

妇产科疾病多学科研究专题

围产期抑郁症辅助诊断预测模型的构建
及机器学习算法的筛选王 妤¹, 褚嘉栋¹, 孙 娜¹, 韩 婧¹, 沈月平¹,
周 磊², 朱新平^{2,3}, 张晓斌⁴, 杨 勇⁴

- [1. 苏州大学苏州医学院公共卫生学院 流行病与卫生统计学系, 江苏 苏州, 215213;
2. 江苏智慧智能软件科技有限公司, 江苏 泰州, 225300;
3. 智慧智能软件科技(湖南)有限公司, 湖南 湘潭, 411100;
4. 江苏省苏州市广济医院/苏州大学附属广济医院, 江苏 苏州, 215137]

摘要: 目的 基于机器学习(ML)算法构建孕产妇围产期抑郁症(PND)辅助诊断预测模型并评估不同 ML 算法模型的性能。方法 采用 9 条目患者健康问卷抑郁量表(PHQ-9)对 5 814 例孕产妇(产前研究对象 4 665 例, 产后研究对象 1 149 例)进行评估, 收集 7 种量表的 19 个量表维度变量和人口学特征作为观察变量。按照年龄分别对产前、产后研究对象进行 1:1 倾向性评分匹配, 利用单因素分析及 Pearson 相关系数确定特征选择变量。基于 Logistic 回归模型、随机森林(RF)、支持向量机(SVM)、极限梯度提升树(XGBoost)、反向传播(BP)神经网络这 5 种 ML 算法分别纳入所有变量和特征选择变量集, 构建产前、产后抑郁的诊断模型。采用 5 折交叉验证方法评估模型的预测性能, 评价指标包括灵敏度、特异度和曲线下面积(AUC)。结果 纳入不同变量的情况下, 通过 5 种 ML 算法分别基于产前研究对象、产后研究对象构建的预测模型的灵敏度、特异度、AUC 均在 0.600~0.900 范围内; RF 算法在产前预测模型(纳入所有变量时, AUC 为 0.834; 纳入特征选择变量集时, AUC 为 0.849)和产后预测模型(纳入所有变量时, AUC 为 0.873; 纳入特征选择变量集时, AUC 为 0.864)的构建中均为最优算法。结论 基于 5 种 ML 算法构建的预测模型均可有效预测孕产妇 PND 风险, 其中以 RF 算法的表现最优, 为开发快速筛查和诊断 PND 的辅助工具提供了参考依据。

关键词: 围产期抑郁症; 人工智能; 机器学习算法; 预测模型; 随机森林; 辅助诊断

中图分类号: R 711; R 74 文献标志码: A 文章编号: 1672-2353(2023)18-093-07 DOI: 10.7619/jcmp.20232044

Construction of a predictive model for auxiliary diagnosis
of perinatal depression and screening of machine
learning algorithmWANG Yu¹, CHU Jiadong¹, SUN Na¹, HAN Qiang¹, SHEN Yueping¹,
ZHOU Lei², ZHU Xinping^{2,3}, ZHANG Xiaobin⁴, YANG Yong⁴

- [1. Department of Epidemiology and Biostatistics, School of Public Health of Suzhou Medical College of Soochow University, Suzhou, Jiangsu, 215123; 2. Jiangsu Smart Intelligence Software Technical Company Limited, Taizhou, Jiangsu, 225300; 3. Smart Intelligence Software Technical Company Limited (Hunan), Xiangtan, Hunan, 411100; 4. Suzhou Guangji Hospital of Jiangsu Province, Guangji Hospital Affiliated to Soochow University, Suzhou, Jiangsu, 215137]

Abstract: Objective To construct a predictive model for assisted diagnosis of maternal perinatal depression (PND) based on machine learning (ML) algorithms and to evaluate the performance of different ML algorithm models. **Methods** A total of 5 814 pregnant women (4 665 prenatal study subjects and 1 149 postnatal study subjects) were evaluated using the 9-item Patient Health Questionnaire Depression Scale (PHQ-9). A total of 19 scale dimension variables of 7 Scales and demographic

收稿日期: 2023-06-27 修回日期: 2023-08-02

基金项目: 江苏省苏州市临床重点病种诊疗技术专项(LCZX201616); 江苏省苏州市姑苏卫生人才计划(GSW2020096); 江苏省苏州市心障碍障碍临床医学中心建设项目(Szlcyczx202109)

