

# 中国成年人高血压患病区域聚集性及危险因素的多水平模型分析

杨永利 付鹏钰 胡东生 张卫东 张梅喜 王重建 平智广

**【摘要】** 目的 分析中国成年人高血压的区域聚集性及危险因素,探讨多水平模型在高血压危险因素研究中的应用。方法 采用多阶段随机抽样方法,于 2000—2001 年从中国 10 地区共抽得年龄在 35~74 岁之间的 15 540 人作为研究对象,采用 MLwiN 2.02 软件对数据进行二水平 logistic 回归模型拟合。结果 高血压的患病存在地区聚集现象,方差成分系数为 3.1%。在调整了年龄和性别的影响后,全身型肥胖人群(BMI $\geq 28$  kg/m<sup>2</sup>)和超重人群(BMI 24~27.9 kg/m<sup>2</sup>)分别为 BMI 正常人群(18.5~23.9 kg/m<sup>2</sup>) 患高血压风险的 4.50(95%CI:4.00~5.06)和 2.26(95%CI:2.07~2.46)倍;中心型肥胖人群(男性腰围 $\geq 85$  cm 或女性 $\geq 80$  cm)为正常腰围人群患高血压风险的 2.62(95%CI:2.42~2.83)倍;甘油三酯、总胆固醇、血糖、低密度脂蛋白胆固醇含量升高或者高密度脂蛋白胆固醇含量降低者患病风险分别为正常者高血压患病风险的 2.10(95%CI:1.89~2.33)、2.08(95%CI:1.84~2.35)、1.85(95%CI:1.60~2.14)、1.58(95%CI:1.38~1.81)和 1.49(95%CI:1.32~1.69)倍,饮酒人群为不饮酒人群患高血压风险的 1.15(95%CI:1.05~1.27)倍。结论 高血压的发生不仅与个体肥胖、血糖升高、血脂异常和饮酒等危险因素有关,还受所居住区域环境因素的影响;在高血压的一级预防中,既要关注高危人群,还要重视以一般人群为基础的群体预防控制工作。

**【关键词】** 高血压;危险因素;区域聚集性;二水平 logistic 回归模型;方差成分系数

**Multilevel regression analysis on region cluster and risk factors of hypertension in the Chinese adult population** YANG Yong-li, FU Peng-yu, HU Dong-sheng, ZHANG Wei-dong, ZHANG Mei-xi, WANG Chong-jian, PING Zhi-guang. Department of Health Statistics, College of Public Health, Zhengzhou University, Zhengzhou 450001, China

Corresponding author: HU Dong-sheng, Email: dongsheng\_hu@hotmail.com

**【Abstract】 Objective** To analyze the region cluster and risk factors of hypertension in the Chinese adult population and to explore the application of multilevel regression model in the risk factors of hypertension. **Methods** Multi-stage random sampling technique was used to choose 15 540 individuals aged 35–74 years from 10 regions in China. Two-level logistic regression models were fitted under MLwiN 2.02 software. **Results** The region cluster of hypertension existed and variance portion coefficient was 3.1%. After adjusting for the age and gender, overall obese people (BMI $\geq 28$  kg/m<sup>2</sup>) were 4.50(95%CI: 4.00–5.06) times, overweight people (BMI=24–27.9 kg/m<sup>2</sup>) were 2.26(95%CI: 2.07–2.46) times more likely to be hypertensive as compared with those of normal BMI (18.5–23.9 kg/m<sup>2</sup>), and those centrally obese people (Waist circumference  $\geq 85$  cm in male or 80 cm in female) were 2.62(95%CI: 2.42–2.83) times more likely to be hypertensive as compared with those of normal WC. The age- and gender-adjusted odds ratios (ORs) of triglyceride (TG), serum cholesterol (TC), glucose, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C) and drinking alcohol were 2.10 (95% CI: 1.89–2.33), 2.08 (95% CI: 1.84–2.35), 1.85 (95% CI: 1.60–2.14), 1.58 (95% CI: 1.38–1.81), 1.49 (95% CI: 1.32–1.69) and 1.15 (95% CI: 1.05–1.27), respectively. **Conclusion** The prevalence of hypertension was not only affected by individual risk factors, such as obesity, drinking alcohol, abnormal glucose and serum lipids profile, but also affected by the geographic environment where people resided in. Population- and risk factors targeted strategies, proved a promising way to reduce individual risk of hypertension in the primary prevention of hypertension.

