

# 病例对照研究

## ——实战操作

吕军 副主任

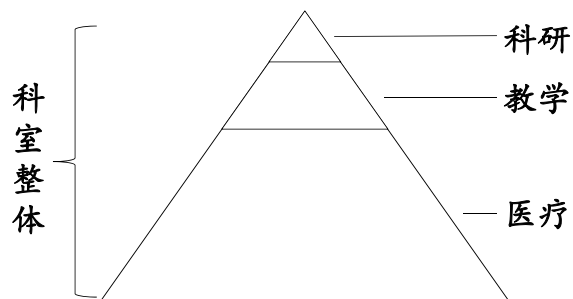
西安交通大学第一附属医院 临床研究中心

2015-10-17

## 临床研究

针对临床医师对患者的诊治过程中出现的临床实践问题所做的研究工作。

## 医疗与科研的关系



## 医院相关临床研究的主要内容

- 观察性研究：
  - 描述性研究：病例报告、横断面研究、纵向研究
  - 分析性研究：队列研究、病例对照研究
- 试验性研究：随机对照研究（RCT）
- 二次研究：系统评价、Meta分析、临床经济学分析

# 病例对照研究

## Case-Control Study

## 基本概念

选择患有某特定疾病的患者作为病例，以不患有该病但具有可比性的个体作为对照，调查他们发病前对某个（些）因素的暴露情况，比较两组中暴露率和暴露水平的差异，研究该疾病与这个（些）因素的关系

## 暴露

研究对象曾经接触过某些因素，或具备某些特征，或处于某种状态，这些因素或状态即为暴露因素，它可以是有害的，也可以是有益的，也叫研究变量

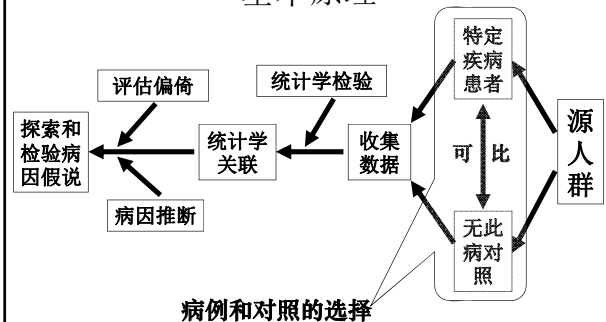
### 举例：

对一组肺癌病人（病例组）和一组未患肺癌但有可比性的人（对照组）调查他们的吸烟（暴露）历史（可包括现在吸烟否，过去吸过烟否，开始吸烟年龄，吸烟年数，最近每天吸烟支数；如已戒烟则为戒烟前每日吸烟支数，已戒烟年数，等等）。其目的为通过比较两组吸烟史的差别，检验吸烟（可疑病因）与疾病（肺癌）有因果联系的假设。

### 病例对照研究的基本原理及特点

- 以现在确诊的患有某特定疾病的病人作为病例，以不患有该病但具有可比性的个体作为对照，通过询问、实验室检查或复查病史，搜集既往各种可能的危险因素的暴露史，测量并比较病例组与对照组中各种因素的暴露比例，经统计学检验，若两组差别有意义，则可认为因素与疾病之间存在着统计学上的关联。在评估了各种偏倚对研究结果的影响之后，再借助病因推断技术，推断出某个或某些暴露因素是疾病的危险因素，而达到探索和检验疾病假说的目的。
- 特点 ④属于观察性研究，不施加干预措施  
⑤设立对照组  
⑥由“果”及“因”，疾病→暴露  
⑦只能判断暴露及疾病之间是否关联及关联程度，不能做因果关系的最终结论。

