

Logistic 回归在建立产前筛查唐氏综合征 风险估计模型中的应用

卓仁杰¹, 李 莺², 张莉娜¹, 沈其君¹

(1. 宁波大学 医学院, 浙江 宁波 315211; 2. 宁波市第一医院 内科, 浙江 宁波 315000)

摘要: 利用 64 772 例 12~20 周筛查孕妇(正常孕妇 64 718 例, 唐氏患者 54 例)的孕妇血清学指标, 采用 Logistic 回归, 建立三联筛查方案和二联筛查方案的风险估计模型. 采用 2 种方案, 计算全部筛查孕妇的风险值, 通过 ROC 曲线下面积、检出率、阳性率评价 2 种筛查方案的效果. 结果表明: 年龄、甲胎蛋白、游离人绒毛膜促性腺激素和游离雌三醇在 Logistic 回归模型中均有统计学意义, 三联方案 ROC 曲线下面积大于二联方案, 但无显著性差异(z 统计量为 1.382, $P=0.1669$); 在同一截断值下, 三联方案检出率高于二联方案, 假阳性率低于二联方案, 差异有统计学意义($t=-9.44$, $P<0.001$). 三联方案根据约登指数计算, 最佳筛查截断值为 1:400, 假阳性率为 4.3%.

关键词: 风险估计模型; Logistic 回归; 孕中期; 唐氏综合征; 筛查方案

中图分类号: R714.15

文献标志码: A

文章编号: 1001-5132 (2016) 02-0108-04

目前在孕中期产前筛查软件中用于估计风险的模型应用较多的是 Palomaki 等^[1]于 1987 年提出的基于似然比计算的多变量正态模型, 该模型要求各项变量数据分布为正态分布, 有诸多限制条件. Logistic 模型(Logistic Regression Analysis)是预测疾病风险最重要统计学模型之一^[2], 近些年已有学者采用 Logistic 回归模型来计算患病风险^[3]. 相比与其他模型, Logistic 回归模型的优点在于: 不需要考虑变量的分布形态、多种不同类型的变量可以直接输入模型进行分析、去除对模型贡献较弱的部分, 可通过标准相关系数分析各个变量对. 目前国内尚未发现有学者利用 Logistic 回归来建立基于中国孕妇人群的个体风险估计模型的报道.

本研究采用 Logistic 模型分析经过孕龄、体重、糖尿病以及双胎等校正后的孕妇血清学数据, 建立了基于中国人群的唐氏综合征产前筛查的风险估计模型, 并利用建立的模型评价唐氏综合征二联筛查方案和三联筛查方案.

1 材料与方法

1.1 研究对象

收集宁波 11 个县区 2007 年 4 月至 2012 年 4 月期间 35 周岁以下进行孕中期产前筛查的孕妇 66 874 例, 空腹抽取静脉血 5 mL, 4 °C 静置 40 min, 离心后取血清, 5 d 内送宁波市妇儿医院产前诊断中心检测血清标记物.

1.2 数据筛选

所有筛查对象均随访至分娩后 6 个月至 4 年. 对可疑婴儿采取外周血进行核型分析. 排除资料中无法估计孕龄以及先天性心脏病等其他先天性疾病后, 得到 64 718 例正常对象和 54 例唐氏患者供分析.

1.3 数据处理

原始数据采用 Epidata3.0 软件, 双机录入数据库, 进行逻辑检错, 以保证数据录入的准确性.