**【Key words】** Hypertension; Risk factor; Region cluster; Two-level logistic regression model; Variance portion coefficient

DOI: 10.3760/cma.j.issn.0254-6450.2009.07.018

基金项目: 亚洲心血管健康多中心合作研究

作者单位: 450001 郑州大学公共卫生学院卫生统计学教研室(杨永利、平智广), 流行病学教研室(胡东生、张卫东、张梅喜、王重建); 河南省疾病预防控制中心(付鹏钰)

通信作者: 胡东生, Email: dongsheng\_hu@hotmail.com

高血压患病率在世界范围内呈上升趋势<sup>[1,2]</sup>,其患病率有地区差异<sup>[3,4]</sup>,原因在于居住在同一地区的居民在社会经济地位、行为和饮食习惯以及健康知识等方面较来自于不同地区的人群更为近似,即表现出血压分布的区域聚集现象<sup>[5]</sup>,而区域聚集性是造成高血压分布地区差异的主要原因。在比较高血压患病的地区差异时,最常用的方法是 $\chi^2$ 检验,但局限于定性层面分析血压分布是否存在地区差异,并不能将造成地区差异的原因中环境因素的作用强度进行量化。此外,分析高血压患病的危险因素也是流行病学的主要研究内容之一,在分析高血压患病的危险因素时,虽然很多学者已经意识到高血压分布的地区差异,在数据分析时,大多仍将来自不同研究区域的研究对象直接合并后采用传统的一水平 logistic 回归模型,该方法未能将环境因素的作用单独分离出来,因此往往高估了危险因素的作用强度。与传统的一水平模型相比,多水平模型能将残差项分解到与其数据层次结构相应的水平上,因此获得的参数估计更为准确<sup>[6,7]</sup>。亚洲心血管病国际合作研究(International Collaborative Study of Cardiovascular Disease in ASIA, InterASIA)资料<sup>[8]</sup>采用多阶段随机抽样方法,有明显的层次结构,即省→城市→县→乡村→户→个体,因此,该研究拟采用多水平 logistic 回归模型,分析环境因素在高血压发病中的作用大小,同时更科学地估计不同危险因素对这些疾病发生作用的强度,研究结果可为制定高血压的群体预防策略提供科学依据。

### 资料与方法

1. 资料来源:第一阶段分别从我国北方和南方地区各抽取 5 个省市。北方省市包括北京、陕西、山东、青海和吉林,南方省市包括上海、湖北、福建、广西和四川。为了充分考虑地理位置和经济水平的代表性,该阶段未采用随机抽样的方式。第二阶段是在第一阶段所选的省市中随机抽取 1 个县和 1 个城区,共抽得 10 个县和 10 个城区。第三个阶段是在每个城区和县的 1 个街道和 1 个镇(或乡),共抽得 10 个街道和 10 个镇(或乡)。第四阶段是以选中街道或镇(或乡)的居民名单将作为样本来源(年龄 35~74 岁),设计每个现场抽得有效样本共 800 人,男、女各半。此次数据分析局限在 35~74 岁,且同时完成问卷调查和人体测量的 15 540 名参与者,其中男性 7526 人(48.4%),女性 8014 人(51.6%)。

2. 研究内容:包括问卷调查、人体测量和实验

室生化检测三部分。血压由经过统一培训的人员使用标准的水银血压计测量,调查者静坐 5 min 后测量 3 次,取后两次 SBP 和 DBP 的平均值为调查所得值<sup>[9]</sup>。高血压诊断标准:SBP $\geq$ 140 mm Hg 或者 DBP $\geq$ 90 mm Hg 或者在调查前两周内接受过抗高血压治疗<sup>[8]</sup>。

3. 多水平模型:根据不同的分类标准,多水平模型可以分成不同的类别。根据研究水平的多少,多水平模型分为二水平模型、三水平模型或四水平模型;根据反应变量的类型,多水平模型分为多水平线性回归模型、多水平 logistic 回归模型以及多水平 poisson 回归模型。从 InterASIA 资料的抽样过程看,有两个层次:个体和研究区域,即 15 540 名研究对象来自 10 个地区,因此定义个体为水平 1 单位,研究区域为水平 2 单位;结合反应变量(是否患高血压)的类型(二分类资料),拟合二水平 logistic 回归模型:

$$f(\pi_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

$$\beta_0 = \beta_0 + u_{0j}$$

$$\text{连接函数 } f(\pi_i) = \log \frac{\pi_i}{1 - \pi_i}$$

式中, $i=1, 2, \dots, n_j$ , 表示水平 1 单位(个体); $j=1, 2, \dots, n_j$  表示水平 2 单位(研究区域); $\beta_0$  为截距, $x_{i1}, x_{i2}, \dots$  分别为第  $j$  个研究区域第  $i$  个观察对象的第  $k$  个解释变量的观察值; $\beta_1, \beta_2, \dots, \beta_k$  分别为第 1, 2,  $\dots, k$  个解释变量的偏回归系数。