通信作者: 杨勇, E-mail: szgjyy_tg@126.com

characteristics were collected as observation variables. Prenatal and postnatal subjects were matched at a 1:1 ratio propensity score according to age. The feature selection variables were determined by single factor analysis and Pearson correlation coefficient. A diagnostic model for prenatal and postnatal depression was constructed based on five ML algorithm, including Logistic regression model, Random Forest (RF), support vector machine (SVM), Limit Gradient Lift Tree (XGBoost) and Backpropagation (BP) neural network. A 5-fold cross-validation method was used to evaluate the predictive performance of the model, including sensitivity, specificity and area under the curve (AUC). **Results** When different variables were included, the sensitivity, specificity and AUC of the prediction model constructed by five ML algorithms based on prenatal and postnatal subjects were all within the range of 0.600 to 0.900. RF algorithm was the optimal algorithm in the construction of both prenatal prediction model (when all variables were included, the AUC was 0.834; when the feature selection variable set was included, the AUC was 0.849) and postnatal prediction model (when all variables were included, the AUC was 0.873; when the feature selection variable set was included, the AUC was 0.864). **Conclusion** The prediction model based on five ML algorithms can effectively predict the risk of PND in pregnant women, and the performance of RF algorithm is the best, which provides a reference for the development of auxiliary tools for rapid screening and diagnosis of PND.

Key words: perinatal depression; artificial intelligence; machine learning algorithm; predictive mode; random forest; auxiliary diagnosis

围产期抑郁症(PND)是指妇女在怀孕、分娩及产后康复期间产生抑郁情绪并达到一定严重程度^[1]。研究^[2]表明,低中等收入水平国家的孕产妇PND患病率高达48.5%,远高于高等收入水平国家。中国孕产妇的PND总体患病率为16.3%(产前抑郁症为19.7%,产后抑郁症为14.8%),且近10年呈明显上升趋势^[3]。若未及时治疗,PND可能引起不良妊娠结局、低体质量儿、母婴关系不佳或亲情缺失、自杀或杀婴等不良后果^[4]。随着大数据时代的来临,基于机器学习(ML)算法的人工智能(AI)模型越来越多地被用于疾病的诊断、分型及预后预测等方面^[5-6]。基于ML算法,联合多个临床量表以及人口学特征、环境因素构建诊断模型,并开发界面友好、易于操作的辅助数字医疗APP或小程序,不仅无创、成本低廉,而且有利于推广,但目前仅有少量研究^[7-8]构建并评估了基于量表开发的孕产妇抑郁诊断模型。本研究对产前抑郁和产后抑郁诊断算法进行探索与验证,以期明确诊断孕产妇抑郁的最优算法。

1 对象与方法

1.1 研究对象

本研究中的研究对象资料均来源于苏州市临床诊疗专项研究数据库,共纳入符合条件的孕产

妇5 814例,其中产前研究对象4 665例,产后研究对象1 149例。本研究经苏州市广济医院伦理委员会审核批准,伦理批件号为苏广伦审2016-010。

1.2 方法

1.2.1 资料收集:①人口学变量。收集研究对象的人口学特征资料,包括年龄、是否独生子女(DSYN)、教育水平(JYSP)、婚姻状况(HYZT)、家族精神病史(qrmh)、婆媳关系(PXMYD)、是否首次妊娠(HRCS)、居住条件(JZTJ)、向孩子发脾气的冲动(EPDSII)。②量表维度变量。本研究调查量表包括艾森克个性量表(EPQ)、社会支持量表(SSRS)、特质应对方式量表(TCSQ)、睡眠质量量表(PSQI)、轻躁狂症评估量表(HCL-32)、心理障碍问卷(MDQ)、广泛性焦虑问卷(GAD)、患者健康问卷抑郁量表(PHQ-9)。将各量表分别分解为不同的维度变量进行分析,最终共得到19个量表维度变量,可从不同角度评估孕产妇的心理健康状况,见表1。

1.2.2 结局判定:第5版《精神疾病诊断与统计手册》(DSM-5)推荐将9条目患者健康问卷抑郁量表(PHQ-9)^[9]作为临床抑郁筛查及抑郁严重程度评估的首选工具。PHQ-9具有良好的内部一致性,Conbach's α 系数为0.8~0.9^[10]。本研究根据PHQ-9评分判定孕产妇抑郁结局,PHQ-9评分 ≥ 10 分为抑郁阳性,<10分为抑郁阴性。

