

橡胶职业暴露与肺癌关系的前瞻性研究

III肺癌危险因素联合作用方式

李 克 俞顺章

摘 要 目的 探讨橡胶工人中亚硝胺暴露、吸烟和经济收入低等主要危险因素对肺癌联合作用的模式。方法 应用广义相对危险度模型拟合某橡胶厂队列随访资料。结果 联合作用下,亚硝胺暴露与吸烟对肺癌危险度为相乘作用形式,亚硝胺暴露与经济收入为相加形式,吸烟与经济收入为相加形式。结论 受样本限制,不同模型之间拟合优度的差别较小,但是,在公共卫生实践和生物学机理的认识方面,确定因素对疾病危险度的联合作用方式具有重要意义。

关键词 橡胶职业暴露 肺癌 队列研究 广义相对危险度模型

A Prospective Study on Relationship between Occupational Exposure to Rubber and Lung Cancer III Combined Effects of Risk Factors for Lung Cancer Li ke*, Yu Shun-zhang*. Shantou University School of Medicine. 515031

Abstract **Objective** To explore the mode of combined effects of main risk factors, such as exposure to nitrosoamine, smoking, low income, et al, on lung cancer in rubber workers. **Methods** Generalized relative risk model was fitted with data collected from the cohort followed-up in a rubber factory. **Results** Combined effects of exposure to nitrosoamine and smoking on lung cancer in rubber workers presented as multiplicative model, and those of exposure to nitrosoamine and low income, and smoking and low income appeared additive. **Conclusion** There was only a little difference in goodness-of-fit between the models due to the limitation of sample size, but determination of the mode of combined effects of different risk factors played important roles in practice of public health and recognition of its biological mechanism.

Key words Rubber occupational exposure, Lung cancer, Cohort study, Generalized relative risk model

前面分析表明,除年龄、性别外,亚硝胺暴露、吸烟、经济收入低是本次研究中肺癌的3个主要危险因素。这些结果是通过流行病学多因素研究中常用的对数线性模型获得的。此类模型假定:(1)各个回归变量的作用是以某个因子乘以疾病发病率(危险度);(2)在危险度上几个变量的联合作用是它们单独作用的乘积^[1]。然而,这种相乘结果并不一定能够较

好地描述暴露与结果间的关系^[2]。公共卫生方面有许多问题属相加作用形式^[3]。由于常常在事先不清楚实际的数据符合何种模型结构,所以需要有一个通用的模型结构来评判因素间作用形式。为此,需要扩展相乘模型为相加或更为一般的相对危险度模型^[4,5]。广义相对危险度模型(General relative risk models)就是一种包括了相加相乘等一系列模型的通用方程式,其回归变量与相对危险度的连接形式范围依据一个参数从相加到相乘的模型结构。本部分将对亚硝胺、吸烟、经济收入等3个主要危险

作者单位: 515031 汕头大学医学院(李克),上海医科大学(俞顺章)

大学(俞顺章)

因素进行三者间作用关系系列模型分析,探讨肺癌主要危险因素间的联合作用模式。

1 材料和方法

1.1 材料

应用大中华橡胶厂工人队列 1973~1995 年间随访资料,将年龄、日历年作为动态时间分组变量,性别作为固定分组变量,按它们之间的交叉分组(即格子)分配人年,仅这 3 个分组变量之间的交叉分组就可组合为 $7 \times 5 \times 2 = 70$ 个格子,有时,还需加上亚硝酸胺暴露、吸烟、经济收入、随访时年龄以及自暴露以来时间等的固定或动态分组变量,这时交叉组成的格子可多达几万个,人年的分配极为复杂。通过这种方式形成的队列成组数据,可用于 Poisson 回归模型^[1,6]。

1.2 统计方法

应用时间分段的 Poisson 回归模型,对以上过程生成的队列人年分组数据进行生存分析,分组是指队列成员的数据按一个或一个以上时间变量(如年龄)与其他变量之间的交叉分组之间进行分配,在每个交叉分组(格子)中,其基本数据为危险时间合计,称为人年,还有病例数及其他协变量等数据。

