系列讲座

现代实用流行病学方法

第二讲　病例对照研究（续）

章扬熙

成组（非配比）的病例对照研究统计推断

一、不分层资料：病例对照研究中所比较的是病例组和对照组中总体暴露率的差异。由于病例对照研究系抽样研究，样本暴露率存在抽样误差，所以，首先要进行χ^2检验，检验假设 Ho 为$\pi_1=\pi_2$，即两组的总体暴露率相等，设检验水准 $\alpha=0.05$。如果检验结果接受Ho，$P>0.05$，则说明该因素不是与发病有关的因素，无需求比数比OR值。如果检验结果拒绝Ho，$P\leqslant 0.05$，则说明该因素是与发病有关的因素（危险因素或保护因素），所以要进一步计算样本比数比OR值及总体比数比的100（1－α）%可信区间。对有多暴露水平的成组资料，还可进而作剂量反应分析，各水平均可与非暴露作χ^2检验，求OR值及总体OR值的100（1－α）%可信区间，作线性趋势检验并求出其斜率。有关的计算公式列于表1中。

表1　不分层资料的表式、统计指标及统计推断

2×2表

表格形式

	暴露	未暴露	
病例	a	c	M_1
对照	b	d	M_0
	N_1	N_0	T

假设检验

χ^2检验

$$\chi^2_{MH}=\frac{[a-E(a)]^2}{V(a)} \tag{7}$$

$$E(a)=\frac{N_1M_1}{T} \tag{8}$$

$$V(a)=\frac{N_1N_0M_1M_0}{T^2(T-1)} \tag{9}$$

df＝1

统计指标或与其100(1－α)%可信区间

$$OR=\frac{ad}{bc} \tag{10}$$

Miettinen法

$$OR^{(1\pm U_\alpha/\chi)} \tag{11}$$

Woolf法

$$\ln^{-1}[\ln OR\pm U_\alpha\sqrt{V(\ln OR)}] \tag{12}$$

$$V(\ln OR)=\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d} \tag{13}$$

2×k表

	暴露水平				
	Y_0	Y_1	$Y_2\cdots$	$Y_j\cdots$	
病例	$a_0(=c)$	a_1	$a_2\cdots$	$a_j\cdots$	M_1
对照	$b_0(=d)$	b_1	$b_2\cdots$	$b_j\cdots$	M_0
	N_0	N_1	$N_2\cdots$	$N_j\cdots$	T

线性趋势检验

$$\chi=\frac{\Sigma aY-\frac{M_1}{T}\Sigma NY}{\left\{\frac{M_1M_0}{T^2(T-1)}\left[T\Sigma NY^2-(\Sigma NY)^2\right]\right\}^{\frac{1}{2}}} \tag{14}$$

$$b=\frac{\Sigma aY-\frac{M_1}{T}\Sigma NY}{\Sigma NY^2-\frac{(\Sigma NY)^2}{T}} \tag{15}$$

〔注：公式10、13中若a、b、c、d中有零值，计算OR及V（lnOR）时，a、b、c、d各值均加0.5〕

〔例4〕　应用成组的病例对照研究来考察吸烟史与肺癌的关系，病例组108人中68人吸烟，对照组

本文作者单位：辽宁省卫生防疫站　110005　沈阳市

108人中49人吸烟（表2），问两组吸烟率有无差别？若有差别进而求比数比和总体比数比的95%可信限。

表2　吸烟与肺癌关系的病例对照研究

	吸烟	不吸烟	合计
病例	68(a)	40(c)	108(M_1)
对照	49(b)	59(d)	108(M_0)
合计	117(N_1)	99(N_0)	216(T)

1.检验假设：病例组与对照组的总体吸烟率相等，$\alpha=0.05$。

2.应用Mantel-Haenszel法求χ^2值：首先用公式8求a的理论值E（a），再用公式9求其方差V（a），最后用公式7求χ^2值。

$$E(a)=\frac{108\times117}{216}=58.50$$

$$V(a)=\frac{108\times108\times117\times99}{216\times216\times215}=13.47$$

$$\chi^2=\frac{(68-58.5)^2}{13.47}=6.70$$

3.求χ^2的界值：自由度df＝（行－1）（列－1）＝（2－1）（2－1）＝1，查χ^2值表知，当df＝1时，$\chi^2_{0.05}=3.84$，$\chi^2_{0.01}=6.63$。

