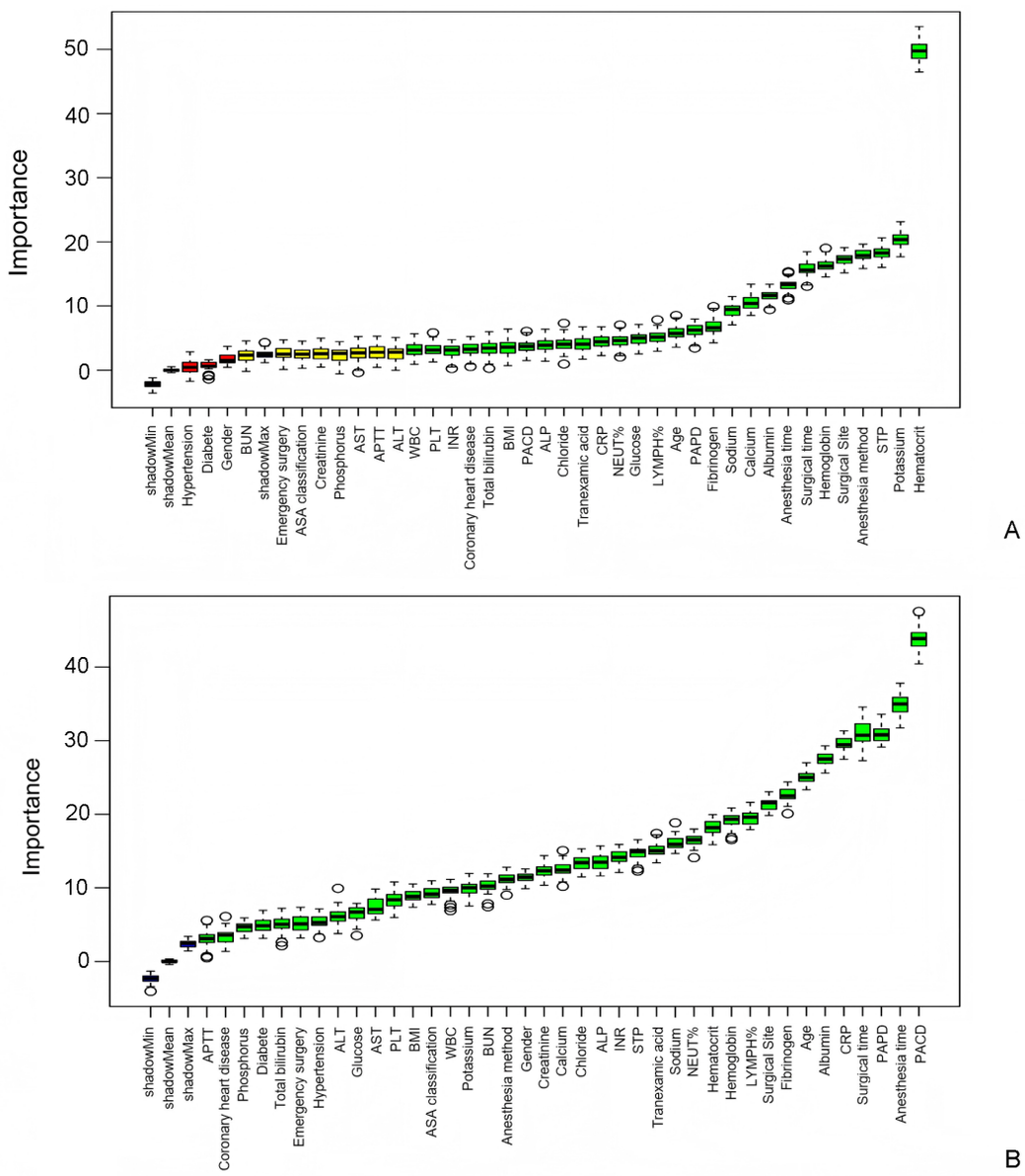
表3 插补后的数据与原始数据的数据分布情况[$M(P_{25},P_{75})$,例(%)]

Table 3 Data distribution of interpolated data and original data[$M(P_{25},P_{75})$, n(%)]

