

原发性高血压影响因素的逐步线性回归 和多元Logistic模型分析

安徽医科大学流行病学教研室

张衍文 薛英杰 倪进发

本文应用逐步线性回归和多元Logistic模型, 比较分析了一宗高血压流行病学调查资料, 提示高血压发生与某些危险因素的定量关系。

一、样本: 1983年3~5月。对合肥市中市区花园巷街道全体居民(661人)进行了血压普查, 同时测量体重、身高及盐味阈。其中7~19岁青少年152人, 20~60岁以上者509人。

随机抽查7~19岁组60人, 20岁以上者153人, 检查血脂水平, 并以此抽样人群作为多因素分析的样本。

二、检测方法: ①血压: 按WHO规定的方法测量; ②血脂: 按常规技术检测; ③盐味阈: 分别以0.3~1.8g% NaCl六种不同浓度溶液和无盐液(对照)滴舌, 以能鉴别溶液的最低浓度作为盐味阈。

三、统计分析方法: 将样本分为 ≥ 20 岁成人和 < 20 岁青少年两组, 分别进行统计分析。分析的变量有性别、年龄(岁)、体重(W, kg)、身高(H, Cm)、过重指数(W/H^3 , Kg/m^3)、总胆固醇(TCL, mg%)、高密度脂蛋白胆固醇(HDL-c, mg%)、甘油三脂(TG, mg%)、盐味阈(g%)。

采用逐步线性回归和多元Logistic模型分析各变量与血压的关系, 由Cromenco微型计算机执行FORTRAN语言完成参数估计。

四、结果:

1. 成人:

①逐步回归分析的结果: 当 $\alpha=0.3$ 时, SBP的逐步回归仅筛选出年龄因素; DBP仅筛选出 W/H^3 , 其余变量均未能引入。最后的回归方程是:

$$SBP=88.35+1.02(\text{年龄, 岁})$$

$$DBP=39.74+2.33(W/H^3, kg/m^3)$$

②多元Logistic模型分析结果: 当 $\alpha=0.05$ 时, 拟合Logistic模型。以 $BP \geq 160/95$ mmHg作为高血压, 逐步筛选出的因素有年龄、 W/H^3 、TG和盐味阈, 其中盐味阈系数为负值。以 $BP \geq 140/90$ 作为“高血压”, 筛选出的因素有年龄、 W/H^3 、HDL-c、盐味阈。其中HDL-c、盐味阈系数为负值, 其余变

量未能进入模型。筛选后拟合的Logistic模型为:

$$\text{当以 } BP \geq 160/95 \text{ mmHg 为病例时, } P=[1+\exp(-9.97-0.09\text{年龄, 岁}-0.38W/H^3, kg/m^3+0.013\text{盐味阈}-0.004TG)]^{-1};$$

$$\text{当以 } BP \geq 140/90 \text{ mmHg 为病例时, } P=[1+\exp(-8.88-0.10\text{年龄, 岁}-0.19W/H^3, kg/m^3+0.013\text{盐味阈}+0.013HDL-c)]^{-1}.$$

2. 青少年:

①逐步线性回归分析结果: 当 $\alpha=0.3$ 时, 不论SBP还是DBP仅选入体重; 其余变量均未进入模型。回归方程是:

$$SBP=76.05+0.39(\text{体重, kg})$$

$$DBP=53.96+0.47(\text{体重, kg})$$

②多元Logistic模型分析结果: 多元Logistic模型筛选出体重和TCL, 其余变量未能选入。拟合Logistic模型是:

$$P=[1+\exp(-9.0-0.1\text{体重}-0.023TCL)]^{-1}$$

以上结果表明, 通过逐步线性回归和多元Logistic模型筛选出的因素并不完全一致。这种差别来自于两种统计模型不同的基本原理。多元线性回归模型可以评价各因素对BP总体平均水平的影响; 而多元Logistic模型则可估计这些因素对产生高血压危险的作用。如果两种模型结合分析, 既可评价某些因素对总体均数的影响, 又可估计高血压患病的危险度。在本项研究中, 成人组的逐步线性回归和多元Logistic, 都筛选出年龄、 W/H^3 , 可见两因素不仅对血压总体水平有影响, 且有增加高血压患病危险的作用。血脂仅被多元Logistic模型筛选出, 而未能进入逐步线性回归, 说明对血压总体水平影响较小, 而对高血压患病危险有一定的影响, 且提示甘油三脂与增加高血压危险有关, HDL-c与降低高血压危险有关。在青少年组, 两种模型都选出了体重, 而 W/H^3 并未能引入。青少年体重升高血压随之上升, 可能是心血管系统发育与体格增长的一种伴随现象。本研究还提示: 总胆固醇与青少年血压升高有关。但这种现象很难作出恰当的解释, 待今后研究。