

基于机器学习产后压力性尿失禁风险 预测模型的构建及验证

王丽云¹,高天勤²,刘雨佳¹,陈青³,陈柳³,沙凯辉¹

(1.滨州医学院护理学院,山东 滨州 256600; 2.滨州医学院附属医院第一介入导管室,山东 滨州 256600;

3.滨州医学院附属医院产后康复中心,山东 滨州 256600)

摘要:目的 基于最新的产后压力性尿失禁(postpartum stress urinary incontinence, PPSUI)定义,采用5种机器学习算法构建风险预测模型,筛选最优模型并评估其临床应用价值。方法 采用前瞻性研究设计,纳入1208例产妇,基于问卷与电子病历收集数据,并按8:2比例随机划分训练集和测试集。采用单因素分析和随机森林算法筛选特征,构建基于Logistic回归、决策树、随机森林、支持向量机和极限梯度提升(extreme gradient boosting, XGBoost)的PPSUI预测模型,并通过网格搜索优化超参数。模型训练与验证采用Bootstrap方法和十折交叉验证,以提升稳定性和泛化能力。最终,从分类性能、临床适用性及预测可靠性等方面综合评估各模型,并筛选最优预测模型。结果 XGBoost预测模型为最优预测模型,测试集AUC值为0.993(95%CI: 0.985~0.998, $P<0.01$)。决策曲线分析显示,该模型在中等阈值范围内净收益最高,校准曲线接近理想状态,预测可靠性较优,具备较高的临床应用价值。结论 XGBoost适用于PPSUI高危人群的早期筛查和风险评估,为精准医学和产后健康管理提供科学依据。

关键词:压力性尿失禁;产后;风险预测;机器学习;预测模型

中图分类号:R714.46

文献标志码:A

Development and validation of a postpartum stress urinary incontinence risk prediction model based on machine learning

WANG Liyun¹, GAO Tianqin², LIU Yujia¹, CHEN Qing³, CHEN Liu³, SHA Kaihui¹

(1. School of Nursing, Binzhou Medical University, Binzhou 256600, Shandong, China;

2. First Interventional Catheterization Laboratory, Affiliated Hospital of Binzhou Medical University,

Binzhou 256600, Shandong, China; 3. Postpartum Rehabilitation Center, Affiliated Hospital of Binzhou Medical University,

Binzhou 256600, Shandong, China)

Abstract: Objective To develop risk prediction models for postpartum stress urinary incontinence (PPSUI) using five machine learning algorithms based on the latest PPSUI definition, identify the optimal model, and evaluate its clinical applicability. **Methods** This study adopted a prospective design and included 1,208 postpartum women. Data were collected from questionnaires and electronic medical records, and the dataset was randomly divided into a training set and a test set in an 8:2 ratio. Feature selection was performed using univariate analysis and the random forest algorithm. Five PPSUI prediction models were developed based on Logistic regression, decision tree, random forest, support vector machine, and extreme gradient boosting (XGBoost), with hyperparameter optimization conducted through grid search. Model training and validation were performed using the bootstrap method and ten-fold cross-validation to enhance stability and generalizability. The models were comprehensively evaluated in terms of classification performance, clinical applicability, and predictive reliability to identify the optimal prediction model. **Results** XGBoost was identified as the optimal prediction model, achieving the highest AUC in the test set (AUC = 0.993, 95% CI: 0.985-0.998, $P<0.01$).

Decision curve analysis showed that XGBoost provided the highest net benefit within the intermediate threshold range, while the calibration curve was closest to the ideal state, indicating superior predictive reliability and significant clinical applicability. **Conclusion** XGBoost is suitable for early screening and risk assessment of high-risk PPSUI populations, providing a scientific basis for precision medicine and postpartum health management.

Key words: Stress urinary incontinence; Postpartum; Risk prediction; Machine learning; Predictive model

产后压力性尿失禁(postpartum stress urinary incontinence, PPSUI)是在产后12个月内首次出现因用力或体力活动(如运动)、打喷嚏或咳嗽而导致不自主遗尿的现象^[1]。其发生与妊娠和分娩引起的盆底损伤、激素水平变化密切相关,这些因素可导致盆底肌力下降、韧带松弛及尿道括约肌功能障碍^[2-3]。既往研究显示,PPSUI的患病率为6.36%~57.45%,并随年龄增长而升高^[4-5]。随着我国“全面三孩”政策的实施,高龄及肥胖孕产妇比例上升,PPSUI的患病风险进一步增加。然而,由于对该疾病认知不足、社会羞耻感以及经济压力,许多患者未能及时寻求医疗帮助或与亲属沟通^[6-8]。因此,如何实现PPSUI的早期识别和干预,降低其患病率及改善产妇生活质量,成为当前的重要议题。

机器学习在处理大规模高维医疗数据方面具有独特优势,尤其在疾病风险预测模型的构建中展现出巨大的应用潜力^[9-10]。本研究从早期筛查与预防角度出发,基于最新的PPSUI定义,采用5种常见机器学习算法——Logistic回归、弹性网络回归、随机森林、支持向量机(support vector machine, SVM)和极限梯度提升(extreme gradient boosting, XGBoost)构建PPSUI风险预测模型。通过比较各模型的性能、临床适用性和预测可靠性,筛选出最佳预测模型,旨在为PPSUI的早期筛查提供精准工具,为临床和护理实践中的个性化健康管理提供理论依据,最终改善产妇生活质量,保障其身心健康。

1 资料与方法

1.1 研究对象

本研究为前瞻性观察性研究,采用便利抽样法选取2023年10月至2024年5月在山东省某三甲医院产后康复门诊复查的产妇1208例,19~42(27.9±4.5)岁。该研究获得滨州医学院伦理审查委员会的批准[伦研批第(2023-381)],并遵循《赫尔

辛基宣言》相关伦理原则,所有参与者均已知情同意并签署知情同意书。纳入标准:①产后12个月内复查;②年龄≥18岁;③单胎妊娠;④无严重分娩并发症;⑤自愿参与本研究。

排除标准:①阴道炎及泌尿系统感染者;②妊娠前有尿失禁病史;③既往接受过尿失禁手术治疗;④产后恢复异常者。

1.2 方法

1.2.1 PPSUI的诊断

PPSUI的诊断由医生通过国际尿失禁咨询委员会尿失禁问卷简表(international consultation on incontinence questionnaire short form, ICIQ-UI-SF)对患者进行评估^[11]。若患者选择“咳嗽或打喷嚏时漏尿”或“进行体力活动或锻炼时漏尿”,则诊断为压力性尿失禁(stress urinary incontinence, SUI)。目前,ICIQ-UI-SF的中文版已被广泛应用,其有效性和可靠性已得到充分验证^[12]。

