

# 基于人工神经网络的早发型子痫前期患者预测模型构建研究

吕伯瀚<sup>1</sup>, 魏丽丽<sup>2\*</sup>, 李培宇<sup>3</sup>, 苑广慧<sup>1</sup>, 王凯选<sup>1</sup>, 于水<sup>1</sup>, 王文远<sup>2</sup>

<sup>1</sup>青岛大学医学部, 山东 青岛

<sup>2</sup>青岛大学附属医院, 山东 青岛

<sup>3</sup>青岛大学医学院松山医院, 山东 青岛

Email: \*13573828157@163.com

收稿日期: 2021年9月11日; 录用日期: 2021年10月2日; 发布日期: 2021年10月14日

## 摘要

目的: 建立基于人工神经网络的子痫前期预测模型, 为疾病早期筛查提供依据。方法: 前瞻性收集2020年3月至2021年6月在青岛大学附属医院进行产前检查并分娩的741名孕妇资料, 通过对所得资料进行单因素logistic回归以及多因素logistic回归筛选出子痫前期的独立危险因素。将所得独立危险因素运用人工神经网络算法进行预测模型的拟合, 运用受试者工作特征曲线(ROC曲线)对模型进行评估。结果: 所调查孕妇中共有71例(9.5%)发生早发型子痫前期, 670例(90.5%)未发生早发型子痫前期, 多因素logistic回归显示早发型子痫前期的独立危险因素有孕前BMI、孕次、孕前是否吸烟、孕前是否饮酒、平均动脉压(MAP)、葡萄糖、AST/ALT、血清游离三碘甲状腺原氨酸(FT3)、甲胎蛋白(AFP), 9项指标; 所得预测模型预测早发型PE的ROC曲线下面积为0.945。结论: 基于人工神经网络的早发型PE预测模型不仅为早发型PE提供了理论和方法的支持, 为疾病的早发现、早诊断、早治疗争取了时间, 具有广阔的应用前景。

## 关键词

早发型子痫前期, 预测模型, 人工神经网络

# Prediction of Early-Onset Preeclampsia Based on Artificial Neural Network

Bohan Lv<sup>1</sup>, Lili Wei<sup>2\*</sup>, Peiyu Li<sup>3</sup>, Guanghui Yuan<sup>1</sup>, Kaixuan Wang<sup>1</sup>, Shui Yu<sup>1</sup>, Wenyuan Wang<sup>2</sup>

<sup>1</sup>School of Medicine, Qingdao University, Qingdao Shandong

<sup>2</sup>The Affiliated Hospital of Qingdao University, Qingdao Shandong

<sup>3</sup>Songshan Hospital, Qingdao University Medical College, Qingdao Shandong

\*通讯作者。

文章引用: 吕伯瀚, 魏丽丽, 李培宇, 苑广慧, 王凯选, 于水, 王文远. 基于人工神经网络的早发型子痫前期患者预测模型构建研究[J]. 临床医学进展, 2021, 11(10): 4451-4457. DOI: 10.12677/acm.2021.1110652

## Abstract

**Objective:** To establish a prediction of early-onset preeclampsia based on artificial neural network, so as to provide a basis for the early screening of the disease. **Methods:** A prospective study was conducted on the data of 741 pregnant women who underwent prenatal examination and delivery in The Affiliated Hospital of Qingdao University from March 2020 to June 2021. Univariate logistic regression and multivariate logistic regression were used to screen the independent risk factors for preeclampsia. The independent risk factors were fitted to the prediction model by artificial neural network algorithm, and the receiver operating characteristic curve (ROC curve) was used to evaluate the model. **Results:** A total of 71 (9.5%) pregnant women developed early-onset preeclampsia, while 670 (90.5%) did not. Multivariate logistic regression showed that the independent risk factors for early-onset preeclampsia were BMI before pregnancy, pregnancy times, smoking before pregnancy, drinking before pregnancy, MAP, glucose, AST/ALT, serum free triiodothyronine (FT3), alpha-fetoprotein (AFP). The predictive model predicted that the area under ROC curve of early-onset PE was 0.945. **Conclusion:** The prediction model of early-onset PE based on artificial neural network not only provides theoretical and method support for early-onset PE, but also buys time for early detection, early diagnosis and early treatment of disease, which has broad application prospect.