### 基本原理



### 病例对照研究的设计和实施

- 研究对象的选择
- 研究因素的收集与测量
- 资料的整理和分析
- 主要偏倚及控制
- 病例对照研究结果的解释

### 研究对象的选择（1）

- 基本原则：病例足以代表总体人群中该病的患者，对照足以代表产生病例的人群总体。
- 病例的选择：
  - ✓ 病例内外部特征的限定
  - ✓ 病例类型：新发病例、现患病例和死亡病例
  - ✓ 病例的来源：医院/社区

## 研究对象的选择（2）

- 对照的选择：
  - ✓ 对照选择的原则：代表性&可比性
  - ✓ 对照的类型：成组对照、配比对照（群体匹配&个体匹配）---防止匹配过头，不超过1: 4
  - ✓ 对照的来源：当地全人口/医院其他病例/病例的配偶、同胞、亲戚、同事或邻居

## 研究对象的选择（3）

- 样本大小的估计：
  - ✓ 取决于下列4个因素：
    - a) 欲研究因素在对照人群中的估计暴露率 $P_0$ ;
    - b) 预期与该暴露有关的RR/OR;
    - c) 所希望达到的检验显著性水平 $\alpha$ ;
    - d) 所希望达到的检验把握度 $(1-\beta)$
  - ✓ 公式计算或从样本含量表中查得需要的病例和对照数。

$$p_1 = p_0 RR [1 \mp P_0 (RR - 1)], \bar{p} = 0.5(p_1 + p_0), \bar{q} = 1 - \bar{p}$$

## 研究对象的选择（4）

- 样本大小的估计1：
 

成组对照样本含量的估计

$$n = 2 \bar{p} \bar{q} (U_{\alpha} + U_{\beta})^2 / (p_1 - p_0)^2 \quad \text{公式1}$$
- 式中 $n$ 为病例组或对照组人数， $U_{\alpha}$ 和 $U_{\beta}$ 分别为与 $\alpha$ 及 $\beta$ 值对应的标准正态分布的分位数，可以从相应表格中查出。 $p_1$ 与 $p_0$ 分别为病例组及对照组估计的某因素暴露率。

## 研究对象的选择（4）

- 样本大小的估计1：
 

病例组及对照组估计的某因素暴露率 $p_1$ 与 $p_0$ 的计算如下：

$$p_1 = p_0 RR [1 + P_0 (RR - 1)], \bar{p} = 0.5(p_1 + p_0), \bar{q} = 1 - \bar{p} \quad \text{公式2}$$

## 研究对象的选择（4）

- 现设计一个研究吸烟与肺癌关系的调查，估计对照组吸烟史比例，即暴露比 $p_0$ 为20%，相对危险度RR约为2，要求 $\alpha=0.05$ ， $\beta=0.1$ ，求样本大小 $N$ 。
- 病例的暴露史 $p_1$ 可用公式2估计
- $$p_1 = (0.2 \times 2) / (1 + 0.2(2-1)) = 0.3333$$
- $$\bar{p} = (0.2 + 0.333) / 2 = 0.267$$
- $$\bar{q} = 1 - 0.267 = 0.733$$
- 从相应表格中查得 $U_{\alpha}=1.96$ ， $U_{\beta}=1.282$ 代入公式1
- $$n = 2 \times 0.267 \times 0.733 (1.96^2 + 1.282^2 / 0.3333 - 0.2)^2$$
- $$= 232$$
- 即得每组需要调查约232人。

## 研究对象的选择（4）

正态分布的分位数表

$\alpha$ 及 $\beta$	$U_{\alpha}$ (单侧检验) $U_{\beta}$ (单侧和双侧检验)	$U_{\alpha}$ (双侧检验)
0.001	3.090	3.290
0.002	2.878	3.090
0.005	2.576	2.807
0.010	2.326	2.576
0.020	2.058	2.326
0.025	1.960	2.242
0.050	1.645	1.960
0.100	1.282	1.645
0.200	0.842	1.282

