·其他肝病·

DOI：10.12449/JCH260822

基于机器学习的肝移植术后糖尿病风险评估模型的构建与验证

王 浩[1,2]，樊永强[1]，史志勇[1]，张 瑞[1]，王 妍[1]，徐 钧[1]

1 山西医科大学第一医院肝胆外科及肝脏移植中心，太原 030001

2 山西医科大学第一临床医学院，太原 030001

通信作者：徐钧，junxuty@163.com（ORCID：0000-0003-3755-9660）

摘要：目的 构建并验证基于多种机器学习算法的肝移植术后糖尿病（PTDM）风险评估模型，以实现PTDM的早期识别。**方法** 回顾性收集山西医科大学第一医院2020年4月1日—2024年12月31日行同种异体肝移植患者的临床资料。按照7∶3比例随机分为训练集和测试集，采用Lasso回归分析结合5折交叉验证进行特征选择。在训练集中分别构建逻辑斯谛回归（LR）、决策树（DT）、朴素贝叶斯（NB）、随机森林（RF）、K-近邻（KNN）、极端梯度提升（XGBoost）和自适应提升（AdaBoost）共7种机器学习模型。在测试集中采用准确率、精确率、召回率、特异度、F1分数、受试者操作特征曲线下面积（AUROC）、精确率-召回率曲线下面积（AUPRC）等方法比较各模型的预测性能，并采用Brier评分与决策曲线分析评估模型的校准度和临床实用价值，使用沙普利加性解释（SHAP）方法进行特征重要性分析。计量资料两组间比较采用成组t检验或Wilcoxon秩和检验。计数资料两组间比较采用χ^2检验或Fisher精确检验。**结果** 共纳入135例肝移植受者，其中26例发生PTDM；训练集94例，测试集41例。特征提取筛选得到性别、超重或肥胖、无肝期、手术时长、住院时长、术后早期低镁血症以及空腹血糖受损（IFG）共7项特征。在测试集中，XGBoost模型表现最佳，其AUROC为0.907（95%置信区间：0.807～0.989），AUPRC为0.649（95%置信区间：0.339～0.955），准确率为0.878，精确率为0.667，召回率为0.750，F1分数为0.706，特异度为0.909，Brier评分为0.104。决策曲线分析显示，当阈值概率低于0.667时，使用XGBoost模型进行临床决策可获得较高的净获益。SHAP分析显示，特征重要性排序依次为住院时长、手术时长、超重或肥胖、性别、IFG、无肝期及术后早期低镁血症。**结论** 基于XGBoost算法构建的肝移植受者PTDM风险评估模型表现优异，能够有效识别潜在的高风险人群，但其泛化能力仍需多中心、大样本数据进一步验证。

关键词： 肝移植；糖尿病；机器学习；模型，统计学

基金项目： 国家自然科学基金（82470693）；山西省重点实验室项目（202204010931008）；山西省重点研发计划（202302130501013）

Construction and validation of a machine learning-based risk assessment model for post-transplant diabetes mellitus

Wang Hao[1,2], *Fan Yongqiang*[1], *Shi Zhiyong*[1], *Zhang Rui*[1], *Wang Yan*[1], *Xu Jun*[1]

1. Department of Hepatobiliary Surgery and Liver Transplant Center, The First Hospital of Shanxi Medical University, Taiyuan 030001, China; 2. The First Clinical Medical College of Shanxi Medical University, Taiyuan 030001, China

Corresponding author: Xu Jun, junxuty@163.com (ORCID: 0000-0003-3755-9660)

Abstract: Objective To construct and validate a risk assessment model for post-transplant diabetes mellitus (PTDM) using multiple machine learning algorithms, and to realize the early identification of PTDM. **Methods** A retrospective analysis was performed for the clinical data of the patients who underwent allogeneic liver transplantation in The First Hospital of Shanxi Medical University from April 1, 2020 to December 31, 2024, and they were randomly divided into a training set and a validation set at a ratio of 7∶3. The LASSO regression analysis combined with 5-fold cross-validation was used for feature selection. Seven machine learning models were developed in the training set, i.e., logistic regression (LR), decision tree (DT), Naive Bayes (NB), random forest (RF), K-nearest neighbor (KNN), extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost). In the