## Keywords

Early-Onset Preeclampsia, Prediction Model, Artificial Neural Network

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

子痫前期(Preeclampsia, PE)是一种妊娠期特发性高血压疾病,指孕妇在妊娠20周后出现高血压、尿蛋白、水肿等临床表现,PE是引起孕妇死亡的主要原因之一[1]。其中于34周前发病者称为早发型子痫前期,这类患者于妊娠早中期即出现血压迅速升高且难以控制,同时合并有严重的并发症及多个靶器官损伤,母儿死亡率较晚发型显著增高,母儿近期及远期预后差[2]。因此,早期筛选出可能发生早发型PE的高危孕妇,尽早采取积极有效的预防措施,可以有效降低发病率、延缓疾病的进展、改善母儿预后。人工神经网络(Artificial neural networks, ANN)是一种应用类似于大脑神经突触联接的结构进行信息处理的数学模型,目前已广泛应用于医疗领域[3] [4] [5],其具有强大的整合、分析和自主学习能力,能有效提高疾病诊断治疗的准确率。本研究拟通过运用人工神经网络技术构建早发型子痫前期患者预测模型,为临床决策提供理论依据。

## 2. 材料与方法

### 2.1. 研究对象

本研究前瞻性收集2020年3月~2021年6月在青岛大学附属医院行产前检查并分娩的741例妊娠期

女性作为研究对象,本研究已获得青岛大学附属医院伦理委员会批准,所有资料的收集均在研究对象知情同意的前提下进行。本研究的纳入标准:① 人绒毛膜促性腺激素(HCG)检测示妊娠试验阳性;② 病历资料完备、定期在院内产检;③ 对本研究知情同意;排除标准:① 孕前即诊断为高血压;② 合并其他严重疾病(如肝、肾、造血系统等严重原发性疾病、严重感染、免疫性疾病、恶性肿瘤等)。

## 2.2. 早发型 PE 的诊断标准

本研究参照中华医学会妇产科学分会妊娠期高血压疾病学组《妊娠期高血压疾病诊治指南(2020)》[6]的标准:妊娠 20 周后孕妇出现收缩压  $\geq 140$  mmHg 和(或)舒张压  $\geq 90$  mmHg,伴有下列任意 1 项:尿蛋白定量  $\geq 0.3$  g/24h,或尿蛋白/肌酐比值  $\geq 0.3$ ,或随机尿蛋白  $\geq (+)$ (无条件进行蛋白定量时的检查方法);无蛋白尿但伴有以下任何 1 种器官或系统受累:心、肺、肝、肾等重要器官,或血液系统、消化系统、神经系统的异常改变,胎盘-胎儿受到累及等;发病孕周  $< 34$  周。

## 2.3. 相关预测指标的收集

根据研究早期相关文献的查阅,与产科专家共同确定预测指标的大体收集范围。在孕妇初次建档产检时(即孕 13 周末前)行人口统计学资料和病史资料的采集,主要包括:年龄、孕前 BMI、平均动脉压(MAP)、孕次、产次、家族史、母亲及姐妹既往产科病史、孕前 3 月内是否吸烟、孕前 3 月内是否饮酒、妊娠早期是否出血、是否有不良孕产史等。同时收集相关实验室检查资料:血常规、尿常规、生化检查和甲状腺功能检查等。于孕 20 周前收集孕妇产前筛查结果,主要有甲胎蛋白(AFP)、绒毛膜促性腺激素  $\beta$  (HCG- $\beta$ )、游离三醇(uE3),检测值均转化为中位倍数值(MoM)并通过孕期和母体因素进行校正。

## 2.4. 数据预处理

本研究的所有数据均应用 SPSS 20.0 进行统计描述和统计分析。对于缺失值  $> 10\%$  的数据选择直接删除,剩余缺失数据采用期望最大化方法(Expectation maximization, EM),进行缺失值的填充。同时删除数据中存在的异常值以及不合理的数据。通过单因素 logistic 回归方法进行预测指标的初步筛选,随后运用多因素 logistic 回归实现预测指标的最终筛选。