## 研究因素的收集与测量

- 暴露因素的规定
  - ✓ 必须在调查前有明确的规定；
  - ✓ 可将暴露因素分级，可能提供因素与疾病间的剂量反应关系
- 暴露因素的收集
  - ✓ 专门制定的调查表，病例组与对照组使用相同的调查表
- 信息资料的来源
  - ✓ 医疗记录、登记报告、职业史记录、访问调查记录、通信调查表等

## 资料的整理和分析

- 描述性分析
  - ✓ 描述研究对象的一般特征
  - ✓ 均衡性检验: 可比性
- 推断性分析
  - ✓ 主要是分析暴露与疾病有无统计学关联，以及关联强度的大小。
    - ◆ 不匹配不分层的资料分析
    - ◆ 匹配资料的分析

## 主要偏倚 (1)

- 选择偏倚：
  - ✓ 由于选择研究对象的方法有问题或缺点，导致入选者与未入选者的某些特征有系统差别而产生的误差
  - ✓ 病例对照研究常常未能随机抽样，故易产生选择偏倚。特别在医院选择病例与对照时更易产生偏倚
  - ✓ 由于诊断标准不明确，或标准不够详细，使病例组内构成不一致

## 主要偏倚 (2)

- 信息偏倚：
  - ✓ 在调查时对两组的暴露史采取了不同的标准或收集手段可引起信息偏倚
  - ✓ 在病例组详细查阅病历或其他记录，而调查对照时则多依据对照口头提供资料
  - ✓ 观察者在调查或测量时收集的资料在两组间准确性不一致或者被调查者提供不准确的信息
  - ✓ 回忆偏倚

## 主要偏倚 (3)

- 混杂偏倚：
  - ✓ 由于混杂因子所造成的偏倚
  - ✓ 混杂因子是指既和研究的疾病有联系（即这个因子必须是一个危险因子）又和研究的暴露有联系的因子
  - ✓ 年龄、性别和许多疾病与许多暴露都有联系，所以是最常见的混杂因子

## 偏倚的控制 (1)

- 选择偏倚：
  - ✓ 在选择对象时，尽可能采取随机抽样原则；
  - ✓ 注意研究对象的代表性；
  - ✓ 如果在医院选择病例，则尽可能多选几所医院；
  - ✓ 对无应答的对象，要设法补救并在分析时对无应答的影响作出特别分析。

## 偏倚的控制 (2)

- 信息偏倚：
  - ✓ 进行检查或调查时尽可能采取盲法；
  - ✓ 调查的变量尽可能采取客观性强的指标；
- 混杂偏倚：
  - ✓ 研究设计阶段：限制和匹配
  - ✓ 分析阶段：分层分析/标准化处理/多因素分析

## 病例对照研究结果的解释

- 暴露因素和疾病之间存在的统计学关联有三种可能：
  - ✓ 机会的作用：抽样误差
 

注意：如果结果无显著性差异，不能轻易地肯定该因素与疾病之间不存在关联，可能由于因素对疾病的作用较小，样本量没能达到分析所要求的精度和把握水平而造成的。此时，应扩大样本含量，再进行研究。
  - ✓ 偏倚的作用：选择偏倚和回忆偏倚
  - ✓ 因果联系
 

病例对照研究仅能借助于逻辑推理即病因推断技术判断是否存在因果联系，它对因果关系仅限于是一种判断而不是因果联系的证明。

## 研究实例

### 实例1 吸烟与肺癌

Do11和Hi11于1950年报道了他们应用病例对照研究的方法研究了吸烟与肺癌的关系。

709例肺癌病例（伦敦20所医院）

709例非肺癌对照病例（同一间医院，同性别、年龄差别不超过5岁）

肺癌病例对照研究吸烟史比较

吸烟史	病例	对照	合计
吸烟	688 (97.04%)	650 (91.68%)	1338
不吸烟	21 (2.96%)	59 (8.32%)	80
合计	709	709	1418

OR:  $(a/c) / (b/d) = ad/bc = (688 \times 59) / (650 \times 21) = 2.97$   
P<0.01

