·论著·

基于机器学习的距骨骨折术后缺血性坏死风险预测模型构建和验证

王珂，申成春，柴晓颖，俞嘉锆

【摘要】**目的**　基于机器学习构建并验证距骨骨折患者术后缺血性坏死（AVN）风险的预测模型。**方法**　本研究为回顾性病例对照研究。纳入 2018 年 1 月至 2023 年 12 月在宁波市第六医院接受手术治疗的距骨骨折患者 207 例，收集其人口学特征、骨折相关变量、手术相关参数及随访信息。按 7∶3 比例随机分为训练集和验证集。采用单因素及多因素 Logistic 回归筛选术后 AVN 的关键预测指标，并应用六种机器学习算法构建预测模型。通过受试者工作特征（ROC）曲线下面积（*AUC*）、校准曲线、决策曲线及多项性能指标评估模型表现，并采用 SHAP 方法对最优模型进行全局与局部解释。**结果**　207 例距骨骨折患者中，45 例（21.74%）发生术后 AVN。多因素 Logistic 回归显示体质量指数（BMI）、骨折分型、同侧足踝合并骨折、吸烟、骨折复位质量及骨折类型均为距骨骨折术后 AVN 的独立预测因素（均 $P < 0.05$）。XGBoost 模型在训练集和验证集中均表现出最优的 *AUC*、特异度与敏感度。**结论**　基于 BMI、骨折分型、同侧足踝合并骨折、吸烟、骨折复位质量及骨折类型构建的 XGBoost 模型，能够有效识别高危患者，为临床个体化防治提供有力工具。

【关键词】　机器学习；危险因素；预测模型；缺血性坏死；距骨骨折
doi:10.3969/j.issn.1671-0800.2026.08.004
【中图分类号】　R683.42　【文献标志码】　A　【文章编号】　1671-0800(2026)08-0837-06

Construction and validation of a risk prediction model for ischemic necrosis after talus fracture surgery based on machine learning

Wang Ke, Shen Chengchun, Chai Xiaoying, Yu Jiapei (Ningbo No.6 Hospital, Ningbo 315040, Zhejiang, China)

【**Abstract**】**Objective**　A predictive model for the risk of postoperative ischemic necrosis (AVN) in patients with talus fractures was constructed and verified based on machine learning. **Methods**　This study is a retrospective case-control study. A total of 207 patients with talus fractures who underwent surgical treatment at Ningbo No.6 Hospital from January 2018 to December 2023 were included. Their demographic characteristics, fracture related variables, surgery-related parameters and follow-up information were collected. They were randomly divided into the training set and the validation set in a 7∶3 ratio. Univariate and multivariate Logistic regression were used to screen the key predictive indicators of postoperative AVN, and six machine learning algorithms were applied to construct the predictive model. The model performance was evaluated by the area under the receiver operating characteristic (ROC) curve (*AUC*), calibration curve, decision curve and multiple performance indicators, and the SHAP method was used to globally and locally interpret the optimal model. **Results**　Among the 207 patients with talus fractures, 45 cases (21.74%) developed postoperative AVN. Multivariate Logistic regression showed that body mass index (BMI), fracture classification, ipilateral ankle combined fracture, smoking, fracture reduction quality and fracture type were all independent predictors of AVN after talus fracture surgery (all $P < 0.05$). The XGBoost model demonstrates the best *AUC*, specificity and sensitivity in both the training set and the validation set. **Conclusions**　The XGBoost model, constructed based on BMI, fracture classification, ipsilateral ankle combined fractures, smoking, fracture reduction quality and fracture type, can effectively identify high-risk patients and provide a powerful tool for individualized prevention and treatment in clinical practice.