表 1 各调查量表的维度变量名称及评分说明

量表名称	量表维度	变量名称	评分说明
EPQ	艾森克个性测试交际维度(e)	<i>epq_e</i>	评分高表示爱交际
	艾森克个性测试情绪维度(n)	<i>epq_n</i>	评分高表示情绪不稳定
	艾森克个性测试关心他人维度(p)	<i>epq_p</i>	评分高表示不关心他人
	艾森克个性测试情感掩饰性维度(l)	<i>epq_l</i>	掩饰性,高低不确定
SSRS	社会支持客观支持分	<i>SSRS_K</i>	—
	社会支持主观支持分	<i>SSRS_Z</i>	—
	对支持的利用度	<i>SSRS_L</i>	—
TCSQ	特质应对方式消极总分	<i>TCSQ_NC</i>	—
	特质应对方式积极总分	<i>TCSQ_PC</i>	—
PSQI	睡眠质量主观感受	<i>PSQI_14</i>	0~3分,评分越高表示睡眠质量主观感受越差
	睡眠时间	<i>NPSQI24</i>	0~3分,评分越高表示睡眠时间越短
	睡眠持续性	<i>PSQI_3</i>	0~3分,评分越高表示睡眠持续性越差
	睡眠效率	<i>NPSQI13</i>	0~3分,评分越高表示睡眠效率越差
	睡眠障碍	<i>NPSQI513</i>	0~3分,评分越高表示睡眠障碍越高
	催眠药物	<i>PSQI_15</i>	0~3分,评分越高表示催眠药物用得越多
	睡眠日间功能	<i>NPSQI617</i>	0~3分,评分越高表示睡眠日间功能越差
HCL-32	有无轻躁狂症异常	<i>NHCL32</i>	二分类变量,1为有轻躁狂症异常(总分≥14分)
MDQ	有无心理障碍异常	<i>NMDQ</i>	二分类变量,1为有心理障碍异常(总分≥6分)
GAD	有无广泛性焦虑	<i>NGAD</i>	二分类变量,1为有广泛性焦虑(总分≥10分)

1.2.3 数据预处理:由于实际人群的抑郁阳性率较低(10%~15%),模型拟合过程中容易产生灵敏度较低的现象。鉴于此,本研究采用倾向性评分匹配法,以年龄作为匹配变量,对病例组进行1:1匹配,产前研究对象成功匹配201对,产后研究对象成功匹配52对。在此基础上构建模型,可更准确地筛选较优的算法。

1.2.4 特征选择变量:采用单因素分析法和去除多重共线性方法进行特征选择。分别对产前、产后的抑郁与非抑郁人群的8个人口学特征、19个量表维度变量的分布情况进行描述及比较,探讨相关因素。正态分布的计量资料采用($\bar{x} \pm s$)描述,非正态分布的计量资料采用 $[M(P_{25}, P_{75})]$ 描述,计数资料采用 $[n(\%)]$ 描述,组间比较分别采用 t 检验、秩和检验、卡方检验或Fisher确切概率法。所有分析为双侧,显著性水准为 $\alpha = 0.05$ 。计算19个量表维度变量间的Pearson相关系数,将2个变量间的 $|r| > 0.7$ 定义为高度线性相关。若2个变量呈高度线性相关,则选择保留临床可解释性更高的其中1项特征变量。

1.2.5 模型构建及算法筛选:本研究基于5种ML算法即Logistic回归模型、随机森林(RF)、支持向量机(SVM)、极限梯度提升树(XGBoost)和反向传播(BP)神经网络^[11],分别纳入全部变量及特征选择变量集构建产前抑郁、产后抑郁预测模型。绘制受试者工作特征(ROC)曲线评估模

型性能,采用5折交叉验证方法计算评价指标,指标包括灵敏度、特异度和曲线下面积(AUC)。此外,基于5种算法对所有变量进行变量重要性排序。所有分析采用SAS 9.4软件和R 4.3.0软件进行分析。

2 结果

2.1 产前、产后研究对象抑郁状态的单因素分析
产前研究对象中,抑郁阳性者在*JYSP*、*qrmh*、*HRCS*、*PXMYD*、*EPDS11*方面与抑郁阴性者比较,差异有统计学意义($P < 0.05$);产后研究对象中,抑郁阳性者在*DSYN*、*EPDS11*方面与抑郁阴性者比较,差异有统计学意义($P < 0.05$),见表2。

产前研究对象中,抑郁阳性者的17个量表维度变量(除*NHCL32*、*PSQI_15*外)与抑郁阴性者比较,差异有统计学意义($P < 0.05$);产后研究对象中,抑郁阳性者的15个量表维度变量(除*NHCL32*、*NPSQI13*、*PSQI_15*、*NMDQ*外)与抑郁阴性者比较,差异有统计学意义($P < 0.05$),见表3。

2.2 不同预测模型的构建与评价

2.2.1 特征变量选择结果:在产前研究对象中,单因素分析得到22个差异有统计学意义的变量,量表维度变量间的相关性分析发现*epq_n*与*TCSQ_NC*呈高度相关,两者中保留*epq_n*,故最

表 2 抑郁阴性或抑郁阳性的产前、产后研究对象的人口学特征比较 [n(%)]

