

肿瘤死亡率与年龄、时间和出生队列关系的定量分析

李会庆¹ 金世宽¹ 吴凡¹ 杨晓梅¹ 孙培洪² 石成山² 范万滕²

摘要 肿瘤死亡率与年龄、时间和出生队列之间存在一定的函数关系,用数学公式表达为 $m_{ij} = e^{a_i} \cdot e^{b_j} \cdot e^{c_k}$, 其中 m_{ij} 为第 i 年龄组第 j 时间的死亡率, a_i 为第 i 层年龄, b_j 为第 j 层时间, c_k 为第 k 层出生队列。多数肿瘤死亡率与年龄、时间和出生队列呈对数线性关系,运算中将依变量 y 和自变量 a 、 b 、 c 都取自然对数,用线性回归的方法求参数 α 、 $\beta_{(a)}$ 、 $\beta_{(b)}$ 、 $\beta_{(c)}$,用 EXP (β) 评价各个参数的作用大小。作者以一个监测点1970~1989年宫颈癌和胃癌的资料进行了上述分析,结果与观察的危险因素变动情况相符合。

关键词 肿瘤 死亡率 出生队列定量分析

An Analysis of Relationship between Mortality of Cancer and Age—Period—Cohort Li Hui-qing, Jin Shi-kuan, Wu Fan, et al. Shandong Academy of Medical Sciences, Jinan, 250001

The APC (age—period—cohort analysis) model makes the assumption that the mortality m_{ij} in a given age—group and year is the (simple) product of three factors: an age—related factor a_i , one period—related factor b_j and one (birth) cohort—related factor C_k : $m_{ij} = e^{a_i} \cdot e^{b_j} \cdot e^{c_k}$.

By taking the natural logarithm, a linear model results which can be treated with fairly standard statistics techniques: $\ln(m_{ij}) = a_i + b_j + c_k$. The m_{ij} data from observation data are used to estimate α 、 $\beta_{(a)}$ 、 $\beta_{(b)}$ 、 $\beta_{(c)}$ by linear regression method. The value of EXP (β) is to estimate the effect of each variable. The APC model is consistent with carcinogenesis modeling of molecular biology on tumour. The results of APC analysis for cervical cancer, male and female stomach cancer show that the risk factors have been changed. In fact, the risk factors of cervical cancer have been decreased since liberation, and the risk factor of stomach was increased in the first period of sixty's. The analytic method used in this paper will benefit the study on epidemiology and etiology of cancer.

Key words Cancer Mortality Age—period—cohort analysis

慢性非传染性疾病或死亡率受年龄、时间和出生队列的影响,对这些影响因素的分析称为出生队列分析(birth cohort analysis)或称 APC 分析(age—period—cohort analysis)。笔者依据栖霞县一个监测点20多年的肿瘤死亡资料进行了出生队列与死亡率之间关系的探讨,报告如下。

资料与方法

一、资料来源:以山东省栖霞县1970~1989年女性宫颈癌、男女胃癌死亡率数据为资料,分1970~1974、1975~1979、1980~1984、1985~1989四个时期整理年龄别死亡

1 山东省医学科学院基础医学研究所 250001
济南

2 山东省栖霞县卫生局

表1 栖霞县1970~1989年宫颈癌死亡率 (/10万)

年龄组 (岁)	不同时期 (年) 观察值			
	1970~ 1974	1975~ 1979	1980~ 1984	1985~ 1989
15~	0.00	0.00	0.00	0.59
20~	0.70	0.00	0.00	0.00
25~	0.00	1.46	0.71	0.70
30~	1.22	1.60	0.78	0.77
35~	3.89	3.95	3.84	0.95
40~	18.59	15.19	3.69	6.07
45~	18.60	15.16	6.70	2.64
50~	70.13	17.52	17.04	9.80
55~	71.27	18.07	20.76	11.02
60~	78.78	71.26	40.42	13.28
65~	71.02	60.57	63.43	40.17
70~	100.60	36.40	52.33	38.02
75~	31.90	58.52	82.59	38.36
80~	11.45	97.94	75.88	27.07
合计	13.00	10.22	8.57	4.90

率 (表1和表2)。其中1970~1974年为全国统一组织调查资料; 1975~1979年为全国肿瘤流行病学专家在该县综合考察的调查资料; 1980~1984年为山东省医科院省75科研课题胃癌高发现场的调查资料; 1985~1989年和1990~1992年为全省两次全死因死亡回顾性调查资料。历次调查都是全死因调查, 其中漏报率、老衰、死因不明所占百分比、婴儿死亡率和总死亡率等质控措施得到当时专家的认可。恶性肿瘤的诊断等级水平构成与全省同期的水平相接近, 故可认为引用资料是可靠的。