Indicators	Original data(n=11 433)	Imputed data(n=11 727)	P	Missing rate/%
BMI/(kg/m^2)	24.24(21.80, 26.78)	24.24(21.80, 26.78)	0.971	0.79
Creatinine/(mg/dL)	0.80(0.68, 0.92)	0.80(0.68, 0.92)	0.990	0.06
BUN/(mg/dL)	15.00(12.00, 18.00)	15.00(12.00, 18.00)	0.989	0.39
Phosphate/(mg/dL)	3.60(3.20, 3.90)	3.60(3.20, 3.90)	0.946	4.25
Chloride/(mmol/L)	103.00(102.00, 106.00)	103.00(102.00, 106.00)	0.953	0.37
Potassium/(mmol/L)	4.00(3.60, 4.40)	4.00(3.60, 4.40)	0.984	0.10
Sodium/(mmol/L)	139.00(137.00, 141.00)	139.00(137.00, 141.00)	0.975	0.11
ALT/(UI/L)	20.00(13.00, 30.00)	20.00(13.00, 30.00)	0.976	0.46
AST/(UI/L)	22.00(18.00, 28.00)	22.00(18.00, 28.00)	0.977	0.44
ALP/(UI/L)	68.00(55.00, 86.00)	68.00(55.00, 86.00)	0.992	0.44
TBil/(mg/dL)	0.50(0.50, 0.80)	0.50(0.50, 0.80)	0.510	8.96
Albumin/(g/dL)	4.20(4.00, 4.40)	4.20(4.00, 4.40)	0.944	0.49
STP/(g/dL)	7.00(6.80, 7.50)	7.00(6.80, 7.50)	0.977	0.37
Calcium/(mg/dL)	9.20(8.90, 9.60)	9.20(8.90, 9.60)	0.964	0.38
Fibrinogen/(mg/dL)	313.00(276.00, 376.00)	313.00(276.00, 376.00)	0.488	6.50
APTT/s	30.60(29.10, 33.00)	30.60(29.10, 33.00)	0.950	0.61
INR	0.98(0.93, 1.03)	0.98(0.93, 1.03)	0.975	0.43
Glucose/(mg/dL)	105.00(96.00, 120.00)	105.00(96.00, 126.00)	0.968	4.95
LYMPH/%	31.30(22.10, 37.80)	31.30(22.10, 37.80)	0.912	1.28
NEUT/%	61.00(53.10, 69.60)	61.00(53.10, 69.60)	0.935	1.31
PLT/(×10^9)	245.00(200.00, 291.00)	245.00(200.00, 291.00)	0.971	0.47
Hemoglobin/(g/dL)	13.20(12.20, 15.00)	13.20(12.20, 15.00)	0.987	0.32
WBC/(×10^9)	6.67(5.58, 8.38)	6.67(5.58, 8.38)	0.864	0.38
CRP/(mg/dL)	0.11(0.02, 0.50)	0.11(0.02, 0.50)	0.566	4.16
ASA classification			0.615	2.71
Ⅰ	3 936(34.42)	3 936(33.57)		
Ⅱ	6 474(56.63)	6 768(57.7)		
Ⅲ	983(8.60)	983(8.38)		
Ⅳ	38(0.33)	38(0.32)		
Ⅴ	2(0.02)	2(0.02)		

BMI: Body mass index; BUN: Blood urea nitrogen; APTT: Activated partial thromboplastin time; INR: International normalized ratio; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; ALP: Alkaline phosphatase; TBIL: Total bilirubin; STP: Serum total protein; LYMPH%: Lymphocyte percentage; NEUT%: Neutrophil percentage; PLT: Platelet count; WBC: White blood cell count; CRP: C-reactive protein

纤维蛋白原、手术部位、淋巴细胞百分比、血红蛋白、红细胞比容、中性粒细胞百分比、钠离子、TXA、总蛋白、凝血酶原时间国际标准化比值、碱性磷酸酶、氯离子、钙离子、肌酐、性别、麻醉方式、尿素氮、钾离子、白细胞、ASA分级、体质指数、血小板、丙氨酸氨基转移酶、血葡萄糖、天门冬氨酸氨基转移酶、高血压、是否急诊、血清总蛋白、糖尿病、磷离子、冠心病和活化部分凝血酶时间,见图3。