4. 统计学分析:采用 MLwiN 2.02 软件进行统计分析<sup>[10]</sup>。根据 InterASIA 资料的抽样过程,首先拟合无解释变量的二水平 logistic 回归零模型。以判断是否存在高水平的聚集作用,如果存在区域聚集性,计算方差成分系数( $VPC = \frac{\sigma_w^2}{\sigma_w^2 + 3.29}$ )<sup>[10, 11]</sup>以衡量区域聚集性的大小,之后分别纳入解释变量以获得各危险因素的参数估计。参数估计采用二阶预测性拟似然法(2<sup>nd</sup> predictive or penalized quasi-likelihood, PQL)<sup>[10]</sup>,调整后的 OR 值及 95% CI 用来评价各危险因素的作用强度,检验水准 $\alpha=0.05$ 。

### 结 果

1. 不同地区高血压患病率:不同地区高血压的患病率不同, $\chi^2=286.9, P<0.01$ 。其中山东省居民高血压患病率最高,而四川省居民高血压患病率最低(图 1)。

2. 区域聚集性的度量:首先拟合二水平 logistic 回归零模型,判断在高血压的发病中,是否存在高水

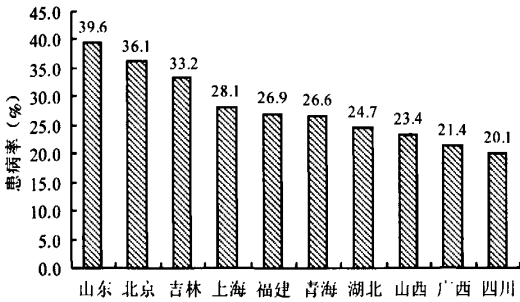


图 1 我国 10 个地区高血压患病率

平的聚集现象。水平 2 方差  $\sigma_{\omega}^2$  的假设检验  $\chi^2 = 5.10$ , 根据自由度为 1 的  $\chi^2$  分布, 认为高血压患病存在地区聚集现象,  $VPC = 3.1\%$ , 说明在高血压的发病中, 环境因素的作用占 3.1%。因此在影响高血压患病的危险因素分析时, 应拟合二水平 logistic 回归模型。

3. 高血压患病影响因素的二水平 logistic 回归模型: 纳入解释变量后的二水平 logistic 回归模型显示(表 1): 调整了年龄和性别的影响后, 饮酒、BMI 和腰围(WC)增加, TG、TC、FPG 和 LDL-C 含量升高, 或者 HDL-C 含量降低是高血压的独立危险因素(表 2)。

表 1 研究变量赋值及纳入模型方式

研究变量	赋值	纳入模型方式
年龄组(岁)	1=35~44.9, 2=45~54.9, 3=55~64.9, 4=65~74	哑变量(参照=1)
性别	1=女, 2=男	原始数据
吸烟状况	1=不吸烟, 2=吸烟*	原始数据
饮酒状况	1=不饮酒, 2=饮酒*	原始数据
BMI(kg/m <sup>2</sup> )	1=<18.5(偏瘦), 2=18.5~23.9(正常), 3=24~27.9(超重), 4= $\geq 28$ (全身型肥胖)	哑变量(参照=2)
WC(cm)	男性: 1=<85(正常), 2= $\geq 85$ (向心性肥胖); 女性: 1=<80(正常), 2= $\geq 80$ (向心性肥胖)	原始数据
TC(mg/dl)	1=<200(正常), 2=200~239(临界值), 3= $\geq 240$ (异常)	哑变量(参照=1)
TG(mg/dl)	1=<150(正常), 2=150~199(临界值), 3= $\geq 200$ (异常)	哑变量(参照=1)
FPG(mg/dl)	1=<110(正常), 2=110~125(临界值), 3= $\geq 126$ (异常)	哑变量(参照=1)
LDL-C(mg/dl)	1=<130(正常), 2=130~159(临界值), 3= $\geq 160$ (异常)	哑变量(参照=1)
HDL-C(mg/dl)	1=<40(异常), 2=40~59(临界值), 3= $\geq 60$ (正常)	哑变量(参照=3)

