

病例-对照研究中1:R 配比适度的应用研究

周 伦¹ 刘希永¹ 余 海¹ 陈 坤² 焦登整² 郑 树¹

摘要 为了探讨及验证 Pitman 指数作为最适配比观察指标的理论在病例对照研究中的实际应用,我们利用在浙江省嘉善县进行的大肠癌人群序贯筛检数据库作实例验证,对检出的41例大肠癌,在62 667例筛检人群中,随机抽取1:1至1:50的配比作病例对照研究。经过统计学处理,结果显示:随着R值的增大,具有显著性意义的因子检出率及与1:50配比研究结果的一致率随之增高;离差系数变化趋势越来越小;稳定性也相对提高,但当R值大于4时则增加值就较为缓慢。因此笔者认为:结合配比研究的灵敏度及稳定性,当在研究经费并不宽裕,且仅限于对某病作探索性研究时,最适配比以1:4或1:5左右为宜。

关键词 病例对照研究 Pitman 指数 1:R 配比

The Practical Application on the Numbers of Control in 1:R Matched Case-control Study Zhou Lun, Liu Xi-yong, Yu Hai, et al. Cancer Institute Zhejiang Medical University, Hangzhou 310009

Abstract We applied the data base of a screening program for colorectal cancer in the general population of Jiashan county, Zhejiang province to verify the theory of Pitman Efficiency in 1:R matched case-control study. Forty one detected cases of colorectal cancer were matched with normal controls as 1:R (R from 1 to 50) who were randomly selected from the screened population including 62 667 subjects. Results showed that the number of risk factors with statistical significance and its coincidence rate with results of 1:50 match were increased along with the increasing of R value; meanwhile the trends of standard deviation with 1:R over 1:50 weakened and the stability of results increased accordingly. The change of these values, however, reached their plateau when $R > 4$. The authors suggest that when taking consideration of sensitivity and stability of matched study, 1:4 of 1:5 ratio might be the most suitable case-control match, particularly when the resources are limited or the study has a exploratory nature.

Key words Case-control study Pitman efficiency 1:R match

在病例对照研究中经常采用1:R 配比方式,该方式不仅能控制某些混杂因素的干扰,而且更为重要的是能有效地提高研究效力^[1,2]。一般而言,R 值越大,则研究效力也较大,但当在R 值增大的同时,研究所需的人力、物力和财力也将随之增大,而且随着配比对子数的增加,研究效力的增加值将越来越小,因此R 的取值应有一定的限制。Ury 曾提出了 Pitman 指数 $[2R / (R + 1)]$ 作为最

适配比的观察指标^[3],为了探讨及验证这一理论在病例-对照研究中的实际应用,我们利用在浙江省嘉善县进行的大肠癌人群序贯筛检数据库,对检出的41例大肠癌病例作1:R 配比病例对照研究,现将分析结果报告如下。

材料与方法

一、研究对象:

1. 病例:来自1989~1990年间浙江省嘉善县30岁以上的大肠癌序贯筛检人群62 667人,以筛检结束时发现的41例大肠癌组成病例组,具体部位、病理及分期见以往报道^[4,5]。

1 浙江医科大学肿瘤研究所 杭州 310009

2 浙江医科大学流行病学教研室

2. 对照：以大肠癌病例为指示病例，根据同性别、同年龄组（以5岁为一龄组）为配对条件，按照1:R 配比，R 分别取1、2、3、4、5...10、15、20、25、30、50，随机从62 667 人群数据库中抽取无结、直肠疾病者为对照。

二、研究项目：所列入的研究项目为19 项，分别为：职业（V1）、教育程度（V2）、喜吃肥肉（V3）、经常吃新鲜蔬菜（V4）、吃红烧鱼（V5）、不洁饮水史（V6）、吸烟（V7）、饮酒（V8）、慢性腹泻（V9）、便秘（V10）、粘液血便（V11）、阑尾炎（V12）、阑尾手术史（V13）、肠息肉史（V14）、哮喘史（V15）、家族癌肿史（V16）、钩虫病史（V17）、精神刺激史（V18）及免疫法粪便潜血试验（V19），均为计数资料。

三、资料的收集：在病例和对照确定之前，对所有目标人群（62 667人），均进行了序贯筛检，初筛方法为免疫法粪便潜血试验结合计算机大肠癌危险隶属度的模式^[4]，在初筛开始时均作了调查表调查，所有上述19 项因素均被列入调查表中。调查员经统一培训，考核合格。调查方式为面访，并对部分调查表作了抽样质检。

四、统计和分析方法：所有筛检资料均以文本形式在计算机上建立数据库，各1:R 配比病例-对照研究分析应用 EGERT 软件包，进行 Logistic 回归分析，各1:R 配比组分析结果的各因子 OR 值及 P 值以 FOXBASE 形式建立数据库，进行计算及数据转换，用 SPSS/PC⁺ 软件包作中位数及四分位数的计算，并作符号秩和检验。