validation set, various methods were used to assess the predictive performance of each model, such as accuracy, precision, recall rate, specificity, F1 score, area under the receiver operating characteristic curve (AUROC), and area under the precision-recall curve (AUPRC). The Brier score and decision curve analysis were used to assess the calibration and clinical practicability of the models, and the SHAP method was used to analyze feature importance. The independent-samples *t* test or the Wilcoxon rank-sum test was used for comparison of continuous data between two groups, and the chi-square test or the Fisher's exact test was used for comparison of categorical data between two groups. **Results** A total of 135 liver transplant recipients were enrolled, among whom 26 developed PTDM, and there were 94 patients in the training set and 41 in the validation set. Feature extraction and screening identified 7 key features of sex, overweight or obesity, anhepatic phase, time of operation, length of hospital stay, early postoperative hypomagnesemia, and impaired fasting glucose (IFG). In the validation set, the XGBoost model showed the best predictive performance, with an AUROC of 0.907 (95% confidence interval [*CI*]: 0.807—0.989), an AUPRC of 0.649 (95%*CI*: 0.339—0.955), an accuracy of 0.878, a precision of 0.667, a recall rate of 0.750, an F1-score of 0.706, a specificity of 0.909, and a Brier score of 0.104. The decision curve analysis showed that when the threshold probability was below 0.667, application of the XGBoost model in clinical decision-making provided relatively high net benefit. The SHAP analysis showed that the length of hospital stay ranked first in terms of feature importance, followed by time of operation, overweight or obesity, sex, IFG, anhepatic phase, and early postoperative hypomagnesemia. **Conclusion** The risk assessment model for PTDM in liver transplant recipients based on XGBoost algorithm has excellent performance and can effectively identify high-risk individuals; however, multicenter large-sample data are needed for further validation.

Key words: Liver Transplantation; Diabetes Mellitus; Machine Learning; Models, Statistical

Research funding: National Natural Science Foundation of China (82470693); Shanxi Provincial Key Laboratory Project (202204010931008); Shanxi Key Research and Development Program (202302130501013)

肝移植是目前终末期肝病的重要治疗方法。随着围术期管理和免疫抑制策略的不断优化，肝移植受者的长期生存率显著提高[1]，术后远期并发症逐渐成为影响预后的关键因素。其中，移植后糖尿病（post-transplantation diabetes mellitus，PTDM）是最常见的代谢并发症之一[2]，对患者的生存质量和远期结局具有重要影响。文献报道，肝移植受者PTDM的发生率为15%～25%[3]，国内一项纳入超过10 000例受者的大样本研究亦报告其发生率高达24.3%[4]。PTDM会显著增加患者心血管疾病、感染等不良结局的风险，从而降低患者的生存率[5-7]。因此，早期识别PTDM对于提高肝移植受者远期生存具有重要意义。

近年来，机器学习因能够处理复杂的非线性关系和多维临床数据，在肝移植术后风险预测领域受到关注，并在疾病预测和预后风险评估方面显示出潜在的应用价值[8]。然而，目前大多数肝移植受者PTDM的预测模型是基于传统的逻辑回归方法而建立，基于机器学习的预测工具仍然有限。因此，本研究旨在应用多种机器学习算法，构建并验证肝移植受者PTDM的风险评估模型，以期提高对PTDM的早期识别能力。

1 资料与方法

1.1 *研究对象* 收集2020年4月1日—2024年12月31日于山西医科大学第一医院行同种异体肝移植术患者的临床资料。纳入标准：(1)年龄≥18岁；(2)临床资料完整，且术后随访时间>1个月。排除标准：(1)术前已诊断为糖尿病；(2)活体肝移植或二次肝移植；(3)术后1个月内死亡；(4)失访或临床数据缺失率>30%。

1.2 *数据收集及定义* 本研究共收集23项临床指标。术前资料包括年龄、性别、血型、超重或肥胖、空腹血糖受损（impaired fasting glucose，IFG）、原发疾病（乙型肝炎肝硬化、丙型肝炎肝硬化、酒精性肝硬化、自身免疫性肝硬化、其他）、腹腔积液、是否合并肿瘤、天冬氨酸氨基转移酶、丙氨酸氨基转移酶、总胆红素、白蛋白、低镁血症、尿素、肌酐，以及乙型肝炎病毒、丙型肝炎病毒和巨细胞病毒的感染情况；术中资料包括手术时长和无肝期；术后资料包括住院时长、术后早期低镁血症和术后早期感染。

超重或肥胖定义为体重指数≥24 kg/m^2。IFG定义为空腹血糖5.6～6.9 mmol/L，且在不同日内重复测量2次或以上均符合此标准。术后早期低镁血症定义为术后

1周内血清镁离子浓度最低值<0.75 mmol/L。术后早期感染定义为术后1个月内，血培养、痰培养、引流液培养、胸腹水培养等任一结果为阳性。PTDM的定义及诊断标准参照2013年国际共识会议[9]，并以移植术后45 d作为诊断起始点[10]。本研究的主要结局为肝移植术后45天新发现的糖尿病。

1.3 免疫抑制方案 肝移植受者术后均采用三联免疫抑制方案：他克莫司+吗替麦考酚酯+激素，并持续检测他克莫司谷浓度，移植后3个月内维持在8～12 ng/mL，3～6个月维持在7～10 ng/mL，6～12个月维持在6～8 ng/mL，12个月以后维持在5 ng/mL左右[11]。

