

分析流行病学

- 又称：分析性研究
- 是在选择的人群中检验疾病病因假设或流行因素的一类方法。
- 通过专门设计的不同组间的比较，分析研究因素的作用的观察性研究方法。
- 包括：队列研究、病例对照研究

队列研究

(cohort study)

第一节 概述

- 一、概念
- 二、基本原理
- 三、队列研究定义
- 四、研究目的
- 五、特点
- 六、分类

一、概念

（一）队列（cohort）：



原意是指古罗马军团中的一个分队



- 日常生活中的队列
形式上：排成队
本质上：
有共同特点或特征

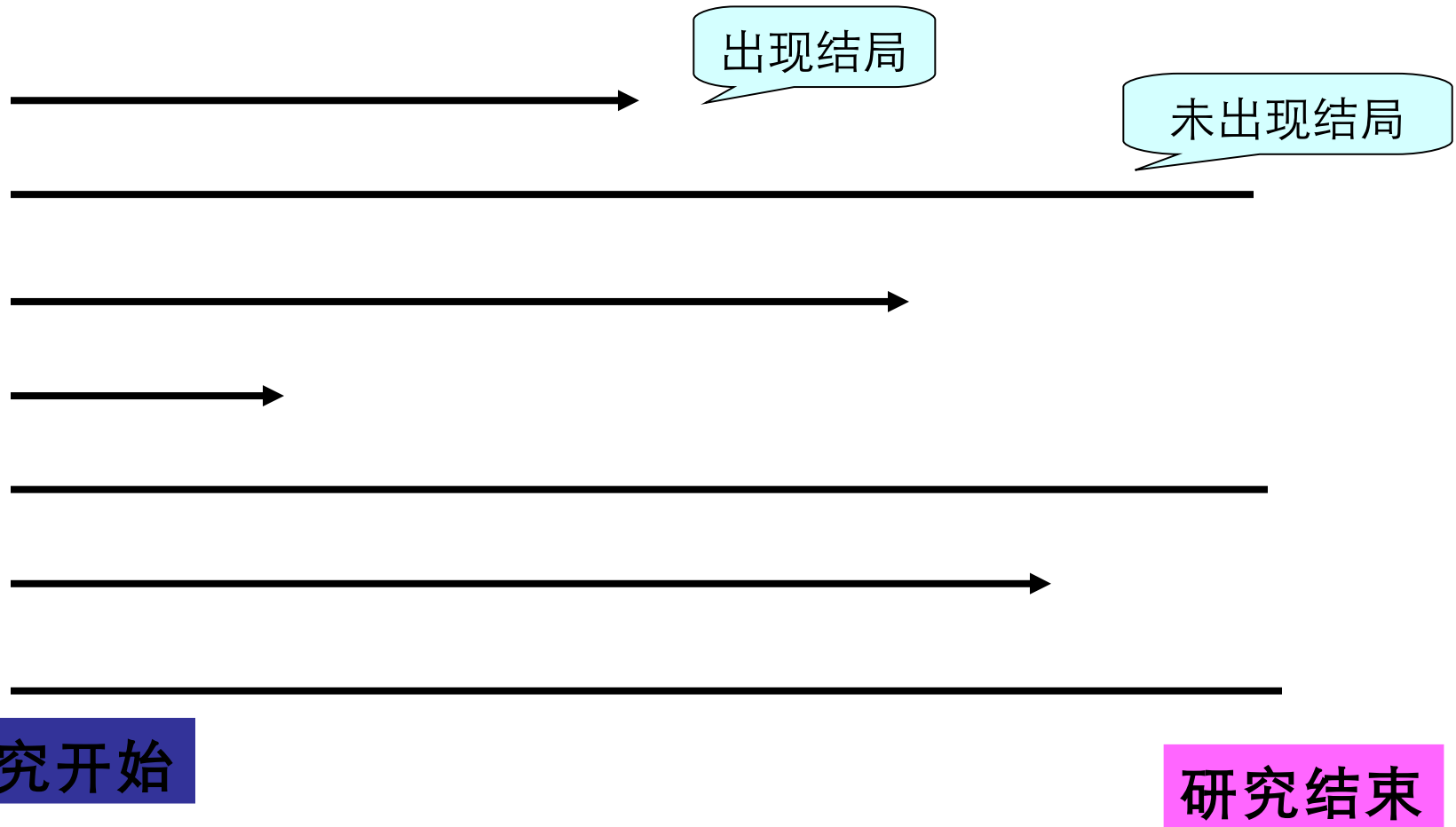


- **流行病学队列：** cohort

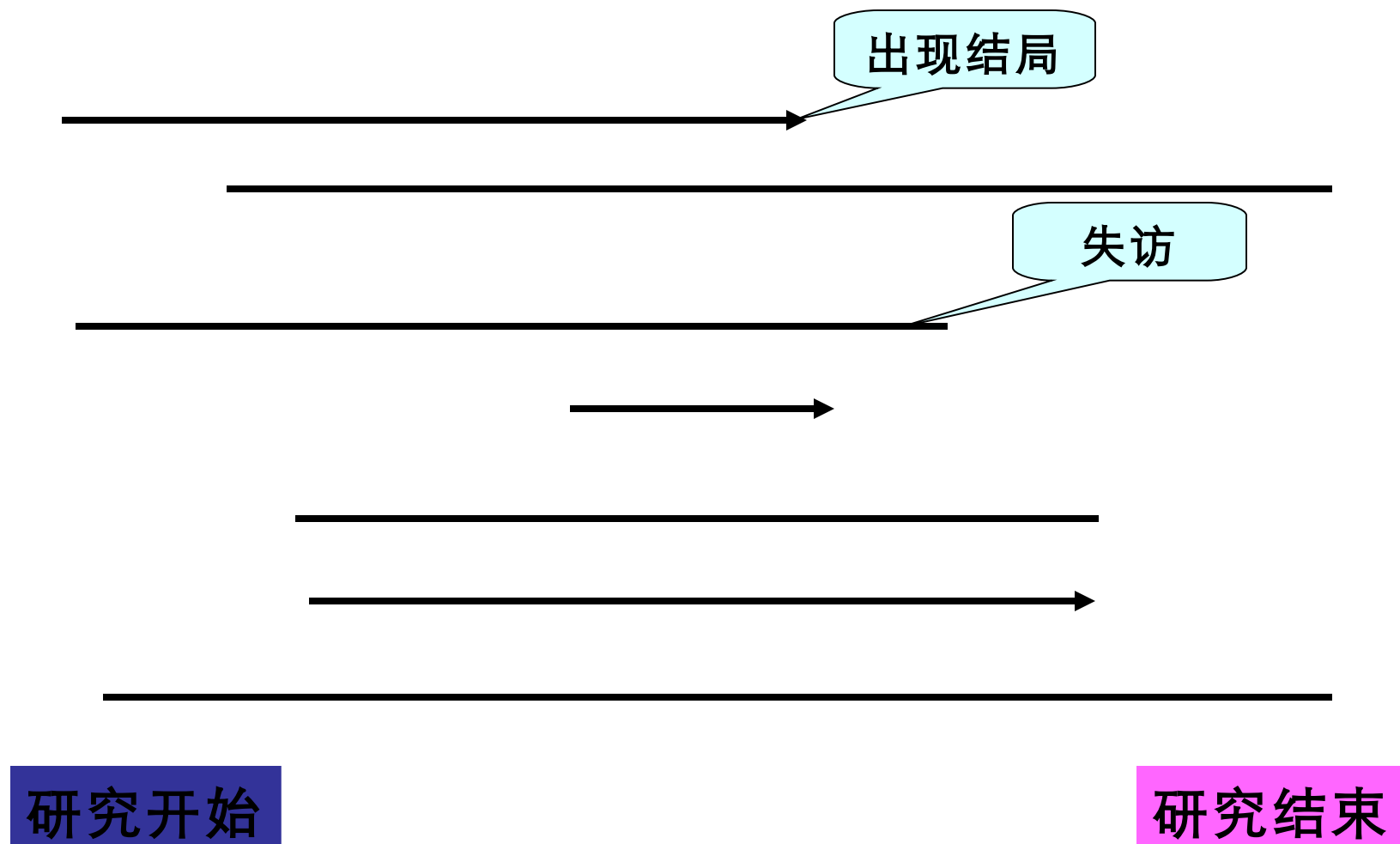
指有共同特征或共同经历一群人。

固定队列和动态队列

固定队列 Fixed Cohort (同时进入、无失访和进入)