1.4 血清学指标处理

孕妇的孕龄采用 Leung 等^[4]公式计算. 甲胎蛋

白(alpha-fetoprotein, AFP)、游离人绒毛膜促性腺激素(free β -subunit of human chorionic gonadotrophin, free β -HCG)和游离雌三醇(unconjugated estriol, uE3) 3种血清学指标采用文献[5]中公式进行孕龄、体重校正, 获得血清学指标的中位数倍数(multiple of the median, MoM); 对 MoM 值进行吸烟、糖尿病、双胎校正(采用文献[6]中方法)。

1.5 统计分析

把经过校正后的 AFP、Free β -HCG、uE3 的 MoM 值作对数变换, 孕妇年龄和 $\lg(\text{AFP_MoM})$ 、 $\lg(\text{HCG_MoM})$ 、 $\lg(\text{uE3_MoM})$ 作为回归模型中独立变量, 采用二分类 Logistic 回归分析。三联筛查方案和二联筛查方案比较采用 ROC 曲线, 比较在不同截断值下检出率和假阳性率的大小。

三联筛查方案为年龄+AFP+Free β -HCG+uE3, 二联筛查方案为年龄+AFP+Free β -HCG。所有统计采用 SAS 9.0.1 和 SPSS 18.0 统计软件完成。

2 结果

2.1 受检孕妇基本情况

64 718 例正常孕妇年龄均值为(26.32 $\pm$ 3.27)周岁, 最大者 35.10 周岁, 最小者 16.12 周岁; 孕龄均值为(119.2 $\pm$ 6.8) d, 最大值为 140 d, 最小值为 105 d; 体重均值为(53.8 $\pm$ 7.2) kg, 最大值为 112 kg, 最小值为 31 kg。54 例唐氏患者孕妇年龄均值为(27.64 $\pm$ 3.86)周岁, 最大者 35.70 周岁, 最小者 20.10 周岁; 孕龄均值为(120.4 $\pm$ 5.6) d, 最大值为 140 d, 最小值为 105 d; 体重均值为(55.4 $\pm$ 6.7) kg, 最大值为 118 kg, 最小值为 43 kg。2 组间经 t 检验, 年龄、体重、孕龄均无统计学意义($P>0.05$)。

2.2 三联筛查方案的 Logistic 回归分析结果

对孕妇年龄、AFP、HCG、uE3 的 MoM 值进行回归分析, 结果见表 1。

表 1 三联筛查方案 Logistic 模型系数

变量	自由度	参数估计	标准误	卡方	P 值
Intercept	1	-11.353 7	1.610 3	54.3	<0.000 1
Age	1	0.108 4	0.053 8	4.1	0.048 7
$\lg\text{AFP}$	1	-5.347 1	0.881 1	8.9	0.003 1
$\lg\text{HCG}$	1	5.297 7	0.518 9	89.3	<0.000 1
$\lg\text{uE3}$	1	-1.788 4	0.836 2	21.5	<0.000 1

模型整体检验、常数项、各回归系数均有统计学意义, 孕妇年龄项的 P 值为 0.048 7, 接近于 0.05。但经国外有关学者研究, 年龄因素是较重要的威胁因素, 因此本研究中没有剔除年龄项。

2.3 二联筛查方案的 Logistic 回归分析结果

对孕妇年龄、AFP、HCG 的 MoM 值进行回归分析, 结果见表 2。

表 2 二联筛查方案 Logistic 模型系数

变量	自由度	参数估计	标准误	卡方	P 值
Intercept	1	-11.864 3	1.541 2	62.5	<0.000 1
Age	1	0.118 9	0.054 3	4.8	0.025 3
$\lg\text{AFP}$	1	-4.711 1	0.951 4	24.8	<0.000 1
$\lg\text{HCG}$	1	5.141 2	0.532 1	95.1	<0.000 1

整体模型检验、常数以及各回归系数均有统计学意义。

2.4 二联方案和三联方案筛查效果比较

正常组和病例组用三联 Logistic 风险估计模型和二联风险估计模型重新计算孕妇患病风险率, 采用 ROC 曲线进行比较, 结果如图 1 所示。

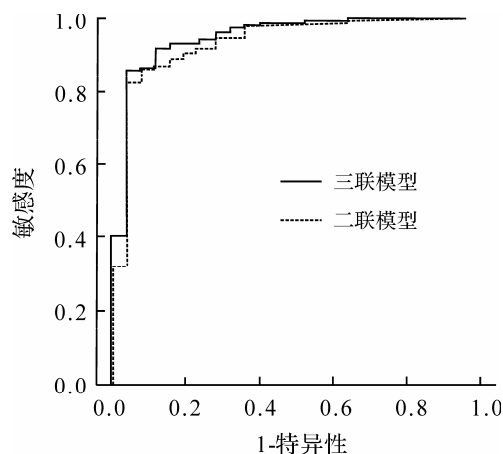


图 1 三联和二联风险估计模型的 ROC 曲线

三联模型 ROC 下面积为 0.964, 标准误 0.017; 二联模型 ROC 下面积为 0.951, 标准误 0.021。经检验无显著性差异(z 统计量为 1.382, $P=0.166 9$)。