【**key Words**】　Machine learning; Risk factors; Prediction model; Ischemic necrosis; Talus fractures
[Modern Practical Medicine,2026,38(8):837-842]

基金项目：　鄞州区农业与社会发展科技项目（2025AS058）；宁波市骨科与运动康复临床医学研究中心（2024L004）
作者单位：　315040 宁波，宁波市第六医院
通信作者：　俞嘉锆，Email：yu13858867726@163.com

距骨骨折较为罕见，占全身骨折的 0.1%～ 2.5%，占足踝部骨折的 3%～ 5%[1-2]。距骨骨折多由高处坠落、交通事故等高能量损伤导致，临床通常需要手术实现解剖复位与坚强固定，以恢复关节面平整性及生物力学功能[3]。距骨解剖结构特殊、血供脆弱，术后出现缺血性坏死（avascular necrosis，AVN）风险高达 31.2%[4]，严重影响预后。

目前国内外针对距骨骨折术后 AVN 的研究多为单维度分析[5]，未能系统整合患者基线资料、骨折及手术相关多维度变量，在临床应用中价值有限。机器学习作为人工智能的重要分支，在医学预测模型中的应用已从学术探索走向临床转化，成为精准医疗的重要技术支撑[6]。近年来，机器学习技术逐渐被用于骨科术后并发症的早期预测，显示出较好的应用价值[7]。然而，目前鲜有研究将机器学习技术应用于距骨骨折患者术后 AVN 的早期预测。本研究旨在利用六种机器学习算法，整合包括患者临床、影像学和手术变量在内的多维度数据，开发并验证可解释的距骨骨折患者术后AVN的预测模型，通过评估模型性能以确定最优机器学习算法，并结合 SHAP 分析以提高模型的可解释性，为骨科医生制定个性化的预防和治疗策略提供指导，现报道如下。

1　对象与方法

1.1　研究对象　本研究为回顾性病例对照研究。纳入 2018 年 1 月至 2023 年 12 月在宁波市第六医院接受手术治疗的距骨骨折患者。纳入标准：（1）新鲜距骨骨折（受伤至手术时间＜ 3 周）；（2）行内固定治疗；（3）年龄≥18 岁；（4）临床资料完整，随访时间＞ 12 个月。排除标准：（1）一期行踝关节融合或截肢；（2）既往足踝手术史；（3）合并严重足部神经病变或血管功能不全；（4）临床资料、随访信息不完整。本研究获得宁波市第六医院伦理委员会批准（甬六医伦审 2025 论第 41 号），免除知情同意。

1.2　数据收集与处理　收集所有纳入患者的人口学资料、骨折相关变量、手术相关参数和随访信息。（1）人口学资料：性别、年龄、美国麻醉医师协会（ASA）分级、体质量指数（BMI）及高血压、糖尿病、心脏病、吸烟、饮酒史；（2）骨折相关变量：致伤机制、骨折侧别、骨折分型（距骨颈骨折采用 Hawkins 分型，距骨体骨折采用 Sneppen 分型）、骨折类型（开放/闭合）、同侧足踝合并骨折；（3）手术相关参数：受伤至手术时间、手术策略、固定方式、手术入路、内外踝是否截骨、术中出血量、手术时长、骨折复位质量；（4）随访信息：随访时间、内固定取出情况。AVN 的诊断主要依据放射学标准，指术后影像学检查（X 线、CT、MRI）观察到患处的硬化、囊性改变或距骨塌陷。

首先计算各候选变量的缺失比例。缺失率＞ 20% 的变量不纳入后续建模，缺失率≤20%的变量进行插补 [8]。连续变量采用以随机森林回归器为基学习器的迭代插补[9]，分类变量采用训练集众数填补[10]。为避免信息泄漏，插补模型仅在训练集中拟合，并将相同的插补规则应用于验证集。

1.3　模型的构建与评估　纳入患者按 7∶3 的比例划分为训练集和验证集。首先在训练集中进行单因素 Logistic 回归分析，并采用 Benjamini-Hochberg 方法对单因素检验的 P 值进行校正，以控制多重比较导致的错误发现率（FDR）[11]。FDR 校正后 $P < 0.05$ 的变量作为多因素 Enter 法 Logistic 回归的候选变量，多因素分析中 $P < 0.05$ 的变量被确定为关键预测指标[12]。基于上述关键指标，采用 Python 3.11.4 构建随机森林（RF）、朴素贝叶斯（NB）、梯度提升机（GBM）、K 近邻（KNN）、极限树（ET）和极端梯度提升（XGBoost）六种机器学习模型。