特征	分类	产前研究对象				产后研究对象			
		抑郁阴性 (n=201)	抑郁阳性 (n=201)	χ^2	P	抑郁阴性 (n=52)	抑郁阳性 (n=52)	χ^2	P
独生子女	否	129(64.18)	140(69.65)	1.360	0.244	29(55.77)	39(75.00)	4.248	0.039
	是	72(35.82)	61(30.35)			23(44.23)	13(25.00)		
教育水平	小学及以下	3(1.49)	0	—	0.005	1(1.92)	1(1.92)	—	0.752
	初中	23(11.44)	40(19.90)			5(9.62)	9(17.31)		
	高中/中专	40(19.90)	51(25.37)			7(13.46)	10(19.23)		
	大专	62(30.85)	64(31.84)			12(23.08)	11(21.15)		
	本科	63(31.34)	43(21.39)			23(44.23)	19(36.54)		
	研究生	10(4.98)	3(1.49)			4(7.69)	2(3.85)		
婚姻状况	未婚	2(1.00)	3(1.49)	—	0.510	1(1.92)	1(1.92)	—	0.135
	初婚	191(95.02)	184(91.54)			48(92.31)	45(86.54)		
	离异	3(1.49)	9(4.48)			1(1.92)	5(9.62)		
	再婚	4(1.99)	4(1.99)			0	1(1.92)		
	丧偶	1(0.50)	1(0.50)			2(3.85)	0		
家族精神病史	无	90(44.78)	173(86.07)	75.755	<0.001	51(98.08)	50(96.15)	—	1.000
	有	111(55.22)	28(13.93)			1(1.92)	2(3.85)		
婆媳关系	极其不满意	2(1.00)	3(1.49)	—	<0.001	1(1.92)	1(1.92)	—	0.068
	非常不满意	3(1.49)	4(1.99)			0	2(3.85)		
	有些不满意	4(1.99)	11(5.47)			2(3.85)	8(15.38)		
	一般	37(18.41)	67(33.33)			7(13.46)	11(21.15)		
	有些满意	27(13.43)	35(17.41)			6(11.54)	9(17.31)		
	非常满意	88(43.78)	59(29.35)			24(46.15)	14(26.92)		
	极其满意	40(19.90)	22(10.95)			12(23.08)	7(13.46)		
首次妊娠	否	89(44.28)	122(60.70)	—	0.001	30(57.69)	29(55.77)	0.039	0.843
	是	112(55.72)	79(39.30)			22(42.31)	23(44.23)		
居住条件	租住	52(25.87)	57(28.36)	—	0.766	8(15.38)	13(25.00)	—	0.566
	有房	146(72.64)	140(69.65)			43(82.69)	38(73.08)		
	其他	3(1.49)	4(1.99)			1(1.92)	1(1.92)		
向孩子发脾气的冲动	无	192(95.52)	132(65.67)	57.265	<0.001	47(90.38)	36(69.23)	7.220	0.007
	有	9(4.48)	69(34.33)			5(9.62)	16(30.77)		

表 3 抑郁阴性或抑郁阳性的产前、产后研究对象的全量表维度变量比较 ($\bar{x} \pm s$) [M(P₂₅, P₇₅)] [n(%)]