动态队列 Dynamic Cohort（随时进入或退出）



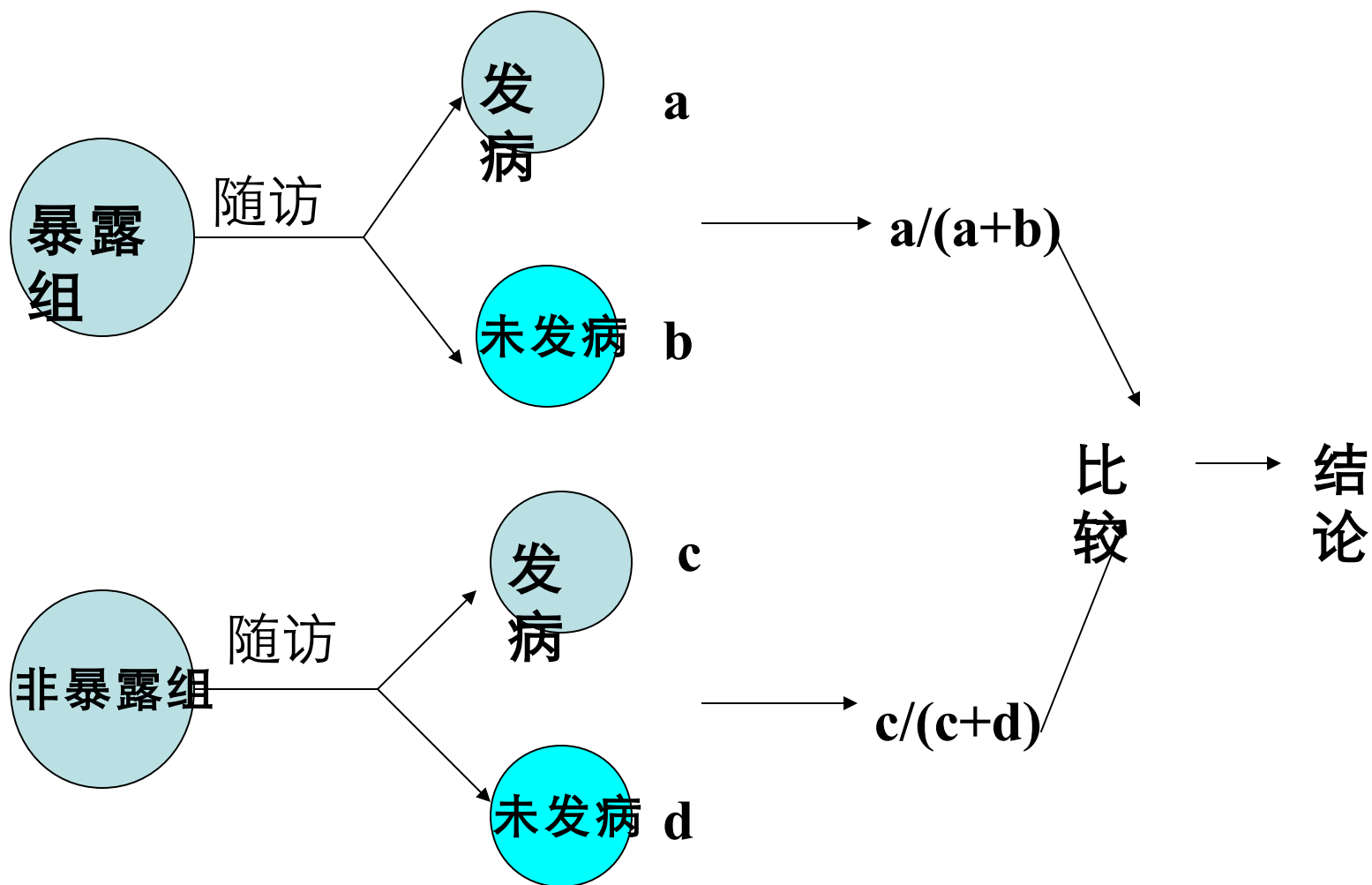
(二) 暴露：指接触过某种因素或具备某种特征。

(三) 危险因素：指能使疾病发病概率增加的因素。

(四) 保护因素：指能使疾病发病概率降低的因素。

二、基本原理

前瞻性



三、队列研究定义

选定暴露及未暴露于某因素的两种人群，追踪其各自的某种疾病结局，比较两者疾病结局的差异，从而判定暴露因子与发病或死亡有无因果关联及关联大小的一种观察研究方法。

四、队列研究的特点：

1. 属于观察法
2. 设立对照组
3. 观察方向由“因”及“果”
4. 能确证暴露与疾病的因果关系

五、队列研究的目的

- 1. 检验病因假设**
- 2. 描述疾病的自然史**
- 3. 评价自发预防的效果**

五、队列研究的类型

- 1. 前瞻性队列研究**