二、模拟和参数估计方法: 肿瘤死亡率可看做是年龄、时间和出生队列三者作用的幂函数之积, 用公式 (1) 表示^[1]:

$$m_{ij} = e^{a_i} \cdot e^{b_j} \cdot e^{c_k} \quad (1)$$

表2 栖霞县1970~1989年男性和女性胃癌死亡率 (/10万)

年龄组 (岁)	各时期 (年) 观察值 (男)				各时期 (年) 观察值 (女)			
	1970~1974	1975~1979	1980~1984	1985~1989	1970~1974	1975~1979	1980~1984	1985~1989
15~	0.00	0.58	0.00	0.00	0.57	0.00	0.67	0.00
20~	2.68	0.00	0.81	0.00	0.68	2.79	0.80	0.70
25~	2.84	2.27	0.86	2.76	0.75	0.33	4.43	2.79
30~	8.13	2.89	7.66	4.42	5.64	23.28	5.53	4.59
35~	15.82	19.26	24.62	9.74	11.90	9.82	14.00	12.27
40~	28.37	39.04	42.95	14.06	25.57	19.79	13.54	14.53
45~	54.26	78.52	76.64	37.18	26.32	48.89	46.44	30.32
50~	131.21	124.25	105.91	104.75	54.25	40.85	48.30	54.49
55~	273.02	219.66	201.13	200.84	112.21	63.49	95.80	73.86
60~	338.51	371.81	435.34	290.91	126.43	109.75	162.90	98.47
65~	437.97	493.39	673.55	414.44	145.64	170.65	345.55	164.87
70~	481.82	517.23	812.31	578.67	173.97	147.19	241.96	281.55
75~	228.47	486.80	643.09	619.69	80.39	282.03	409.81	234.03
≥80	139.07	471.44	753.91	575.87	35.00	171.51	295.10	290.48
合计	57.41	63.14	78.60	61.08	23.61	26.97	38.31	29.85

式中 m_{ij} 为第 i 年龄组第 j 时间的死亡率, a_i 为第 i 层年龄, b_j 为第 j 层时间, c_k 为第 k 层出生队列, 出生队列对应的死亡率数值为对

角线上的数值 (表1和表2)。

对 (1) 式两边取对数, 得

$$\ln(m_{ij}) = a_i + b_j + c_k \quad (2)$$

(2) 式用一般线性回归方程式表示为

$$y_{ij} = a + \beta a_i + \beta b_j + \beta c_k + \epsilon \quad (3)$$

多数肿瘤的死亡率与年龄、时间和出生队列呈对数线性关系,运算中将依变量 y 和自变量 a 、 b 、 c 都取自然对数,用线性回归的方法求参数 α 、 $\beta_{(a)}$ 、 $\beta_{(b)}$ 和 $\beta_{(c)}$,用 $EXP(\beta)$ 评价各个参数的作用大小。

三、年龄、时间和出生队列的数值量化:年龄变量值为各相应年龄组的组中值,如15~岁年龄组赋值为17.5,小于15~岁组无肿瘤死亡则不参入运算。出生队列变量的赋值方法为:以1970~1974年时17.5岁组的出生队列值赋值为1,22.5岁组为2, ..., 82.5岁则为14;在1975~1979年22.5岁组是由1970~1974年时17.5岁组转变来的,故该组赋值为1,此后出生队列号顺延;同理,在1985~1989年27.5岁组赋值为1,1990~1992年32.5岁组赋值为1。在出生队列序号下延过程中,小年龄组的出生队列序号全赋值1,因在回归模型中所有变量取对数后参入运算,1的对数为0,即癌死亡率很低的队列的作用可忽略。时间变量赋予相应时间间距的组中值,如1970~1974年取 $(0+4)/2=2$,1975~1979年取值为 $(5+9)/2=7$,1980~1984年取值为 $(80-70) + (0+4)/2=12$,1985~1989年取值为17。

结 果

一、宫颈癌死亡率的 APC 分析:从表1可以看出,1970~1974年35~至60~岁年龄组相应的各出生队列宫颈癌死亡率在后三个时期都下降。用1970~1984年的数据分析 $EXP(\beta_a)$ 为3.34, $EXP(\beta_b)$ 为1.10, $EXP(\beta_c)$ 为5.16;用1970~1989年的数据分析三者分别为2.78、1.02和4.73,显著性检验只有 $EXP(\beta_c)$ 有显著性意义。

二、男性胃癌死亡率的 APC 分析:从表2可以看出,1970~1974年30~至65~岁年龄组相应的各出生队列死亡率在1975~1979年和1980~1984年呈上升趋势,到1985~1989

年则开始有所下降。用1970~1984年的数据分析 $EXP(\beta_a)$ 为2.72, $EXP(\beta_b)$ 为1.79, $EXP(\beta_c)$ 为10.55;用1970~1989年数据分析 $EXP(\beta_a)$ 为8.15, $EXP(\beta_b)$ 为1.55, $EXP(\beta_c)$ 为5.31。三个参数都有显著性意义。出生队列的作用后一时期有所下降,而年龄的作用有所上升。