变量	产前研究对象				产后研究对象				
	抑郁阴性(n=201)	抑郁阳性(n=201)	t/ χ^2 /Z	P	抑郁阴性(n=52)	抑郁阳性(n=52)	t/ χ^2 /Z	P	
epq_e/分	10.00(7.00, 13.00)	9.00(6.00, 13.00)	2.069	0.039	10.69±4.01	10.69±4.01	2.860	0.005	
epq_n/分	8.00(4.00, 12.00)	17.00(13.00, 20.00)	-12.051	<0.001	8.00(4.00, 12.50)	18.00(13.50, 20.00)	-6.098	<0.001	
epq_p/分	3.00(2.00, 5.00)	5.00(3.00, 7.00)	-4.721	<0.001	4.00(2.00, 5.00)	5.00(3.00, 8.00)	-2.955	0.003	
epq_l/分	14.00(12.00, 16.00)	12.00(9.00, 14.00)	6.285	<0.001	14.00(11.00, 16.00)	12.00(9.50, 14.00)	2.956	0.003	
SSRS_K/分	10.00(8.00, 12.00)	9.00(7.00, 11.00)	4.118	<0.001	10.71±3.40	8.90±2.32	3.160	0.002	
SSRS_Z/分	23.00(19.00, 27.00)	21.00(17.00, 24.00)	4.243	<0.001	23.00(19.00, 27.00)	19.00(18.00, 21.00)	3.122	0.002	
SSRS_L/分	8.00(7.00, 9.00)	7.00(6.00, 8.00)	4.548	<0.001	7.50(7.00, 9.00)	6.00(5.00, 8.00)	3.192	0.001	
TCSQ_NC/分	22.17±5.87	29.00(25.00, 33.00)	-10.553	<0.001	21.69±7.58	21.69±7.58	-6.490	<0.001	
TCSQ_PC/分	35.00(31.00, 38.00)	31.02±5.69	6.009	<0.001	36.00(31.00, 40.50)	28.29±5.51	4.810	<0.001	
PSQL_3/分	2.00(0, 3.00)	0(0, 2.00)	5.312	<0.001	0(0, 2.00)	1.00(0, 2.50)	-2.395	0.017	
PSQL_14/分	1.00(0, 1.00)	2.00(1.00, 2.00)	-9.583	<0.001	1.00(0, 1.00)	2.00(1.00, 2.00)	-4.389	<0.001	
PSQL_15/分	0(0, 0)	0(0, 0)	-2.005	0.450	0(0, 0)	0(0, 0)	-1.015	0.310	
NPSQI24/分	1.00(0, 1.00)	1.00(1.00, 2.00)	-6.418	<0.001	0.50(0, 1.00)	1.50(1.00, 2.00)	-4.270	<0.001	
NPSQI3/分	3.00(3.00, 3.00)	3.00(3.00, 3.00)	-2.337	0.019	3.00(3.00, 3.00)	3.00(3.00, 3.00)	—	—	
NPSQI513/分	1.00(1.00, 1.00)	2.00(1.00, 2.00)	-9.138	<0.001	1.00(1.00, 1.00)	2.00(1.00, 2.00)	-5.300	<0.001	
NPSQI617/分	1.00(1.00, 2.00)	2.00(1.00, 2.00)	-6.437	<0.001	1.00(0, 2.00)	2.00(2.00, 2.00)	-5.320	<0.001	
NHCL32	0	140(69.65)	123(61.19)	3.178	0.075	32(61.54)	29(55.77)	0.357	0.550
	1	61(30.35)	78(38.81)			20(38.46)	23(44.23)		
NMDQ	0	184(91.54)	152(75.62)	18.563	<0.001	47(90.38)	43(82.69)	1.321	0.251
	1	17(8.46)	49(24.38)			5(9.62)	9(17.31)		
NGAD	0	201(100.00)	141(70.15)	70.526	<0.001	51(98.08)	22(42.31)	38.650	<0.001
	1	0	60(29.85)			1(1.92)	30(57.69)		

终得到 21 个特征选择变量。在产后研究对象中,单因素分析得到 17 个差异有统计学意义的变量,量表维度变量间的相关性分析发现 epq_n 与 $TCSQ_NC$ 呈高度相关,两者中保留 epq_n , 故最终得到 16 个特征选择变量。

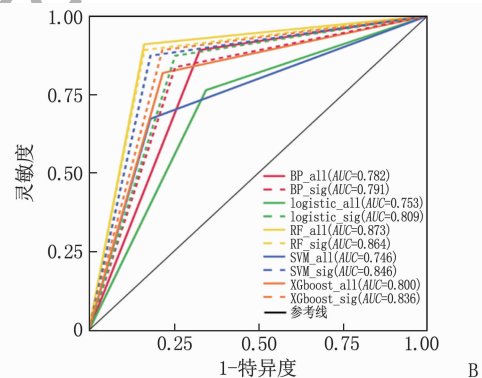
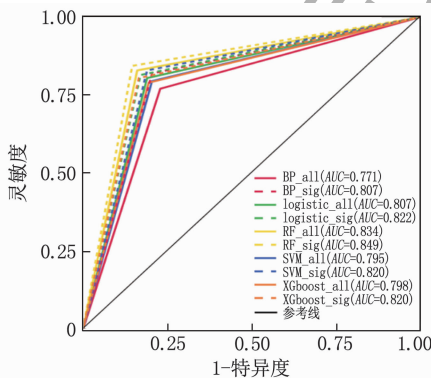
2.2.2 预测模型评价结果: 纳入不同变量的情况下, 产前预测模型、产后预测模型 5 折交叉验证的灵敏度、特异度、 AUC 均在 0.600 ~ 0.900 范围

内; RF 算法在产前预测模型(纳入所有变量时, AUC 为 0.834; 纳入特征选择变量集时, AUC 为 0.849)和产后预测模型(纳入所有变量时, AUC 为 0.873; 纳入特征选择变量集时, AUC 为 0.864)中均为最优算法; 相较于纳入所有变量构建的诊断模型, 大部分算法基于特征选择变量集构建的诊断模型性能明显提升, 但仍有个别算法提升不明显甚至下降, 见表 4、图 1。

表 4 以 PHQ_9 评分为分界标准的产前、产后预测模型的 5 折交叉验证结果

模型	模型算法	纳入所有变量			纳入特征选择变量集*		
		灵敏度	特异度	AUC	灵敏度	特异度	AUC
产前预测模型(201 对)	Logistic	0.805	0.810	0.807	0.815	0.829	0.822
	SVM	0.795	0.795	0.795	0.829	0.810	0.820
	RF	0.829	0.839	0.834	0.844	0.854	0.849
	XGboost	0.790	0.805	0.798	0.805	0.834	0.820
	BP	0.771	0.771	0.771	0.820	0.795	0.807
产后预测模型(52 对)	Logistic	0.764	0.655	0.753	0.873	0.746	0.809
	SVM	0.673	0.818	0.746	0.873	0.818	0.846
	RF	0.909	0.836	0.873	0.891	0.836	0.864
	XGboost	0.818	0.782	0.800	0.891	0.782	0.836
	BP	0.891	0.673	0.782	0.836	0.746	0.791

