

· 基础理论与方法 ·

列线图法预测人群归因分值

邱宏 余德新

【导读】 人群归因分值(人群归因危险度百分比,PAF)是广大流行病学工作者熟悉的公共卫生学指标。PAF 的计算主要根据某个危险因素对某病的相对危险度(RR)和人群中该危险因素的暴露比例(P_e)。文中介绍由 RR 和 P_e 估计 PAF 列线图的制作方法,以便快速简便地估算 PAF。

【关键词】 人群归因分值; 人群暴露比例; 相对危险度; 列线图

Making nomogram to estimate the population attributable fraction QIU Hong, Ignatius Tak-sun YU.
Department of Community and Family Medicine, School of Public Health, Chinese University of Hong Kong, HKSAR, China

Corresponding author: Ignatius Tak-sun YU, Email: iyu@cuhk.edu.hk

【Introduction】 Population attributable fraction (PAF) is an important public health parameter that is familiar to epidemiologists. The calculation of PAF is frequently based on the relative risk (RR) of a risk factor for a certain disease and the prevalence of exposure to the risk factor in the total population (P_e). We introduce here the method for making a nomogram linking PAF to RR and P_e . The nomogram will be very handy for public health professionals in estimating PAF quickly and accurately when information on RR and P_e is available.

【Key words】 Population attributable fraction; Prevalence of exposure in total population; Relative risk; Nomogram

人群归因分值(PAF)或称人群归因危险度百分比(population attributable risk proportion, PAR%)是定量描述暴露危险因素对人群致病作用大小的统计指标,表示总体人群中某病归因于某种因素引起的发病(或死亡)占人群全部发病(或死亡)的比例,亦可理解为消除某种危险因素后可使人群中该病的发病(或死亡)降低的比重。因此,PAF 在慢性病防治研究中常用于评价因素在人群中的危害程度,为卫生防病决策的制定提供依据,具有重要的流行病学和公共卫生意义。

PAF 的计算基于这样的考虑:假设人群中没有某因素的暴露,则人群总发病率(I_p)和未暴露人群发病率(I_o)应该相等,因此人群中归因于某种暴露引起的发病占人群总发病的比例(即 PAF)可由 $(I_p - I_o)/I_p$ 计算。但因为未暴露人群的发病率 I_o 常常未知,实际运用中多根据人群暴露率(P_e)和病例对照研究获得的相对危险度(RR)的估计值用公式 $PAF = \frac{P_e \times (RR - 1)}{1 + P_e \times (RR - 1)}$ 推算^[1],两个公式等

价。第二个公式相对难记,易于出错,为方便广大公共卫生专业人员分析人群中不同危险因素对各种疾病所造成的负担、比较不同干预手段的有效性,本文根据 PAF 的计算公式制作了列线图,以求更简便直观地估计 PAF。

列线图亦称诺模图,一种古老的计算工具,在计算机高速发展的今天,仍以其简捷、直观、方便的特性深受欢迎^[2-6],在医药卫生领域特别是临床医学上也得到较多的关注^[7-11]。我们尝试用列线图估计 PAF,得到很好的效果。因为相关文献往往偏重于介绍列线图的使用,缺乏列线图做图方法的介绍,而且现有统计软件 SPSS、SAS、S-PLUS 等均不能直接制作列线图,故本文以 PAF 的计算公式为例,介绍列线图的做图方法^[12],用于快速简便地估算 PAF。

基本原理

制作诺模图的目的是根据已知指标的数值,在图上直接读出与之有关联的未知指标的数值。诺模图有多种形式,为带有刻度的直线或曲线组成,其中最简单的一种即三条垂直直线构成的列线图,三指标之间函数关系的基本形式为 $W = U \times V$ 。对方程两端取对数后成为 $\log W = \log U + \log V$,制作列线