A: Boruta analysis of intraoperative bleeding; B: Boruta analysis of postoperative length of hospital stay. Green indicates importance, yellow indicates temporary importance, and red indicates unimportance. The black shadow represents shadow features, which are "false" features generated by randomly shuffling the original features to aid in evaluating the importance of the original features. If the importance of an original feature is significantly higher than the maximum importance of a shadow feature (shadowMax), then that feature is considered important. BMI: Body mass index; BUN: Blood urea nitrogen; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; ALP: Alkaline phosphatase; STP: Serum total protein; LYMPH%: Lymphocyte percentage; NEUT%: Neutrophil percentage; PLT: Platelet count; WBC: White blood cell count; CRP: C-reactive protein; APTT: Activated partial thromboplastin time; INR: International normalized ratio; PACD: Preoperative use of anticoagulants; PAPD: Preoperative use of antiplatelet drugs

图3 特征选择Boruta分析

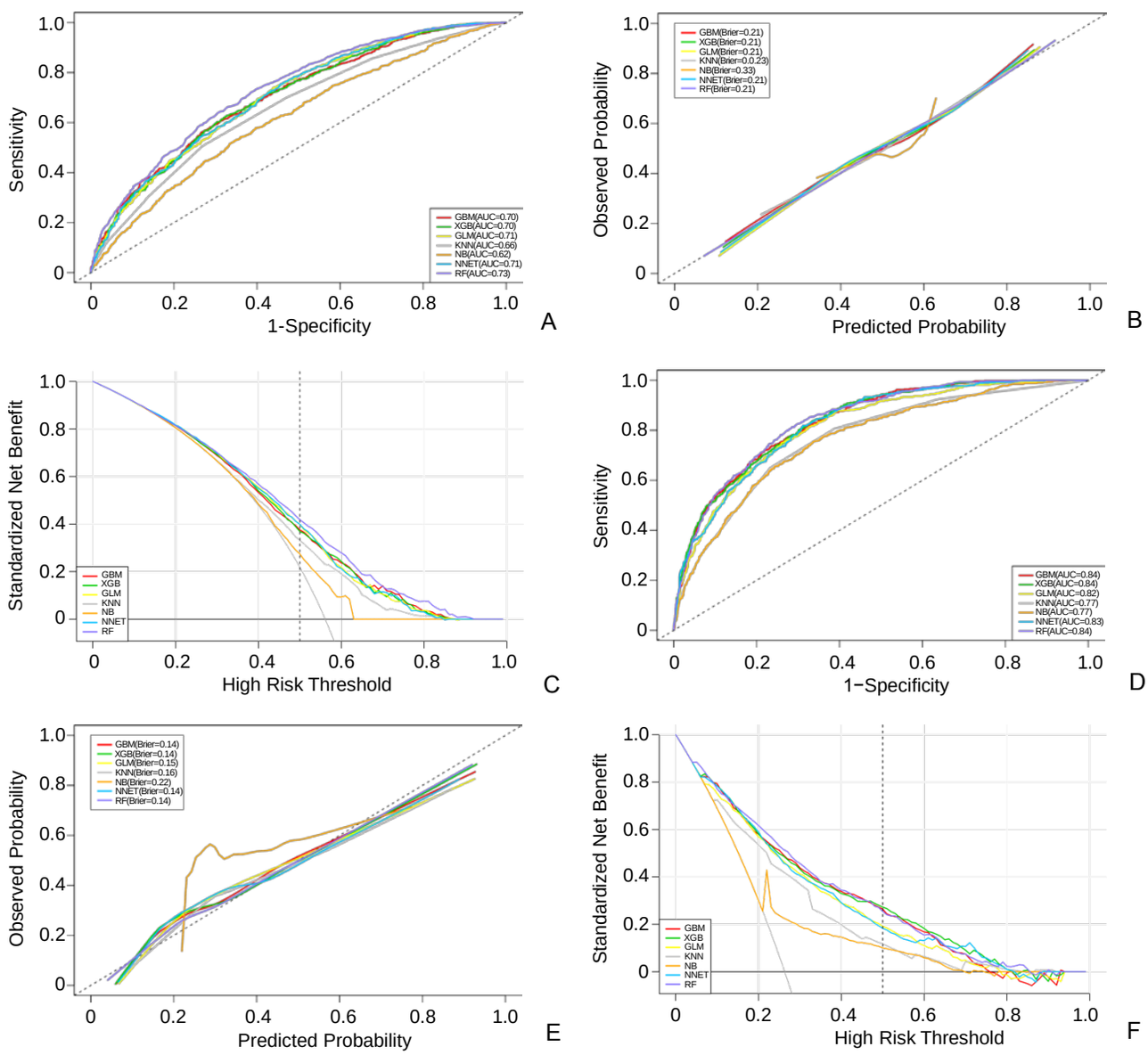
Figure 3 Feature selection by Boruta analysis