在截断值为 1:270~1:400 时, 2 种模型的检出率相差不大, 但是二者的假阳性率有明显差别, 经配对 t 检验, 有统计学意义($t=-9.44$, $P<0.001$), 三联回归的风险估计模型假阳性率比二联的风险估计模型小。利用约登指数, 最佳截断值为 1:400, 假阳性率为 4.3%。

表 3 不同截断值三联模型和二联模型的
检出率和假阳性率

截断值	检出率/%		假阳性率/%	
	三联模型	二联模型	三联模型	二联模型
1:270	70.37	70.37	2.8	3.5
1:350	74.07	72.22	3.2	4.9
1:400	77.78	77.78	4.3	5.5
1:450	79.63	77.78	5.0	5.9
1:500	79.63	79.63	5.6	6.6
1:550	83.33	81.48	6.2	7.1
1:650	85.19	81.48	7.2	8.3

2.5 自建 Logistic 模型和自建多变量正态模型三
联方案的比较

把 AFP、free β -HCG、uE3 这 3 种血清学指标, 经过 MOM 值变化, 然后再经过对数转换, 数据符合正态分布, 可作多变量正态模型. 结合年龄风险率进行孕妇的风险计算, 在 5% 的假阳性率下, 筛查结果跟 Logistic 模型进行比较, 结果见表 4.

表 4 5% 的假阳性率下检出率和阈值

检验结果变量	阈值	检出率/%
多变量正态三联模型	330	78.10
Logistic 三联模型	450	79.63

3 讨论

综上所述, 在三联风险估计模型中, 在截断值为 1:270 时, 检出率为 70.37%, 阳性率为 2.8%. 在截断值为 1:450 时, 检出率为 79.68%, 阳性率为 5.0%, 这比 Johnson 等^[7]报道的研究效果要好(检出率为 74.0%, 阳性率为 5%). 二联和三联模型比较中, 在截断值为 1:270~1:400 时, 2 种方案的检出率相差不大, 但是二者的假阳性率有明显的差别, 平均相差 1.3%, 三联筛查方案效果要好于二联筛查方案, 这个结果与 Hwa 等^[3]相一致. Hwa 同时从群体成本效益方面对 Logistic 模型的三联和二联筛查进行分析, 认为三联方案群体成本更低.

从表 1 中可以看出, 年龄、AFP、Free β -HCG 和 uE3 这 4 种风险因素中, AFP 和 uE3 的回归系数分别为 -5.347 1 和 -1.700 8, 说明在孕妇年龄和其他血清学指标 MoM 值不变的情况下, 随着 AFP

浓度的 MoM 值增加, 孕妇胎儿的患唐氏综合征风险率逐渐减小. Free β -HCG 的回归系数为 5.297 7, Free β -HCG 浓度的 MoM 值从 1 增加到 2, 患病风险率提高 4.9 倍; MoM 值增加到 3, 患病风险率提高 12.5 倍. 风险率的增加倍数, 以指数形式增加, 孕妇胎儿患病风险率也增加. 从回归系数看, AFP、Free β -HCG 的 MoM 值变化对患病风险率的影响要比 uE3 大.

Johnson 等^[7]应用多变量正态模型的筛查软件, 采用三联方案, 在血清学指标筛查中发现, 在 5.1% 阳性率下, 检出率为 74%. Sahota 等^[8]在研究中发现, 在 5.0% 阳性率下, 检出率为 72%. 而本研究用 Logistic 模型的检出率要高于目前采用多变量正态模型的国外筛查软件.