为了评估模型性能，将使用最优超参数在训练集和测试集上绘制受试者工作特征曲线（ROC）并计算曲线下面积（*AUC*），评估模型在新数据上的泛化能力。同时，通过决策曲线分析（DCA）评估模型在实际临床中的净收益，利用校准曲线验证预测可靠性。其他评估指标包括准确度、灵敏度、特异度、阳性预测值（PPV）、阴性预测值（NPV）、召回率和 F1 值，全面评估六个预测模型识别距骨骨折患者术后 AVN 的预测价值[13]。

1.4　性能最佳预测模型解释　采用 SHAP 分析为机器学习模型提供全局解释和局部解释。

1.5　统计方法　使用 SPSS 26.0 和 Python 3.11.4 软件进行统计分析。正态分布的计量资料以均数±标准差表示，组间比较采用两独立样本 t 检验；计数资料以例（%）表示，组间比较采用 χ^2 检验或 Fisher 确切概率法。单因素 Logistic 回归的 P 值采用 Benjam-

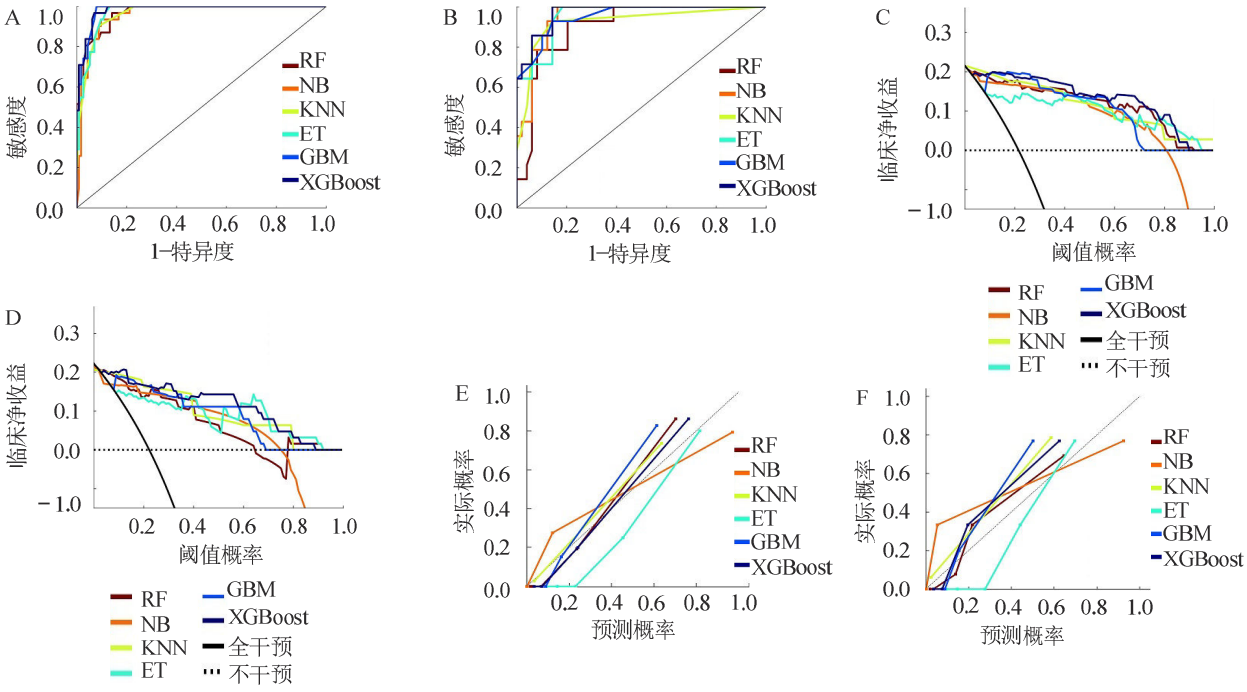
ini-Hochberg 法进行 FDR 校正，$P < 0.05$ 的变量进入多因素 Enter 法 Logistic 回归；多因素分析采用 Logistic 回归分析。$P < 0.05$ 表示差异有统计学意义。