结 果

一、1:R 配比组与1:50配比组统计结果的一致性：为了比较各1:R 配比组的研究效力，我们把1:50配比组作为标准，经条件 Logistic 回归分析结果，1:50组具有统计学意义（ $P < 0.05$ ）的因子共有8个，分别为：喜吃肥肉（V3）、慢性腹泻（V9）、便秘（V10）、粘液血便（V11）、肠息肉（V14）、家

族癌肿史（V16）、精神刺激史（V18）及免疫法粪便潜血试验（V19）。各1:R 配比病例对照研究组分析结果与1:50配比组具有相同的统计结果的因子数（当1:50组 OR 值大于1.0时，1:R 也大于1.0因子数，及当1:50组 OR 值小于1.0时，1:R 组的 OR 值也小于1.0的因子数）及8个有显著意义的因子中分析结果有显著性意义的因子数见表1，其中当 $P = 0.05$ 的因子计为0.5个因子。

表1 配比病例对照研究中各1:R 配比组与1:50配比组符合率

R	显著性意义因子数(A)	A/8×100% (%)	与1:50具有相同统计结果因子数	与1:50组的一致率(%)	相对工作量(1:R组/1:1组)
1	2	25.0	9*	64.3	1.0
2	2.5	31.3	13	68.4	1.5
3	3	37.5	14	73.7	2.0
4	5	62.5	16	84.2	2.5
5	5	62.5	16	84.2	3.0
6	5.5	68.8	16	84.2	3.5
7	5.5	68.8	16	84.2	4.0
8	6	75.0	17	89.5	4.5
9	6	75.0	17	89.5	5.0
10	6	75.0	17	89.5	5.5
15	7	87.5	18	94.7	8.0
20	7	87.5	17	89.5	10.5
25	7	87.5	18	94.7	13.0
30	7	87.5	18	94.7	15.5

* 其中5个因子因叠代次数超过20次，但仍未出结果而中止运算，所以1:1组仅14个因子有结果

由表1可见，随着R 值的增大，具有显著性意义的因子检出率及与1:50配比研究结果的一致率也随之增高，但这种增加主要在1:4之前，当R 值大于4时，则增加越来越有限。

二、R 值与研究结果的稳定性：为了比较各1:R 配比组的分析结果的稳定性，我们把 $|OR_{1:R} - OR_{1:50}| / OR_{1:50}$ 的值确定为离差系数（S），并作各配比组19个因子的离差系数中位数和四分位数，各配比组与1:30组的

离差系数(S值)作配对符号秩和检验,结果见表2。

表2 各1:R 配比组稳定系数的比较

R	S 值中位数	四分位数	与1:30组比较 P 值
1	0.270	0.160~0.510	<0.001
2	0.320	0.130~0.550	<0.001
3	0.250	0.170~0.300	<0.001
4	0.260	0.090~0.390	<0.01
5	0.170	0.060~0.340	<0.001
6	0.180	0.100~0.370	<0.001
7	0.180	0.110~0.380	<0.001
8	0.180	0.070~0.220	<0.001
9	0.110	0.030~0.230	0.049
10	0.160	0.090~0.300	<0.001
15	0.090	0.040~0.170	>0.05
20	0.070	0.040~0.150	>0.05
25	0.050	0.010~0.080	>0.05
30	0.040	0.030~0.070	

由表2可见,随着R值增大,离差系数变化趋势越来越小,即R值越大则1:R 配比病例对照研究的各因子的OR值越接近1:50配比组研究结果,当R值≥15时,则配比研究达到较为稳定水平。

讨 论

随着1:R 配比多元统计方法及软件包的广泛应用^[6,7],1:R 配比设计目前为流行病学研究者所重视,对最适配比方式在理论上进行了一些探讨^[8]。卡恩^[9]通过采用病例和对照的P值(具有危险因素的比例)之差的标准误(SE)来测定,结果表明,随着配比数的增加,研究的精度将提高,而在R=2以后则这种SE的减少就不明显了,因此卡恩认为:当R值>2时可能没有什么必要。而Ury^[2]则提出了Pitman指数(Pitman Efficiency)作为最适配比的观察指标,Pitman指数的公式为 $2R/(R+1)$,该值在1:1时为1,当R值达到 $+\infty$ 时则为2。

Pitman指数随着R值的增加而从1.00

到2.00之间变动,当R值>4时则增加值就较为缓慢,因此Ury等认为在1:R 配比研究中,1个病例的配对不宜超过4个对照。但以上这些结论仅为理论上的推断。分别随机抽取1:1至1:50的配比,并对它们的研究效力作实际应用验证和评价,本文尚属首次报道。

本次研究的调查员在大肠癌筛检过程中对受检者进行了调查,调查时调查员及被调查者均尚未知筛检结果,可以认为是一次双盲试验,较为有效地排除了由调查引起的偏倚。在对照选择上按与相应病例同性别同年龄组为条件,从大肠癌筛检建立起的数据库中随机挑选对照,较好地控制了选择偏倚。从稳定性来看,当R值为15时,其研究结果已达较为稳定的水平,因此可以认为1:50配比研究结果相对较为可靠,可以作为评价的参照指标。