- 2. 历史性队列研究**
- 3. 双向性队列研究**

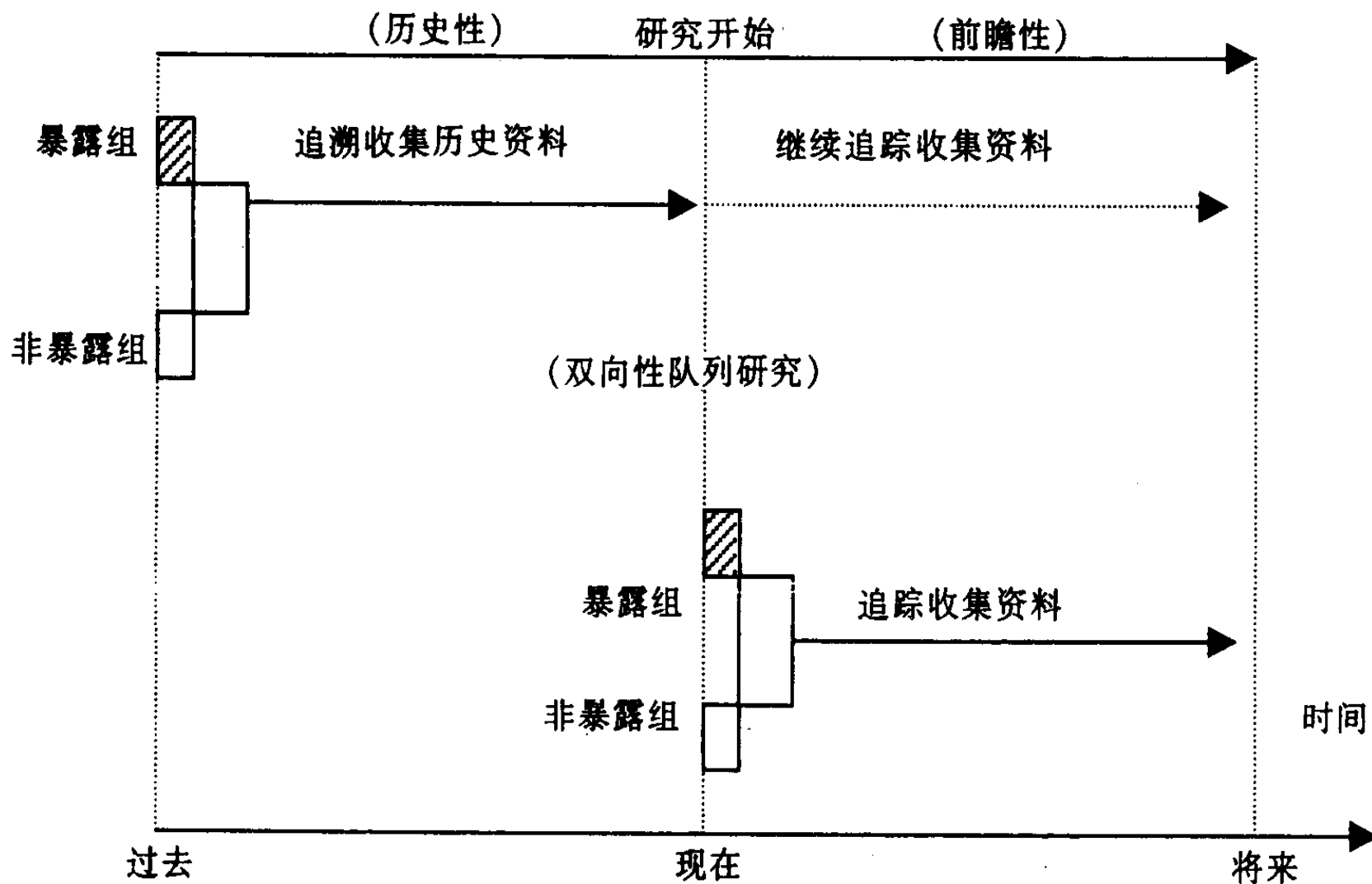


图 6-2 三类队列研究示意图

前瞻性队列研究

Prospective Cohort Study

- 研究队列的确定是现在 (concurrent)
- 根据研究对象现在暴露分组
- 需要随访 (follow-up)
- 结局在将来某时刻出现

前瞻性队列研究

Prospective Cohort Study

优点

- 时间顺序增强病因推断可信度
- 直接获得暴露与结局资料，结果可信

缺点

- 花费大，时间长

应用条件

- 明确地检验假设
- 所研究疾病的发生率较高，一般不低于5‰
- 明确规定暴露因素和结局变量
- 可靠的测量手段
- 足够的观察人群和暴露情况
- 能完成随访的人群
- 足够的人、财、物力