## 2.5. 基于 ANN 的早发型 PE 预测模型的构建

运用 SPSS 软件自带的 Bernoulli 函数将 741 例研究对象按照 7:3 的比例进行无偏随机划分,其中 522 例(70%)作为训练集,进行模型训练,219 例作为测试集,进行模型验证。将多因素 logistic 回归所得最终预测指标纳入 ANN 模型中进行模型训练。本研究采用 SPSS 中 ANN 的多层感知器模块,具有单向传播、单隐层的结构、上下层全连接以及每层间无连接的特点[7]。具体操作为:输入参数为孕前 BMI、孕次、产次、孕前吸烟、孕前饮酒、MAP、AST/ALT、血清游离三碘甲腺原氨酸(FT3)、AFP\_MOM 值、葡萄糖,共 10 个变量,所有输入的参数按照“连续型变量”和“分类变量”进行区分后,进行归一化,处理到[0, 1]区间。输出参数设置为是否发生早发型 PE,其中 0 为未发生早发型 PE,1 为发生早发型 PE。网络设置为自动体系结构选择,隐层节点数 1~50,训练类型为“批处理”,优化算法为“调整的共轭梯度”,初始 Lambda 值为 0.000005,初始 Sigma 值 0.00005,间隔中心点为 0,间隔偏移量为  $\pm 0.5$ 。通过运行后可获得“早发型 PE”的预测拟概率,同时绘制受试者工作特征曲线以及计算曲线下面积 AUC。

## 3. 结果

### 3.1. 影响早发型 PE 的单因素 Logistic 回归分析

本研究最终纳入 741 例研究对象,其中有 71 例(9.5%)发生早发型 PE,670 例(90.5%)未发生早发型

PE。运用单因素 logistic 回归进行预测指标的初步筛选, 所得孕前 BMI、孕次、产次、孕前 3 月内是否吸烟、孕前 3 月内是否饮酒、妊娠早期是否出血、MAP、白细胞计数、红细胞计数、血小板、白蛋白、葡萄糖、AST/ALT、FT3、AFP\_MOM 值, 15 项指标有统计学意义( $P < 0.05$ ), 见表 1。

**Table 1.** Single factor logistic regression analysis of early-onset PE  
**表 1.** 早发型 PE 单因素 logistic 回归分析

变量	B	SE	OR	95% CI		P
				下限	上限	
年龄(岁)	0.11	0.08	1.12	0.95	1.31	0.17
孕前 BMI	0.38	0.10	1.46	1.19	1.78	<0.001
孕次(次)	1.36	0.57	3.88	1.27	11.87	0.02
产次(次)	-3.56	0.94	0.03	0.00	0.18	<0.001
有无糖尿病家族史	-0.99	0.98	0.37	0.05	2.53	0.31
有无高血压家族史	0.10	0.83	1.10	0.22	5.66	0.91
孕前 3 月内是否吸烟	3.16	1.05	23.61	3.00	185.61	<0.001
孕前 3 月内是否饮酒	1.59	0.76	4.91	1.12	21.57	0.04
与上次怀孕间隔是否超过十年	-1.51	1.09	0.22	0.03	1.87	0.17
月经周期是否规律	0.51	0.78	1.66	0.36	7.69	0.52
妊娠早期是否出血	1.15	0.68	3.17	0.84	12.02	0.09
流产史	0.73	0.91	2.08	0.35	12.30	0.42
MAP	0.25	0.06	1.28	1.14	1.44	<0.001
白细胞计数( $\times 10^9/L$ )	0.56	0.19	1.76	1.21	2.55	<0.001
红细胞计数( $\times 10^{12}/L$ )	2.26	1.12	9.56	1.07	85.76	0.04
血红蛋白(g/L)	0.03	0.03	1.03	0.97	1.10	0.36
血小板( $\times 10^9/L$ )	-0.02	0.01	0.98	0.97	1.00	0.02
白蛋白(g/L)	0.15	0.08	1.17	1.00	1.35	0.05
总胆红素( $\mu\text{mol/L}$ )	0.00	0.09	1.00	0.84	1.20	1.00
葡萄糖(mmol/l)	1.23	0.41	3.42	1.53	7.65	<0.001
AST/ALT	2.36	0.79	10.60	2.24	50.17	<0.001
尿隐血	0.32	0.67	1.37	0.37	5.12	0.64
尿蛋白	-0.13	0.72	0.88	0.21	3.64	0.86
尿酮体	0.20	0.62	1.22	0.36	4.12	0.75
尿 PH	0.44	0.49	1.56	0.59	4.10	0.37
尿白细胞	0.62	0.42	1.86	0.82	4.26	0.14
尿维生素 C	-2.02	2.35	0.13	0.00	13.43	0.39
FT3	1.80	0.49	6.02	2.32	15.62	<0.001
FT4	0.04	0.08	1.04	0.89	1.21	0.63
TSH	0.06	0.13	1.07	0.83	1.37	0.62
AFP_MOM 值	1.94	0.62	6.97	2.09	23.31	0.00
HCG $\beta$ _MOM 值	-0.15	0.51	0.86	0.32	2.33	0.77
uE3_MOM 值	-0.17	0.80	0.84	0.18	4.00	0.83