*: 产前预测模型纳入特征选择变量 21 个, 包括 $JYSP$ 、 $PXMYD$ 、 $HRCS$ 、 $EPDS11$ 、 $qrmh$ 、 epq_e 、 epq_n 、 epq_p 、 epq_l 、 $SSRS_K$ 、 $SSRS_Z$ 、 $SSRS_L$ 、 $TCSQ_PC$ 、 $PSQI_3$ 、 $PSQI_14$ 、 $NPSQI24$ 、 $NPSQI13$ 、 $NPSQI513$ 、 $NPSQI1617$ 、 $NMDQ$ 、 $NGAD$;
产后预测模型纳入特征选择变量 16 个, 包括 $DSYN$ 、 $EPDS11$ 、 epq_e 、 epq_n 、 epq_p 、 epq_l 、 $SSRS_K$ 、 $SSRS_Z$ 、 $SSRS_L$ 、 $TCSQ_PC$ 、 $PSQI_3$ 、 $PSQI_14$ 、 $NPSQI24$ 、 $NPSQI513$ 、 $NPSQI1617$ 、 $NGAD$ 。



A: 产前预测模型(all: 纳入所有变量; sig: 纳入特征选择变量集);

B: 产后预测模型(all: 纳入所有变量; sig: 纳入特征选择变量集)。

图 1 纳入不同变量情况下基于 5 种算法的产前、产后预测模型的 ROC 曲线

2.3 变量重要性排序

基于 5 种 ML 算法, 分别对产前、产后研究对象进行变量重要性排序(将 27 项特征变量按照重要性分数从大到小排列), 见图 2。以 RF 算法为例: 产前研究对象中, 重要性排序前 5 名的变量分别是 epq_n 、 $PSQI_14$ 、 $qrmh$ 、 $TCSQ_NC$ 、 $NGAD$; 产后研究对象中, 重要性排序前 5 名的变量分别是 $NGAD$ 、 $TCSQ_NC$ 、 epq_n 、 $NPSQI513$ 、 $NPSQI1617$ 。

3 讨论

孕产妇的心理健康一直是临床中备受关注的课题, 但目前国内外尚无特异性的生物学指标, 且由于精神专科医师的缺乏及相应访谈评估技术的限制, PND 患者很难在首诊的产科及社区及时获得评估与确诊。一项综述^[12]指出, 30% ~ 70% 的 PND 病例未被发现, 且仅 15% 的病例接受了适当治疗, 故及时筛查并发现 PND 仍然是医疗保健工