历史性队列研究

Historical (Retrospective) Cohort Study

- 根据研究开始时研究者掌握的有关研究对象在过去某时刻的暴露情况的历史材料分组
- 不需要随访，研究开始时结局已出现

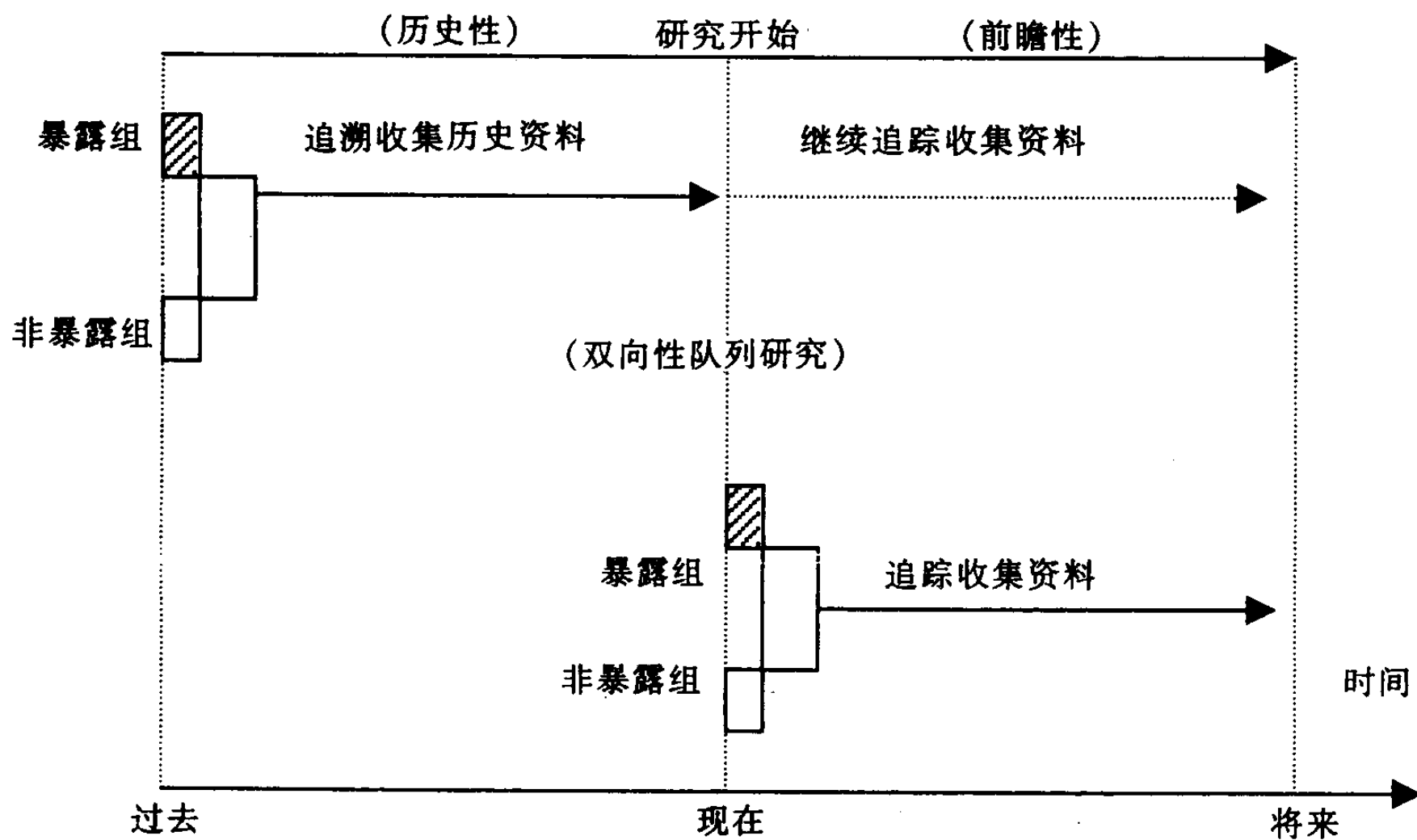


图 6-2 三类队列研究示意图

历史性队列研究

优点

- 短期内完成资料的收集和分析
- 时间顺序仍是由因到果
- 省时、省力、出结果快

缺点

- 资料积累时未受到研究者的控制，内容上未必符合要求
- 需要足够完整可靠的过去某段时间有关研究对象的暴露和结局的历史记录或档案材料

第二节 队列研究的设计与实施

- 一 确定研究因素**
- 二 确定研究结局**
- 三 确定研究现场和研究人群**
- 四 确定队列大小**
- 五 资料收集与随访**
- 六 质量控制**