2.3 “单一结局”模型构建及表现

对于不同的目标结局,构建了7种模型,分别是GBM、XGBoost、GLM、KNN、NB、NNET以及RF,模型在测试集中的表现见表4以及图4。

对于术中出血风险预测,在测试集中,RF模型其AUC达到了0.73,这一结果显著优于其他模型,其特异度、灵敏度、阳性和阴性预测值为0.79、0.52、0.68和0.66。此外,RF模型的Kappa值和Brier值分别为0.32和0.21,优于其他模型。对测试集的7个预测模型的决策曲线进行了比较,RF模型的决策曲线比其他6个模型的优势大。

对于术后住院时间延长预测,在测试集中,GBM模型、XGBoost模型和RF模型的AUC最高达到了0.84。但XGBoost模型的Kappa值为0.45,优于其他所有模型。此外,对决策曲线进行分析,XGBoost模型较其他模型更有优势。XGBoost模型的特异度、灵敏度、阳性和阴性预测值为0.48、0.93、0.70和0.83。



A~C: ROC curve, calibration curve, and decision curve of the 7 prediction models for intraoperative bleeding; D~F: ROC curve, calibration curve, and decision curve of the 7 prediction models for postoperative length of hospital stay. ROC curve: Receiver operating characteristic curve

图4 术中出血预测模型和术后住院时间延长预测7种模型表现

Figure 4 Predictive performance of the 7 models for intraoperative bleeding and postoperative length of hospital stay

表4 预测7种模型结果

Table 4 Results of the 7 prediction models

Models	GBM	XGBoost	GLM	KNN	NB	NNET	RF
Intraoperative bleeding risk prediction							
AUC	0.704	0.703	0.705	0.660	0.619	0.707	0.731
Accuracy	0.651	0.649	0.662	0.625	0.588	0.662	0.674
Sensitivity	0.790	0.768	0.794	0.703	0.856	0.800	0.793
Specificity	0.473	0.497	0.493	0.526	0.245	0.485	0.522
PPV	0.657	0.661	0.667	0.655	0.592	0.665	0.68
NPV	0.637	0.626	0.652	0.581	0.571	0.656	0.664
F1	0.717	0.711	0.725	0.678	0.700	0.727	0.732
Kappa	0.270	0.271	0.295	0.231	0.108	0.294	0.323
Brier	0.214	0.214	0.213	0.232	0.325	0.212	0.207
Prolonged postoperative hospital stay prediction							
AUC	0.837	0.838	0.818	0.77	0.768	0.828	0.842
Accuracy	0.805	0.808	0.787	0.768	0.763	0.785	0.803
Sensitivity	0.446	0.477	0.405	0.305	0.19	0.462	0.465
Specificity	0.933	0.927	0.924	0.933	0.968	0.901	0.924
PPV	0.703	0.698	0.655	0.618	0.68	0.624	0.687
NPV	0.825	0.832	0.813	0.79	0.77	0.824	0.829
F1	0.546	0.566	0.501	0.408	0.297	0.531	0.555
Kappa	0.429	0.449	0.375	0.284	0.205	0.395	0.434
Brier	0.138	0.137	0.145	0.159	0.221	0.142	0.137

AUC: Area under the receiver operating characteristic curve; PPV: Positive predictive value; NPV: Negative predictive value