本研究中建立的多变量正态模型的三联筛查方案, 在同样 5% 假阳性率下, 检出率为 78.1%, 略低于 Logistic 模型的三联筛查方案. 这可能是由于个体差异, 检出率相差 1.53% 并不能完全说明两个模型的效率.

参考文献:

[1] PALOMAKI G E, HADDOW J E. Maternal serum alpha-fetoprotein, age, and Down syndrome risk[J]. Am J Obstet Gynecol, 1987, 156(2):460-463.

[2] LEMESHOW S, HOSMER D W J. A review of goodness of fit statistics for use in the development of logistic regression models[J]. Am J Epidemiol, 1982, 115(1):92-106.

[3] HWA H L, MING T K, HSIEH F J, et al. Risk prediction for Down syndrome in young pregnant women using maternal serum biomarkers: determination of cut-off risk from receiver operating characteristic curve analysis[J]. Journal of Evaluation in Clinical Practice, 2007, 13: 254-258.

[4] LEUNG T N, PANG M W, DALJIT S S. Fetal biometry in ethnic Chinese: Biparietal diameter, head circumference, abdominal circumference and femur length[J]. Ultrasound Obstet Gynecol, 2008, 31(3):321- 327.

[5] 卓仁杰, 秦小岸, 张檀, 等. 孕中期唐氏综合征筛查中血清标志物中位数倍数校正模型[J]. 中国医科大学学报, 2011, 40(1):67-69.

[6] WALD N J, GEORGE L, SMITH D, et al. Serum screening for Down syndrome between 8 and 14 weeks of pregnancy[J]. Br J Obstet Gynecol, 1996, 103(5):407-

- 412.
- [7] JOHNSON J, PASTUCK M, METCALFE A, et al. First-trimester Down syndrome screening using additional serum markers with and without nuchal translucency and cell-free DNA[J]. *Prenat Diagn*, 2013, 33(11):1044-1049.
- [8] SAHOTA D S, LEUNG W C, CHAN W P, et al. Prospective assessment of the Hong Kong hospital authority universal down syndrome screening programme [J]. *Hong Kong Med J*, 2014, 19(2):101-108.

Predicting Risk of Prenatal Screening for Down Syndrome Using Logistic Regression Model

ZHUO Ren-jie¹, LI Ying², ZHANG Li-na¹, SHEN Qi-jun¹

(1.School of Medicine, Ningbo University, Ningbo 315211, China;

2.Internal Medicine, Ningbo First Hospital, Ningbo 315000, China)

Abstract: In this study, a total of 64 772 pregnant woman in gestational timeframe between 12 to 20 weeks are selected for analytical purposes, including 64 718 normal pregnant woman and 54 suffering Down's syndrome. The maternal serum markers are converted to serum with multiple of median (MoM) adjusted by grouping the gestational age and weight. The MoM is adjusted by twin, smoking and diabetes in the Logistic regression model. Risk estimation is calculated for each pregnant woman using two different screening methods. Area under the curve of receiver operator characteristic (AUC), detection rate and false positive rate are calculated individually to assess the effect of these two different methods. Results show that age, alpha-fetoprotein (AFP), free human chorionic gonadotrophin (HCG) and unconjugated estriol (uE3) are significantly dominant in the Logistic regression model. The AUC of tripe serum screening is larger than that of double serum screening, but manifests no significant difference ($z=1.382$, $P=0.1669$). With the same cutoff value, the detection rate of triple serum screening is higher than that of double serum screening, and the false positive rate of triple serum screening is lower than that of double serum screening. The significant difference is detected between these two different models ($t=-9.44$, $P<0.001$). Triple serum screening is found to be the superior to double serum screening compared with other screenings. Also noted is that when the false positive rate is 4.3%, the optimal cutoff value turns out to be 1:400.

key words: risk assessment model; Logistic regression; second trimester; Down's syndrome; screening

(责任编辑 史小丽)