1.2.2 数据收集

本研究数据来源于结构化问卷和电子病历两部分。问卷用于收集参与者的主观报告信息,电子病历则提供客观临床数据。未参与本研究的专业护士通过结构化问卷进行数据收集,问卷内容涵盖一般人口学信息,如居住地、职业、受教育程度、人均家庭月收入、孕前体质量指数(body mass index, BMI)、便秘史、慢性咳嗽史、妊娠期SUI、妊娠期是否盆底肌锻炼。

与妊娠及分娩相关的数据通过电子病历提取,主要包括此次分娩年龄、初产年龄、分娩前BMI、孕次、产次、流产次数及妊娠期体质量增加。同时,还记录了分娩前子宫高度和腹围、妊娠期并发症(如妊娠期高血压、妊娠期糖尿病)、孕周、双顶径、催产药物使用情况及分娩镇痛。

1.2.3 数据预处理

为确保模型构建和验证过程中数据分割的独立性,防止数据泄露,采用随机划分方法,以8:2的比例将数据集分为训练集和测试集。所有数据预处理步骤均在训练集上完成,测试集仅应用训练

集学习到的插补参数、编码规则和标准化参数,以保证特征转换的一致性并提升模型的泛化能力。本研究的缺失值为数值型变量,且缺失比例低于30%,为降低缺失值对模型的潜在偏倚影响,采用多重插补法填充缺失值^[13]。所有分类变量均采用独热编码进行处理,通过将每个类别映射为二进制特征矩阵,以便模型能有效识别类别信息;数值型变量则进行了标准化处理,转化为均值为0、标准差为1的标准正态分布,从而确保所有特征处于统一的量纲范围内,避免不同量纲的特征对模型训练产生不良影响。

1.2.4 特征选择

采用两阶段特征选择方法,以优化模型输入变量并确保特征筛选的科学性和合理性。第一阶段,采用单因素分析筛选与PPSUI具有统计学关联性的变量($P<0.05$),以减少冗余特征并提高特征筛选的可解释性。第二阶段,基于随机森林算法计算候选变量的重要性评分,并筛选排名前十的变量用于模型构建。随机森林能够处理非线性关系、高维数据、变量交互效应及噪声,确保筛选出的变量更具代表性,充分反映数据中的关键信息^[14]。相较于单一特征选择方法,本研究采用的两阶段特征选择策略结合了统计学筛选和数据驱动方法的优势,既确保纳入变量的统计合理性,又避免完全依赖统计假设或机器学习模型特征重要性排序的局限性,从而提升变量筛选的稳健性,为PPSUI预测模型的构建提供了更加可靠的特征选择方案。

1.2.5 模型构建及参数优化

基于训练集数据,采用Logistic回归、弹性网络回归、随机森林、SVM和XGBoost五种机器学习算法构建预测模型,并通过Bootstrap重采样、十折交叉验证(10-Fold Cross-Validation)和网格搜索进行优化。Bootstrap重采样通过随机抽样生成多个训练子集,提高模型的稳健性和泛化能力。十折交叉验证通过循环训练和验证,减少评估偏差,确保模型在不同数据集上的稳定性。网格搜索在超参数空间内系统搜索最优参数组合,最大化模型性能并降低过拟合风险。所有模型的训练、优化与评估严格基于训练集进行,确保不同模型的公平比较和结果的可重复性。

1.2.6 模型评估与临床适用性分析

采用多项指标对模型性能进行综合评估。绘制受试者工作特征(receiver operating characteristic,

ROC)曲线,计算曲线下面积(area under the curves, AUC),同时结合准确率、精确率、召回率、特异性、F1分数全面衡量分类模型的预测能力。为提高评估结果的稳健性,采用Bootstrap方法对AUC进行1000次重采样,并计算其95%CI,确保结果的稳定性和统计学可靠性。此外,为评估模型的临床应用价值,采用决策曲线分析(decision curve analysis, DCA)和校准曲线分析。DCA主要用于评估模型在不同决策阈值下的净收益,以衡量其临床决策价值;校准曲线则用于分析模型预测概率与实际发生概率之间的一致性,以验证其预测可靠性。所有评估均在独立测试集验证,以验证模型的泛化能力和临床适用性。最终,基于AUC、F1分数、精确率、召回率等综合评估结果,选择预测能力最优且具有较强泛化性和临床可解释性的模型,以支持精准筛查和个体化干预策略的制定。

1.3 统计学处理

采用R 4.3.2和Python 3.8软件进行分析。正态分布的连续性变量使用 $\bar{x}\pm s$ 表示,组间比较采用两独立样本 t 检验。非正态分布的数据以 $M(P_{25}, P_{75})$ 表示,组间比较采用Wilcoxon秩和检验。计数资料以 $n(\%)$ 表示,组间比较采用 χ^2 检验。检验水准 $\alpha=0.05$ 。

2 结果

2.1 基线特征分析

在研究期间,共有2657例产妇前往产后康复门诊进行复查。经筛选,1449例产妇因不符合纳入标准被排除(图1),最终纳入1208例参与者。研究群体中,PPSUI组406例,非PPSUI组802例,PPSUI的发生率为33.61%。PPSUI组21~42(31.50 ± 5.05)岁,非PPSUI组19~33(26.05 ± 2.97)岁。PPSUI组与非PPSUI组的基线特征详见表1。

2.2 特征选择

单因素分析结果显示,新生儿体质量、妊娠期体质量增加、此次分娩年龄等变量与PPSUI风险显著相关($P<0.05$),见表1。基于上述结果,将相关变量纳入候选特征集,并使用随机森林算法对特征重要性进行排序。最终,从中筛选出对PPSUI风险预测贡献最大的10项关键特征,包括新生儿体质量、妊娠期体质量增加、分娩前BMI、此次分娩年龄、孕前BMI、孕次、产次、妊娠期SUI、初产年龄及分娩前宫高。见图2。