2　结果

2.1　一般资料比较　本研究共纳入 207 例距骨骨折手术患者，其中 45 例（21.74%）发生术后 AVN，162 例（78.26%）未发生。患者被随机分配至训练集（n=144，70%）和验证集（n=63，30%）。训练集男 117 例，女 27 例；年龄（43.0±13.4）岁；ASA 分级 1 级 21 例，2 级 123 例；BMI（23.95±3.42）kg/m^2；合并高血压 23 例、糖尿病 4 例、心脏病 8 例，吸烟 42 例，饮酒 23 例；验证集男 48 例，女 15 例；年龄（46.1±14.6）岁；ASA 分级 1 级 13 例，2 级 50 例；BMI（23.76±3.62）kg/m^2；合并高血压 15 例、糖尿病 3 例、心脏病 2 例，吸烟 22 例，饮酒 11 例。两组基线资料差异均无统计学意义（均 $P > 0.05$），见表 1。

2.2　单因素、多因素 Logistic 回归分析　单因素 Logistic 回归的 P 值经 Benjamini-Hochberg FDR 校正后，BMI、吸烟、骨折分型、骨折类型、同侧足踝合并骨折、手术策略、手术入路、术中出血量、手术时长、骨折复位质量及内固定取出共 11 个变量的 $P < 0.05$，见表 2。将上述变量纳入多因素 Enter 法 Logistic 回归，显示 BMI（OR=1.018，95%CI：1.006 ～ 1.031）、骨折分型（OR=1.029，95%CI：1.009 ～ 1.050）、同侧足踝合并骨折（OR=1.120，95%CI：1.022 ～ 1.229）、吸烟（OR=1.175，95%CI：1.063 ～ 1.298）、骨折复位质量（OR=1.419，95%CI：1.239 ～ 1.624）及骨折类型（OR=1.395，95%CI：1.169 ～ 1.667）均为距骨骨折术后 AVN 的独立预测因素（均 $P < 0.05$）。

2.3　模型构建和性能评估　基于多因素 Logistic 回归分析确定的六个关键预测指标，采用六种机器学习算法构建预测模型，并评估六种预测模型性能。结果显示这六种模型的 AUC 值均超过 0.960，预测性能良好。其中 XGBoost 在训练集和验证集中均表现出最高的 AUC 值（图 1A ～ B），且在 DCA 曲线和校准曲线方面也表现最佳（图 1C ～ F）。六个机器学习预测模型在训练集和测试集中的准确率、灵敏度、特异性、阳性预测值（PPV）、阴性预测值（NPV）、精确率、召回率和 F1 分数等指标见封三图 2。综合所有预测性能指标评估，XGBoost 模型预测距骨骨折术后 AVN 最为准确可靠。使用 10 折交叉验证对 XGBoost 模型进行了评估，其平均 AUC 为 0.941，95%CI：0.905 ～ 0.977，交叉验证 AUC 较验证集 AUC 有所下降且存在一定波动，提示模型可能存在一定过拟合倾向，但总体仍保持了良好的区分能力。