4.统计推断：本例$\chi^2=6.70$，若检验假设Ho成立，$\pi_1=\pi_2$，亦即总体比数比$\psi=1$，得到χ^2值为6.70或比6.70还大的概率$P<0.01$，表式为$Pr(\chi^2\geqslant 6.70\mid Ho)<0.01$，按$\alpha=0.05$水准拒绝检验假设，接受备择假设Ha，说明两组吸烟率有差别，即病例组吸烟率较对照组为高。应当说明，用公式7计算得χ^2值比一般统计书上χ^2公式所得的结果略小，为其χ^2值乘以（T－1）/T。

在病例对照研究中，用比数比（Odds Ratio）对危险因素的危险程度进行估计，本例用公式10计算OR，得：

$$OR=\frac{68\times59}{49\times40}=2.05$$

这个结果说明吸烟者患肺癌的危险度为不吸烟者的2.05倍。

以上求得的是样本的比数比OR，总体比数比ψ的可信限，常用Miettinen法或Woolf法求得。本例100（1－α）＝95，$\alpha=0.05$，查U值表（或t值表中df＝∞时的值），$U_{0.05}=1.96$，应用Miettinen法公式11得ψ的95%可信限为：

$$2.05^{(1\pm1.96/\sqrt{6.70})}=1.19\sim3.53$$

或用Woolf法，先用公式13求lnOR的方差得：

$$V(\ln OR)=\frac{1}{68}+\frac{1}{49}+\frac{1}{40}+\frac{1}{59}=0.077063$$

用公式12求ψ的95%可信限为：

$$\ln^{-1}[\ln2.05\pm1.96\times\sqrt{0.077063}]=1.19\sim3.53$$

两法所得结果相同。这结果说明总体比数比在1.19～3.53范围内的概率为95%。

把人群定性地分为“暴露”和“未暴露”两种情况是初步评价暴露于某因素与研究疾病之间联系的方法。当发现有联系后，还要进一步考察暴露剂量（水平）与反应的关系，做2×k表的分析。

〔例5〕对例4资料按不同吸烟量分等级统计，得表3，试进行剂量反应分析。

表3　不同吸烟量与肺癌的关系

组段	每日吸烟量（支）					合计
	0	1～	5～	15～	25～	
组中值	0(Y_0)	3(Y_1)	10(Y_2)	20(Y_3)	30(Y_4)	
病例	40(a_0=c)	14(a_1)	30(a_2)	12(a3)	12(a_4)	108(M_1)
对照	59(b_0=d)	18(b_1)	22(b_2)	8(b_3)	1(b_4)	108(M_0)
合计	99(N_0)	32(N_1)	52(N_2)	20(N_3)	13(N_4)	216(T)
OR（相对于不吸烟）	1.00	1.15	2.01	2.21	17.70	

为了检验剂量反应关系的线性趋势，应用公式14，进行χ^2检验。

1.检验假设：总体回归系数$\beta=0$，即无剂量反应关系。设$\alpha=0.05$。

2.将有关数值代入公式14，得：

$$\chi=\frac{\left(40\times0+\cdots+12\times30-\frac{108}{216}(99\times0+\cdots+13\times30)\right)}{\left\{\frac{108\times108}{216^2\times215}\left[216(99\times0^2+\cdots+13\times30^2)-(99\times0+\cdots+13\times30)^2\right]\right\}^{1/2}}=3.77$$

3.求χ的界值：χ分布即U分布，查U值表，$U_{0.05}=1.96$，故$Pr(\chi\geqslant3.77\mid H_0)<0.05$，按检验水准$\alpha=0.05$拒绝检验假设，说明吸烟量越大，发生肺癌越多，两者存在线性剂量反应关系。可进而应用公式15求出回归系数b，得：

$$b=\frac{(40\times0+\cdots+12\times30)-\frac{108}{216}(99\times0+\cdots+13\times30)}{(99\times0^2+\cdots+13\times30^2)-\frac{(99\times0+\cdots+13\times30)^2}{216}}=0.0149$$