一、确定研究因素

研究因素：暴露因素、其它影响结局的因素

暴露因素

选择：建立在大量研究基础上。

定义：

测量：定性、定量

敏感、精确、简单、可靠的方法

询问、实验室测定、查阅记录

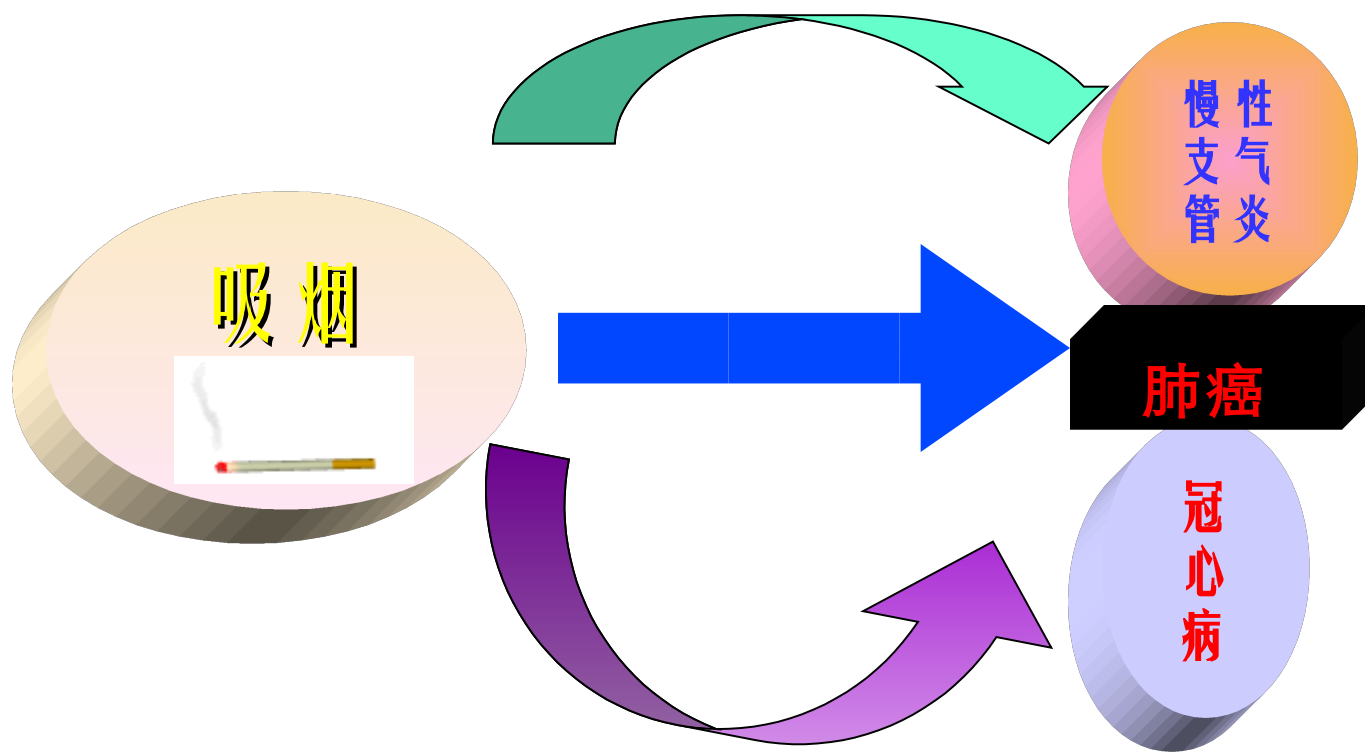
其它影响结局的因素

二、确定研究结局

结局指研究者预期的结果事件。

- 发病或死亡
- 某些血清指标
- 分子标志的变化
- 定性或定量

❖ 一次研究可有多个结局



❖ 结局的测量

采用国际或国内通用的标准

结果具备可比性

三 确定研究现场和人群

（一）确定研究现场

- 有足够符合条件的研究对象
- 领导重视、群众支持
- 医疗条件较好，交通较便利
- 发病率较高

（二）研究对象的选择

1. 选择研究对象的两种基本方式

- （1）按规定的暴露选择一组有暴露的人群，然后另外选择一组非暴露人群；这组非暴露人群称为外对照组。**
- （2）选择一个暴露因素分布不均匀的人群，将其一分为二，即分成暴露组和非暴露组，这个非暴露组叫内对照组。**

2. 暴露人群的选择

- (1) 职业人群
- (2) 特殊暴露人群
- (3) 一般人群
- (4) 有组织的人群团体

3. 对照人群的选择：

可比性原则

对照人群除未暴露于所研究的因素外，其它各种因素或人群特征应尽可能与暴露人群相同。

对照形式：

- （1）内对照：在同一研究人群中，无暴露或暴露水平最低的人员。
- （2）特设对照（外对照）
- （3）总人口对照
- （4）多重对照

四、确定样本大小

（一）影响样本大小因素

- ❖ 抽样方法
- ❖ 暴露组与非暴露组的比例
- ❖ 失访

(二) 决定影响样本含量参数：

- (1) 一般人群中所研究疾病的发病率水平 p_0 :
 p_0 越接近 0.5 , 所需观察的人数越少。
- (2) 暴露人群的发病率 p_1 : $d = p_1 - p_0$, d
值越大, 所需观察人数越少。
- (3) 显著性水平: α : 0.05 或 0.01
- (4) 把握度: $1-\beta$: 0.9 , 0.8 , 0.7

公式

$$N = \frac{\left(Z_{\alpha} \times \sqrt{2\bar{P}(1-\bar{P})} + Z_{\beta} \times \sqrt{P_1(1-P_1) + P_0(1-P_0)} \right)^2}{(P_1 - P_0)^2}$$

\bar{P} : 两个发病率的平均

P_1 值暴露组预期发病
率

P_0 : 对照组预期发病

条件 暴露组和对照组样本含量相等

例： $p_0=0.007$, $RR=2.5$, $\alpha=0.05$, $\beta=0.1$, 计算样本量

根据上述条件

$$Z_{\alpha}=1.96 , \quad Z_{\beta}=1.282 , \quad q_0=0.993$$

$$p_1=p_0 \times RR=0.0175 , \quad q_1=0.9825$$

$$\text{代入公式得} \quad N=2310$$

考虑到失访，实际上每组需扩大 10% ，即 2541 人

- 查表法

条件

已知 α 、 β 、 p_0 (对照组发病率, 表中为 p_2) 、 RR

群组研究的样本量(1)

 $\alpha = 0.05$ (双侧), $\beta = 0.10$

RR	P_2								
	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
1.1	*****	*****	732510	548803	438579	365096	312609	273243	242625
1.2	570732	288047	191819	143705	114837	96591	81844	71534	63515
1.3	267961	133825	89113	66757	53344	44401	38014	33223	29497
1.4	157273	78541	52297	39175	31302	26053	22304	19492	17305
1.5	104843	52355	34859	26111	20862	17363	14863	12989	11531
1.6	75715	37807	25171	18853	15063	12535	10730	9376	8323
1.7	57764	28842	19201	14381	11489	9560	8183	7150	6347
1.8	45861	22897	15243	11415	9119	7588	6495	5674	5037
1.9	37527	18735	12471	9339	7460	6207	5313	4641	4119
2.0	31443	15697	10448	7824	6249	5199	4450	3887	3450
2.5	16299	8134	5413	4052	3235	2691	2302	2011	1784
3.0	10475	5226	3476	2601	2076	1727	1477	1289	1143
3.5	7539	3760	2500	1870	1493	1241	1061	926	821
4.0	5815	2899	1927	1441	1150	955	816	712	631
4.5	4698	2342	1556	1163	928	770	658	574	509
5.0	3923	1954	1298	970	773	642	548	478	424
6.0	2927	1457	967	722	576	478	408	355	314
7.0	2321	1155	766	572	455	377	322	280	248
8.0	1917	953	632	471	375	311	265	230	203
9.0	1630	810	536	400	318	263	224	195	172
10.0	1416	703	465	346	275	228	194	168	148
11.0	1250	620	410	305	242	200	170	148	130
12.0	1118	555	367	273	216	179	152	132	116
13.0	1011	501	331	246	195	161	137	118	104
14.0	923	457	302	224	177	146	124	107	94
15.0	848	420	277	205	162	134	113	98	86
16.0	784	388	256	189	150	123	104	90	79
17.0	729	360	237	176	139	114	97	84	73
18.0	682	336	221	164	129	106	90	78	68
19.0	639	315	207	153	121	99	84	72	63
20.0	602	297	195	144	114	93	79	68	59