1.3 统计模型

为解决不同的问题, Poisson 模型有多种形式,分述如下:

(1) 相乘(相对)危险度模型(multiplicative or relative risk models): 即为标准的相乘模型。设有 j 层和 k 个暴露水平,模型方程为

$$\log \lambda_{jk} = T_j + U_k \quad (1)$$

在此, $\lambda_{jk} = \theta_j \cdot \dot{\lambda}_k$, $T_j = \log \theta_j = \log \lambda_{j1}$, $U_k = \log \dot{\lambda}_k$ 。 $\dot{\lambda}_k$ 表示 k 水平暴露的疾病相对危险(与基准线水平相比, $j_1 = 1$)。可推广至多个暴露变量的情况^[6]。

(2) 相加相对危险度模型(additive relative risk models): 相乘模型〔式(1)〕的欠缺之一,至少当暴露变量为定量时,就会导致相对危险函数随暴露呈指数增加: $RR(x) = \exp(xU)$, 很显然,在许多流行病学研究中,率的增加不

会这么明显,解决方法之一是相加 RR 模型

$$\lambda_{jk} = \exp(\alpha_j) \{1 + x_{jk}U\} \quad (2)$$

这个模型中,要求 $x_{jk}U \geq -1$,否则会出现负的 RR 值,所以对于 x_{jk} 的编码有限制。

(3) 广义相对危险度模型簇: 式(1)和式(2)之间有一系列模型,可概括为

$$\lambda_{jk} = \exp(\alpha_j) r(x_{jk}; U) \quad (3)$$

相对危险度函数由参数 d 来定义

$$\log r(x_{jk}; U) = \begin{cases} \frac{(1 + x_{jk}U)^d - 1}{d}, & d \neq 0 \\ \log(1 + x_{jk}U), & d = 0 \end{cases}$$

当 $d = 0$ 时,为相加危险度模型〔式(2)〕,当 $d = 1$ 时为相乘模型〔式(1)〕。如果在模型中(相乘 RR 或相加 RR 模型)引进的交互作用项有统计意义时,应该重新检查暴露与疾病间的相乘关系,正规的评价是通过拟合广义相对危险度模型。

(4) 模型拟合优度分析: 各类模型主要使用 GLIM 软件和 EPICURE 软件^[7,8]拟合。应用对数似然比统计量^[1]评价模型拟合优度。比较观测值和拟合值之间一致性的计算公式为

$$G^2 = 2 \left\{ \sum_{j=1}^J \sum_{k=1}^K d_{jk} \log(d_{jk}/a_{jk}) + (a_{jk} - d_{jk}) \right\} \quad (4)$$

G^2 在 GLIM 中称偏差度(Deviance)。

拟合优度不佳的最常见原因是暴露变量与相对危险度之间关系过于简单,可以通过拟合添加变量(包括交互作用项)的嵌套模型来解决,两个模型的偏差度(G^2)之差即为自由度等于添加变量个数的卡方分布,也可用于检验添加变量的显著性。另有 2 个检验统计量为回归系数检验和效率计分检验。在相加相乘系列模型中,联合作用的结构以及危险度函数是由幂关联 d 来决定的。偏差度衡量数据中不可解释的变异,其值越低,表明拟合数据的模型越适宜,以此作为统计量,选择最佳模型。

2 分析结果

2.1 肺癌相对危险度、超额相对危险度的定量估计

表显示,在相乘模型下,亚硝酸胺暴露每增加一个单位(年),肺癌相对危险度增加约 4%。

相加 RR模型下，每增加一个暴露单位（年），估计吸烟（年数）和经济收入的肺癌危险度定相对危险度增加量约为 0.06 以此类推，还可量水平。

表 1 亚硝酸胺、吸烟及经济收入等暴露的肺癌危险度评价

因素及单位	相乘模型			相加 RR模型		
	U	SE	95% CI	U	SE	95% CI
亚硝酸胺暴露（年数）	0.037	0.010	0.017, 0.057	0.064	0.031	0.002, 0.126
吸烟（年数）	0.051	0.013	0.024, 0.077	0.733	1.002	- 1.231, 2.697
经济收入（元）	- 0.080	0.022	- 0.123, 0.037	- 0.007	0.001	- 0.008, - 0.005