然后，再用公式10计算不同吸烟量者与不吸烟者之比数比，结果列于表3的第末行。比如，每日吸5～14支烟者，应用公式10得：

$$OR=\frac{30\times59}{22\times40}=2.01,\ \text{余类推。}$$

从各OR值可明显看出，吸烟量越大患肺癌的危险度也越大。

病例对照研究分析的重点是估计暴露与疾病的联系及其强度。但是，对于公共卫生问题而言，还要研究暴露对人群的影响程度或减少暴露对人群的社会效益。为此，还要计算特异危险度百分比（attributable risk proportion，简称AR%）和人群特异危险度百分比(population attributable risk proportion，简称PAR%）。这两个指标原为定群研究的统计指标，AR%表示在暴露者中由于暴露于某因素所致发病占暴露者发病的百分比，PAR%表示人群中由于暴露在某因素而致发病占人群发病的百分比，计算公式分别为：

$$AR\%=\frac{Ie-Iu}{Ie}\times100\% \qquad (16)$$

由于RR＝Ie/Iu，上式可导出：

$$AR\%=\frac{RR-1}{RR}\times100\% \qquad (17)$$

在病例对照研究中，用OR来估计RR，故：

$$AR\%=\frac{OR-1}{OR}\times100\% \qquad (18)$$

$$PAR\%=\frac{It-Iu}{It}\times100\% \qquad (19)$$

由于It＝PeIe＋（1－Pe）Iu，RR＝Ie/Iu，从上式可导出：

$$PAR\%=\frac{Pe(RR-1)}{Pe(RR-1)+1}\times100\% \qquad (20)$$

在病例对照研究中，用OR来估计RR，故：

$$PAR\%=\frac{Pe(OR-1)}{Pe(OR-1)+1}\times100\% \qquad (21)$$

上式中Pe为人群中的暴露率，未知时可用对照组的暴露率来代替。由于PAR%等于病例组的暴露率Pc与AR%的乘积，故PAR%也可用下式计算：

$$PAR\%=Pc\cdot\frac{OR-1}{OR}\times100\% \qquad (22)$$

诸上式中，Ie为暴露组的发病率，Iu为未暴露组的发病率，It为人群的发病率，RR为相对危险度。如果PAR%的计算公式不乘以100%，则称为病因分值（etiologic fraction，简称EF)；同样，AR%的计算公式不乘以100%，则称为暴露人群的病因分值（EFe）。应当说明，在病例对照研究中，计算PAR%的条件是病例组为目标人群中全部新发病例或为其一个随机样本，对照组为目标人群中未发病者的一个随机样本。

总体AR%的100（1－α）%可信区间可先求出OR的可信区间上限值和下限值代入公式18得到。总体PAR%的100（1－α）%可信区间用下式计算：

$$PAR\%\pm U\alpha\sqrt{V(PAR\%)} \qquad (23)$$

式中V（PAR%）是PAR%的方差，为：

$$V(PAR\%)=\left(\frac{cM_0}{dM_1}\right)^2\left(\frac{a}{cM_1}+\frac{b}{dM_0}\right) \qquad (24)$$

〔例6〕 若例4资料符合计算PAR%的条件，求AR%、PAR%及其95%可信区间。

从例4知，OR＝2.05，OR的95%可信区间下限值为1.19，上限值为3.53，将有关数值代入公式18，得：

$$AR\%=\frac{2.05-1}{2.05}\times 100\%=51.22\%$$

这说明在吸烟的人群中，由于吸烟导致肺癌的占肺癌患者总数的51.22%。

再将有关上下限数值代入公式18，得AR%的95%可信区间为：

$$\frac{1.19-1}{1.19}\times 100\%\sim\frac{3.53-1}{3.53}\times 100\%=15.97\%\sim 71.67\%$$