注：A 为训练集 ROC 曲线，B 为验证集 ROC 曲线，C 为训练集 DCA 曲线，D 为验证集 DCA 曲线，E 为训练集校准曲线，F 为验证集校准曲线

图 1　六种机器学习预测模型的预测效能

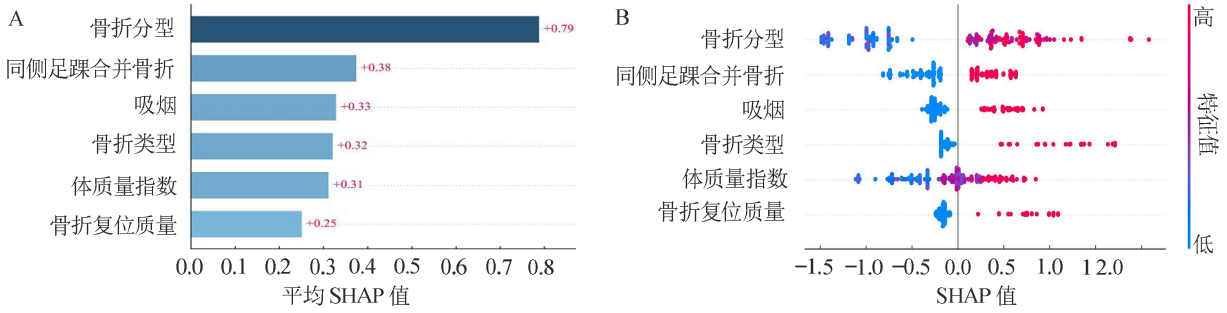
表 1　训练集和验证集临床资料比较

指标	训练集（n=144）	验证集（n=63）	χ^2（t）值	P 值
损伤机制[例（%）]			0.15	＞0.05
高能量	63（43.75）	25（39.68）		
低能量	81（56.25）	38（60.32）		
骨折侧别（左/右，例）	72/72	26/37	1.01	＞0.05
骨折分型[例（%）]				
Hawkins Ⅰ	4（2.78）	1（1.59）	11.75	＞0.05
Hawkins Ⅱ	21（14.58）	4（6.35）		
Hawkins Ⅲ	9（6.25）	11（17.46）		
Hawkins Ⅳ	7（4.86）	2（3.17）		
Sneppen Ⅱ	29（20.14）	17（26.98）		
Sneppen Ⅲ	3（2.08）	3（4.76）		
Sneppen Ⅳ	26（18.06）	8（12.70）		
Sneppen Ⅴ	34（23.61）	13（20.63）		
距骨颈合并距骨体骨折	11（7.64）	4（6.35）		
骨折类型[例（%）]			0.78	＞0.05
闭合骨折	131（90.97）	54（85.71）		
开放骨折	13（9.03）	9（14.29）		
同侧足踝合并骨折[否/是，例]	79/65	41/22	1.49	＞0.05
受伤至手术时间（d）	7.47±4.19	7.21±4.65	（0.40）	＞0.05
手术策略[例（%）]			4.81	＞0.05
一期闭合固定	1（0.69）	3（4.76）		
一期切开固定	138（95.83）	56（88.89）		
分期固定	5（3.47）	4（6.35）		
固定方式[例（%）]			0.97	＞0.05
克氏针	12（8.33）	7（11.11）		
空心钉	89（61.81）	39（61.90）		
可吸收棒	14（9.72）	4（6.35）		
钢板	5（3.47）	2（3.17）		
钢板联合空心钉	24（16.67）	11（17.46）		
手术入路[例（%）]			0.01	＞0.05
单一入路	91（63.19）	41（65.08）		
联合入路	53（36.81）	22（34.92）		
外踝是否截骨[否/是，例]	139/5	62/1	0.09	＞0.05
内踝是否截骨[否/是，例]	131/13	56/7	0.05	＞0.05
术中出血量（mL）	53.54±50.07	50.32±40.00	（0.45）	＞0.05
手术时长（min）	91.25±37.86	87.86±31.56	（0.62）	＞0.05
骨折复位质量[例（%）]			1.49	＞0.05
良好	122（84.72）	58（92.06）		
不佳	22（15.28）	5（7.94）		
随访时间（个月）	16.55±5.25	16.49±5.53	（0.74）	＞0.05
内固定取出[否/是，例]	127/17	55/8	0.00	＞0.05