肺癌病例对照研究吸烟史比较(男性)

吸烟史	病例	对照	合计
吸烟	647 (99.69%)	622 (95.84%)	1269
不吸烟	2 (0.31%)	27 (4.16%)	29
合计	649	649	1298

OR:  $(a/c) / (b/d) = ad/bc = (647 \times 27) / (622 \times 2) = 14.04$   
P<0.01

### 肺癌病例对照研究吸烟史比较(女性)

吸烟史	病例	对照	合计
吸烟	41 (68.33%)	28 (46.67%)	69
不吸烟	19 (31.67%)	32 (53.33%)	51
合计	60	60	120

OR:  $(a/c) / (b/d) = ad/bc = (41 \times 32) / (28 \times 19) = 2.47$   
P<0.05

### 实例2 己烯雌酚激素与阴道腺癌

#### 研究背景

- 美国波士顿Vincent纪念医院妇产科医生Herbst 1966~1969年收治7例阴道腺癌患者,发现:
  - 7例全是腺癌(通常阴道癌占女性生殖系统癌的2%,阴道腺癌仅占阴道癌的5%~10%,非常罕见)
  - 7例年龄15~22岁(过去腺癌年龄均大于25岁)

### 研究方法

- Herbst对阴道腺癌危险因素(阴道局部刺激史、性生活史、避孕药物、母亲妊娠期间的情况)进行探索
  - 7例患者加上另一个医院1例患者作为病例组
  - 每个病人配4个对照,共32个对照
  - 调查员用标准调查表对病例、对照与她们的母亲进行了调查,经统计学处理后的主要结果见下表。

#### 阴道腺癌病例与对照的母亲主要暴露因素的比较

病例号	母亲年龄		母亲吸烟		此次怀孕出血		以往流产史		此次怀孕时使用过雌激素		母亲哺乳		此次怀孕时照射过x线	
	病例	四个对照平均	病例	对照	病例	对照	病例	对照	病例	对照	病例	对照	病例	对照
1	25	32	有	2/4	否	0/4	有	1/4	有	0/4	否	0/4	否	1/4
2	30	30	有	3/4	否	0/4	有	1/4	有	0/4	否	1/4	否	0/4
3	22	31	有	1/4	有	0/4	否	1/4	有	0/4	有	0/4	否	0/4
4	33	30	有	3/4	有	0/4	有	0/4	有	0/4	有	2/4	否	0/4
5	22	27	有	3/4	否	1/4	否	1/4	否	0/4	否	0/4	否	0/4
6	21	29	有	3/4	有	0/4	有	0/4	有	0/4	否	0/4	否	1/4
7	30	27	否	3/4	有	0/4	有	1/4	有	0/4	有	0/4	否	1/4
8	26	28	有	3/4	有	0/4	有	0/4	有	0/4	否	0/4	有	1/4
合计			7/8	21/32	3/8	1/32	6/8	5/32	7/8	0/32	3/8	3/32	1/8	4/32
平均	26.1	29.3												
$\chi^2$	自由度为1 <sup>a</sup>		0.53		4.52		7.16		23.22		2.35		0	
P	(不显著 <sup>b</sup> )		0.50		<0.05		<0.01		<0.0001		0.20		(不显著)	
OR			5.7		8.0		10.5		28.0		10.0		3.0	

注 ① 用Fisher与Morrow的配对对照 $\chi^2$ 检验公式。

② 配对t检验,  $S_{\bar{x}}=1.7$ 岁

### 研究结果

- 在比较的因素中,只有三个因素差别有显著性
  - 母亲怀孕期间使用过己烯雌酚激素治疗(P<0.00001)
  - 母亲以前流产史(P<0.01)
  - 此次怀孕阴道出血史(P<0.05)
- 因有后两个因素存在才使用己烯雌酚治疗
- 结论: 母亲在妊娠早期服用己烯雌酚使她们在子宫中的女儿以后发生阴道腺癌的危险性增加。