三、女性胃癌死亡率的 APC 分析:从表2还可看出,1970~1974年30~至65~岁年龄组各出生队列在后三个时期中变化不大。用1974~1984年的数据分析 $EXP(\beta_a)$ 为31.77, $EXP(\beta_b)$ 为1.43, $EXP(\beta_c)$ 为1.52;用1970~1989年的数据分析三者的大小分别为26.35、1.37和1.74,显著性检验年龄因素有非常显著性意义,时间因素有显著性意义,出生队列因素没有显著性意义。

讨 论

本文分析结果表明,宫颈癌死亡率与出生队列的作用非常显著,与已报道的宫颈癌死亡率下降与解放初消灭性病、控制性混乱等有关^[2],即解放后妇女暴露于宫颈癌的危险因素程度下降。

栖霞县为胃癌高发区,胃癌死亡男女性别比为2:1。男性出生队列的作用明显。据报道该地三年自然灾害期间(1959~1961)居民饮食质量差,食入霉变食品,蛋白质和维生素类防癌物摄入量少,当时易感性的出生队列组易产生癌基因突变,随时间推移,癌发病死亡增加,而到80年代末这种作用开始下降。女性出生队列的作用不明显,可能男性吸烟、饮酒等行为有助于致癌物的致癌作用^[3]。

该数学模拟回归模型出自 Gompertz 模型,适用于多数发病率或死亡率与年龄呈对数线性关系的肿瘤,只是增加了出生队列参数,而出生队列参数可以看作是年龄和时间两因素的交互作用项增加于模型之中^[1]。该模型与肿瘤细胞的生长机理和分子流行病学机理相符合^[4]。

参 考 文 献

1 Kupper LL, Janis JM, Karmous A, et al. Statistics age-period cohort analysis: a review and critique. J Chron Dis, 1985 38 (10): 811, 830.

2 李会庆, 金世宽, 胡烈民, 等. 山东省宫颈癌死亡率下降的流行因素研究. 中华流行病学杂志, 1992, 13 (4):

202.

3 李会庆. 山东省胃癌高低发区胃癌死亡率变化趋势的出生队列分析. 中华流行病学杂志, 1990, 11 (6): 337.

4 Moolgavkar SH. Carcinogenesis modeling: from molecular biology to epidemiology. Ann Rev Public Health. 1986, 7: 151.

(收稿: 1994-11-24 修回: 1995-1-12)

黑龙江省1949~1988年流行性脑脊髓膜炎疫情分析

刘盛庭¹ 许清凯¹ 安 辉²

对黑龙江省1949~1988年40年的流行性脑脊髓膜炎(简称流脑)疫情资料的统计分析,本省流脑疫情的消长大致可分为三个阶段:第一阶段(1949~1964年)呈散发状态;第二阶段(1965~1982年)呈流行状态;第三阶段(1983~1988年)由于普遍接种了流脑菌苗使之呈稳步下降状态,但病死率与发病率呈反比关系,即发病率高的年代病死率较低,而发病率较低的年代病死率反而较高。流脑在本省也有周期性、季节性及地区性流行的特点。每8~10年为一个流行周期,且流行期、下降年、稳定年、上升年均均为2~3年左右,但广泛接种流脑菌苗后可改变这一规律。流行期(45%~71%)在2~4月,流行前期(6.0%~15%)为11~1月,流行后期(13%~31%)为5~7月,间歇期(1.0%~11%)为8~10月。这一特点在流行年和非流行年无显著性差异。流行强度大的地区多在佳木斯市、牡丹江市、鸡西市等本省东部,即在东经129°~134°之间;流行强度居中的地区多在伊春市、大庆市、绥化地区、哈尔滨市等本省中部,即在东经126°~130°之间;流行强度小的地区多在大兴安岭地区、黑河地区、齐齐哈尔市等本省西

部,即在东经122°~126°之间。

我省流脑发病历年来皆以儿童为主,特别是婴幼儿。据对全省1983~1988年的15273例流脑病例分析,1岁以内婴儿发病最多(14.11%),并随年龄增长而递减。3岁以下婴儿占总例数的39.49%,8岁以下儿童占67.92%,15岁以下人群占86.61%,15岁以上人群仅占13.39%。所以说流脑防治的重点人群是15岁以下的人群,特别是婴幼儿。

我们曾对293名流脑患者或密切接触者进行过病原学检查,检出65株流脑杆菌,其中A群64株(占98.46%),可见引起我省流脑流行的菌群属于A群。但健康人群带菌菌群则以B群为主,占55.34%(1756/3173),次为A群(11.28%)和C群(9.2%)。这一点应引起我们的注意,进行必要的观察研究。

(收稿: 1994-10-17 修回: 1994-11-08)

1 哈尔滨市道里区中医院 150016

2 黑龙江省卫生防疫站