1.4 模型构建及验证 首先，基于原始数据进行描述性统计分析，比较PTDM组与非PTDM组的基线特征。随后，将肝移植受者按照7∶3比例随机分为训练集和测试集。为避免数据泄露，训练集和测试集内部分别进行多重插补以填补缺失值。鉴于机器学习算法在不平衡数据中的性能较差，本研究使用ADASYN算法对训练集进行过采样，以平衡数据。随后在训练集中采用Lasso回归分析结合5折交叉验证进行特征选择，筛选出稳定特征用于后续模型构建。基于选定特征，在训练集中分别构建7种机器学习模型，包括逻辑斯谛回归（logistic regression，LR）、决策树（decision tree，DT）、朴素贝叶斯（naive Bayes，NB）、随机森林（random forest，RF）、K-近邻（K-nearest neighbour，KNN）、极端梯度提升（eXtreme gradient boosting，XGBoost）和自适应提升（adaptive boosting，AdaBoost）算法。

在测试集中对各模型的性能进行评估，评价指标包括准确率、精确率、召回率、特异度、F1分数、受试者操作特征曲线下面积（area under the receiver operating characteristic，AUROC）和精确率-召回率曲线下面积（area under the precision-recall curve，AUPRC），并对测试集进行1 000次bootstrap重复抽样，计算AUROC和AUPRC的均值及95%置信区间（95% confidence interval，95%*CI*）。此外，计算Brier评分以评估模型的校准度，并使用决策曲线评估模型在不同阈值概率下的临床净收益。最后采用沙普利加性解释（SHapley Additive exPlanations，SHAP）方法进行可解释性分析。

1.5 统计学方法 采用R 4.4.2统计软件完成数据预处理，应用Python 3.12.7软件进行模型构建及验证。符合正态分布的计量资料以$\bar{x}\pm s$表示，两组间比较采用成组t检验；不符合正态分布的计量资料以$M(P_{25}\sim P_{75})$表示，两组间比较采用Wilcoxon秩和检验。计数资料两组间比较采用χ^2检验或Fisher精确检验。P<0.05为差异有统计学意义。

2 结果

2.1 基线临床特征 本研究共纳入135例肝移植受者，其中26例（19.3%）发生PTDM。患者中男100例，女35例，平均年龄（51.0±9.9）岁。PTDM组26例患者中，男16例，女10例，平均年龄（52.5±8.2）岁；非PTDM组109例患者中，男84例，女25例，平均年龄（50.6±10.3）岁。与非PTDM组相比，PTDM组患者超重或肥胖、术前IFG和术后早期低镁血症比例较高，且手术时间和住院时间均明显延长（P值均<0.05）；两组其他指标比较，差异均无统计学意义（P值均>0.05）（表1）。按照7∶3比例划分训练集和测试集，其中训练集94例，测试集41例。训练集与测试集各项临床指标比较，差异均无统计学意义（P值均>0.05）。

2.2 模型构建与评估 通过Lasso回归和5折交叉验证对候选变量进行筛选，当λ=0.025 8时，模型的均方误差最小，筛选出的稳定特征包括性别、超重或肥胖、无肝期、手术时长、住院时长、术后早期低镁血症以及IFG共7个变量（图1）。基于上述特征，本研究构建了7种机器学习模型。图2和图3展示了各模型在测试集中的AUROC和AUPRC，表2汇总了主要性能指标。在所有模型中，XGBoost模型表现最佳。该模型在测试集中的AUROC为0.907（95%*CI*：0.807～0.989），AUPRC为0.649（95%*CI*：0.339～0.955），准确率为0.878，精确率为0.667，召回率为0.750，F1分数为0.706，特异度为0.909。因此，将XGBoost作为评估肝移植受者PTDM风险的最优模型。进一步评价显示，该模型的Brier评分为0.104，提示其整体预测校准较好。决策曲线分析表明，当阈值概率低于0.667时，采用XGBoost模型进行临床决策具有显著的正向净收益（图4）。

2.3 SHAP可解释性分析 SHAP值分析用于解释XGBoost模型的特征贡献度。特征重要性分析结果显示，7个特征的重要性由高到低依次为：住院时长、手术时长、超重或肥胖、性别、IFG、无肝期以及术后早期低镁血症（图5）。

3 讨论

PTDM是肝移植术后常见的并发症，严重影响患者的长期生存[12]。本研究中肝移植受者PTDM的发生率为19.3%，与既往研究一致[3]。本研究通过对肝移植受者

表1 PTDM组与非PTDM组患者特征比较

Table 1 Comparison of baseline characteristics between patients in the PTDM and non-PTDM groups