$Pe=\frac{b}{Mo}=\frac{49}{108}=0.4537$，代入公式21得：

$$PAR\%=\frac{0.4537\times(2.05-1)}{0.4537\times(2.05-1)+1}\times 100\%$$

$$=32.27\%$$

或者，$P_c=\frac{a}{M_1}=\frac{68}{108}=0.62963$，代入公式22得：

$$PAR\%=\frac{0.62963\times(2.05-1)}{2.05}\times 100\%$$

$$=32.25\%$$

两公式的计算结果一致，尾数差是由于取位数时四舍五入造成的。这结果说明，在总人群的患肺癌者中有32.25%是由于吸烟导致患病的。

将有关数值代入公式24得

$$V(PAR\%)=\left(\frac{40\times 108}{59\times 108}\right)^2\left(\frac{68}{40\times 108}+\frac{49}{59\times 108}\right)=0.01077$$

查U值表，$U\alpha=1.96$，将有关数值代入公式23得PAR%的95%可信区间为：

$$0.3225\pm 1.96\times\sqrt{0.01077}=0.1191\sim 0.5259$$
$$=11.91\%\sim 52.59\%$$

另外，Taylor（1977）证明，PAR%还可用简单的公式求出

$$PAR\%=1-\frac{cM_0}{dM_1}\qquad(25)$$

读者可自行计算，结果与上法一致。

二、分层资料：分层分析的目的在于把混杂因素的影响排除。通常需要对年龄、性别、种族等可能混杂因素加以考虑，按可能混杂因素分层归纳成多个2×2表进行分析，各层的2×2表分析方法同简单2×2表，还可进而计算分层资料的总χ^2值。如果各层的比数比经非均匀性χ^2检验认为基本一致，可求总比数比及其可信区间，否则计算SMR及其可信区间。若暴露水平为计量资料或等级资料，还可进一步考察暴露剂量与反应的关系，做趋势性扩展的χ检验，当说明有剂量反应关系后，可进而求回归系数及SRR。以上的有关指标、假设检验及总体参数的可信区间计算公式详见表4。

〔例7〕 用病例对照研究来考察肺癌与接触毒物的关系，结果如表5。由于吸烟为可能混杂因素，遂把表5按吸烟分层，得到两个2×2表（表6、7），试求OR并给予解释。

不分层资料求得粗比数比OR＝22.94，但按吸烟分层后，吸烟层的调整比数比$OR_1=1$，不吸烟层$OR_2=1$，说明肺癌与接触毒物无联系，而不分层的OR＝22.94是一个假象，是由吸烟这个混杂因素造成的。

〔例8〕 Doll和Hill应用病例对照研究吸烟与肺癌的关系，结果如表8，由于不同性别人群的暴露率和暴露水平均有不同，性别可能是混杂因素，对该资料按性别分层整理，结果得表9，试进行分析。

从χ^2检验结果知$\chi^2_{MH}=19.12$，$P<0.001$，两组吸烟率有差别，吸烟为肺癌的危险因素。为了消除性别的影响，应用公式26进行分层资料汇总χ^2_{MH}检验，得：

$$\chi^2_{MH}=\frac{(688-669)^2}{14.4864}=24.92$$

$P<0.001$，说明消除了性别的影响，吸烟仍为肺癌的危险因素，本例各层的比数比是否一致，可进行非均匀性卡方检验。检验假设为各层总体比数比相等，设$\alpha=0.05$。首先要应用公式39计算总OR，得

总$OR=\frac{24.39}{5.39}=4.53$，然后应用迭代法计算总比数比的渐近极大似然估计$OR_{ML}$，使求得的$OR_{ML}$满足公式34、35的条件，可用总OR值作为$OR_{ML}$迭代的初值。公式34、35中的E(ai)是$OR_{ML}$为其比数比时ai的期望值。本例得$OR_{ML}=4.25$，将此值与表9中男性2×2表中有关值代入公式35，得：