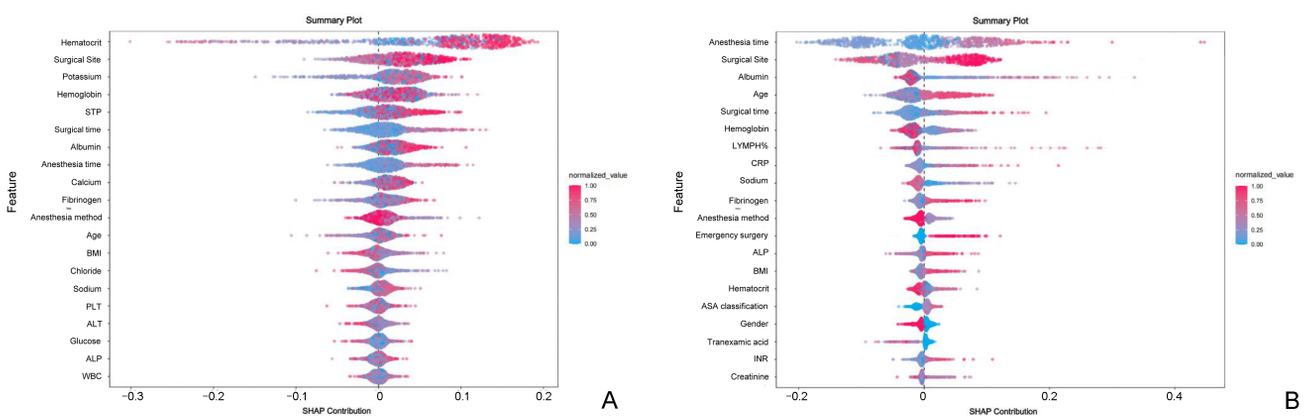
2.4 特征重要性

分别对两个结局的最优模型进行了特征重要性分析,采用SHAP方法揭示了各特征对模型输出的贡献程度。对于术中出血风险预测的RF模型,特征重要性排序前十的特征(见图5A)依次为红细胞比容、手术部位、钾离子、血红蛋白、血清总蛋白、手术时间、白蛋白、麻醉时间、钙离子以及纤维蛋白原。对于术后住院时间延长预测的XGBoost模型,特征重要性排序前十的特征(见图5B)依次为麻醉时间、手术部位、白蛋白、年龄、手术时间、血红蛋白、淋巴细胞百分比、C反应蛋白、钠离子以及纤维蛋白原。

2.5 模型集成和决策推荐

针对术中出血风险的预测任务,RF模型展现出了最为优异的预测性能;而在术后住院时间延长预测任务中,XGBoost模型则取得了最佳的预测效果。基于加权集成策略,本研究将RF模型与XGBoost模型进行集成,且在集成过程中,为RF模型和XGBoost模型分别赋予了同等权重,构建集成决策模型。

基于集成模型,进行TXA策略推荐,输入变量数值,对结局进行预测,选择可以产生更好结局的策略。并对符合模型推荐策略和不符合策略的患者进行结局的统计,包括术中出血量和术后住院时间,见表5。符合模型推荐策略的患者,术中出血量更少,住院时间更短(P<0.001)。



A: SHAP bee swarm plot of the RF model for predicting intraoperative bleeding risk; B: SHAP bee swarm plot of the XGBoost model for predicting postoperative length of hospital stay. BMI: Body mass index; ALT: Alanine aminotransferase; ALP: Alkaline phosphatase; STP: Serum total protein; PLT: Platelet count; WBC: White blood cell count; CRP: C-reactive protein; INR: International normalized ratio

图5 特征重要性SHAP蜂群图

Figure 5 SHAP beeswarm plot of feature importance

表5 集成模型决策推荐效果

Table 5 Decision recommendation effectiveness of the integrated model

Indicator	Doctor's decision	Consistent with model decision	*P*
Number of decisions	2 344	251	-
Amount of bleeding/mL	235.49 (0.00, 479.15)	183.47 (0.00, 436.50)	0.003
The average postoperative hospitalization time/min	7 775.00(4 825.00,11 915.00)	5 385.00(3 550.00, 9 925.00)	<0.001

The doctor's decision is the decision based on the original medication data in the testing dataset (n=2 344). The model's decision is the decision recommended by the model based on the input variables, which is consistent with the decision based on the original data. Consistent with the model decision indicates the consistency between integrated model decision-making and doctor's decision-making

3 讨论

多目标决策是指在医疗场景中,面对多个相互冲突或竞争的治疗目标时,通过科学方法选取最优治疗方案或策略的过程[14]。多目标决策为临床医生应对复杂、多维度的医疗挑战提供了关键决策框架。通过同时考量并优化相互关联的多个临床目标(如患者安全、治疗效果、舒适性及资源效率),医生能够制定出更优的、基于证据且个体化的治疗方案,避免了单纯追求单一目标而忽视其他重要结局指标。

骨科手术是应用多目标决策的典型场景,控制术中及术后出血,对提升围手术期安全性、缩短住院时间及优化医疗资源利用至关重要。TXA作为有效的抗纤溶止血药物,虽能显著降低出血风险,但也伴随深静脉血栓形成、肝肾功能损伤及胃肠道不适等潜在风险[5, 15-17]。这凸显了在出血控制与用药安全风险之间寻求最优平衡的核心挑战。为应对这一挑战,本研究构建了一个基于机器学习的多目标决策模型,旨在优化骨科手术围手术期TXA的使用。该模型通过同时优化关键目标(如最小化出血量、降低相关并发症风险、缩短住院时间),辅助医生制定个体化的TXA用药方案。应用结果表明,该模型有效减少了术中术后出血量,缩短了住院时间,同时将用药相关风险控制在可接受水平,从而综合提升了骨科手术的安全性与整体疗效。