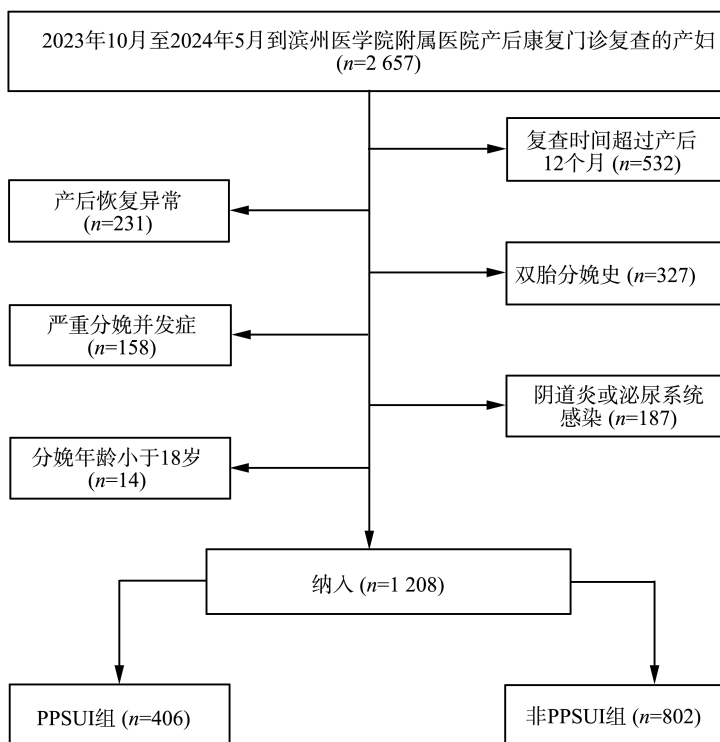


图1 纳入人群筛选和纳入流程图

Figure 1 Flowchart for screening and inclusion of the inclusion population

表1 PPSUI组和非PPSUI组的基线特征分析

Table 1 Baseline characterization of the PPSUI and Non-PPSUI groups

项目	例数	非 PPSUI 组 (n=802)	PPSUI 组 (n=406)	$t/Z/\chi^2$	P
此次分娩方式				0.137	0.712
平产	592	390(48.6)	202(49.8)		
剖宫产	616	412(51.4)	204(50.2)		
此次分娩年龄 ^a /岁		26.05±2.97	31.50±5.05	-23.562	<0.001
初产年龄 ^a /岁		25.53±2.62	26.78±2.91	-7.549	<0.001
妊娠期体质量增加 ^a /kg		8.89±2.08	13.82±3.88	-28.775	<0.001
孕前 BMI ^b		20.03(18.67,22.21)	19.05(18.38,20.82)	-6.239	<0.001
分娩前 BMI ^a		25.43±2.98	25.91±3.32	-2.549	0.011
孕次 ^b /次		1(1,1)	3(1,4)	-16.868	<0.001
产次 ^b /次		1(1,1)	2(1,3)	-19.273	<0.001
流产次数 ^b /次		0(0,0)	1(0,2)	-15.958	<0.001
妊娠期 SUI				333.482	<0.001
否	957	757(94.4)	200(49.3)		
是	251	45(5.6)	206(50.7)		
分娩前宫高 ^b /cm		29(28,30)	28(27,30)	-3.341	0.001
分娩前腹围 ^a /cm		105.21±6.61	105.42±6.35	-0.522	0.602
新生儿体质量 ^a /g		3 176.22±222.63	3 724.01±343.85	-33.373	<0.001
孕周 ^a /周		38.26±0.80	38.28±0.78	-0.275	0.783
双顶径 ^a /cm		9.37±0.27	9.40±0.20	-1.667	0.096
胎位不正				1.189	0.228
否	1 186	785(97.9)	401(98.8)		
是	22	17(2.1)	5(1.2)		
第二产程时间/min		0(0,15)	4.5(0,12)	-1.101	0.271

续表

项目	例数	非 PPSUI 组(<i>n</i> = 802)	PPSUI 组(<i>n</i> = 406)	<i>t</i> / <i>Z</i> / χ^2	<i>P</i>
急产				<0.001	>0.999
否	1 202	798(99.5)	404(99.5)		
是	6	4(0.5)	2(0.5)		
催产药物使用				31.305	<0.001
否	1 096	701(87.4)	395(97.3)		
是	112	101(12.6)	11(2.7)		
分娩镇痛				59.470	<0.001
否	938	570(71.1)	368(90.6)		
是	270	232(28.9)	38(9.4)		
会阴切开术				7.720	0.005
否	1 165	765(95.4)	400(98.5)		
是	43	37(4.6)	6(1.5)		
产钳助娩				0.121	0.728
否	1 207	802(100)	405(99.8)		
是	1	0(0)	1(0.2)		
胎吸				0.067	0.796
否	1 206	800(99.8)	406(100)		
是	2	2(0.2)	0(0)		
会阴撕裂				34.244	<0.001
无	988	619(77.2)	369(90.9)		
一度	218	181(22.6)	37(9.1)		
二度	2	2(0.2)	0(0)		
三度及以上		0(0)	0(0)		
妊娠期高血压				1.953	0.162
否	1 083	726(90.5)	357(87.9)		
是	125	76(9.5)	49(12.1)		
妊娠期糖尿病				50.236	<0.001
否	1 002	709(88.4)	293(72.2)		
是	206	93(11.6)	113(27.8)		
宫颈机能不全史				34.089	<0.001
否	1 178	797(99.4)	381(93.8)		
是	30	5(0.6)	25(6.2)		
居住地				0.721	0.396
城市	616	402(50.1)	214(52.7)		
农村	592	400(49.9)	192(47.3)		
职业				0.207	0.976
脑力劳动者	300	197(24.6)	103(25.4)		
体力劳动者	304	204(25.4)	100(24.6)		
部分脑力和部分体力劳动者	302	202(25.2)	100(24.6)		
未参加工作者	302	199(24.8)	103(25.4)		
受教育程度				8.054	0.045
初中及以下	339	212(26.4)	127(31.3)		
高中、职业高中或中专	354	228(28.4)	126(31)		
大专或本科	366	251(31.3)	115(28.3)		
硕士及以上	149	111(13.8)	38(9.4)		
人均家庭月收入/元				0.500	0.919
<5 000	308	209(26.1)	99(24.4)		
5 000~10 000	298	198(24.7)	100(24.6)		
10 001~20 000	300	198(24.7)	102(25.1)		
>20 000	302	197(24.6)	105(25.9)		

续表

项目	例数	非 PPSUI 组($n=802$)	PPSUI 组($n=406$)	$t/Z/\chi^2$	P
便秘史				2.084	0.149
否	1 067	716(89.3)	351(86.5)		
是	141	86(10.7)	55(13.5)		
慢性咳嗽史				0.020	0.887
否	1 200	796(99.3)	404(99.5)		
是	8	6(0.7)	2(0.5)		
盆腔器官脱垂史				126.562	<0.001
否	1 131	796(99.3)	335(82.5)		
是	77	6(0.7)	71(17.5)		
妊娠期是否盆底肌锻炼				7.835	0.005
否	1 141	747(93.1)	394(97)		
是	67	55(6.9)	12(3)		

注:^a 数据以 $\bar{x}\pm s$ 形式表示;^b 数据以 $M(P_{25}, P_{75})$ 形式表示;其他数据均以 $n(\%)$ 形式表示。