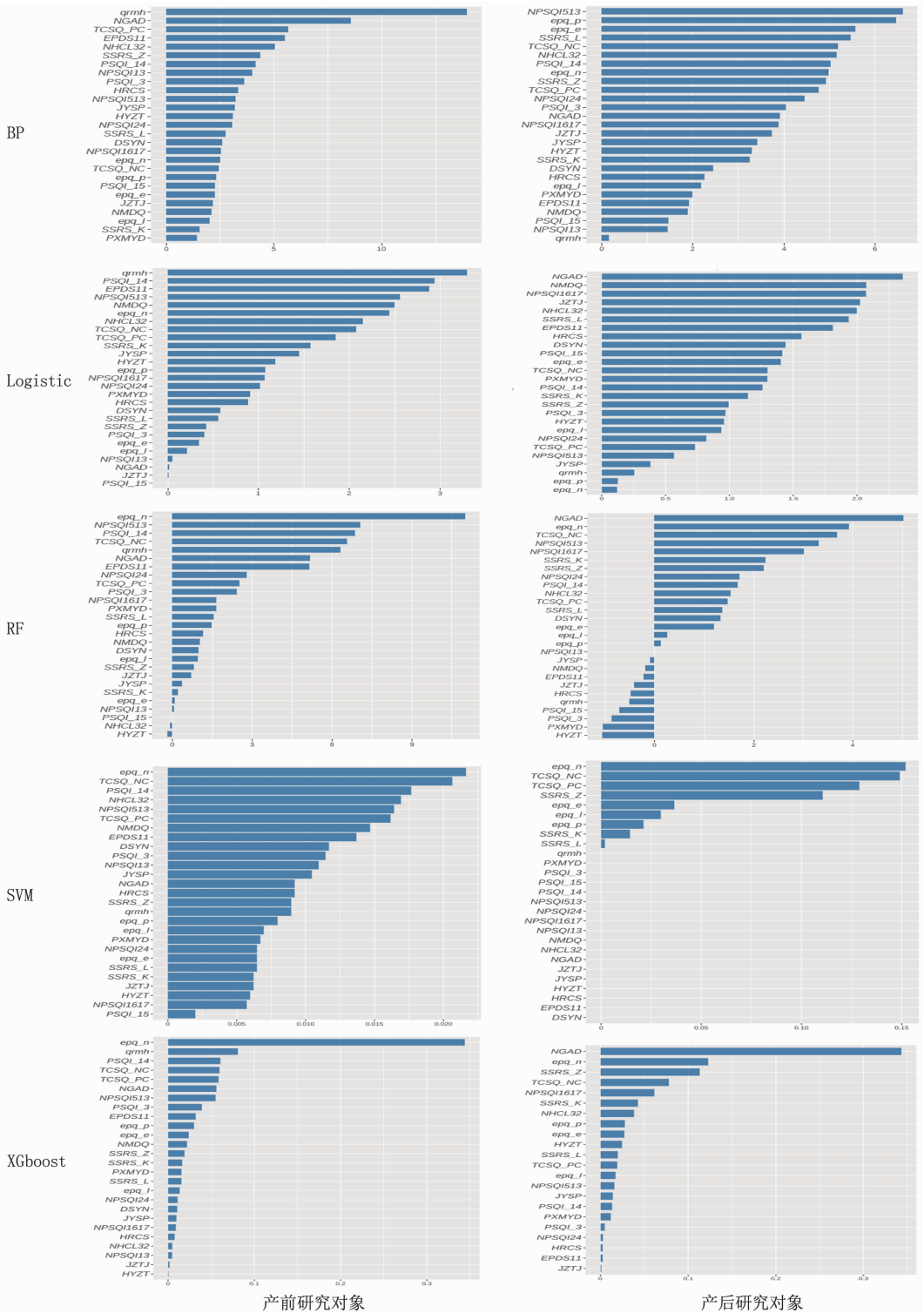


图 2 产前、产后研究对象基于 5 种算法的变量重要性排序

作的重要挑战之一。目前,基于 ML 算法构建孕产妇抑郁诊断预测模型仍然是一个相对不成熟的领域。MATSUO S 等^[13] 基于 Logistic 回归和另 4 种 ML 算法构建预测模型,但预测变量仅选用了基本的临床信息。钟雨婷等^[8] 开发了一种基于 Android 平台的产后抑郁心理护理干预系统,该系

统仅应用了爱丁堡产后抑郁自评量表,而单一的量表难以反映和解释疾病的全部变异性,故仍需进一步研究。

本研究基于 5 种 ML 算法构建产前模型、产后模型,预测性能均良好,其中 RF 算法表现最优。一方面说明,基于不同 ML 算法构建的 PND 诊断

模型均预测性能比较稳定,且本研究模型的预测效能相较于其他研究^[14]明显提升;另一方面说明,RF 算法对数据的拟合情况更好,对存在多种危险因素疾病患者的预后预测具有独特优势。本研究发现,对于大部分模型而言,纳入特征选择变量集的预测性能相较于纳入全部变量时明显提高。由此提示,用于 ML 算法的特征需要具备良好的差异性 & 较低的特征间相关性,这在提升性能的同时还能兼顾临床可解释性。值得注意的是,在产后研究对象中,表现最优的 RF 算法纳入特征选择变量集所构建的预测模型性能略差于纳入所有变量的预测模型。由此提示,RF 算法对于特征选择相对不敏感^[15],且产后抑郁预测模型仍需尽可能多地纳入特征信息。

本研究筛选出的特征选择变量集与既往研究结论基本一致,例如低社会经济地位、经济压力^[16]、人际关系、感知支持^[17]与 PND 相关。本研究中,变量重要性排序结果显示(以 RF 算法为例),排序前 5 名的重要变量首先均是单因素分析中差异有统计学意义的变量,其次在既往研究中均被视为重要的预测因素。*epq_n* 反映情绪稳定状态,*TCSQ_NC* 反映应对问题时情绪波动,*NGAD* 反映焦虑情绪变化趋势,这 3 个变量均与情绪状态有关。张雪莲等^[18]发现,情绪调节能力与抑郁水平显著相关,情绪调节能力低者更倾向于消极应对,更容易出现焦虑、抑郁等负向情绪。*PSQI_14*、*NPSQI513*、*NPSQI1617* 均与睡眠状态有关。李洋等^[19]研究发现,孕期抑郁情绪与睡眠异常等因素有关。*Qrmh* 反映亲人的精神疾病史,一项综述^[20]发现家族精神疾病史对于早期识别 PND 高风险妇女至关重要。