注: 孕前 BMI: 孕前身体指数; MAP: 平均动脉压; AST/ALT: 谷草转氨酶比谷丙转氨酶; 尿 PH: 尿酸碱度; FT3: 血清游离三碘甲状腺原氨酸; FT4: 血清游离甲状腺素; TSH: 促甲状腺激素; AFP: 甲胎蛋白; HCG $\beta$ : 绒毛膜促性腺激素  $\beta$ ; uE3: 游离三醇。

### 3.2. 影响早发型 PE 的多因素 Logistic 回归分析

将上述单因素 logistic 回归中有意义的因素作为自变量, 是否发生早发型 PE 作为因变量, 进行多因素 logistic 回归分析, 最终得出孕前 BMI、孕次、孕前是否吸烟、孕前是否饮酒、MAP、葡萄糖、AST/ALT、FT3、AFP\_MOM 值, 9 项指标有统计学意义( $P < 0.05$ ), 为早发型 PE 的独立影响因素, 见表 2。

**Table 2.** Multivariate Logistic regression analysis of early-onset PE

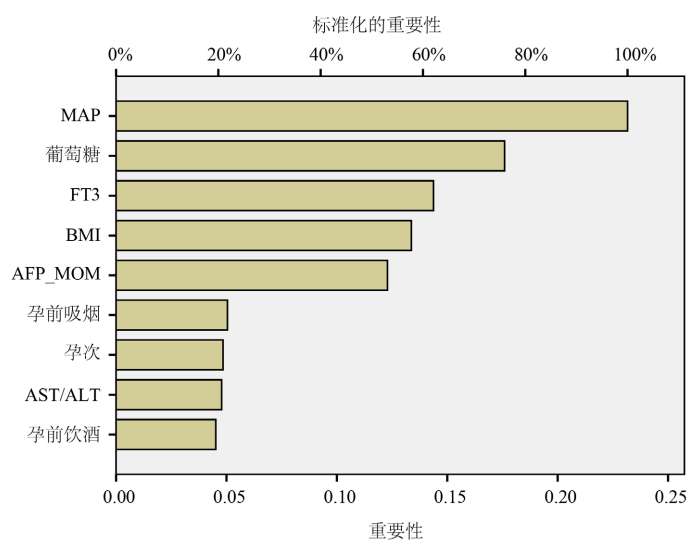
**表 2.** 早发型 PE 多因素 Logistic 回归分析

变量	<i>B</i>	<i>SE</i>	<i>OR</i>	95% CI		<i>P</i> 值
				下限	上限	
孕前 BMI	0.28	0.06	1.33	1.19	1.48	<0.001
孕次(次)	1.10	0.26	2.99	1.79	5.00	<0.001
孕前 3 月内吸烟	2.76	0.65	15.77	4.40	56.49	<0.001
孕前 3 月内饮酒	1.69	0.53	5.42	1.93	15.24	<0.001
MAP	0.22	0.04	1.25	1.16	1.34	<0.001
AST/ALT	1.26	0.41	3.51	1.57	7.85	<0.001
FT3	1.19	0.24	3.29	2.07	5.24	<0.001
AFP_MOM 值	1.48	0.36	4.39	2.16	8.92	<0.001
葡萄糖	0.99	0.25	2.70	1.67	4.36	<0.001

注: 孕前 BMI: 孕前身体指数; MAP: 平均动脉压; AST/ALT: 谷草转氨酶比谷丙转氨酶; FT3: 血清游离三碘甲腺原氨酸; AFP: 甲胎蛋白。