项目	非PTDM组（n=109）	PTDM组（n=26）	统计值	P值
年龄[例(%)]			χ^2=0.50	0.479
<45岁	31(28.4)	5(19.2)		
≥45岁	78(71.6)	21(80.8)		
性别[例(%)]			χ^2=1.89	0.169
男	84(77.1)	16(61.5)		
女	25(22.9)	10(38.5)		
超重或肥胖[例(%)]			χ^2=3.99	0.046
否	68(62.4)	10(38.5)		
是	41(37.6)	16(61.5)		
高血压[例(%)]			χ^2=0.08	0.784
否	93(85.3)	21(80.8)		
是	16(14.7)	5(19.2)		
术前IFG[例(%)][1)]			χ^2=9.09	0.003
否	78(83.0)	11(50.0)		
是	16(17.0)	11(50.0)		
血型[例(%)]				0.691
A型	29(26.6)	5(19.2)		
B型	33(30.3)	11(42.3)		
AB型	8(7.3)	2(7.7)		
O型	39(35.8)	8(30.8)		
原发疾病[例(%)]				0.866
乙型肝炎肝硬化	50(45.9)	13(50.0)		
丙型肝炎肝硬化	4(3.7)	0(0.0)		
酒精性肝硬化	18(16.5)	3(11.5)		
自身免疫性肝硬化	8(7.3)	3(11.5)		
其他	29(26.6)	7(26.9)		
是否合并肿瘤[例(%)]			χ^2=0.50	0.479
否	78(71.6)	21(80.8)		
是	31(28.4)	5(19.2)		
腹腔积液[例(%)][1)]			χ^2=0.38	0.827
无	35(33.0)	9(34.6)		
轻度	36(34.0)	10(38.5)		
中-重度	35(33.0)	7(26.9)		
乙型肝炎病毒感染[例(%)]			χ^2=0.00	>0.05
否	58(53.2)	14(53.8)		
是	51(46.8)	12(46.2)		
丙型肝炎病毒感染[例(%)][1)]				>0.05
否	104(96.3)	25(96.2)		
是	4(3.7)	1(3.8)		
巨细胞病毒感染[例(%)][1)]				0.400
否	90(93.8)	23(88.5)		
是	6(6.3)	3(11.5)		
总胆红素（μmol/L）	53.3(28.0～129.9)	66.7(36.8～156.6)	Z=−0.60	0.547
白蛋白[例(%)]			χ^2=0.74	0.391
<30 g/L	30(27.5)	10(38.5)		
≥30 g/L	79(72.5)	16(61.5)		

表1（续）

Table 1（continued）

项目	非PTDM组（n=109）	PTDM组（n=26）	统计值	P值
术前低镁血症[例（%）][1)]			χ^2=0.35	0.556
否	64（78.0）	13（68.4）		
是	18（22.0）	6（31.6）		
尿素（mmol/L）	5.2（3.9～7.3）	5.7（3.8～7.6）	Z=−0.59	0.558
肌酐（μmol/L）	59.3（49.6～74.0）	59.5（46.5～73.5）	Z=0.29	0.770
手术时长[例（%）][1)]			χ^2=6.65	0.010
<10 h	71（68.3）	10（38.5）		
≥10 h	33（31.7）	16（61.5）		
无肝期[例（%）][1)]			χ^2=2.66	0.103
<1 h	88（84.6）	17（68.0）		
≥1 h	16（15.4）	8（32.0）		
住院时长[例（%）][1)]				0.002
<25 d	54（50.5）	4（15.4）		
≥25 d	53（49.5）	22（84.6）		
术后早期低镁血症[例（%）][1)]				0.043
否	22（22.0）	1（3.8）		
是	78（78.0）	25（96.2）		
术后早期感染[例（%）][1)]			χ^2=0.00	>0.05
否	37（34.6）	9（34.6）		
是	70（65.4）	17（65.4）		

注：1）受回顾性研究的限制，指标存在缺失值，但缺失率<30%。PTDM，移植后糖尿病；IFG，空腹血糖受损。

表2 测试集中7种机器学习模型的性能比较

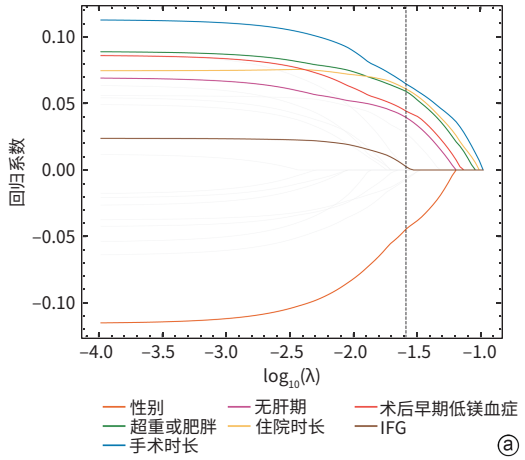
Table 2 Performance comparison of seven machine learning models in the test set

模型	准确率	精确率	召回率	F1分数	特异度	AUROC（95%*CI*）	AUPRC（95%*CI*）
LR	0.610	0.333	1.000	0.500	0.515	0.915（0.806～0.989）	0.733（0.396～0.965）
DT	0.854	0.625	0.625	0.625	0.909	0.892（0.777～0.976）	0.552（0.264～0.900）
NB	0.634	0.348	1.000	0.516	0.546	0.899（0.786～0.976）	0.694（0.356～0.921）
RF	0.829	0.571	0.500	0.533	0.909	0.911（0.805～0.987）	0.656（0.333～0.958）
KNN	0.585	0.286	0.750	0.414	0.546	0.769（0.564～0.922）	0.431（0.171～0.731）
XGBoost	0.878	0.667	0.750	0.706	0.909	0.907（0.807～0.989）	0.649（0.339～0.955）
AdaBoost	0.829	0.546	0.750	0.632	0.849	0.896（0.777～0.986）	0.687（0.376～0.960）