在方法学层面,选择适合的机器学习技术至关重要。目前基于机器学习的临床决策模型(尤其是强化学习模型)常需大量高频采样的数据进行训练和优化。然而,患者术前评估数据通常稀疏且分散,难以满足此要求。再者,强化学习模型在单一目标优化方面表现优异[18],但在多目标优化场景下难以同时兼顾多个相互关联或冲突的临床结局指标,存在局限性。鉴于围手术期术前数据的上述特点以及多目标优化的核心需求,本研究选择采用基于简单机器学习模型的加权集成策略。该方法首先针对单一结局构建独立的预测模型,再通过加权集成实现多结局的协同优化。这一策略的优势在于能够在有限的数据条件下,有效整合多个结局指标,为临床决策提供更为全面和优化的解决方案,从而克服了强化学习模型在当前场景下的适用性限制。

本研究基于INSPIRE数据库的大规模围手术期数据展开。通过Boruta算法进行特征选择,分别筛选出36个影响术中出血风险的关键变量和39个影响术后住院时间延长的关键变量,变量涵盖患者基线、术前检验、手术麻醉等多维度信息,充分体现了骨科手术结局的复杂性。模型构建结果显示,随机森林(RF)在预测术中出血风险上表现最优(AUC=0.73),而XGBoost在预测术后住院时间延长上更具优势(AUC=0.84)。基于SHAP蜂群图深入分析关键特征发现:对于出血风险,红细胞比容升高(可能通过增加血液黏稠度影响凝血功能)、下肢关节手术部位以及血钾浓度升高(可能通过改变红细胞形态导致血液流动性降低)是主要风险因素;对于住院时间延长,麻醉时间延长(增加并发症风险)、下肢手术部位以及术前低白蛋白水平(影响伤口愈合和感染控制)是核心预测因子,这些发现与既有文献相互印证[19-26]。

针对当前研究多局限于构建单一结局预测模型而未转化临床决策[27-29],且TXA随机对照试验的严格纳排标准常与真实临床存在差距[30-32]。本研究创新性地整合上述最优预测模型(RF用于出血风险,XGBoost用于住院时间),利用真实世界数据构建了多目标TXA用药决策推荐模型。通过对比模型推荐策略与医生实际决策,验证了模型推荐可有效降低术中出血风险并缩短住院时间。其次,模型决策显然比医生决策更加积极使用TXA来预防出血。使用TXA会给患者带来血栓风险,因此医生决策相对更加保守。然而,多项临床研究已经证明[33],在骨科手术中使用TXA并没有显著增加血栓形成的风险。这表明该模

型能综合考虑患者个体差异(如检验指标)和手术特性(如部位、时长),提供更科学的个性化用药建议,对减少输血需求、控制并发症及促进康复具有重要价值。

尽管取得进展,研究仍存在局限:数据源自单一中心可能引入选择偏倚;纳入变量或未完全覆盖所有影响因素;模型可解释性有待提升以增强临床接受度。未来需纳入多中心数据验证泛化性,探索更丰富变量和先进算法以提升性能与可解释性,并通过前瞻性研究进一步验证模型在真实临床中优化TXA应用的实效。

本研究通过集成机器学习模型,构建了骨科手术围手术期TXA的个性化多目标决策框架。该模型基于真实世界数据,同步优化术中出血控制与缩短术后住院时间两大核心目标,为临床提供个体化用药推荐。本研究为突破传统单一结局预测模型的局限、推动数据驱动的精准用药决策提供了参考与借鉴。

利益冲突 本文所有作者声明不存在利益冲突

参考文献:

[1] FRISCH N B, WESSELL N M, CHARTERS M A, *et al*. Predictors and complications of blood transfusion in total hip and knee arthroplasty[J]. J Arthroplasty, 2014, 29(9 Suppl): 189-192. DOI: 10.1016/j.arth.2014.03.048.

[2] LIU K C, PIPLE A S, RICHARDSON M K, *et al*. Increased risk of venous thromboembolism in patients with postoperative *Anemia* After total joint arthroplasty: are transfusions to blame?[J]. J Bone Joint Surg Am, 2023, 105(17): 1354-1361. DOI: 10.2106/JBJS.23.00146.