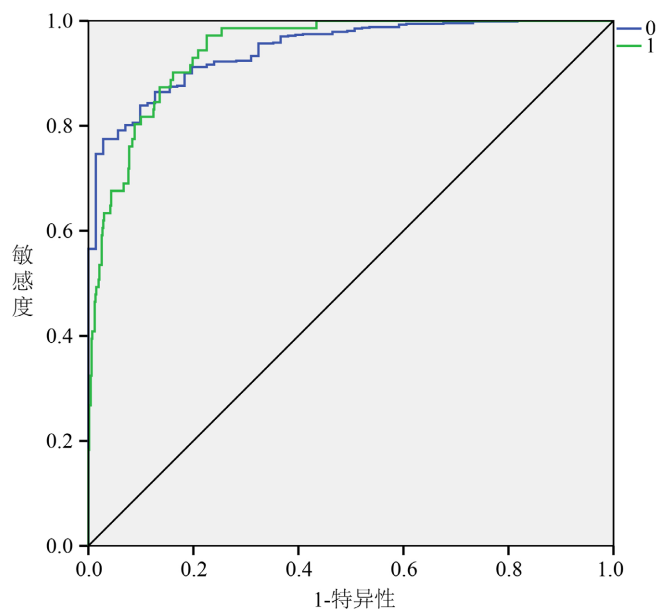
### 3.3. ANN 模型的构建

将上述指标纳入 ANN 模型中进行模型训练, 最终得到输入单元 11 个、单隐层 6 个节点、输出单元 2 个的模型。各项预测指标的重要性排序如图 1 所示, 依次为 MAP、葡萄糖、FT3、孕前 BMI、AFP\_MOM 值、孕前 3 月内吸烟、孕次、AST/ALT、孕前 3 月内饮酒。所得模型训练集的早发型 PE 正确预测率为 93.3%, 测试集的早发型 PE 正确预测率为 94.5%。在训练集和测试集的基础上得到多层向前神经网络模型预测的 AUC 为 0.945, 见图 2。



**Figure 1.** Ranking the importance of each indicator

**图 1.** 各项指标的重要性排序



**Figure 2.** Multi-layer forward neural network model for predicting ROC curve of early-onset PE

**图 2.** 多层向前神经网络模型预测早发型 PE 的 ROC 曲线

#### 4. 讨论

PE 根据发病时间的不同分为早发型 PE (发病孕周  $< 34$  周)和晚发型 PE (发病孕周  $\geq 34$  周), 相关研究指出两者可能存在不同的发病机制, 由于发病时间过早尚未达到足月生产的条件, 早发型 PE 对母婴健康的影响更为严重[8]。目前尚未完全阐明其真正的发病机制, 多因素共同致病是国内外学普遍认同的, 例如由于妊娠早期体内滋养细胞的侵袭能力不足, 进而导致子宫螺旋动脉的重塑障碍, 母体的炎症反应出现异常, 多种细胞因子的分泌紊乱, 血管内皮细胞广泛损伤等, 诸如此类因素共同导致胎盘出现缺血缺氧的症状, 进而影响胎盘的发育以及着床。该病主要病理生理学上的表现为全身小动脉痉挛性的改变, 进而导致各脏器血液灌注量不足, 导致微循环障碍, 组织明显缺血缺氧, 这对母体与婴儿都产生极为严重的影响, 为妊娠期女性的健康生产极为不利, 严重者可至母婴死亡[9]。

由于早发型 PE 的多因素致病特点, 目前尚无有效的治疗手段, 及时终止妊娠仍为主要手段, 因此近些年来国内外学者致力于早期预测早发型 PE 的研究, 以期将疾病的危险程度降到最低。目前在该疾病预测方面已有丰富的产出, 出现了从单一指标预测到结合母体因素、相关实验室检查、辅助检查等多项指标联合共同预测的方法。但尚为出现有效方法应用于临床, 有研究表明多项指标联合预测的方法其准确率明显高于单一指标的简单预测[10]。本研究首次采用 ANN 进行早发型 PE 的预测, 相较于传统统计学方法更能将数据充分利用, 同时能对连续型变量和分类变量的结合预测提供更加精准的手段, 最终所得模型预测准确率高, 可为临床诊断提供依据。

本研究通过临床文献回顾以及与专家共同商讨, 初步确定了孕妇人口统计学资料以及相关实验室检查资料的收集范围, 后将所获数据进行单因素 logistic 回归分析以及多因素 logistic 回归分析最终确定孕前 BMI、孕次、孕前是否吸烟、孕前是否饮酒、MAP、葡萄糖、AST/ALT、FT3、AFP\_MOM 为早发型 PE 的独立影响因素, 这与前人研究结果类似[11] [12] [13], 同时也证明了早发型 PE 是多种因素共同致病的结论。根据 ANN 模型的重要性排序得知 MAP 与疾病关系最为密切, 考虑可能为妊娠早期已出现相关血流动力学改变, 这可为后续疾病的预测和治疗提供思路。而孕早期的血糖值重要度排序同样较高, 间