注：LR，逻辑斯谛回归；DT，决策树；NB，朴素贝叶斯；RF，随机森林；KNN，K-近邻；XGBoost，极端梯度提升；AdaBoost，自适应提升；AUROC，受试者操作特征曲线下面积；AUPRC，精确率-召回率曲线下面积；95%*CI*，95%置信区间。

进行回顾性分析，构建并验证了LR、DT、NB、RF、KNN、XGBoost和AdaBoost共7种机器学习模型，最终得到基于XGBoost算法的肝移植受者PTDM风险评估模型，能较为准确地识别肝移植受者发生PTDM的风险。

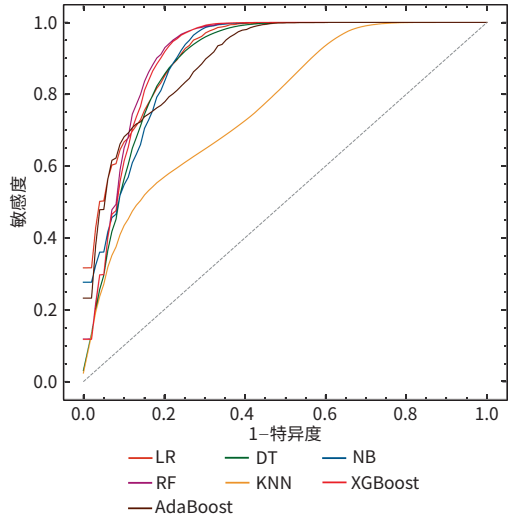
在特征选择方面，本研究采用Lasso回归分析筛选出7个稳定特征，包括性别、超重或肥胖、无肝期、手术时长、住院时长、术后早期低镁血症以及IFG。一项Meta分析显示，男性是肝移植受者PTDM发生的独立危险因素，其发生风险较女性增加约0.5倍（比值比=1.53，95%*CI*：1.24～1.90）[13]。超重或肥胖患者发生PTDM的风险显著高于体重正常者，这与既往研究一致[14]。肥胖常伴有血脂异常，而甘油三酯升高等脂代谢紊乱也被证实与PTDM相关[15-16]。PTDM患者的住院时间明显延长，可能与术后并发症和整体病程更复杂有关。既往研究表明，年龄、术前蔡尔德-皮尤评分、住院期间并发症均可影响肝移植患者住院时间，从而增加血糖紊乱的风险[17]。此外，低镁血症在本研究中亦是重要的影响因素，PTDM患者术后1周血镁水平显著低于非PTDM患者，而术前血镁水平在两组间差异不显著，低镁血症可能通过影响胰岛素受体或介导他克莫司的致糖尿病效



注：a，基于5折交叉验证的Lasso回归系数路径图；b，5折交叉验证确定Lasso回归的最优λ值。IFG，空腹血糖受损。

图1 Lasso回归分析

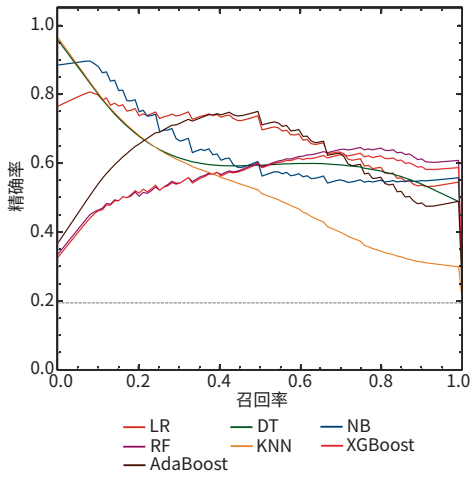
Figure 1 Lasso regression analysis



注：LR，逻辑斯谛回归；DT，决策树；NB，朴素贝叶斯；RF，随机森林；KNN，K-近邻；XGBoost，极端梯度提升；AdaBoost，自适应提升；ROC曲线，受试者操作特征曲线。

图2 7种机器学习模型在测试集中的ROC曲线

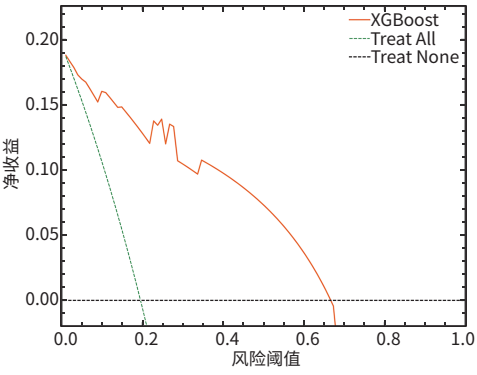
Figure 2 ROC curves of seven machine learning models in the test set