2.4　模型解释　采用 SHAP 分析对 XGBoost 模型进行全局解释和局部解释，全局特征重要性显示，骨折分型、同侧足踝合并骨折和吸烟是对模型预测贡献最大的 3 个预测指标（图 2A），SHAP 摘要图提供了每个指标如何影响预测的更详细视图（图 2B）。显示吸烟、BMI=28.197 kg/m^2、Hawkins Ⅱ型骨折将该患者的预测结局推向出现“距骨坏死”类别，而同侧足踝不合并骨折、复位质量良好、闭合骨折将该患者的预测结局推向不出现“距骨坏死”类别。

3　讨论

本研究基于机器学习成功构建了距骨骨折术后 AVN 风险预测模型，确定了 BMI、骨折分型、同侧足踝合并骨折、吸烟、骨折复位质量及骨折类型等关键



注：A 为 SHAP 分析全局特征，对六个关键预测指标进行重要性排序；B 为 SHAP 摘要图，提供了每个指标如何影响预测结果

图 2　采用 SHAP 分析对 XGBoost 模型进行全局解释和局部解释

表 2　训练集单因素分析

指标	*OR* 值	95%*CI*	*P* 值	FDR *P* 值
年龄	1.000	0.995 ～ 1.004	＞ 0.05	＞ 0.05
性别	0.880	0.761 ～ 1.017	＞ 0.05	＞ 0.05
美国麻醉师分级	0.974	0.829 ～ 1.145	＞ 0.05	＞ 0.05
体质量指数	1.038	1.021 ～ 1.054	＜ 0.05	＜ 0.05
高血压	1.003	0.858 ～ 1.171	＞ 0.05	＞ 0.05
糖尿病	0.801	0.567 ～ 1.133	＞ 0.05	＞ 0.05
心脏病	1.037	0.809 ～ 1.331	＞ 0.05	＞ 0.05
吸烟	1.495	1.335 ～ 1.674	＜ 0.05	＜ 0.05
饮酒	1.112	0.952 ～ 1.298	＞ 0.05	＞ 0.05
损伤机制	1.016	0.906 ～ 1.140	＞ 0.05	＞ 0.05
骨折侧别	0.986	0.880 ～ 1.105	＞ 0.05	＞ 0.05
骨折分型	1.082	1.060 ～ 1.105	＜ 0.05	＜ 0.05
骨折类型	2.001	1.680 ～ 2.382	＜ 0.05	＜ 0.05
同侧足踝合并骨折	1.400	1.261 ～ 1.554	＜ 0.05	＜ 0.05
受伤至手术时间	0.996	0.982 ～ 1.009	＞ 0.05	＞ 0.05
手术策略	1.704	1.298 ～ 2.239	＜ 0.05	＜ 0.05
固定方式	1.033	0.986 ～ 1.082	＞ 0.05	＞ 0.05
手术入路	1.182	1.052 ～ 1.327	＜ 0.05	＜ 0.05
外踝是否截骨	0.984	0.720 ～ 1.344	＞ 0.05	＞ 0.05
内踝是否截骨	1.311	1.078 ～ 1.594	＜ 0.05	＞ 0.05
术中出血量	1.003	1.001 ～ 1.004	＜ 0.05	＜ 0.05
手术时长	1.003	1.001 ～ 1.004	＜ 0.05	＜ 0.05
骨折复位质量	1.830	1.600 ～ 2.094	＜ 0.05	＜ 0.05
随访时间	1.002	0.991 ～ 1.012	＞ 0.05	＞ 0.05
内固定取出	1.336	1.124 ～ 1.587	＜ 0.05	＜ 0.05