综上所述,基于 5 种 ML 算法构建的预测模型均可有效预测孕产妇 PND 风险,其中以 RF 算法的表现最优,这为开发快速筛查和诊断 PND 的辅助工具提供了参考依据。但本研究尚存在一定局限性:① 本研究以 PHQ-9 评分作为有无 PND 的分界标准,然而临床中诊断 PND 的金标准是经过专业训练并取得精神专科医师执照的临床医生的访谈评估结论;② 受样本量和时间的限制,本研究未进行外部验证。本研究团队将在后续实际临床试验中引入精神障碍临床定式结构化访谈(SCID)、简明国际神经精神障碍访谈检查(MINI)等诊断量表并结合临床诊断作为金标准,进一步筛选合适的量表和相应的临床变量。

参考文献

[1] O'HARA M W, WISNER K L. Perinatal mental illness: definition, description and aetiology[J]. *Best Pract Res Clin Ob-*

stet Gynaecol, 2014, 28(1): 3-12.

[2] DOSANI A, ARORA H, MAZMUDAR S. mHealth and perinatal depression in low-and middle-income countries: a scoping review of the literature[J]. *Int J Environ Res Public Health*, 2020, 17(20): 7679.

[3] NISAR A, YIN J, WAQAS A, *et al.* Prevalence of perinatal depression and its determinants in Mainland China: a systematic review and meta-analysis[J]. *J Affect Disord*, 2020, 277: 1022-1037.

[4] ZHANG Z Y, YU J J, ZENG W T, *et al.* Association between antenatal depression and adverse perinatal outcomes: a prospective cohort study[J]. *J Affect Disord*, 2023, 323: 490-495.

[5] 王嘉旭,薛鹏,江宇,等.基于机器学习算法的宫颈瘤辅助诊断模型[J].*中华肿瘤防治杂志*,2023,30(1):48-53.

[6] 王艺任,刘艾文,詹翔,等.基于机器学习算法筛选鼻咽癌诊断基因标志物的研究[J].*实用临床医药杂志*,2023,27(7):6-11.

[7] 周华,唐剑叶,秦志强,等.网络平台应用对产后抑郁筛查及干预效果的探析[J].*现代预防医学*,2018,45(20):3719-3722.

[8] 钟雨婷,刘月月,左莹莹,等.基于 Android 平台的产后抑郁心理护理干预系统研究[J].*电脑知识与技术*,2021,17(16):58-60.

[9] SPITZER R L, KROENKE K, WILLIAMS J B. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire* [J]. *JAMA*, 1999, 282(18): 1737-1744.

[10] 秦泽慧,梁列新.中文版 PHQ-9 在不同人群筛查抑郁的最佳截止值的研究分析[J].*临床消化病杂志*,2019,31(5):333-336.

[11] 尹畅.基于机器学习的癌症辅助诊断应用研究[D].成都:电子科技大学,2020.

[12] BRÄNN E, FRANSSON E, WIKMAN A, *et al.* Who do we miss when screening for postpartum depression A population-based study in a Swedish region[J]. *J Affect Disord*, 2021, 287: 165-173.

[13] MATSUO S, USHIDA T, EMOTO R, *et al.* Machine learning prediction models for postpartum depression: a multicenter study in Japan[J]. *J Obstet Gynaecol Res*, 2022, 48(7): 1775-1785.

[14] ZHONG M H, ZHANG H, YU C, *et al.* Application of machine learning in predicting the risk of postpartum depression: a systematic review[J]. *J Affect Disord*, 2022, 318: 364-379.

[15] 李磊磊.基于随机森林算法的 IVF-ET 治疗结局预测模型建立及预测特征分析研究[D].长春:吉林大学,2022.

[16] DOLBIER C L, RUSH T E, SAHADEO L S, *et al.* Relationships of race and socioeconomic status to postpartum depressive symptoms in rural African American and non-Hispanic white women[J]. *Matern Child Health J*, 2013, 17(7): 1277-1287.

[17] KETTUNEN P, HINTIKKA J. Psychosocial risk factors and treatment of new onset and recurrent depression during the post-partum period[J]. *Nord J Psychiatry*, 2017, 71(5): 355-361.

[18] 张雪莲,汪燕妮,李英,等.正念与情绪调节对孕产妇产后临床抑郁症状的影响[J].*中国健康教育*,2022,38(12):1072-1076.

[19] 李洋,赵斌,袁佳英,等.依托孕产妇身心健康服务新模式的孕期抑郁情绪影响因素研究[J].*中国卫生标准管理*,2023,14(7):17-21.

[20] KJELDSSEN M Z, BRICCA A, LIU X Q, *et al.* Family history of psychiatric disorders as a risk factor for maternal postpartum depression: a systematic review protocol[J]. *Syst Rev*, 2022, 11(1): 68.