群组研究的样本量(2)

$\alpha = 0.05$ (双侧), $\beta = 0.10$

RR	P_2								
	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.045	0.050
1.1	218131	144649	107907	85862	71166	60668	52795	46672	41773
1.2	57100	37854	28231	22458	18608	15859	13797	12193	10910
1.3	26517	17574	13103	10420	8632	7354	6396	5651	5055
1.4	15555	10307	7682	6108	5058	4308	3746	3308	2958
1.5	10364	6865	5116	4066	3366	2866	2491	2200	1968
1.6	7481	4954	3690	2932	2427	2066	1795	1584	1416
1.7	5704	3776	2812	2234	1848	1573	1368	1205	1077
1.8	4526	2995	2230	1771	1465	1246	1082	954	852
1.9	3702	2449	1823	1447	1196	1017	883	779	695
2.0	3100	2050	1525	1210	1000	850	738	651	581
2.5	1602	1058	786	623	514	436	378	332	296
3.0	1027	677	502	397	327	277	239	210	187
3.5	737	485	359	283	233	197	170	149	132
4.0	567	372	275	217	178	150	129	113	100
4.5	456	299	221	173	142	120	103	90	79
5.0	380	249	183	144	117	99	85	74	65
6.0	282	184	135	105	86	72	61	53	46
7.0	222	144	105	82	66	55	47	40	35
8.0	182	118	86	66	53	44	37	32	28
9.0	154	99	72	55	44	37	31	26	22
10.0	133	85	61	47	38	31	26	22	19
11.0	116	74	53	41	32	26	22	18	15
12.0	103	66	47	36	28	23	19	16	13
13.0	93	59	42	32	25	20	16	13	11
14.0	84	53	37	28	22	17	14	11	9
15.0	77	48	34	25	20	15	12	10	8
16.0	70	44	31	23	18	14	11	9	7
17.0	65	40	28	21	16	12	10	7	6
18.0	60	37	26	19	14	11	8	6	5
19.0	56	35	24	17	13	10	7	6	4
20.0	53	32	22	16	12	9	7	5	

注意事项：

- （1）暴露组与对照组的比例：通常采用两组等量的方法。**
- （2）通常按 10% 来估计失访率，故计算出来的样本量再加 10% 做为实用样本量。**

五、资料的收集——随访或追踪

❖ 基线资料

❖ 随访

❖ 资料收集

❖ 基线 (base-line) 资料

随访研究对象之前，研究人员要获得有关研究对象的多种必要的确实材料做为判断研究终点和分析比较的基础，这称为基线资料。

1. 基线资料主要内容：

1) 研究对象暴露于研究因素的情况。

2) 与判断终点有关的资料：

例如观察终点是糖尿病，则全部对象要测空腹血糖或糖耐量等以排除糖尿病患者。

3) 有关协变量的资料：

即暴露于其它与患病有关的危险因素的情况。

2. 获取基线资料的方式

- 1) 查阅记录或档案
 - 2) 访问研究对象或其它能够提供信息的人
 - 3) 对研究对象进行测定或检查
 - 4) 有时要对环境作调查与检测
- 目的是确定暴露，如对水质进行化验。

❖ 随访 (follow-up)

1. 随访对象、内容和方法
2. 观察终点
3. 观察终止时间
4. 随访问隔

1. 随访的对象、内容和方法

对象：

暴露组和对照组同等对待。

随访内容：

- 1) 确定研究对象是否仍处于观察之中。
- 2) 确定研究人群中的结局事件。
- 3) 进一步收集有关暴露和混杂因素的资料，以备分析资料时用。

方法：

利用记录或档案；特殊随访：定期体检、访问等；对环境做调查或检测

2. 观察终点和终止时间

（1）观察终点：

指观察对象出现了预期的结局。

（2）观察终止时间：

指整个研究工作观察的截至时间。即随访期，可根据潜隐期确定。

3. 随访问隔

观察期长短

结局变化的速度

人力、物力

例如： 1-2 年

六 质量控制

❖ 调查员选择

❖ 调查员培训

❖ 制定调查员手册

❖ 监督

第三节 资料分析

- 一、资料整理表
- 二、率的计算
- 三、率的显著性检验
- 五、效应估计

一、资料整理表

	病例	非病例	合计
暴露组	a	b	a+b
非暴露组	c	d	c+d
合计	a+c	b+d	a+b+c+d

暴露组发病率 = $a / (a+b)$

对照组发病率 = $c / (c+d)$

如果二者差异显著，并且无明显偏倚，则暴露因素与疾病之间有可能存在因果关系

二、 率的计算

- ❖ 累积发病率
- ❖ 发病密度
- ❖ 标化死亡比
- ❖ 标化比例死亡比

累积发病率 (cumulative incidence)

$$CI = \frac{\text{观察期内发病（或死亡）人数}}{\text{观察开始时的人口数}}$$

适用条件： 样本大，人口稳定，整齐的资料
固定队列

发病密度 (incidence density)

$$ID = \frac{\text{观察期内发病（或死亡）人数}}{\text{观察人时}}$$

适用条件： 观察时间长，人口不稳定，动态队列

人时的计算

❖ 精确法

❖ 近似法

❖ 寿命表法

1. 以个体为单位计算暴露人年（精确法）

表 6-6 3 例出生日期与进出研究时间

编号	出生日期	进入研究时间	退出研究时间
1	1927.03.21	1966.07.19	1977.09.14 (迁居外地)
2	1935.04.09	1961.11.11	1973.12.01 (死亡)
3	1942.11.12	1970.02.01	1981.01.01 (观察结束时健在)