预测指标。XGBoost 模型在泛化能力、临床适用性和校准度方面均表现优异，可为临床距骨骨折患者术后 AVN 个体化防治提供有力工具。

本研究结果显示，吸烟是最重要的关键预测指标，其通过诱导血管痉挛、血栓形成及微循环障碍，损害局部血供，增加 AVN 风险[14]。本研究 SHAP 分析进一步量化了吸烟对个体预测风险的贡献值，提示术前戒烟应作为距骨骨折患者围术期管理的重要环节。高 BMI 患者常伴有肥胖相关的代谢紊乱和慢性炎症，同时还可能增加距骨局部机械应力，进一步加剧骨折移位风险，对高 BMI 距骨骨折患者应加强术前评估与体重管理[15]。同侧足踝合并骨折多提示高能量损伤，往往造成距骨周围软组织及血供的严重破坏，术后 AVN 风险显著增加[16]。骨折类型与分型反映了距骨骨折损伤严重程度和解剖结构的破坏程度[17]。本研究中距骨颈合并距骨体骨折患者的 AVN 发生率显著高于单一部位骨折，Hawkins 和 Sneppen 分型系统作为距骨颈骨折经典分型，在本研究构建的机器学习预测模型中仍然保持了重要预测价值，验证了其临床实用性。骨折复位质量是 AVN 的重要可干预因素。解剖复位可最大限度恢复距骨周围血供，减少关节面压力异常分布[17]。本研究将复位质量纳入模型，为术中追求解剖复位提供了量化依据，对于复位质量不佳的患者，应加强术后影像学随访，早期发现 AVN 征象。

本研究对比了六种常用机器学习算法，结果显示 XGBoost 综合所有预测性能指标评估最为准确可靠。XGBoost 作为一种基于梯度提升框架的集成学习算法，通过迭代优化损失函数、引入正则化项防止过拟合，在处理小样本、高维数据时具有明显优势[18]。本研究通过将数据预处理和变量筛选限定在训练集、减少最终输入变量数量，并采用 10 折交叉验证，对模型稳定性进行了初步评估。然而，本研究仅纳入 207 例患者，其中 AVN 事件 45 例，且验证集仅 63 例；在此条件下，复杂机器学习模型仍可能学习到样本特异性模式，导致性能高估。10 折交叉验证 *AUC* 低于单次验证集 *AUC*，也支持这一谨慎判断。后续研究应扩大多中心、前瞻性样本，依据候选参数数量和事件率进行样本量设计[19]。传统机器学习模型常因缺乏可解释性而难以被临床医生接受，本研究引入 SHAP 分析，从全局层面揭示了各危险因素对模

型预测的贡献度，从个体层面展示了每个患者风险预测的形成过程为医患沟通和个体化干预提供了直观依据，有助于模型在临床实践中的落地应用。

本研究存在以下局限性：(1)回顾性研究的数据质量受电子病历记录完整性影响。尽管采用预设缺失率界值和随机森林插补，插补结果仍依赖于缺失机制及模型设定，可能存在残余信息偏倚；(2)本研究为单中心研究，仅纳入 207 例患者，其中 AVN 事件 45 例，验证集样本量亦较小。10 折交叉验证 *AUC* 低于单次验证集 *AUC*，提示模型可能存在一定程度的乐观偏倚和过拟合，当前结果不能替代多中心外部验证。尽管 XGBoost 已内置正则化项并采用早停策略，但小样本下过拟合风险仍难完全规避，模型泛化能力需谨慎解读；(3) 尽管采用 Benjamini-Hochberg FDR 校正降低了多重单因素检验导致的假阳性风险，基于单因素筛选后再进行多因素建模的策略在小样本中仍可能不稳定，且该筛选基于线性逻辑回归，所选特征未必是树模型的最优子集，本研究未采用 LASSO 等替代筛选进行敏感性分析，关键因子的稳健性有待验证。未来可结合惩罚回归、Bootstrap 或嵌套交叉验证进行变量选择和模型收缩，以提高预测因子的稳健性。

利益冲突　所有作者声明无利益冲突

作者贡献声明　王珂：酝酿和设计实验，实施研究，分析/解释数据，起草文章，统计分析；申成春：实施研究，采集数据，分析/解释数据，起草文章，统计分析；柴晓颖：实施研究，采集数据，行政、技术或材料支持；俞嘉锫：酝酿和设计实验，分析/解释数据，对文章的知识性内容作批评性审阅，统计分析，获取研究经费

参　考　文　献

[1] Schwartz A M, Runge W O, Hsu A R, et al. Fractures of the talus: Current concepts[J]. Foot Ankle Orthop, 2020, 5(1): 2473011419900766.