接反映出孕期血糖的变化同样与早发型 PE 的发病密切相关,这也与妊娠期糖尿病为 PE 危险因素的结论一致[14]。

## 5. 结论

综上所述,本研究运用多层向前神经网络构建的早发型 PE 预测模型拥有较高的疾病预测能力(AUC = 0.945),可为临床预测及诊断提供相应的支持,同时通过分析得出的 MAP、孕前 BMI 等预测指标可为临床疾病筛查的侧重点提供依据。但本研究受样本量与地区的限制,实际运用到临床中还需反复验证。后续考虑进一步扩大样本量以及样本收集范围,实现单中心、大样本的研究格局,在此基础上反复调整模型,使模型更适用于我国临床。ANN 预测模型不仅为早发型 PE 提供了理论和方法的支持,更有利于早期发现患者 PE 的倾向,为疾病治疗争取宝贵时间,具有广阔应用前景。

## 参考文献

- [1] Chen, H., Zhang, Y., Jia, J., *et al.* (2021) Macrophage-Stimulating Protein Is Decreased in Severe Preeclampsia and Regulates the Biological Behavior of HTR-8/SVneo Trophoblast Cells. *Placenta*, **103**, 33-42. <https://doi.org/10.1016/j.placenta.2020.10.004>
- [2] Tranquilli, A.L., Brown, M.A., Zeeman, G.G., *et al.* (2013) The Definition of Severe and Early-Onset Preeclampsia. Statements from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*, **3**, 44-47. <https://doi.org/10.1016/j.preghy.2012.11.001>
- [3] 吴妍, 姚蕾, 盛文丽, 等. 基于人工神经网络的卵巢早衰预测模型研究[J]. 中国全科医学, 2017, 20(27): 3410-3415.
- [4] Bugajski, A., Lengerich, A., Koerner, R., *et al.* (2021) Utilizing an Artificial Neural Network to Predict Self-Management in Patients with Chronic Obstructive Pulmonary Disease: An Exploratory Analysis. *Journal of Nursing Scholarship*, **53**, 16-24. <https://doi.org/10.1111/jnu.12618>
- [5] Mai, R., Lu, H., Bai, T., *et al.* (2020) Artificial Neural Network Model for Preoperative Prediction of Severe Liver Failure after Hemihpatectomy in Patients with Hepatocellular Carcinoma. *Surgery*, **168**, 643-652. <https://doi.org/10.1016/j.surg.2020.06.031>
- [6] 中华医学会妇产科学分会妊娠期高血压疾病学组. 妊娠期高血压疾病诊治指南(2020) [J]. 中华妇产科杂志, 2020, 55(4): 227-228.
- [7] 方建红, 梁金丽, 郑剑兰, 等. 基于人工神经网络的剖宫产术后再次妊娠阴道分娩的预测研究[J]. 中国妇幼保健, 2019, 34(12): 2680-2683.
- [8] 何韵婷, 王筱金, 王炳顺. 子痫前期孕早期预测的研究进展[J]. 国际妇产科学杂志, 2020, 47(4): 378-383.
- [9] Rana, S., *et al.* (2019) Preeclampsia: Pathophysiology, Challenges, and Perspectives. *Circulation Research*, **124**, 1094-1112. <https://doi.org/10.1161/CIRCRESAHA.118.313276>
- [10] Al-Rubaie, Z., *et al.* (2016) The Performance of Risk Prediction Models for Pre-Eclampsia Using Routinely Collected Maternal Characteristics and Comparison with Models That Include Specialised Tests and with Clinical Guideline Decision Rules: A Systematic Review. *BJOG: An International Journal of Obstetrics and Gynaecology (BJOG)*, **123**, 1441-1452. <https://doi.org/10.1111/1471-0528.14029>
- [11] 史春, 张美龄, 韩燕媚, 等. 早发型子痫前期发病危险因素分析[J]. 临床军医杂志, 2021, 49(6): 666-667.
- [12] 杜娟, 朱一麟, 高雪梅, 等. 外周血 sFlt-1、 $\beta$ -hCG、PAPP-A、AFP 的表达与早发型子痫前期的关系[J]. 疑难病杂志, 2016, 15(4): 394-397.
- [13] 周佳任, 杜鹃, 乔宠, 等. 早发型重度子痫前期患者甲状腺功能和 24 h 尿蛋白特征[J]. 广东医学, 2015, 36(13): 2056-2059.
- [14] 高娜. 早发型子痫前期合并妊娠期糖尿病母婴结局分析[D]: [硕士学位论文]. 济南: 山东大学, 2021.