作者单位:香港中文大学公共卫生学院社区及家庭医学系

通讯作者:余德新,Email:iyu@cuhk.edu.hk

图的关键是根据此函数关系计算代表三指标列线的图尺系数和列线之间的距离。

假设三条列线的图尺系数分别为 m_U 、 m_V 和 m_W ，列线 W 和列线 U 、 V 之间的距离分别为 D_1 和 D_2 ，依三指标的函数关系由图 1 可知：

$$\frac{m_W}{m_V} = \frac{D_1}{D_1 + D_2}, \quad \frac{m_W}{m_U} = \frac{D_2}{D_1 + D_2}$$

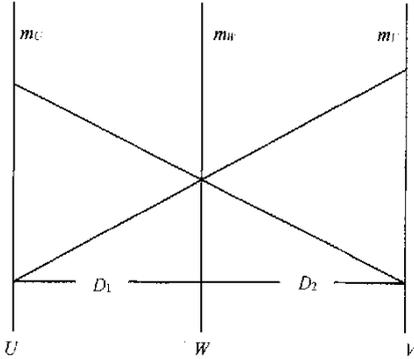


图1 三条列线之间的关系示意图

并由此可得： $\frac{m_U}{m_V} = \frac{D_1}{D_2}$ 和 $m_W = \frac{m_U \times m_V}{m_U + m_V}$

如下以 PAF 的计算公式为例，介绍图尺系数和列线距离的计算以及列线图的制作。

实例分析

1. PAF 的计算公式：

$$PAF = \frac{I_e \cdot I_o}{I_p} = \frac{Pe \times (RR - 1)}{1 + Pe \times (RR - 1)}$$

式中： $I_p = I_e \times Pe + I_o \times (1 - Pe)$ ；暴露人群发病率 (incidence in the exposed, I_e)； $RR = I_e / I_o$ 。

设 $A = Pe \times (RR - 1)$ ，两边取对数得： $\log A = \log Pe + \log (RR - 1)$ ，可绘制由 Pe 和 $RR - 1$ 推算 A 的、由三条平行列线图尺组成的诺模图。再根据 $PAF = \frac{A}{1 + A}$ 将列线 A 的刻度转换成 PAF，同时将列线 $RR - 1$ 的刻度转换成 RR ，就可得到由 Pe 和 RR 推算 PAF 的列线图。

2. 计算各列线图尺的系数 (m)：假设各变量的取值范围为： $Pe = 0.01 \sim 0.9$ ， $RR = 1.5 \sim 10.0$ ，则 $A = 0.005 \sim 8.1$ 。

设各列线图尺长度均为 15 cm，则图尺系数分别为：

$$m_{Pe} = \frac{15}{\log 0.9 - \log 0.01} = 7.676$$

$$m_{(RR-1)} = \frac{15}{\log 9.0 - \log 0.5} = 11.950$$

$$m_{Pe \times (RR-1)} = \frac{15}{\log 8.1 - \log 0.005} = 4.674 = \frac{m_{Pe} \times m_{(RR-1)}}{m_{Pe} + m_{(RR-1)}}$$

3. 计算列线图尺各标值点长度 (L) 的标值方程为各图尺定刻度：

$$L_{Pe} = 7.676 \times (\log Pe - \log 0.01)$$

$$L_{(RR-1)} = 11.950 \times [\log (RR - 1) - \log 0.5]$$

$$L_{Pe \times (RR-1)} = 4.674 \times [\log \{Pe \times (RR - 1)\} - \log 0.005]$$