注：LR，逻辑斯谛回归；DT，决策树；NB，朴素贝叶斯；RF，随机森林；KNN，K-近邻；XGBoost，极端梯度提升；AdaBoost，自适应提升；PRC曲线，精确率-召回率曲线。

图3 7种机器学习模型在测试集中的PRC曲线

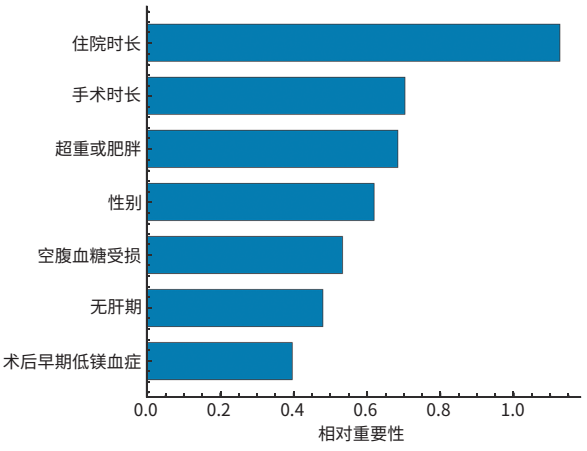
Figure 3 PRC of seven machine learning models in the test set



注：XGBoost，极端梯度提升。

图4 XGBoost模型在测试集中的决策曲线分析

Figure 4 Decision curve analysis of the XGBoost model in the test set



注：XGBoost，极端梯度提升。

图5 XGBoost模型的特征重要性分析

Figure 5 Feature importance analysis of the XGBoost model

应，从而促进PTDM的发生[18]。术前IFG已被研究证实是肝移植受者PTDM的独立危险因素[17]，提示肝移植受者在术前可能存在潜在的糖代谢紊乱，在手术应激、免疫抑制剂、激素等多种因素作用下更容易进展为PTDM[19]。值得注意的是，尚未有研究显示无肝期和手术时长与肝移植受者PTDM相关，但在本研究中被Lasso回归筛选为重要特征。无肝期延长可能导致更严重的短暂性肝缺血再灌注损伤，进而加重全身炎症反应和代谢应激；而手术时长反映了手术复杂程度及围手术期应激水平，两者均可能通过增加应激性高血糖和影响胰岛素敏感度，从而间接促进PTDM的发生。

LR模型因建模简洁、结果易于解释，既往一直被广泛应用。然而，其属于线性模型，难以充分捕捉变量之间复杂的非线性特征，可能导致预测性能受限。Bai等[17]纳入年龄、术前IFG、术后空腹血糖及住院时长构建了肝移植受者的多因素LR模型，在验证集中该模型的敏感度为0.792，特异度为0.681，AUROC为0.782（95%*CI*：0.708～0.856）。本研究测试集中LR模型的AUROC为0.915（95%*CI*：0.806～0.989），召回率为1.000，但特异度仅为0.515，F1分数为0.500，提示该模型在不平衡数据中存在大量假阳性，其实际应用价值受限。相比之下，机器学习模型能够自动学习复杂的特征模式，在预测准确性和泛化能力方面更具优势。娄连玉等[20]基于RF算法构建了肝移植受者PTDM预测模型，其召回率为0.825，F1分数为0.722，模型整体的预测性能较高。Loosen等[21]构建的肝移植受者PTDM预测模型中，RF和XGBoost模型性能也优于传统LR模型。因此，本研究采用多种机器学习算法构建模型，综合比较各模型性能后，发现XGBoost模型表现最优，在测试集中的AUROC为0.907（95%*CI*：0.807～0.989），AUPRC为0.649（95%*CI*：0.339～0.955），召回率为0.750，F1分数为0.706，特异度为0.909，其在小样本不平衡数据中仍能保持良好的泛化性能。最后采用SHAP方法对模型进行可解释性分析，量化各特征的重要性，以提升模型的应用价值。

在本研究中，为克服样本量及不平衡数据带来的挑战，笔者采用多种方法提升模型的性能及稳定性。在模型选择中，优先选择XGBoost、NB等在小样本条件下表现良好的模型[22]，并通过多种模型比较确定最优模型。在数据预处理中，应用ADASYN算法对训练集进行过采样处理，通过合成数据增加少数类样本，缓解数据不平衡问题。在特征选择中，结合Lasso回归分析和交叉验证筛选出表现稳定的特征，用于后续的模型构建。在模型评估中，采用AUPRC、F1分数等在不平衡数据中更具参考价值的指标，全面评估模型对肝移植受者PTDM的识别能力。