表4　分层资料的表式、统计指标及统计推断

表格形式

分层2×2表

第i层

	暴露	未暴露	
病例	a_i	c_i	M_{1i}
对照	b_i	d_i	M_{0i}
	N_{1i}	N_{0i}	T_i

$i=1,2,\cdots,g$

分层2×k表

第i层

暴露水平

Y	Y_0	$Y_1 \cdots Y_j \cdots Y_k$	
病例	$a_{i0}(=c_i)$	$a_{i1} \cdots a_{ij} \cdots a_{ik}$	M_{1i}
对照	$b_{i0}(=d_i)$	$b_{i1} \cdots b_{ij} \cdots b_{ik}$	M_{0i}
	N_{i0}	$N_{i1} \cdots N_{ij} \cdots N_{ik}$	T_i

$i=1,2,\cdots,g \qquad j=1,2,\cdots,k$

假设检验

分层2×2表

χ^2检验

$$\chi^2_{MH}=\frac{[\Sigma a_i-\Sigma E(a_i)]^2}{\Sigma V(a_i)} \quad (26)$$

df = 1

$$E(a_i)=\frac{N_{1i}M_{1i}}{T_i} \quad (27)$$

$$V(a_i)=\frac{N_{1i}N_{0i}M_{1i}M_{0i}}{T_i^2(T_i-1)} \quad (28)$$

非均匀性χ^2检验

$$\chi^2_{HET}=-2\left\{\Sigma\left[a_i\ln\left(\frac{E(a_i)}{a_i}\right)+b_i\ln\left(\frac{E(b_i)}{b_i}\right)+c_i\ln\left(\frac{E(c_i)}{c_i}\right)+d_i\ln\left(\frac{E(d_i)}{d_i}\right)\right]\right\} \quad (33)$$

$$\Sigma a_i-\Sigma E(a_i)=0 \quad (34)$$

$$OR_{ML}=\frac{E(a_i)[M_{0i}-N_{1i}+E(a_i)]}{[M_{1i}-E(a_i)][N_{1i}-E(a_i)]} \quad (35)$$

分层2×k表

趋势性扩展的χ检验

$$\chi_{MEXT}=\frac{\sum\limits_i\left(\sum\limits_j a_{ij}Y_j-\frac{M_{1i}}{T_i}\Sigma N_{ij}Y_j\right)}{\left\{\sum\limits_i\frac{M_{1i}M_{0i}}{T_i^2(T_i-1)}\left[T_i\sum\limits_i N_{ij}Y_j^2-\left(\sum\limits_i N_{ij}Y_j\right)^2\right]\right\}^{\frac{1}{2}}} \quad (40)$$

统计指标或与其100（1－α）%可信区间

分层2×2表

$$OR_i=\frac{a_id_i}{b_ic_i} \quad (29)$$

Miettinen法

$$OR_i^{(1\pm U_\alpha/\chi)} \quad (30)$$

Woolf法

$$\ln^{-1}[\ln OR_i\pm U_\alpha\sqrt{V(\ln OR_i)}] \quad (31)$$

$$V(\ln OR_i)=\frac{1}{a_i}+\frac{1}{b_i}+\frac{1}{c_i}+\frac{1}{d_i} \quad (32)$$

$$SMR=\Sigma a_i\Big/\left(\Sigma\frac{b_ic_i}{d_i}\right) \quad (36)$$

$$SMR\pm U_\alpha\sqrt{V(SMR)} \quad (37)$$

$$V(SMR)=\sqrt{\Sigma a_i}\Big/\left(\frac{b_ic_i}{d_i}\right) \quad (38)$$

$$\text{总}OR=\left(\frac{a_id_i}{T_i}\right)\Big/\left(\Sigma\frac{b_ic_i}{T_i}\right) \quad (39)$$

总OR可信区间计算公式同公式30

分层2×k表

$$b_Y=\frac{\sum\limits_i\left(\sum\limits_j a_{ij}Y_j-\frac{M_{1i}}{T_i}\Sigma N_{ij}Y_j\right)}{\sum\limits_i\left[\sum\limits_j N_{ij}Y_j^2-\frac{(\Sigma N_{ij}Y_j)^2}{T_i}\right]} \quad (41)$$

$$SRR_j=\left(\sum_i\frac{a_{ij}d_i}{b_{ij}}\right)\Big/\left(\sum_i c_i\right) \quad (42)$$

$$\left[SRR_j^{-1}\pm U_\alpha\sqrt{V(SRR_j^{-1})}\right]^{-1} \quad (43)$$

$$V(SRR_j^{-1})=\left(\sqrt{\sum_i c_i}\right)\Big/\left(\sum_i\frac{a_{ij}d_i}{b_{ij}}\right) \quad (44)$$