### 实例3 药物基因组学相关病例对照研究实例



×2



International Journal of  
Molecular Sciences  
(IF: 2.862)

European Journal of Clinical Microbiology & Infectious Diseases  
(IF: 2.668)

Genotype	CHF n (%)	Controls n (%)	HWE $p^*$	Dominant Model $p^*$	Recessive Model $p^*$	Additive Model $p^*$	Allele Contrast $p^*$
TT (W)	140 (40.82)	198 (40.75)					
CT	132 (44.15)	161 (40.45)	0.46	0.44	0.73	0.62	0.67
CC	27 (9.03)	39 (9.8)					
rs2007720							
AA (W)	175 (56.82)	243 (61.52)					
AG	107 (36.03)	133 (33.47)	0.88	0.49	0.88	0.79	0.53
GG	15 (5.05)	19 (4.83)					
rs0311675							
CC	234 (76.26)	319 (79.95)	0.82	0.59	0.43 <sup>b</sup>	0.52	0.43
CT	59 (19.73)	76 (19.05)					
TT	6 (2.0)	4 (1)					
rs8119756							
AA	223 (74.83)	270 (68.01)	0.82	0.0006	0.47	0.143	0.051
AG	68 (22.82)	114 (28.72)					
GG	7 (2.35)	13 (3.27)					
rs244072							
TT	205 (68.56)	272 (68.34)	0.39	0.95	0.59	0.86	0.81
TC	85 (28.43)	111 (27.89)					
CC	9 (3.01)	15 (3.77)					
rs0311889							
TT	228 (76.25)	320 (80.4)	0.08	0.19	0.74	0.33	0.27
GT	69 (22.97)	70 (17.58)					
GG	5 (1.67)	8 (2.01)					
rs220084							
TT	114 (44.97)	189 (47.63)					
TC	129 (43.29)	165 (41.56)	0.44	0.49	0.71	0.78	0.48
CC	35 (11.74)	43 (10.83)					
rs452159							
CC	93 (31.31)	162 (40.7)	0.14	0.011	0.48	0.039	0.004
CA	151 (50.84)	173 (43.47)					
AA	51 (17.85)	63 (15.83)					
rs73508374							
GG	259 (87.21)	361 (90.48)	0.040	0.17	0.83 <sup>b</sup>	0.23	0.26
AG	37 (12.48)	35 (8.77)					
AA	1 (0.34)	3 (0.75)					

Abbreviations: CHF, chronic heart failure; HWE, Hardy-Weinberg equilibrium;  $p^*$  values were calculated from two-sided  $\chi^2$  test.

SNPs	Dominant Model	Recessive Model	Additive Model	Allele Contrast
	Odds Ratio (95% CI) $p$	Odds Ratio (95% CI) $p$	Odds Ratio (95% CI) $p$	Odds Ratio (95% CI) $p$
rs452159	1.537 (1.10–2.16) 0.013	1.005 (0.65–1.55) 0.981	1.224 (0.97–1.54) 0.085	1.229 (0.98–1.55) 0.081
rs8119756	0.775 (0.54–1.10) 0.164	0.912 (0.34–2.47) 0.856	0.816 (0.60–1.12) 0.204	0.814 (0.59–1.12) 0.201

## 精髓

- 分组，找差异
- 自圆其说

## 如何开始——建议套路

- 从CNKI上下载文章
- 专找本专业相关病例对照研究文章
- 硕、博士论文更详细
- 找到自己感兴趣的题目，换一个要素作为自己的题目
- 开整！

## 如何选杂志

- <http://www.medsci.cn/sci/>
- 列表、总结
- 近年来病例对照研究论文发表情况
- 初期IF1-2，逐步提高
- Endnote, noteexpress

## 投稿要点

### 无缝连接

**谢谢聆听！**