[3] HARATIAN A, SHELBY T, HASAN L K, *et al*. Utilization of tranexamic acid in surgical orthopaedic practice: indications and current considerations[J]. Orthop Res Rev, 2021, 13: 187-199. DOI: 10.2147/ORR.S321881.

[4] LUO W, SUN R X, JIANG H, *et al*. The efficacy and safety of topical administration of tranexamic acid in spine surgery: a meta-analysis[J]. J Orthop Surg Res, 2018, 13(1): 96. DOI: 10.1186/s13018-018-0815-0.

[5] PATEL P A, WYROBEK J A, BUTWICK A J, *et al*. Update on applications and limitations of perioperative tranexamic acid[J]. Anesth Analg, 2022, 135(3): 460-473. DOI: 10.1213/ANE.0000000000006039.

[6] GANDHI R, EVANS H M K, MAHOMED S R, *et al*. Tranexamic acid and the reduction of blood loss in total knee and hip arthroplasty: a meta-analysis[J]. BMC Res Notes, 2013, 6: 184. DOI: 10.1186/1756-0500-6-184.

[7] LUO H, SHEN C J, QU T Y, *et al*. Tranexamic acid-induced focal convulsions after spinal surgery: a rare case report and literature review on side effects of accidental spinal administration of tranexamic acid[J]. EFORT Open Rev, 2023, 8(6): 482-488. DOI: 10.1530/EOR-23-0016.

[8] WONG J, GEORGE R B, HANLEY C M, *et al*. Tranexamic acid: current use in obstetrics, major orthopedic, and trauma surgery[J]. Can J Anaesth, 2021, 68(6): 894-917. DOI: 10.1007/s12630-021-01967-7.

[9] NIZAR AHAMED H, MOHAN S, KRISHNAN R. Role of tranexamic acid in controlling blood loss in hemiarthroplasty of the hip: a prospective and observational study[J]. Cureus, 2024, 16(6): e62617. DOI: 10.7759/cureus.62617.

[10] LIM L, LEE H, JUNG C W, *et al*. INSPIRE, a publicly available research dataset for perioperative medicine[J]. Sci Data, 2024, 11: 655. DOI: 10.1038/s41597-024-03517-4.

[11] GAO F Q, LI Z J, ZHANG K, *et al*. Four methods for calculating blood-loss after total knee arthroplasty[J]. Chin Med J, 2015, 128(21): 2856-2860. DOI: 10.4103/0366-6999.168041.

[12] NADLER S B, HIDALGO J H, BLOCH T. Prediction of blood volume in normal human adults[J]. Surgery, 1962, 51(2): 224-232.

[13] DANG T, NGUYEN T T, FRANCISCO MORENO-GARCÍA C, *et al*. Weighted ensemble of deep learning models based on comprehensive learning particle swarm optimization for medical image segmentation[C]//2021 IEEE Congress on Evolutionary Computation (CEC). June 28 - July 1, 2021, Kraków, Poland. IEEE, 2021: 744-751. DOI: 10.1109/CEC45853.2021.9504929.

[14] DAI Z Q, XU S M, WU X, *et al*. Knowledge mapping of multicriteria decision analysis in healthcare: a bibliometric analysis[J]. Front Public Health, 2022, 10: 895552. DOI: 10.3389/fpubh.2022.895552.

[15] FOUCHE P F, STEIN C, NICHOLS M, *et al*. Tranexamic acid for traumatic injury in the emergency setting: a systematic review and bias-adjusted meta-analysis of randomized controlled trials[J]. Ann Emerg Med, 2024, 83(5): 435-445. DOI: 10.1016/j.annemergmed.2023.10.004.

[16] AL-JEABORY M, SZARPAK L, ATTILA K, *et al*. Efficacy and safety of tranexamic acid in emergency trauma: a systematic review and meta-analysis[J]. J Clin Med, 2021, 10(5): 1030. DOI: 10.3390/jcm10051030.

[17] RELKE N, CHORNENKI N L J, SHOLZBERG M. Tranexamic acid evidence and controversies: an illustrated review[J]. Res Pract Thromb Haemost, 2021, 5(5): e12546. DOI: 10.1002/rth2.12546.

[18] LIU J, XIE Y H, SHU X, *et al*. Value function assessment to different RL algorithms for heparin treatment policy of

patients with sepsis in ICU[J]. Artif Intell Med, 2024, 147: 102726. DOI: 10.1016/j.artmed.2023.102726.