表5　接触毒物与肺癌关系的病例对照研究

	接触毒物	不接触毒物	合计
病例	91	19	110
对照	19	91	110
合计	110	110	

$\chi^2_{MH}=93.83$　　$P<0.001$

$OR=(91\times91)/(19\times19)=22.94$

表6　吸烟者接触毒物与肺癌的关系

	接触毒物	不接触毒物	合计
病例	90	9	99
对照	10	1	11
合计	100	10	110

$\chi^2_{MH}=0$　　$P>0.05$

$OR_1=(90\times1)/(9\times10)=1$

表7　不吸烟者接触毒物与肺癌的关系

	接触毒物	不接触毒物	合计
病例	1	10	11
对照	9	90	99
合计	10	100	110

$\chi^2_{MH}=0$　　$P>0.05$

$OR_2=(1\times90)/(9\times10)=1$

表8　吸烟与肺癌关系的病例对照研究

	吸烟	不吸烟	合计
病例	688	21	709
对照	650	59	709
合计	1338	80	1418

$\chi^2_{MH}=19.12$　　$P<0.001$

$OR=2.97$

表9　不同性别肺癌病例与对照吸烟史的比较

层别		吸烟	不吸烟	合计	a	E(a)	V(a)	OR	$\frac{ad}{T}$	$\frac{bc}{T}$	$\frac{bc}{d}$
男	病例	647(643.4)	2(5.6)	649	647	634.5	7.0935	14.04	13.46	0.96	46.07
	对照	622(625.6)	27(23.4)	649	($\chi^2_{MH}=22.03$	$P<0.001$)					
i=1	合计	1269	29	1298							
女	病例	41(44.6)	19(15.4)	60	41	34.5	7.3929	2.47	10.93	4.43	16.63
	对照	28(24.4)	32(35.6)	60	($\chi^2_{MH}=5.71$	$P<0.05$)					
i=2	合计	69	51	120							
合　计					688	669	14.4864	—	24.39	5.39	62.70

$$\frac{E(a_1)[649-1269+E(a_1)]}{[1269-E(a_1)][649-E(a_1)]}=4.25,$$

解$E(a_1)=643.4$，同理，将$OR_{ML}=4.25$值与表9中女性2×2表中有关值代入公式35，得

$$\frac{E(a_2)[60-69+E(a_2)]}{[69-E(a_2)][60-E(a_2)]}=4.25,$$

解$E(a_2)=44.6$，将$E(a_1)$、$E(a_2)$值代入公式34得$647+41-(643.4+44.6)=0$，满足了公式34、35的条件。于是，得到$E(a_1)$、$E(b_1)$、$E(c_1)$、$E(d_1)$、$E(a_2)$、$E(b_2)$、$E(c_2)$、$E(d_2)$各值，列于表9各实际值旁的括号中。再将有关值代入公式33得：

$$\chi^2_{HET}=-2\left\{\left[647\ln\frac{643.4}{647}+\cdots+27\ln\frac{23.4}{27}\right]+\left[41\ln\frac{44.6}{41}+\cdots+32\ln\frac{35.6}{32}\right]\right\}=5.62$$

自由度$df=g-1=2-1=1$，查χ^2值表，$\chi^2_{0.05}=3.84$，故$P<0.05$，拒绝检验假设，说明男、女的总体比数比不等。由于各层比数比不一致，不宜求总OR，而用各层比数比或标化发病率比SMR来分析，

应用公式36得$SMR=\frac{688}{62.70}=10.97$，总体SMR的

95%可信区间应用公式37、38得$V(SMR)=\frac{\sqrt{688}}{62.70}$

$=0.4183$，$10.97\pm1.96\times\sqrt{0.4183}=10.15\sim11.79$，总体SMR的95%可信区间为10.15～11.79。