[2] 关路,赵宏,王振林,等.经内踝截骨入路开放复位内固定治疗Hawkins Ⅱ、Ⅲ型距骨颈骨折的疗效分析[J].现代实用医学,2025,37(5):519-522.

[3] Majeed H, McBride D J. Talar process fractures: An overview and update of the literature[J]. EFORT Open Rev, 2018, 3(3): 85-92.

[4] Kubisa M J, Kubisa M G, Palka K, et al. Avascular necrosis of the talus: Diagnosis, treatment, and modern reconstructive options[J]. Medicina (Kaunas), 2024, 60(10): 1692.

[5] Alley M C, Vallier H A, Tornetta P. Identifying risk factors for osteonecrosis after talar fracture[J]. J Orthop Trauma, 2024, 38(1): 25-30.

[6] Churpek M M, Carey K A, Edelson D P, et al. Internal and external validation of a machine learning risk score for acute kidney injury [J]. JAMA Netw Open, 2020, 3(8): e2012892.

[7] Harris A H, Kuo A C, Bowe T, et al. Prediction models for 30-day mortality and complications after total knee and hip arthroplasties for veteran health administration patients with osteoarthritis[J]. J Arthroplasty, 2018, 33(5): 1539-1545.

[8] Liu F, Yao J, Liu C Y, et al. Construction and validation of machine learning models for sepsis prediction in patients with acute pancreatitis[J]. BMC Surg, 2023, 23(1): 267.

[9] Shah A D, Bartlett J W, Carpenter J, et al. Comparison of random forest and parametric imputation models for imputing missing data using MICE: A CALIBER study[J]. Am J Epidemiol, 2014, 179(6): 764-774.

[10] Afkanpour M, Hosseinzadeh E, Tabesh H. Identify the most appropriate imputation method for handling missing values in clinical structured datasets: A systematic review[J]. BMC Med Res Methodol, 2024, 24(1): 188.

[11] Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing[J]. J R Stat Soc Ser B Stat Methodol, 1995, 57(1): 289-300.

[12] Du S L, Wu Y Z, Tao J R, et al. Development and validation of machine learning models for outcome prediction in patients with poor-grade aneurysmal subarachnoid hemorrhage following endovascular treatment[J]. Ther Clin Risk Manag, 2026, 21: 293-307.

[13] Quan K R, Lin W R, Hong J B, et al. A machine learning approach for predicting radiation-induced hypothyroidism in patients with nasopharyngeal carcinoma undergoing tomotherapy[J]. Sci Rep, 2024, 14(1): 8436.

[14] Patel R A, Wilson R F, Patel P A, et al. The effect of smoking on bone healing: A systematic review[J]. Bone Joint Res, 2013, 2(6): 102-111.

[15] Collins A T, Kulvaranon M L, Cutcliffe H C, et al. Obesity alters the *in vivo* mechanical response and biochemical properties of cartilage as measured by MRI[J]. Arthritis Res Ther, 2018, 20(1): 232.

[16] Srinath A, Southall W G S, Nazal M R, et al. Talar neck fractures with associated ipsilateral foot and ankle fractures have a higher risk of avascular necrosis[J]. J Orthop Trauma, 2024, 38(6): 220-224.

[17] Jordan R K, Bafna K R, Liu J Y, et al. Complications of talar neck fractures by Hawkins classification: A systematic review[J]. J Foot Ankle Surg, 2017, 56(4): 817-821.

[18] Bukowski M, Kurek J, Swiderski B, et al. Custom loss functions in XGBoost algorithm for enhanced critical error mitigation in drill-wear analysis of melamine-faced chipboard[J]. Sensors (Basel), 2024, 24(4): 1092.

[19] Riley R D, Snell K I, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes[J]. Stat Med, 2019, 38(7): 1276-1296.

收稿日期：2026-03-31

（本文编辑：吴迪汉）