由标值方程计算标值点长度 (L)，并将 $RR - 1$ 和 A 换算成 RR 和 PAF，见表 1。

表1 三条列线图尺的标值点长度计算表

Pe	L_{Pe} (cm)	RR	$RR - 1$	$L_{(RR-1)}$ (cm)	PAF	$\frac{Pe \times (RR - 1)}{(RR - 1)}$	$L_{Pe \times (RR-1)}$ (cm)
0.01	0.00	1.5	0.5	0.00	0.005	0.005	0.00
0.02	2.31	2	1	3.60	0.01	0.010	1.41
0.05	5.37	2.5	1.5	5.70	0.02	0.020	2.81
0.1	7.68	3	2	7.20	0.05	0.053	4.79
0.15	9.03	3.5	2.5	8.35	0.08	0.087	5.80
0.2	9.99	4	3	9.30	0.10	0.111	6.29
0.3	11.34	4.5	3.5	10.10	0.20	0.250	7.94
0.4	12.30	5	4	10.79	0.30	0.429	9.04
0.5	13.04	5.5	4.5	11.40	0.40	0.667	9.93
0.6	13.65	6	5	11.95	0.50	1.000	10.76
0.7	14.16	6.5	5.5	12.45	0.60	1.500	11.58
0.8	14.61	7	6	12.90	0.70	2.333	12.48
0.9	15.00	7.5	6.5	13.31	0.80	4.000	13.57
		8	7	13.70	0.85	5.667	14.28
		8.5	7.5	14.05	0.89	8.100	15.00
		9	8	14.39			
		9.5	8.5	14.70			
		10	9	15.00			

4. 计算各列线图尺之间的距离 (D)：设列线 Pe 和列线 ($RR - 1$) 之间的距离为 20 cm，列线 A 和列线 Pe 的距离为 D_1 ，列线 A 和列线 ($RR - 1$) 的距离为 D_2 ，根据 $\frac{m_{Pe}}{m_{(RR-1)}} = \frac{D_1}{D_2}$ ，且 $D_1 + D_2 = 20$ cm 的关系可以得出：

$$D_1 = \frac{m_{Pe}}{m_{Pe} + m_{(RR-1)}} \times 20 = 7.82 \text{ cm}$$

$$D_2 = \frac{m_{(RR-1)}}{m_{Pe} + m_{(RR-1)}} \times 20 = 12.18 \text{ cm}$$

5. 制作列线图：利用上述计算结果，在 Excel 中以距离 D 为横坐标，标值点长度 L 为纵坐标做散点图并连成直线，略加修饰可成三条平行排列的列线图。参照表 1 的对应结果将列线 $RR - 1$ 和列线 A 的刻度分别以 RR 和 PAF 标注，即成所需的由 Pe 、 RR 估算 PAF 的列线图 (图 2)。

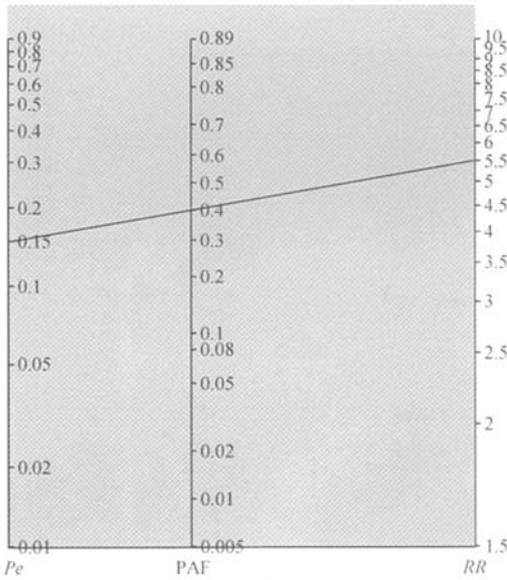


图2 Pe、RR 和 PAF 的列线图(RR = 1.5~10.0)

讨论

在饮酒和食管癌关系的病例对照研究中,假设饮酒者发生食管癌的危险性是不饮酒者的 5.5 倍 (OR = 5.5, 对于发病率低的食管癌而言, 可以作为 RR 的估计值), 且人群危险因素调查表明, 人群中饮酒比例大约为 15% ($Pe = 0.15$)。用列线图解算 PAF: 连接图尺 Pe 的 0.15 点和图尺 RR 的 5.5 点, 交于图尺 PAF 于 0.4 点, 即 $PAF = 0.4$, 说明食管癌的总发病病例中, 有大约 40% 是归因于饮酒所引起的。公式计算 $PAF = 0.4030$, 误差率为 0.74%。

图算法具有直观简捷的优势, 但不如直接用公式计算精确, 实际应用中可根据需要增加图尺的刻度 (即增加表 1 中三图尺的标值点数目) 或改变指标的取值范围来提高使用时估算的精确度。对于诸如 $1.0 < RR < 1.5$ 的某些环境危险因素 (如被动吸烟, RR 较低但 Pe 并不低; 某种空气污染物, RR 很低但 Pe 可达 100%), 可参考图 3 估算 PAF (制作方法同前)。