表 6-7 3 例人年的计算

年龄组	对象 1 1927 年 3 月 21 日出生	对象 2 1935 年 4 月 9 日出生	对象 3 1942 年 4 月 12 日出生	暴露人年
25~		61.11.11~65.04.08 共 3 年 4 个月 27 天, 合 3.41 人年	70.02.01~72.11.11 共 2 年 9 个月 10 天, 合 2.78 人年	6.19
30~		65.04.09 ~ 70.04.08 共 5.00 人年	72.11.12 ~ 77.11.11 共 5.00 人年	10.00
35~	66.07.19~67.03.20 共 8 个月, 合 0.67 人年	70.04.09~73.12.01 共 3 年 7 个月 22 天, 合 3.65 人年	77.11.12~81.01.01 共 3 年 1 个月 20 天, 合 3.14 人年	7.46
40~	67.03.21 ~ 72.03.20 共 5.00 人年			5.00
45~	72.03.21 ~ 77.03.20 共 5.00 人年			5.00
50~54	77.03.21~77.09.14 共 5 个月 24 天, 合 0.48 人年			0.48
累计	66.07.19 ~ 77.09.14 共 11.15 人年	61.11.11 ~ 73.12.01 共 12.06 人年	70.02.01 ~ 81.01.0 共 10.92 人年	34.13 人年

对象 1		67	03	20
	-	66	07	19
		<hr/>		
			08	01

$$0 + 8/12 + 01/30/12 = 0.67 \text{ 人年}$$

2. 近似法（大样本或不知道队员精确的进入、退出时间，精确度要求不高时）

表 6-8 人年计算举例

年龄（岁）	观 察 人 数						人年数
	1951.11.01	1952.11.01	1953.11.01	1954.11.01	1955.11.01	1956.04.01	
35~	8836	9149	9287	9414	9710	9796	41211
45~	7117	7257	7381	7351	7215	7191	32156
55~64	4094	4212	4375	4601	5057	5243	19909
合计	20097	20618	21043	21366	21982	22230	93276

（引自：耿贯一主编《流行病学》第四版，p54）

3. 寿命表法：大样本，精度要求高

$$L_x = I_x + 1/2(N_x - D_x - W_x), \quad I_{x+1} = I_x + N_x - D_x - W_x$$

表 7.6 应用寿命表法计算人年数

观察时间 (年数)	年初人数	年内进人数	年内发病数	年内失访人数	暴露人年数
x (1)	I _x (2)	N _x (3)	D _x (4)	W _x (5)	L _x (6)
1	1403	79	4	30	1425.5
2	1448	45	2	11	1464.0
3	1480	60	3	8	1504.5
4	1529	5	2	19	1521.0
5	1513	10	7	25	1502.0
6	1491	18	8	29	1481.5
7	1472	13	3	73	1440.5
8	1409	12	5	74	1375.5
9	1342	9	2	467	1112.0
10	882	3	1	819	473.5
11	65	0	0	57	36.5

标准化死亡比

(standardized mortality ratio SMR) (1)

$$SMR = \frac{\text{研究人群中观察发病（死亡）数}}{\text{标准人口（全人口）预期发病（死亡）数}}$$

适用条件：结局事件的发生率比较低，无论观察的时间长或短，都不宜直接计算率时用

标化死亡比 (SMR) (2)

预期发病（死亡）数的计算

以全人口某病的发病（死亡）率乘以某单位 人口数

- **例：**某厂 30-40 岁组工人有 500 名，某年有 2 人死于肺癌，已知该年全人口 30-40 岁组肺癌的死亡率为 2%，求 SMR
- 观察死亡数 = 2
- 预期死亡数 = $500 \times 2\% = 1$
- **SMR = $2/1=2$**
- **结论：**该厂 30-40 岁组工人死于肺癌的危险是一般人群的 2 倍。

标准化死亡比 (SMR) (3)

意义

被研究人群发生（死于）某病的危险性是标准

人群的多少倍
 $SMR=1$

研究人群某病发病（死亡）危险等于标准人群

$SMR>1$

研究人群某病发病（死亡）危险大于标准人群，
是标准人群的 SMR 倍

$SMR<1$

研究人群某病发病（死亡）危险小于标准人群

标化比例死亡比（1）

（ standardized proportional mortality ratio, SPMR ）

$$SPMR = \frac{\text{实际死亡数}}{\text{预期死亡数}}$$

适用条件

不能得到历年人口资料，仅有死亡人数、原因、日期和年龄

标化比例死亡比（2）

预期死亡数的计算

以全人口中某病因死亡占全部死亡之比例
乘以某单位实际全部死亡数

- **例：**某厂 30-40 岁组工人死亡总数为 100 名，某年有 5 人死于肺癌，已知该年全人口 30-40 岁组肺癌的死亡占全部死亡的比例为 2.2%，求 SPMR
- 观察死亡数 = 5
- 预期死亡数 = $100 \times 2.2\% = 2.2$
- $SPMR = 5 / 2.2 = 2.27$
- **结论：**该厂 30-40 岁组工人死于肺癌的危险是一般人群的 2.27 倍。

四、 率的显著性检验

- ❖ U 检验
- ❖ 直接概率法
- ❖ 二项分布检验
- ❖ 泊松（ Poisson ） 分布检验
- ❖ 卡方检验
- ❖ 计分检验（ score test ）

五、效应估计

- ❖ 相对危险度
- ❖ 归因危险度
- ❖ 归因危险度百分比
- ❖ 人群归因危险度
- ❖ 人群归因危险度百分比
- ❖ 剂量反应关系

相对危险度， RR (relative risk)

$$RR = \frac{I_e}{I_o} = \frac{a/n_1}{c/n_0}$$

I_e : 暴露组率

I_o : 非暴露组率

意义

- 暴露组发病或死亡的危险是非暴露组的多少倍

RR 与关联强度

RR		<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
0.9~1.0	1.0~1.1	<input type="checkbox"/>
0.7~0.8	1.2~1.4	<input type="checkbox"/>
0.4~0.6	1.5~2.9	<input type="checkbox"/>
0.1~0.3	3.0~9.9	<input type="checkbox"/>
<input type="checkbox"/> 0.1	10~	<input type="checkbox"/> <input type="checkbox"/>