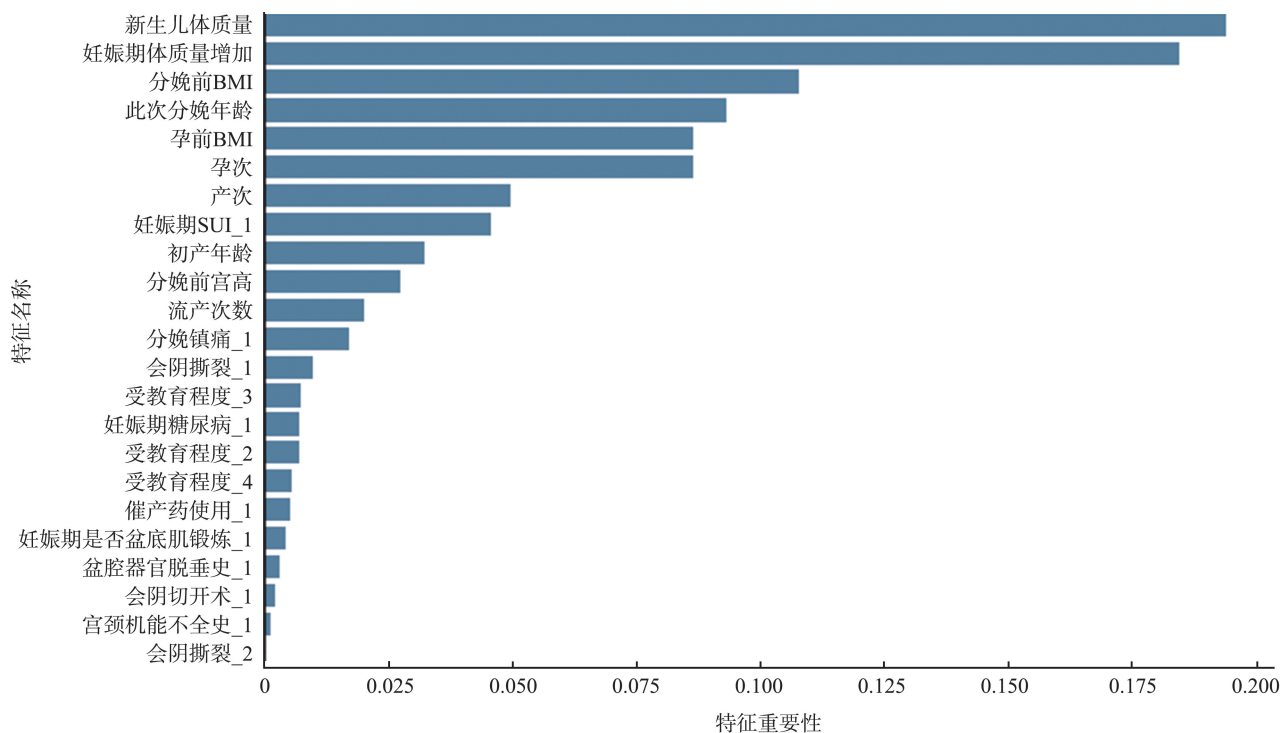


图2 随机森林特征重要性排名

Figure 2 Importance ranking of random forest features

2.3 预测模型的构建与性能评估

基于筛选出的关键特征,分别构建了 Logistic 回归、弹性网络回归、随机森林、SVM 和 XGBoost 5 种 PPSUI 预测模型。所有模型均经过超参数优化后进行训练,并使用准确率、精确率、召回率、特异性、F1 分数、AUC 及其 95% CI 等多项指标对模型性能进行全面评估。训练集十折交叉验证和 Bootstrap 方法评估结果显示, XGBoost 模型分类性能最佳 (AUC = 0.981, 95% CI: 0.975 ~ 0.987, $P < 0.01$), 其次为随机森林 (AUC = 0.979, 95% CI: 0.970 ~ 0.986, $P < 0.01$)。SVM、Logistic 回归和弹性网络回归的 AUC 值分别为 0.973 (95% CI:

0.964 ~ 0.982, $P < 0.01$)、0.963 (95% CI: 0.952 ~ 0.972, $P < 0.01$) 和 0.961 (95% CI: 0.951 ~ 0.971, $P < 0.01$)。见表 2, 图 3。

测试集验证结果进一步支持训练集的发现。XGBoost 模型在测试集中表现最佳 (AUC = 0.993, 95% CI: 0.985 ~ 0.998, $P < 0.01$), 分类性能优于其他模型。随机森林次之 (AUC = 0.989, 95% CI: 0.980 ~ 0.995, $P < 0.01$)。SVM、Logistic 回归和弹性网络回归的 AUC 值分别为 0.981 (95% CI: 0.964 ~ 0.993, $P < 0.01$)、0.964 (95% CI: 0.940 ~ 0.983, $P < 0.01$) 和 0.960 (95% CI: 0.933 ~ 0.982, $P < 0.01$)。见表 3, 图 4。

表 2 5 种机器学习预测模型在训练集的性能比较
Table 2 Comparison of the performance of five machine learning prediction models in the training set

预测模型	准确率	精确率	召回率	特异性	F1 分数	AUC(95%CI)	<i>P</i>
Logistic 回归	0.897	0.879	0.803	0.944	0.840	0.963(0.952~0.972)	<0.01
弹性网络回归	0.897	0.889	0.791	0.950	0.837	0.961(0.951~0.971)	<0.01
随机森林	0.938	0.929	0.883	0.966	0.905	0.979(0.970~0.986)	<0.01
SVM	0.915	0.893	0.849	0.949	0.870	0.973(0.964~0.982)	<0.01
XGboost	0.930	0.898	0.892	0.949	0.895	0.981(0.975~0.987)	<0.01

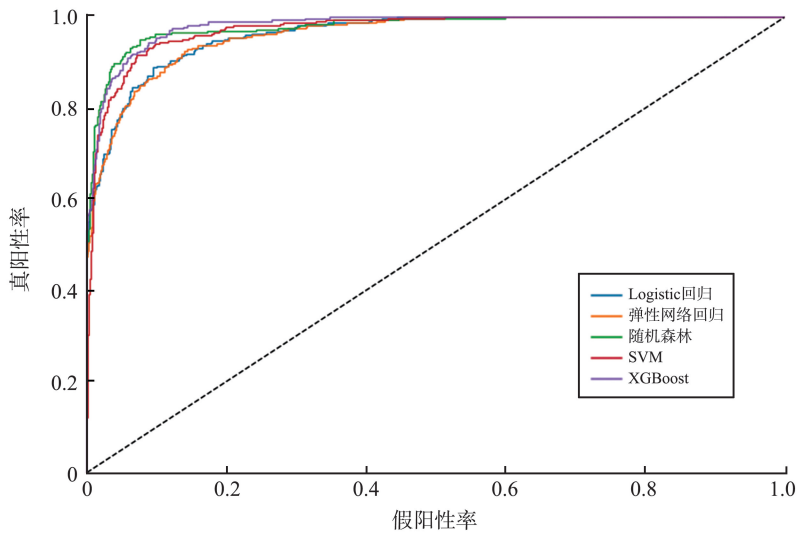


图 3 训练集 ROC 曲线

Figure 3 ROC curve of the training set

表 3 5 种机器学习预测模型在测试集的性能比较
Table 3 Comparison of the performance of five machine learning prediction models in the test set