本研究中1:50配比研究结果有8个因子具有统计学显著性意义,当配比从1:1提高到1:4时,调查例数增加至2.5倍,这几个因子中的显著性检出率也提高到2.5倍,与1:50配比组研究结果的一致率提高了19.9%;而当R值从4增加至30时,调查例数增加了5.2倍,显著性因子检出率仅增加了0.4倍,与1:50配比组研究结果的一致率仅提高了10.5%。因此,从效益角度来看,选择1:4左右配比较为合适。

从研究结果的稳定性来看,当配比数从1:1增加到1:15时,随着配比数增大,稳定性随之相对提高也较为明显,而当R值达到15以上时,则研究结果达到较为稳定的水平,因此当研究的病例较难得到时,为了增加把握度,可适当增加配比数,但配比数没必要超过15。

因此笔者认为:结合配比研究的灵敏度及稳定性,当在研究经费并不宽裕,且仅限于对某病线索作探索性研究时,最适配比以1:4或1:5左右为宜,因为R值在4或5左右时,稳定性及灵敏度较1:1配比均有了较大的提高。而当为了对某些因子精确量化,旨在建

立数学模型,并应用于预测及估计个体危险时,应根据实际情况,适当增加配比对子数是必要的,但最大不宜超过15。

参 考 文 献

- 1 Mckeown-Eyssen GE. Sample size determination in case-control studies: The influence of the distribution of exposure. *J Chro Dis*, 1985, 38:559.
- 2 Schelesselman JJ. Case-control studies design, Conduct, Analysis. Oxford Univ. press, NY, 1982, 105-170.
- 3 Ury HK. Efficiency of case-control studies with mutiple controls per case: Continuous or dichotomous data. *Biometrics*, 1975, 31:643.
- 4 郑树,余海,吴金民,等. 结直肠癌的人群筛检模式及评价. *中华医学杂志*, 1991, 71:381.
- 5 周伦,余海,陈坤,等. 基于筛检人群的大肠癌危险因素病例-对照研究. *中国慢性病预防与控制*, 1993, 1:213.
- 6 Smith J. Planing the size of a matched case-control Study for estimation of the odds ratio. *Am J Epidemiol*, 1985, 122:345.
- 7 武光林. 配比对数固定和变化的1:R 配比病例对照研究. *中华肿瘤杂志*, 1990, 12:104.
- 8 Samuels ML. Matching and design efficiency in epidemiologic studies. *Biometrika*, 1981, 68:577.
- 9 卡恩 HA. 流行病学方法导论(文万青译). 湖南省卫生防疫站情报室资料(长沙). 1982:18.

(收稿:1996-05-03)

唐山市城区儿童肥胖症与高血压关系的流行病学调查

张秀华¹ 张树立² 李秀芹¹ 徐锦屏¹

高血压是最常见的心血管疾病死亡的原因。儿童期肥胖与高血压的发生有直接关系。饮食营养因素对血压调节有重要意义,合理饮食将对高血压的防治起重要作用。

一、对象与方法:对象为1988年6月至1995年6月出生的唐山市城区集体与散居1~7岁小儿1 755名(男952名,女803名)。以1985年全国九省市城区小儿生长发育调查的各年龄组体重平均值为标准;按1985年全国九省市城区体格发育调查的方法进行测身高体重。凡超过身高标准体重20%以上的儿童判定为肥胖。分别测其安静坐位右上臂肱动脉血压,连测三次,取平均值,若血压大于正常年龄血压的第95个百分位数值则诊断高血压。凡高血压的小儿均于第二、第三周重测血压。

二、结果:1 755名儿童只查出肥胖儿童175名,占9.97%。其中1岁组和6~7岁组儿童肥胖症人数比例大。1 755名儿童共查出高血压儿童33名,患病率为1.88%。高血压的出现率在男女之间差别无显

著性。高血压儿童的年龄为3岁1名、4岁2名,其余均在5岁以后并开始增多。收缩压范围在16~20kPa,舒张压范围在10.7~12kPa。175例肥胖儿童中,22例有高血压,占12.57%。而1 580例非肥胖儿童中,只有11例有高血压,占0.7%。差别有显著性($\chi^2=230, P<0.001$)。

三、讨论:国内儿童高血压的发生率为1%~2%。国外资料诊断不一,发病率为0.6%~11%,本文为1.88%,与国内相符。

肥胖导致高血压的机制不明,可能与下列因素有关:①肥胖伴高胰岛素血症,也可导致肾素-血管紧张素-醛固酮系统功能异常;②因细胞膜离子转运功能异常,钠-钾泵活性异常;③血容量和心输出量增加;④神经分泌失调。

儿童肥胖者血压升高开始是一种代偿机制,但长期血压升高会给儿童心脑血管带来永久的损害。儿童肥胖是日后心脑血管疾病的高危因素,为预防和减少靶器官的损害,儿童期重视消除高危因素具有极其重要的意义。

(收稿:1996-06-25)

1 河北省唐山市妇幼保健院 063000

2 河北省唐山市卫校附属医院