注：控制了年龄和性别的混杂，下同。

2.2 亚硝酸胺暴露、吸烟、经济收入间相互作用分析

表 2进一步考虑了亚硝酸胺暴露与其他 2个非职业暴露因素间的交互作用。在相乘模型

表 2 亚硝酸胺暴露、吸烟和经济收入三因素间交互作用分析

交互作用项	U	SE	95% CI	似然比	P 值（ $\nu=1$ ）
亚硝酸胺暴露与吸烟					
相乘模型	0.002	0.001	0.0002, 0.004	2.880	0.0897
相加 RR模型	0.006	0.003	0.0004, 0.011	5.762	0.0164
亚硝酸胺暴露与经济收入					
相乘模型	- 0.002	0.001	- 0.004, 0.0003	4.689	0.0304
相加 RR模型	- 0.002	0.001	- 0.003, - 0.001	1.142	0.2853
吸烟与经济收入					
相乘模型	0.002	0.001	0.001, 0.003	9.039	0.0026
相加 RR模型	0.0002	0.000	- 0.001, 0.001	2.625	0.1052

△与相应不含交互项模型比较,* $P < 0.05$

下,亚硝酸胺暴露与吸烟交互作用没有统计学显著性,在相加 RR模型下,亚硝酸胺与经济收入的交互作用没有统计学显著性,表明亚硝酸胺暴露与吸烟的联合作用关系倾向于相乘关系,而亚硝酸胺与经济收入则倾向于相加关系。吸烟与经济收入间倾向于相加关系。

应用广义相对危险度模型簇〔式 (3)〕检查上述三者因素间的联合作用关系。图 1表明,亚硝酸胺暴露与吸烟联合作用倾向于相乘模型,但 d值的改变所引起的偏差度改变很小,看来,要区分不同模型的影响,尚需加大样本量。就现结果看,相乘模型更适宜。

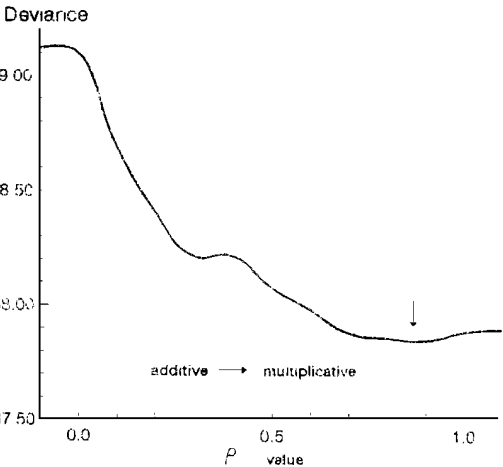


图 1 亚硝酸胺暴露与吸烟联合作用的广义相对危险度模型簇拟合偏差度曲线

图2显示亚硝酸胺暴露与经济收入之间的联合作用关系的比较, d 为0时偏差度最小(238.663), 结果说明可用相加模型来解释。图3显示吸烟与经济收入在 d 小于0.5时偏差度较小, 说明倾向于相加关系。