预测模型	准确率	精确率	召回率	特异性	F1 分数	AUC(95%CI)	<i>P</i>
Logistic 回归	0.909	0.893	0.827	0.950	0.859	0.964(0.940~0.983)	<0.01
弹性网络回归	0.900	0.952	0.741	0.981	0.833	0.960(0.933~0.982)	<0.01
随机森林	0.942	0.913	0.913	0.957	0.913	0.989(0.980~0.995)	<0.01
SVM	0.926	0.870	0.914	0.932	0.892	0.981(0.964~0.993)	<0.01
XGboost	0.950	0.906	0.951	0.950	0.928	0.993(0.985~0.998)	<0.01

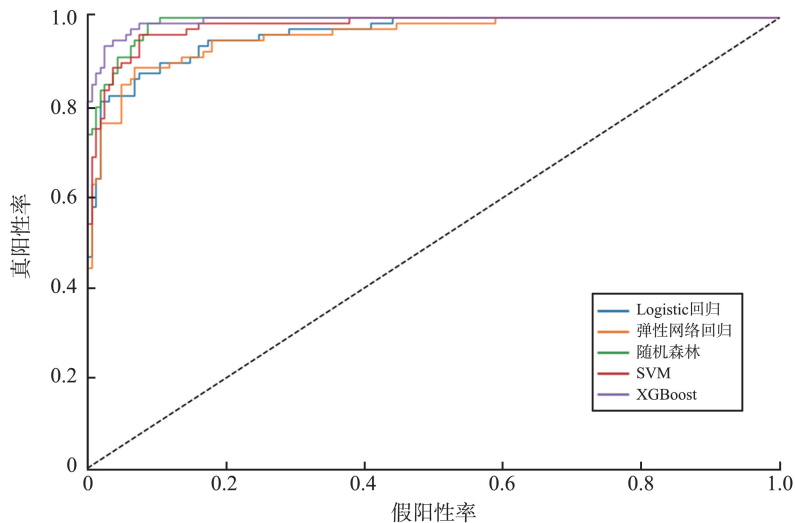


图 4 测试集 ROC 曲线

Figure 4 ROC curve of the test set

2.4 临床适用性与预测概率的可靠性分析

决策曲线分析结果显示,XGBoost 在训练集和测试集的中等阈值范围(0.2~0.6)内提供了最高的净收益,显著优于其他模型,表明其在不同决策阈值下均具有较高的临床适用性和预测价值。随机森林在部分阈值范围(尤其是 0.2~0.6 之间)接近 XGBoost,整体表现稳定,但在高阈值区域(>0.6)净收益略低,提示其在高风险个体识别方面可能存在一定局限性。Logistic 回归在 0.3~0.5 阈值范围内

的净收益较高,表明其在中等风险人群筛查中的一些价值,但在低阈值(<0.3)和高阈值(>0.6)区域的净收益较低,说明其适用于特定的决策场景,但在极端风险情况下预测能力有限。SVM 在测试集中的 DCA 曲线波动较大,特别是在中高阈值范围内存在不稳定性,反映出其对数据分布较为敏感,可能导致泛化能力不足。弹性网络回归整体净收益较低,在所有阈值范围内均未表现出较大的优势,提示其临床适用性较为有限。见图 5。

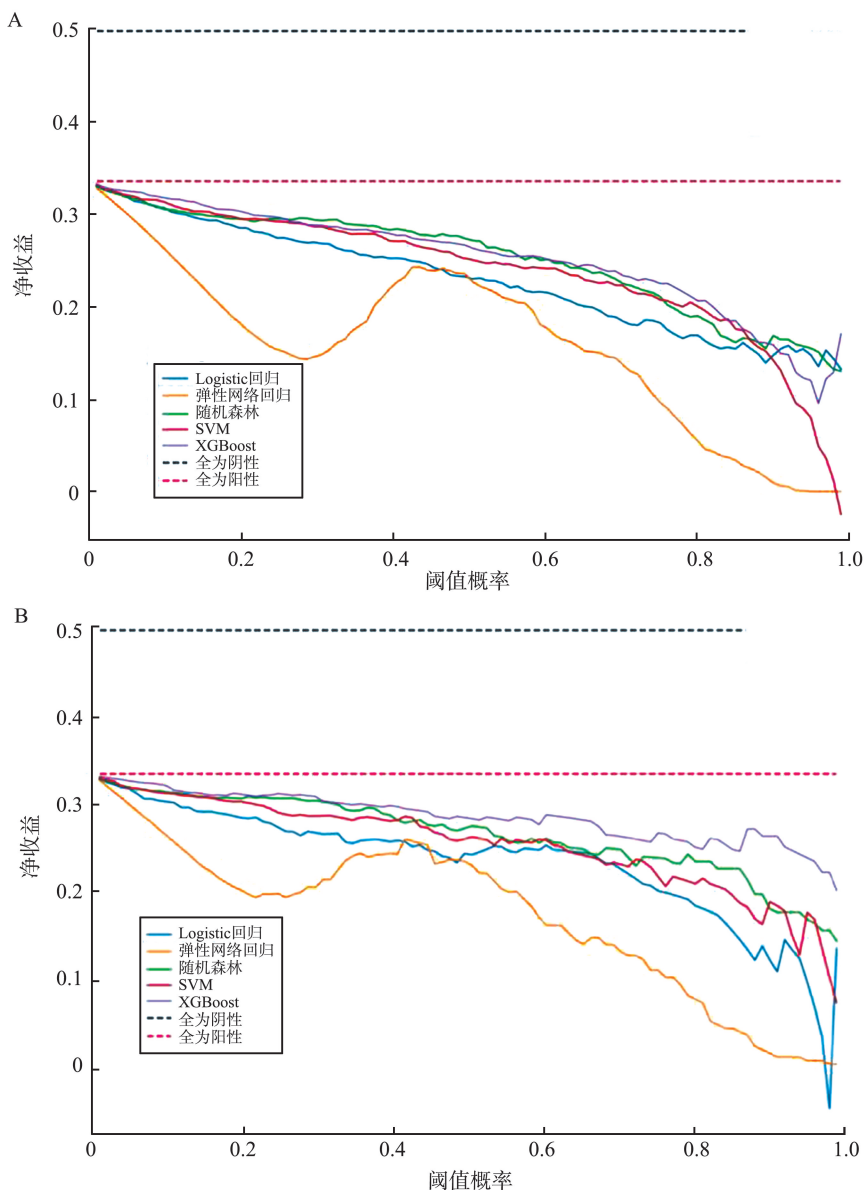


图5 训练集(A)和测试集(B)DCA 曲线

Figure 5 DCA curves for training set (A) and test set (B)