若各层比数比一致，可求总OR及其95%可信区间，以本例数据练习计算方法得总OR＝4.53，总OR 95%可信区间应用公式30得：

$$4.53^{(1\pm1.96/\sqrt{24.92})}=2.50\sim8.20$$

同样地可应用公式45求PAR%，应用公式46、47求PAR%的100（1－α）%可信区间。

$$PAR\%=1-\left[\Sigma M_{0i}\frac{c_i}{d_i\Sigma M_{1i}}\right] \quad (45)$$

本例得：

$$PAR\%=1-\left[\frac{649\times2}{27\times709}+\frac{60\times19}{32\times709}\right]=0.8819$$

$$V(PAR\%)=\frac{1}{(\Sigma M_{1i})^2}\left\{\left|\Sigma\left(\frac{M_{0i}c_i}{d_i}\right)^2\left(\frac{1}{d_i}+\frac{b_i}{M_{0i}d_i}\right)-(\Sigma M_{1i})(1-PAR\%)^2\right|\right\} \quad (46)$$

总体PAR%的100（1－α）%可信区间为：

$$PAR\%\pm U_\alpha\sqrt{V(PAR\%)} \quad (47)$$

本例：

$$V(PAR\%)=\frac{1}{709^2}\left\{\left|\left(\frac{649\times2}{27}\right)^2\left(\frac{1}{27}+\frac{622}{649\times27}\right)+\left(\frac{60\times19}{32}\right)^2\left(\frac{1}{32}+\frac{28}{60\times32}\right)-709\times(1-0.8819)^2\right|\right\}^{\frac{1}{2}}=0.00042849$$

PAR%的95%可信区间为：

$$0.8819\pm1.96\sqrt{0.00042849}=0.8413\sim0.9225$$

（84.13～92.25%）

这说明该人群中有88.19%的肺癌病例归因于吸烟，其95%可信区间为84.13%～92.25%（消除了性别的影响后所得的结果）。

幼儿园学龄前儿童HBV感染标志的血清学研究

张树修　封太昌　王桂芝　邓长英

为了解学龄前儿童的HBV感染率，加强对病毒性乙型肝炎（乙肝）的综合防治，我们对北京市朝阳区五所幼儿园712名儿童进行了HBsAg、抗-HBs、抗-HBc三项血清学研究。结果报道如下。

HBsAg、抗-HBs、抗-HBc三项HBV感染标志阳性率分别为0.42%、12.08%和12.08%。HBV总感染率为20.37%，其中男21.39%（80/374），女19.23%（65/338），男女间无显著性差异（$\chi^2=0.385$，$P>0.05$）。年龄分布3岁10.78%(18/167)，4岁18.18%（22/121），5岁30.5%（61/200），6岁18.03%（44/244），各年龄组间有显著性差异（$\chi^2=23.56$，$P<0.05$）。单位分布总队幼儿园最高（34.76%，57/164），副食公司幼儿园最低(8.43%，14/166），单位间有显著性差异（$\chi^2=32.681$，$P<0.05$）。父（母）有肝炎史者47.61%（10/21）、曾注射过乙型肝炎疫苗者81.81%（9/11），其中抗-HBs阳性72.73%（8/11）。

本研究结果表明，朝阳区五所幼儿园学龄前儿童HBV感染率20.37%，大大低于一般人群水平（40%～50%左右），其年龄分布呈一特殊模式：3～5岁随年龄增长而上升，符合一般规律，但6岁下降，可能与母婴传播此时抗体水平下降（消失）和年龄尚小无更多感染机会有关。父（母）有肝炎史者HBV感染率高，说明母婴及水平传播是重要的。

据此我们认为：应在全国范围内对学龄前儿童进行乙肝疫苗注射。新生儿注射乙肝疫苗后5～6年（入小学前）应加强注射一次，以保持一定免疫水平，这对减少乙肝发病率，提高人民健康水平有着十分重要的意义。

（收稿：1993-01-29 修回：1993-04-24）

本文作者单位：武警北京总队医院 100027