来自英国罗切斯特大学机电工程学院的 Thomas 等编写了一个在线制作列线图的 JAVA 程序^[13], 操作步骤简便, 制作的图形美观实用, 但因其提供的公式形式有限, 在使用上有一定局限性。

本文介绍的列线图法适用于危险因素暴露与否 (暴露因素为两分类变量) 的人群归因分值估算, 多个危险因素并存时宜用校正的 OR 值代入估算。多因素并存且以不同暴露水平分层时欲分析各危

因素的调整人群归因分值 (adjusted PAFs) 和多因素综合的人群归因分值 (summary PAF) 可参考 Paolo 等^[14] 提出的公式, 列线图法不适用。

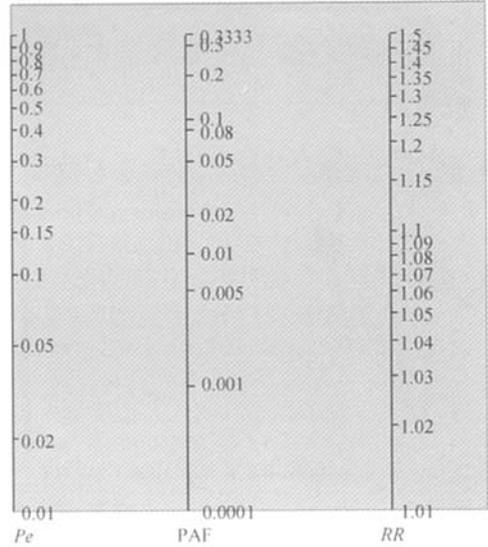


图3 Pe、RR 和 PAF 的列线图(RR = 1.01~1.50)

参考文献

- [1] 沈福民. 流行病学原理与方法. 上海: 复旦大学上海医科大学出版社, 2001: 59, 69.
- [2] 陈开熔. 轴心受压钢构件稳定计算公式的诺模图. 北京建筑工程学院学报, 1998, 14(3): 109-112.
- [3] 桂珍. 速披公式和刀具角度的快速近似计算. 安徽大学学报 (自然科学版), 1994, (1): 92-94.
- [4] 宋春亭, 陈纯江. 湖北省小流域洪峰流量计算诺模图探讨. 中国农村水利水电, 2002, 10: 43-45.
- [5] 崔思严. 用诺模图快速求解梯形断面渠道的临界水深. 水电站设计, 2006, 22(3): 53-56.
- [6] Kang DC, Tian HB, Zhang SH. Study on the forming limit nomogram of tensile stamping operations. J Master Sci Technol, 2004, 20(3): 363-364.
- [7] Li YH, Zhao X, Qing JI, et al. A nomogram for prediction of absorption rate coefficient. Chin Med J, 2004, 117(5): 689-694.
- [8] 姚正行, 张丽宁. 纠正低血钾、高血糖用药量计算列线图. 苏州医学院学报, 1994, 4(3): 252.
- [9] 王平基. 体重指数 (BMI) 列线图日测法. 中国心血管病杂志, 2004, 9(3): 227.
- [10] 王平基, 王新荣. 按需调整血管活性药物静脉点滴量速列线图日测法. 中国实用护理杂志, 2005, 21(8): 37.
- [11] Fagan TJ. Letter: nomogram for Bayes theorem. N Engl J Med, 1975, 293: 257.
- [12] Lee HJ. Nomography and empirical equations. New York: John Wiley & Sons, Inc, 1952: 18-43.
- [13] An example of the nomogram use. An online tool come from <http://www.ece.rochester.edu/~joncs/NomoDevel/>.
- [14] Paolo B, Sylvan BG, David PB, et al. Estimating the population attributable risk for multiple risk factors using case-control data. Am J Epidemiol, 1985, 122(5): 904-914.

(收稿日期: 2007-08-30)

(本文编辑: 张林东)