注: \*当前吸烟且累计吸烟量 $\geq 100$ 支; †过去 1 年中饮酒次数 $> 12$ 次; BMI 和 WC 分类标准参考国际生命科学学会“中国肥胖问题工作组”《中国成人超重和肥胖症预防与控制指南》

表 2 高血压患病危险因素的二水平 logistic 回归模型

解释变量	OR 值(95%CI)	
BMI(kg/m <sup>2</sup> )	<18.5	0.74(0.60~0.90)
	24~27.9	2.26(2.07~2.46)
	$\geq 28$	4.50(4.00~5.06)
WC(cm)	$\geq 85$ (男)或 $\geq 80$ (女)	2.62(2.42~2.83)
TG(mg/dl)	150~199	1.82(1.64~2.02)
	$\geq 200$	2.10(1.89~2.33)
TC(mg/dl)	200~239	1.56(1.43~1.70)
	$\geq 240$	2.08(1.84~2.35)
FPG(mg/dl)	110~125	1.86(1.63~2.12)
	$\geq 126$	1.85(1.60~2.14)
LDL-C(mg/dl)	130~159	1.39(1.26~1.54)
	$\geq 160$	1.58(1.38~1.81)
HDL-C(mg/dl)	<40	1.49(1.32~1.69)
	40~59	1.23(1.10~1.36)
吸烟	-	0.78(0.71~0.86)
饮酒	-	1.15(1.05~1.27)

注: 调整了年龄、性别的影响

### 讨 论

在高血压的一级预防中, 以人群为基础的预防策略由于其经济有效而引起关注, 多水平模型在流行病学研究中的应用为研究疾病的地理分布和环境因素的作用提供了一种无可替代的工具。2004 年 Merlo 等<sup>[5]</sup>将多水平回归模型应用到 SBP 区域聚集性研究中, 对来自 39 个国家、年龄在 35~64 岁之间的 23 796 名男性和 24 986 名女性进行二水平线性回归模型分析, 结果显示在 SBP 的总变异中, 由于区域因素引起的变异达到 7%~8%; 2007 年该学者又对瑞典马尔默城市 17 个地区年龄在 45~73 岁之间的 15 569 名妇女进行 DBP 区域聚集性的度量, 结果表明在 DBP 的变异中, 环境因素起的作用占 0.4%。但我国成年人高血压发病的区域环境因素究竟起了多大作用? 目前未见报道。通过拟合二水平 logistic 回归模型, 发现我国成年人高血压的患病状况也受其所居住区域环境因素的影响, 环境因素的作用占 3.1%, 由于我国人口基数大, 因此在高血压的预防工作中, 更应该考虑以一般人群为基础的预防策略。

本文还在既往研究的高血压各危险因素基础之上, 获得了我国成年人高血压危险因素的更为合适的估计。

1. BMI、WC 与高血压: 肥胖是高血压的独立危险因素, 体形越胖的人群, 患高血压的风险也越大, 全身型肥胖人群为 BMI 正常人群患病风险的 4.50 倍, 超重人群为 BMI 正常人群患病风险的 2.26 倍, 中心型肥胖人群为正常腰围人群患病风险的 2.62

倍,与 Rohrer 等<sup>[12]</sup>的研究结果较为接近,因此,保持合适的体形可能会降低高血压的发生。针对肥胖人群开展有效的减肥治疗是相当困难的,但通过合理膳食或者增加运动量来保持合适的体形,有望降低患高血压的风险。

2. 血糖、血脂与高血压:高血压的发病与血糖和血脂含量异常有关, TG、TC、FPG、LDL-C 含量升高或者 HDL-C 含量降低者患病风险分别为其正常者的 2.10、2.08、1.85、1.58 和 1.49 倍,与文献报道一致<sup>[3]</sup>。进一步研究表明,人们的生活方式会影响到心血管疾病相关危险因素<sup>[14]</sup>,因此欲降低高血压的发生,应以合理的膳食和健康的生活方式为基础。