本研究仍存在一定局限性。首先，回顾性收集临床资料可能出现数据缺失、记录不完整等问题；其次，本研究为单中心研究，虽然通过内部验证评估了模型的稳定性，但模型的泛化能力仍需多中心研究进一步验证；最后，由于肝源短缺等客观因素，肝移植患者总体数量有限，故本研究样本量相对较小，虽然已在算法选择、特征筛选以及数据平衡等方面进行了优化，但小样本条件下仍可能存在模型过拟合或泛化能力不足的风险。因此，本研究构建的机器学习模型仍需在更大规模、多中心的前瞻性队列中进一步验证其可靠性与临床适用性。

综上所述，本研究基于7种机器学习算法构建了肝移植受者PTDM的风险评估模型，其中XGBoost模型的综合性能最为优异。该模型能够在术后早期精准识别PTDM高风险人群，为临床医师提供一种快速、可靠的辅助决策工具，有助于推动个体化监测与早期干预，从而改善肝移植受者的远期预后。尽管本研究的结果具有一定的参考价值，但模型的稳定性与泛化能力仍需在多中心、大样本的前瞻性研究中进一步验证。

伦理学声明：本研究方案于2024年10月17日经山西医科大学第一医院伦理委员会审批，批号：KYLL-2024-281。

利益冲突声明：本文不存在任何利益冲突。

作者贡献声明：王浩负责设计论文框架，分析数据，起草论文；樊永强负责文献查找，数据收集；史志勇、张瑞负责论文修改；王妍负责绘制图表；徐钧负责拟定写作思路，指导撰写文章并最后定稿。

参考文献：

[1] 夏强，冯浩. 肝脏移植肿瘤学的现状与挑战[J]. 中华消化外科杂志，2026，25(1)：69-75. DOI：10.3760/cma.j.cn115610-20251106-00677.
Xia Qiang，Feng Hao. Current status and challenges of liver transplantation oncology[J]. Chin J Dig Surg，2026，25(1)：69-75. DOI：10.3760/cma.j.cn115610-20251106-00677.

[2] 陈以宁，肖云，韩晓雨，等. 肝移植术后排斥合并高血糖的诊治和用药的实践与思考[J]. 中国临床药理学与治疗学，2023，28(5)：550-555. DOI：10.12092/j.issn.1009-2501.2023.05.009.
Chen Yining，Xiao Yun，Han Xiaoyu，et al. Diagnosis and individualized drug therapy for the rejection with hy-perglycemia after liver Transplantation[J]. Chin J Clin Pharmacol Ther，2023，28(5)：550-555. DOI：10.12092/j.issn.1009-2501.2023.05.009.

[3] Chin Y H，Tan H Q M，Ng C H，et al. A time-based meta-analysis on the incidence of new onset diabetes after liver transplantation[J]. J Clin Med，2021，10(5)：1045. DOI：10.3390/jcm10051045.

[4] Ling Q，Xu X，Xie H Y，et al. New-onset diabetes after liver transplantation：A national report from China Liver Transplant Registry[J].