校准曲线分析结果显示,XGBoost 在 0.3~0.8 的概率范围内校准效果最佳,曲线整体最接近理想校准曲线,表明其预测概率与真实发生率匹配度较高,具备良好的校准能力。其在训练集和测试集中均保持较优的校准表现,提示模型在不同数据集上

的稳定性较强,适用于 PPSUI 风险预测的临床应用。随机森林在训练集中校准效果较好,但在测试集中高概率区域(>0.8)出现一定偏离,提示其在外数据上的稳定性可能受影响。Logistic 回归和 SVM 在训练集的低概率区域(<0.3)存在低估趋势,

而在0.4~0.7的中等概率范围内校准较好。然而,在测试集中,这两种模型的校准曲线波动较大,提示其泛化能力可能受限。弹性网络回归在训练集和测

试集的低概率区域均表现出明显的预测偏差,整体校准能力较差,可能影响其在实际临床应用中的可靠性。见图6。

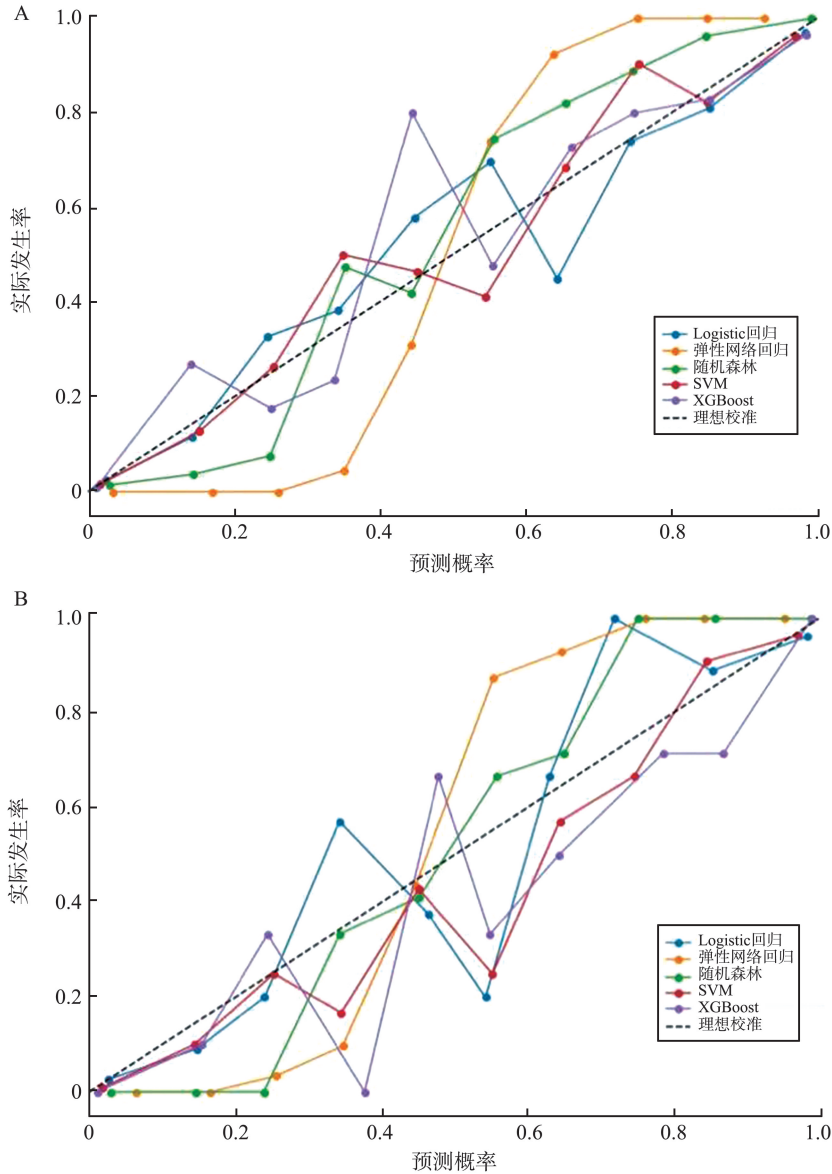


图6 训练集(A)和测试集(B)校准曲线

Figure 6 Calibration curves for training set (A) and test set (B)

2.5 最优预测模型的筛选

综合评估模型的性能、临床适用性及预测概率的可靠性,最终筛选出最优预测模型。在测试集中, XGBoost 的 AUC 值最高为 0.993 (95% CI: 0.985~0.998, $P < 0.01$), 分类性能优于其他模型。决策曲线分析显示, XGBoost 在 0.2~0.6 的中等决策阈值范围内提供了最高的净收益, 临床适用性较优。校准曲线分析表明, XGBoost 在 0.3~0.8 概率范围内校准效果最佳, 尤其在 0.3~0.7 区间, 其预测概率与真实发生率匹配度较高, 显示出稳健的预测可靠性。相比之下, 随机森林的 AUC 为值 0.989 (95% CI:

0.980~0.995, $P < 0.01$) 接近 XGBoost, 在部分决策阈值范围内净收益相近, 但校准曲线在低概率和部分高阈值区域略逊, 提示其在外部验证时可能存在一定不稳定性。Logistic 回归在 0.3~0.5 阈值范围内的净收益较高, 具备较好的稳定性和可解释性, 但在分类性能和临床适用性方面仍不及 XGBoost。SVM 和弹性网络回归的 DCA 曲线波动较大, 适用性相对有限, 特别是在测试集中校准能力较差, 可能影响其泛化能力。综上, XGBoost 在分类性能、临床适用性和校准能力方面均表现最佳, 被确定为 PPSUI 风险预测的最优模型。

3 讨论

PPSUI 是一种常见但易被忽视的产后健康问题,对产妇身体活动、情感关系和心理健康带来多重影响^[15-16]。尽管 PPSUI 并非致命性疾病,但高昂的治疗费用和长期康复过程对患者及其家庭构成了沉重负担^[17]。精准的早期筛查和个体化防治策略对于减少 PPSUI 的发生至关重要^[18]。目前,通过纳入多种变量构建的个体风险预测模型,已成为医疗保健提供者筛查高危人群的重要工具^[19]。国际尿失禁协会(International Continence Society, ICS)/国际泌尿妇科协会(International Urogynecological Association, IUGA)均强调,基于多因素风险评估的预测模型在提高早期筛查效率和优化防治策略方面具有重要应用价值^[1]。然而,目前的 PPSUI 预测工具仍存在局限性:①多数研究仅关注产后 6 周至 7 个月,在长期预测能力上存在不足;②现有模型主要依赖 Logistic 回归等传统统计方法,该类方法受线性假设限制,难以精准刻画复杂变量间的非线性关系,影响高危个体的识别精度。因此,本研究基于最新 PPSUI 定义,引入 XGBoost、随机森林、SVM 等机器学习算法构建预测模型,并优化变量筛选及模型验证策略,以提高预测性能和临床适用性。