3. 吸烟、饮酒与高血压:吸烟曾被认为是心血管疾病的重要危险因素之一,但其与高血压存在何种关系,目前仍不清楚。周超等<sup>[15]</sup>和 Bowman 等<sup>[16]</sup>的研究表明,吸烟是高血压的危险因素,但 Suliburska 和 Duda<sup>[17]</sup>发现吸烟与血压水平无关,该研究表明:吸烟与高血压的发病呈负相关,而饮酒与高血压的发病呈正相关。出现此种不一致结论的原因是多方面的,其中信息的准确性是重要原因之一。由于吸烟和饮酒是难以精确测量的指标,在研究吸烟、饮酒与高血压的关系时,还应尽量考虑每日吸烟量或饮酒量、吸烟或饮酒年限、烟草的种类、酒精度数等更为详细的信息。

总之,本研究采用多水平建模技术,一方面定量评价了环境因素在高血压发病中的作用强度,另一方面将环境因素的效应分离出来,因此对于各危险因素的估计更为准确。研究结果为制定高血压的群体预防策略提供了科学依据。当然本研究也有横断面研究不可避免的局限性,在揭示高血压危险因素的致病作用时说服力较差,建议通过队列研究或干预研究进一步验证这些危险因素。

#### 参 考 文 献

[1] 胡以松,姚崇华,王文志,等. 2002 年中国部分民族高血压患病情况. 卫生研究, 2006, 35(5): 573-575.  
 [2] Bloch MJ, Basile J. Hypertension prevalence increases and hypertension control rates improve in the United States; 2003-2004. J Clin Hypertens (Greenwich), 2007, 9(2): 147-149.  
 [3] Kearney PM, Whelton M, Reynolds K, et al. Worldwide prevalence of hypertension: a systematic review. J Hypertens, 2004, 22(1): 11-19.

[4] Reynolds K, Gu D, Muntner P, et al. Geographic variations in the prevalence, awareness, treatment and control of hypertension in China. J Hypertens, 2003, 21(7): 1273-1281.  
 [5] Merlo J, Asplund K, Lynch J, et al. Population effects on individual systolic blood pressure: a multilevel analysis of the World Health Organization MONICA Project. Am J Epidemiol, 2004, 159(12): 1168-1179.  
 [6] 杨珉, 李晓松. 医学和公共卫生: 研究常用多水平统计模型. 北京: 北京大学医学出版社, 2007.  
 [7] Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: integrating random and fixed effects in multilevel logistic regression. Am J Epidemiol, 2005, 161(1): 81-88.  
 [8] Gu D, Reynolds K, Wu X, et al. Prevalence, awareness, treatment, and control of hypertension in China. Hypertension, 2002, 40(6): 920-927.  
 [9] Matthews KA, Katholi CR, McCreath H, et al. Blood pressure reactivity to psychological stress predicts hypertension in the CARDIA study. Circulation, 2004, 110(1): 74-78.  
 [10] Rasbash J, Steele F, Browne W, et al. A user's guide to MLwiN. Version 2.0. London: Centre for Multilevel Modelling Institute of Education University of London, 2004.  
 [11] Merlo J, Chaix B, Ohlsson H, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health, 2006, 60(4): 290-297.  
 [12] Rohrer JE, Anderson GJ, Furst JW. Obesity and pre-hypertension in family medicine: implications for quality improvement. BMC Health Serv Res, 2007, 7: 212.  
 [13] Boyko EJ, Barr EL, Zimmet PZ, et al. Two-hour glucose predicts the development of hypertension over 5 years: the AusDiab study. J Hum Hypertens, 2008, 22(3): 168-176.  
 [14] Fukuda H, Haruyama Y, Nakade M, et al. Relationship between lifestyle and change of cardiovascular risk factors based on a five-year follow-up of employees in Japan. Ind Health, 2007, 45(1): 56-61.  
 [15] 周超, 雷燕, 杨春霞, 等. 饮酒、吸烟和嗜盐等生活方式与高血压发病关联的 Meta 分析. 现代预防医学, 2006, 33(4): 488-490.  
 [16] Bowman TS, Gaziano JM, Buring JE, et al. A prospective study of cigarette smoking and risk of incident hypertension in women. J Am Coll Cardiol, 2007, 50(21): 2085-2092.  
 [17] Suliburska J, Duda G. Tobacco smoking and the frequency of arterial hypertension in adults. Przegł Lek, 2006, 63(10): 864-866.

(收稿日期: 2009-03-04)

(本文编辑: 张林东)