Liver Int, 2016, 36(5): 705-712. DOI: 10.1111/liv.13042.
[5] Richardson B, Khan M Q, Brown S A, et al. Personalizing diabetes management in liver transplant recipients: The new era for optimizing risk management[J]. Hepatol Commun, 2022, 6(6): 1250-1261. DOI: 10.1002/hep4.1876.
[6] Sarabhai T, Mathew A, Rashidi-Alavijeh J, et al. Diabetes mellitus and osteoporosis after organ transplantation: Frequency, clinical features, and treatment[J]. Dtsch Arztebl Int, 2026, 123(10): 281-288. DOI: 10.3238/arztebl.m2026.0019.
[7] Park S S, Koo B K, Park S, et al. Impact of new-onset diabetes after transplantation on cardiovascular risk and mortality in Korea: A nationwide population-based study[J]. Diabetes Metab J, 2025, 49(1): 117-127. DOI: 10.4093/dmj.2024.0078.
[8] 刘迎春, 杨斌. 人工智能在肝移植中的应用研究进展[J]. 器官移植, 2024, 15(6): 883-888. DOI: 10.3969/j.issn.1674-7445.2024084.
Liu Yingchun, Yang Bin. Research progress on the application of artificial intelligence in liver transplantation[J]. Organ Transplant, 2024, 15(6): 883-888. DOI: 10.3969/j.issn.1674-7445.2024084.
[9] Sharif A, Hecking M, de Vries A P, et al. Proceedings from an international consensus meeting on posttransplantation diabetes mellitus: Recommendations and future directions[J]. Am J Transplant, 2014, 14(9): 1992-2000. DOI: 10.1111/ajt.12850.
[10] 董骏峰, 薛强, 滕飞, 等. 移植后糖尿病危险因素研究进展[J]. 器官移植, 2024, 15(1): 145-150. DOI: 10.3969/j.issn.1674-7445.2023154.
Dong Junfeng, Xue Qiang, Teng Fei, et al. Research progress in risk factors of post-transplantation diabetes mellitus[J]. Organ Transplant, 2024, 15(1): 145-150. DOI: 10.3969/j.issn.1674-7445.2023154.
[11] 中华医学会器官移植学分会. 中国肝移植免疫抑制治疗与排斥反应诊疗规范(2019版)[J]. 器官移植, 2021, 12(1): 8-14, 28. DOI: 10.3969/j.issn.1674-7445.2021.01.002.
Branch of Organ Transplantation of Chinese Medical Association. Diagnosis and treatment specification for immunosuppressive therapy and rejection of liver transplantation in China (2019 edition)[J]. Organ Transplant, 2021, 12(1): 8-14, 28. DOI: 10.3969/j.issn.1674-7445.2021.01.002.
[12] Shimada S, Miyake K, Venkat D, et al. Clinical characteristics of new-onset diabetes after liver transplantation and outcomes[J]. Ann Gastroenterol Surg, 2024, 8(3): 383-393. DOI: 10.1002/ags3.12775.
[13] Li D W, Lu T F, Hua X W, et al. Risk factors for new onset diabetes mellitus after liver transplantation: A meta-analysis[J]. World J Gastroenterol, 2015, 21(20): 6329-6340. DOI: 10.3748/wjg.v21.i20.6329.
[14] Dascal R, Wiebe C, Niazi M, et al. Post-transplant diabetes mellitus in Canadian liver and renal transplant recipients[J]. Can Liver J, 2022, 5(4): 466-475. DOI: 10.3138/canlivj-2022-0010.
[15] Liang J, Yi X L, Xue M J, et al. A retrospective cohort study of preoperative lipid indices and their impact on new-onset diabetes after liver transplantation[J]. J Clin Lab Anal, 2020, 34(5): e23192. DOI: 10.1002/jcla.23192.
[16] 尚莹, 邵非, 孔心涓. 成人原位肝移植术后糖尿病发生的危险因素分析[J]. 精准医学杂志, 2024, 39(4): 341-345, 351. DOI: 10.13362/j.jpmed.202404013.
Shang Ying, Shao Fei, Kong Xinjuan. An analysis of risk factors for the development of diabetes mellitus after orthotopic liver transplantation in adults[J]. J Precis Med, 2024, 39(4): 341-345, 351. DOI: 10.13362/j.jpmed.202404013.
[17] Bai R P, An R, Chen S Y, et al. Risk factors and prediction score for new-onset diabetes mellitus after liver transplantation[J]. J Diabetes Investig, 2024, 15(8): 1105-1114. DOI: 10.1111/jdi.14204.
[18] Chen S J, Bowen D G, Liu K, et al. Hypomagnesaemia, an independent risk factor for the development of post-transplant diabetes mellitus in liver and renal transplant recipients A systematic review[J]. J Hum Nutr Diet, 2024, 37(6): 1407-1419. DOI: 10.1111/jhn.13354.
[19] Porrini E L, Díaz J M, Moreso F, et al. Clinical evolution of post-transplant diabetes mellitus[J]. Nephrol Dial Transplant, 2016, 31(3): 495-505. DOI: 10.1093/ndt/gfv368.
[20] 娄连玉, 李树娟, 张丙良, 等. 基于随机森林的肝移植术后糖尿病风险预测模型的构建与验证[J]. 护理学报, 2026, 33(1): 60-67. DOI: 10.16460/j.issn2097-6569.2026.01.060.
Lou Lianyu, Li Shujuan, Zhang Bingliang, et al. Construction and validation of a random forest-based prediction model for post-liver transplant diabetes mellitus[J]. J Nurs, 2026, 33(1): 60-67. DOI: 10.16460/j.issn2097-6569.2026.01.060.
[21] Loosen S H, Krieg S, Chaudhari S, et al. Prediction of new-onset diabetes mellitus within 12 months after liver transplantation-a machine learning approach[J]. J Clin Med, 2023, 12(14): 4877. DOI: 10.3390/jcm12144877.
[22] 梁智星, 叶林森, 杨扬. 人工智能在肝移植中的应用[J]. 临床肝胆病杂志, 2022, 38(1): 30-34. DOI: 10.3969/j.issn.2095-5332.2024.06.016.
Liang Zhixing, Ye Linsen, Yang Yang. Application of artificial intelligence in liver transplantation[J]. J Clin Hepatol, 2022, 38(1): 30-34. DOI: 10.3969/j.issn.2095-5332.2024.06.016.

收稿日期: 2026-02-03;**录用日期:** 2026-04-10
本文编辑: 葛俊

引证本文: 王浩, 樊永强, 史志勇, 等. 基于机器学习的肝移植术后糖尿病风险评估模型的构建与验证[J]. 临床肝胆病杂志, 2026, 42(8): 1894-1901.
Wang Hao, Fan Yongqiang, Shi Zhiyong, et al. Construction and validation of a machine learning-based risk assessment model for post-transplant diabetes mellitus[J]. J Clin Hepatol, 2026, 42(8): 1894-1901.