本研究基于最新定义构建了 PPSUI 风险预测模型,并在时间维度上将预测范围延展至产后 12 个月。相较于以往仅涵盖产后 6 周至 7 个月的研究^[20-22]。本研究有助于识别晚发型 PPSUI 病例,提高早期筛查的覆盖面,有效填补了现有模型在长期预测能力上的不足,进一步提升了模型的临床适用性。本研究系统纳入新生儿体质量、妊娠期体质量增加等关键相关变量,构建了更全面的 PPSUI 风险预测模型,为医护人员提供了科学的决策支持。基于该预测模型,临床医务人员可以精准评估高危孕产妇的个体风险,制定个性化干预策略,特别是针对可控风险因素的管理,如合理控制孕期体质量、预防高龄妊娠和流产及指导孕期盆底肌锻炼等措施,可以有效降低 PPSUI 的发生风险。尽管已有研究证明孕期盆底肌锻炼能有效降低 PPSUI 的发病率^[23],但本研究中纳入人群的孕期盆底肌锻炼比例较低,提示未来应加强健康教育与产前保健工作,以提升孕产妇对盆底健康管理的重视程度,并促进其在孕期及产后积极开展盆底康复训练。本研究构建的 PPSUI 预测模型为该疾病的早期筛查与精准防治提供了科学依据,可优化医疗资源配置,提高风险分层

管理的精准度,增强个体化筛查和干预的效率,从而推动产后健康管理体的优化与可持续发展。

既往 PPSUI 风险预测模型多基于 Logistic 回归^[20-22],该方法因计算效率高、可解释性强,在医学研究中被广泛应用。然而,在处理复杂非线性特征和高维交互效应时,其适用性可能受限^[24]。近年来,机器学习算法凭借灵活的建模能力,在刻画特征间复杂关系、捕捉非线性模式方面展现出优势,逐渐应用于医学预测领域^[24-25]。本研究基于多种机器学习算法构建 PPSUI 预测模型,并从分类性能、临床适用性及校准能力等方面进行系统评估。结果显示,Logistic 回归在 AUC 方面与部分机器学习模型相近,且在 0.3~0.5 决策阈值范围内 DCA 净收益较高,提示其在中等风险人群筛查中具有一定临床价值。然而,在低(<0.3)和高(>0.6)阈值范围内,其净收益相对较低,且在测试集中校准曲线波动较大,提示泛化能力可能受限。相比之下,XGBoost 在 AUC、DCA 和校准曲线评估中均表现最佳,能够在不同风险水平和决策阈值下保持稳定的预测能力和较高的可靠性,尤其适用于高危人群筛查和个体化风险评估。随机森林在部分指标上接近 XGBoost,但在高阈值区域的临床适用性略逊一筹。SVM 和弹性网络回归在部分评估指标上的表现相对较弱,泛化能力欠佳,在本研究中优势有限。整体而言,Logistic 回归因其良好的可解释性和稳定性,仍适用于部分对模型透明性要求较高的场景。然而,在复杂、非线性、高维交互的数据环境下,XGBoost 在分类性能、决策支持价值和校准稳定性方面更具优势,可为 PPSUI 的精准筛查和个体化干预策略提供科学依据。

2018 年《全国医院信息化建设标准与规范(试行)》提出,要利用人工智能技术对疾病进行风险预测^[26]。随着我国居民健康档案和医院电子信息系统的互联互通,海量医疗数据的电子化和数字化为 PPSUI 风险评估及高危人群管理奠定了坚实基础^[27]。在此背景下,将 XGBoost 模型嵌入医院电子信息系统并输入孕妇数据,可以辅助临床评估,精准识别高风险患者,为后续个性化干预提供科学依据。此外,PPSUI 机器学习模型的透明性和可解释性显著提升,不仅增强了医生和患者对预测结果的信任,为卫生管理者提供了更可靠的决策支持,还为 PPSUI 的精准预测和个性化产后健康管理提供了新思路和研究范式。

PPSUI 的发生受多种因素影响。研究表明,阴道分娩、高龄、高 BMI 指数、妊娠期体质量过度增加

和妊娠期 SUI 等因素与 PPSUI 风险相关^[4-5]。本研究基于随机森林特征重要性分析发现,新生儿体质量、妊娠期体质量增加等变量在预测 PPSUI 风险方面具有较高的模型贡献度,这一结果与既往研究报道^[4,28]一致,提示上述变量在 PPSUI 的发生机制中发挥一定作用。其机制为通过增加腹腔压力,对盆底肌肉和韧带造成额外负担,从而增加 PPSUI 的发生风险。此外,高龄产妇由于盆底肌功能退化及组织弹性降低,更易发生 PPSUI^[29]。值得注意的是,分娩方式在本研究中经单因素分析未显示出差异的统计学意义,这一结果与部分既往研究^[4-5]存在差异,可能与样本中年龄、BMI 和孕次等关键特征分布的均质性有关,这些因素削弱了分娩方式对 PPSUI 风险的独立影响。此外,新生儿体质量、妊娠期体质量增加和分娩前 BMI 等在模型中的权重较高,可能掩盖了分娩方式的作用。现代产科技术的进步,尤其是剖宫产和辅助分娩技术的改进,使阴道分娩与剖宫产对盆底功能的影响趋于一致。同时,“全面三孩政策”下高龄产妇的增加和多胎妊娠的累积效应可能进一步改变了分娩方式与 PPSUI 风险的关系。这些因素综合作用解释了本研究结果与以往研究的差异,并提示未来需深入探讨其机制。

尽管本研究基于多种机器学习算法构建了的 PPSUI 风险预测模型,在分类性能、临床适用性及校准能力方面表现优异,但仍存在一定局限性:①研究数据来源于单一医疗中心,可能影响样本代表性及模型的泛化能力,未来需结合多中心大规模数据进一步验证;②尽管本研究采用了十折交叉验证和 Bootstrap 方法评估模型的泛化能力,并结合 DCA 和校准曲线分析模型的临床适用性,但仍需在独立外部数据集中进行验证,以确保其稳定性和广泛适用性;③本研究构建的模型基于静态数据,未考虑风险因素在时间维度上的动态变化,在反映孕产期风险状态的时序演变方面存在一定局限性。未来可探索引入动态更新机制,以提升模型在连续监测和动态管理中的应用价值。

综上所述,本研究为 PPSUI 高危人群的精准筛查提供了科学工具,未来研究可进一步优化模型并拓展其临床应用,以提高其推广价值和实际可行性。

参考文献:

- [1] Doumouchsis SK, de Tayrac R, Lee J, et al. An International Continence Society (ICS)/International Urogynecological Association (IUGA) joint report on the terminology for the assessment and management of obstetric pelvic floor disorders [J]. *Int Urogynecol J*, 2023, 34 (1): 1-42.
- [2] You C, Zhao YJ, Zhang C, et al. Pelvic floor parameters predict postpartum stress urinary incontinence: a prospective MRI study [J]. *Insights Imaging*, 2023, 14(1): 160. doi:10.1186/s13244-023-01488-5
- [3] Coll CM, Franco EM, Qureal LA, et al. Hormonal influence in stress urinary incontinence during pregnancy and postpartum [J]. *Reprod Sci*, 2022, 29(8): 2190-2199.
- [4] Dai SD, Chen HT, Luo TZ. Prevalence and factors of urinary incontinence among postpartum: systematic review and meta-analysis [J]. *BMC Pregnancy Childbirth*, 2023, 23(1): 761. doi:10.1186/s12884-023-06059-6
- [5] Wang K, Xu XL, Jia GM, et al. Risk factors for postpartum stress urinary incontinence: a systematic review and meta-analysis [J]. *Reprod Sci*, 2020, 27 (12): 2129-2145.
- [6] Rashidi Fakari F, Hajian S, Darvish S, et al. Predictors of help-seeking behaviors in women with urinary incontinence: Based on Iranian women's lens [J]. *PLoS One*, 2023, 18 (8): e0289785. doi: 10.1371/journal.pone.0289785
- [7] Rashidi Fakari F, Hajian S, Darvish S, et al. Explaining factors affecting help-seeking behaviors in women with urinary incontinence: a qualitative study [J]. *BMC Health Serv Res*, 2021, 21 (1): 60. doi:10.1186/s12913-020-06047-y
- [8] 李巧巧, 于霞, 罗映娟, 等. 女性压力性尿失禁疾病负担研究进展 [J]. *中国计划生育和妇产科*, 2022, 14(9): 10-12.
- [9] Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery [J]. *Lancet Oncol*, 2019, 20(5): e262-e273.
- [10] Zhang A, Xing L, Zou J, et al. Shifting machine learning for healthcare from development to deployment and from models to data [J]. *Nat Biomed Eng*, 2022, 6 (12): 1330-1345.
- [11] Mikuš M, Corić M, Matak L, et al. Validation of the UDI-6 and the ICIQ-UI SF - Croatian version [J]. *Int Urogynecology J*, 2020, 31(12): 2625-2630.
- [12] Huang L, Zhang SW, Wu SL, et al. The Chinese version of ICIQ: a useful tool in clinical practice and research on urinary incontinence [J]. *Neurourol Urodyn*, 2008, 27(6): 522-524.
- [13] Legendre B, Cerasuolo D, Dejardin O, et al. How to deal with missing data? Multiple imputation by chained equations: recommendations and explanations for clinical practice [J]. *Nephrol Ther*, 2023, 19(3): 171-179.
- [14] Speiser JL. A random forest method with feature selection for developing medical prediction models with clus-

- tered and longitudinal data[J]. *J Biomed Inform*, 2021, 117: 103763. doi:10.1016/j.jbi.2021.103763
- [15] AIQuaiz AM, Kazi A, AIYousefi N, et al. Urinary incontinence affects the quality of life and increases psychological distress and low self-esteem[J]. *Healthcare*, 2023, 11(12): 1772. doi:10.3390/healthcare11121772
- [16] Suar G, Cevik F, Simal Yavuz N, et al. Urinary incontinence in the postpartum 1-year period: its prevalence and effect on psychosocial status of women[J]. *Low Urin Tract Symptoms*, 2023, 15(5): 191-199.
- [17] Chisholm LP, Sebesta EM, Gleicher S, et al. The burdens of incontinence: quantifying incontinence product usage and costs in women[J]. *Neurourol Urodyn*, 2022, 41(7): 1601-1611.
- [18] Benítez-Andrades JA, García-Ordás MT, Álvarez-González M, et al. Detection of the most influential variables for preventing postpartum urinary incontinence using machine learning techniques[J]. *Digit Health*, 2022, 8: 2055207622111289. doi:10.1177/2055207622111289
- [19] 史可为, 吴亚飞, 方亚. 临床预测模型的可解释性及应用进展[J]. *现代预防医学*, 2023, 50(6): 1122-1127.
SHI Kewei, WU Yafei, FANG Ya. Interpretability and application of clinical prediction models[J]. 2023, 50(6): 1122-1127.
- [20] Xu CC, Guo Y, Chi XL, et al. Establishment and validation of a simple nomogram for predicting early postpartum stress urinary incontinence among women with vaginal delivery: a retrospective study [J]. *BMC Womens Health*, 2023, 23(1): 8. doi:10.1186/s12905-023-02160-2
- [21] Liu W, Qian LX. Establishment and validation of a risk prediction model for postpartum stress urinary incontinence based on pelvic floor ultrasound and clinical data [J]. *Int Urogynecol J*, 2022, 33(12): 3491-3497.
- [22] Eric Jelovsek J, Piccorelli A, Barber MD, et al. Prediction models for postpartum urinary and fecal incontinence in primiparous women[J]. *Female Pelvic Med Reconstr Surg*, 2013, 19(2): 110-118.
- [23] Zhang DF, Bo KR, Montejo R, et al. Influence of pelvic floor muscle training alone or as part of a general physical activity program during pregnancy on urinary incontinence, episiotomy and third- or fourth-degree perineal tear: Systematic review and meta-analysis of randomized clinical trials[J]. *Acta Obstet Gynecol Scand*, 2024, 103(6): 1015-1027.
- [24] Krishnan R, Rajpurkar P, Topol EJ. Self-supervised learning in medicine and healthcare [J]. *Nat Biomed Eng*, 2022, 6(12): 1346-1352.
- [25] Chen RJ, Wang JJ, Williamson DFK, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare[J]. *Nat Biomed Eng*, 2023, 7(6): 719-742.
- [26] 国家卫生健康委员会发布《全国医院信息化建设标准与规范(试行)》:明确医院信息化建设内容和建设要求[J]. *医学信息学杂志*, 2018(4): 94.
- [27] 孙焯祥, 吕筠, 沈鹏, 等. 健康医疗大数据驱动下的疾病防控新模式[J]. *中华流行病学杂志*, 2021, 42(8): 1325-1329.
SUN Yexiang, LYU Jun, SHEN Peng, et al. A new model for disease control and prevention driven by big data in healthcare[J]. *Chinese Journal of Epidemiology*, 2021, 42(8): 1325-1329.
- [28] Sheng XM, Wang J, Shaw JM, et al. Pelvic floor support and symptoms 1 year postpartum in relation to excess body habitus before, during and after first vaginal birth[J]. *Female Pelvic Med Reconstr Surg*, 2022, 28(5): 273-279.
- [29] DeLancey JOL, Masteling M, Pipitone F, et al. Pelvic floor injury during vaginal birth is life-altering and preventable: what can we do about it? [J]. *Am J Obstet Gynecol*, 2024, 230(3): 279-294.

(编辑:房红娟)