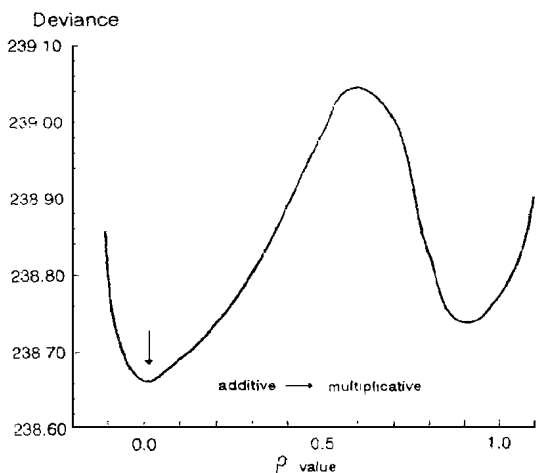


图2 亚硝酸胺暴露与经济收入联合作用的广义相对危险度模型簇拟合偏差度曲线

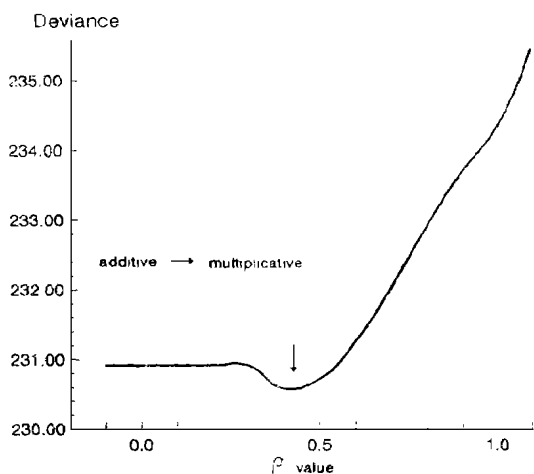


图3 吸烟与经济收入联合作用的广义相对危险度模型簇拟合偏差度曲线

3 讨论

在资料分析中, 混杂作用需要排除或控制, 与之不同, 因素间的效应修饰作用则是需要研究并阐明的。本文对该队列肺癌主要危险因素的单独作用方式和相互作用方式进行了

分析。结果表明, 亚硝酸胺暴露的肺癌危险度适合用相乘也适合用相加危险度模型来描述。这时, 暴露的危险度均呈显著。但如果用偏差度来评价, 相加模型稍优于相乘模型。吸烟的肺癌危险度适合用相乘模型描述。经济收入的肺癌效应仅适合用相加模型描述。

三者对肺癌协同作用的方式为: 亚硝酸胺暴露与吸烟协同作用关系为相乘形式, 偏向于相乘结构; 亚硝酸胺暴露与经济收入的协同作用方式趋向于相加结构; 吸烟与经济收入的协同作用方式为相加结构形式。但总的看来, 不同模型的拟合优度差别均很小, 这可能受到了样本大小的限制。然而, 广义模型簇之间的选择是有很意义的, 在资料分析时均有必要探索因素的单独作用危险度或协同作用危险度的结构形式。不同结构形式的相对危险度模型, 代表着不同的因素间相互作用关系, 这对于公共卫生实践和生物学机理方面的理解和认识具有重要意义^[1]。

4 参考文献

- 1 Breslow NE, Day NE. Statistical methods in cancer research. Vol II. The design and analysis of cohort studies. Lyon: IARC, 1987.
- 2 Berry G. Dose-response in case-control studies. J Epidemiol Community Health, 1980, 34: 217.
- 3 Thomas DC. General relative risk models for survival time and matched case-control analysis. Biometrics, 1981, 37: 673.
- 4 Thompson R, Baker R. Composite link functions in generalized linear models. Appl Stat, 1981, 30: 125.
- 5 Pierce DA, Preston DL, Ishimaru T. A method for analysis of cancer incidence in Japanese A-bomb survivors, with application to acute leukemia. (RERF Technical Report). Hiroshima Radiation Effects Research Foundation, 1985.
- 6 李克, 俞顺章. Poisson模型与队列研究——理论与应用. 中国卫生统计, 1997, 14 (1): 46.
- 7 Baker RJ, Nelder JA. The GLIM system: release 3.77 manual. Oxford: Numerical Algorithms Group, 1985.
- 8 Preston DL, Lubin JH, Pierce DA. Epicure user's guide. Seattle: HiroSoft International Corp, 1992.

(收稿 1997-09-29)