[19] WALKER I S, VLOK A J, ESTERHUIZEN T M, *et al*. Prediction of hematocrit decline and the impact of perioperative fluid use in lumbar spinal fusion surgery[J]. Eur Spine J, 2024, 33(1): 307-313. DOI: 10.1007/s00586-023-07977-x.

[20] RANUCCI M, LADDOMADA T, RANUCCI M, *et al*. Blood viscosity during coagulation at different shear rates [J]. Physiol Rep, 2014, 2(7): e12065. DOI: 10.14814/phy2.12065.

[21] TAN Z L, SUN Z J, MU ER TI ZHA M E A L M, *et al*. Personalized tourniquet pressure versus uniform tourniquet pressure in trauma orthopedic surgery of extremities: a prospective randomized controlled study[J]. J Orthop Surg Res, 2024, 19(1): 721. DOI: 10.1186/s13018-024-05217-6.

[22] HOFFMAN J F. Chapter 5 Functional Roles of lntracellular Potassium in Red Blood Cells [M]//GIEBISCH G. Current Topics in Membranes and Transport. Academic Press. 1987: 87-96.

[23] AL BALUSHI H, HANNEMANN A, REES D, *et al*. The effect of antioxidants on the properties of red blood cells from patients with sickle cell *anemia* [J]. Front Physiol, 2019, 10: 976. DOI: 10.3389/fphys.2019.00976.

[24] TRUONG A, HANNA M H, MOGHADAMYEGHANEH Z, *et al*. Implications of preoperative hypoalbuminemia in colorectal surgery[J]. World J Gastrointest Surg, 2016, 8(5): 353-362. DOI: 10.4240/wjgs.v8.i5.353.

[25] GRIFFIN F S, STEAD T S, ZEYL V G, *et al*. Low preoperative albumin levels significantly associated with increased risk of wound infection and bleeding after panniculectomy[J]. Plast Surg (Oakv), 2024: 22925503241292350. DOI: 10.1177/22925503241292350.

[26] WANG Y F, GONG Y, CHEN D, *et al*. C-reactive protein/albumin ratio is associated with mortality in patients with moderate to severe traumatic brain injury [J]. World Neurosurg, 2023, 173: e234-e240. DOI: 10.1016/j.wneu.2023.02.037.

[27] GUO J L, HE Q H, LI Y H. Development and validation of machine learning models to predict perioperative transfusion risk for hip fractures in the elderly[J]. Ann Med, 2024, 56(1): 2357225. DOI: 10.1080/07853890.2024.2357225.

[28] MORA D, MATEO J, NIETO J A, *et al*. Machine learning to predict major bleeding during anticoagulation for venous thromboembolism: possibilities and limitations [J]. Br J Haematol, 2023, 201(5): 971-981. DOI: 10.1111/bjh.18737.

[29] HUI V, LITTON E, EDIBAM C, *et al*. Using machine learning to predict bleeding after cardiac surgery[J]. Eur J Cardiothorac Surg, 2023, 64(6): ezad297. DOI: 10.1093/ejcts/ezad297.

[30] WANG H H, OUYANG Y H, CHEN Y L, *et al*. Clinical application of tranexamic acid in arthroscopic rotator cuff repair surgery: a randomized controlled trial[J]. Medicine, 2024, 103(26): e38515. DOI: 10.1097/MD.0000000000038515.

[31] CHOI W S, IRWIN M G, SAMMAN N. The effect of tranexamic acid on blood loss during orthognathic surgery: a randomized controlled trial [J]. J Oral Maxillofac Surg, 2009, 67(1): 125-133. DOI: 10.1016/j.joms.2008.08.015.

[32] ALVAREZ J C, SANTIVERI F X, RAMOS I, *et al*. Tranexamic acid reduces blood transfusion in total knee arthroplasty even when a blood conservation program is applied[J]. Transfusion, 2008, 48(3): 519-525. DOI: 10.1111/j.1537-2995.2007.01564.x.

[33] GAUSDEN E B, QUDSI R, BOONE M D, *et al*. Tranexamic acid in orthopaedic trauma surgery: a meta-analysis[J]. J Orthop Trauma, 2017, 31(10): 513-519. DOI: 10.1097/BOT.0000000000000913.

(收稿:2025-07-07;修回:2025-11-11)

(编辑 邓强庭)