RR95%CI

Woolf 法

$$Var(InRR) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}$$

lnRR □ 95%CI

$$l_n RR \pm 1.96\sqrt{Var(\ln RR)}$$

反自然对数即为 RR 95%CI

归因危险度, AR (Attributable risk)

$$AR = I_e - I_0 = \frac{a}{n_1} - \frac{c}{n_0} \quad \text{或}$$

$$AR = RR \times I_0 - I_0 = I_0 (RR - 1)$$

意义

- 暴露与非暴露人群比较, 所增加的疾病发生数量
- 在暴露组人群中完全由暴露因素造成的率
- AR 值越大, 暴露因素消除后所减少的疾病数量越大

RR 与 AR 的区别

1/10

疾病			RR	AR
肺癌	50.12	4.69	10.7	45.43
心血管疾病	296.75	170.32	1.7	126.43

意义

- RR 吸烟对肺癌的病因学意义较大
- AR 戒烟对心血管疾病的预防作用较大，
即公共卫生意义较大

归因危险度百分比 AR%

(病因分值 EF)

$$AR\% = \frac{I_e - I_o}{I_e} \times 100\% \quad \square$$

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

意义

暴露人群中的发病或死亡归因于
暴露的部分占全部发病或死亡百
分比

人群归因危险度

population attributable risk PAR

$$PAR = I_t - I_o$$

I_t : 总人群率

I_o : 非暴露组率

意义

- 暴露人群与一般人群比较，所增加的疾病发生率的大小
- PAR 值越大，暴露因素消除后所减少的疾病数量越大

人群归因危险度百分比 PAR%

(人群病因分值)

$$PAR\% = \frac{I_t - I_0}{I_t} \times 100\% \quad \text{或}$$

$$PAR\% = \frac{P_e(RR - 1)}{P_e(RR - 1) + 1} \times 100\%$$

Pe : 总人群的暴露比例

意义

PAR 占总人群全部发病（或死亡）的百分比

剂量反应关系

分析方法

- 列出不同暴露水平下的发病率
- 以最低暴露水平组为对照，计算各暴露水平的
RR 和 AR
- 趋势性检验

血清胆固醇	人数	病例数	危险度	平均年发病率	RR	AR
<210	454	16	0.0352	0.0059	1.00	0.0000
210~	455	29	0.0637	0.0106	1.81	0.0285
>245	424	51	0.1203	0.0200	3.39	0.0851
合计	1333	96	0.0720	0.0120		

结果

随着血清胆固醇水平的升高，其患冠心病的 RR 增大，说明存在**剂量效应关系**

例：吸烟者肺癌年死亡率 $I_e = 0.96\%$

非吸烟者肺癌年死亡率 $I_o = 0.07\%$

全人群肺癌年死亡率 $I_t = 0.56\%$

$$\text{RR} = I_e / I_o = 0.96\% / 0.07\% = 13.7$$

$$\text{AR} = I_e - I_o = 0.96\% - 0.07\% = 0.89\%$$

$$\text{AR}\% = (I_e - I_o) / I_e \times 100\% = (0.96\% - 0.07\%) / 0.96\% = 92.7\%$$

$$\text{PAR} = I_t - I_o = 0.56\% - 0.07\% = 0.49\%$$

$$\text{PAR}\% = (I_t - I_o) / I_t \times 100\% = (0.56\% - 0.07\%) / 0.56\% = 87.5\%$$

五、得出结论：

有无联系

危险因素、保护因素

联系强度

剂量反应关系

第四节 常见偏倚及其控制

选择偏倚

信息偏倚

混杂偏倚

选择偏倚

selection bias

**研究人群在一些重要因素方面与一般人群
或待研究的总体人群存在的差异**

产生原因

- 选择对象的方法不当
- 最初选定参加研究的对象中有人拒绝参加
- 历史性队列研究中部分档案丢失或记录不全
- 志愿者队列
- 研究开始时未能发现早期病人等

控制

预防为主，抽样方法正确，严格按照规定标准选择对象

失访偏倚

lost to follow-up

研究对象因迁移、外出、死于非终点疾病
或拒绝继续参加观察而退出队列所引起的
偏倚

控制

设计时

选择便于随访的人群

在计算的研究样本的基础上扩大 10%

实施时

加强对随访员的管理

制定随访计划和监测措施

整理资料时

对于有缺项或漏项的对象进行补查

失访率 $> 10\%$ 时分析失访的作用；

失访率 $> 20\%$ ，结果真实性可疑

信息偏倚

information bias

在获取暴露、结局或其他信息时所出现的系统误差或偏差

产生原因

- 疾病、暴露标准不明确
- 检验仪器不精确、检验技术不熟
- 询问技巧不佳记录错误，造假等

控制

- 提高临床诊断技术、明确各项标准
- 选择精确稳定的测量方法
- 事前调准仪器
- 严格实验操作规程
- 同等对待每个研究对象
- 培训调查员，提高技巧，统一标准

混杂偏倚

confounding bias

与所研究因素和结果均有联系的第三因素在暴露组与对照组的分布不均衡，混淆了研究因素和结果间的真实联系

控制

研究设计阶段

限制研究对象，匹配

分析阶段

分层分析、标准化或多因素分析

第四节 优缺点

优 点

- ❖ 直接获得暴露组和非暴露组的发病率或死亡率
- ❖ 直接估计危险度
- ❖ 符合时间顺序，验证病因的能力较强
- ❖ 获得一种暴露与多种结局的关系
- ❖ 收集的资料完整可靠，不存在回忆偏倚
- ❖ 可研究疾病的自然史

缺 点

- ❖ 不适于发病率很低的疾病的病因研究
- ❖ 易发生失访偏倚
- ❖ 耗时，耗人力、物力、财力
- ❖ 设计要求严密，资料的收集和分析难度较大
- ❖ 随访过程中，已知变量的变化或未知变量的引入增加分析难度

下课了

