

# 全国高等学校预防医学专业第七轮 卫生部规划教材目录

## 1. 卫生统计学 第7版

主编：方积乾 副主编：徐勇勇 陈峰

## 2. 流行病学 第7版

主审：李立明 主编：詹思延 副主编：叶冬青 谭红专

## 3. 卫生化学 第7版

主编：郭爱民 副主编：杜晓燕 毋福海 和彦玲

## 4. 职业卫生与职业医学 第7版

主编：孙贵范 副主编：邬堂春 牛侨

## 5. 环境卫生学 第7版

主编：杨克敌 副主编：郑玉建

## 6. 营养与食品卫生学 第7版

主编：孙长颢 副主编：凌文华 黄国伟

## 7. 儿童少年卫生学 第7版

主编：季成叶 副主编：陶芳标 武丽杰

## 8. 毒理学基础 第6版

主审：庄志雄 主编：王心如 副主编：孙志伟 陈雯

## 9. 卫生微生物学 第5版

主编：张朝武 副主编：邱景富

## 10. 社会医学 第4版

主编：李鲁 副主编：吴群红

## 11. 卫生事业管理学 第3版

主编：梁万年 副主编：饶克勤 王亚东

## 12. 卫生经济学 第3版

主编：程晓明 副主编：罗五金 刘国祥

### （三）疾病病因和危险因素的研究

这是为了达到预防疾病的目的而必须进行的工作。因为只有透彻地了解疾病发生、多发或流行的原因才能更好地防制乃至消灭某一疾病，也就是说流行病学必定要有发掘病因及疾病危险因素的工作。

有些疾病的病因是单一的，如传染病中的麻疹，利刃作用的割伤等。有些却不这样单纯，非传染性疾病就是由多种因素综合作用的结果。如高血压、高血脂、吸烟、肥胖等，这些都是冠心病的危险因素。其实，对于很多传染病也是如此。流行病学的主要用途之一就是尽量逐个澄清这些危险因素。有时，真正的病因尚未完全被阐明，而诸多危险因素已被发掘出来，据此防制疾病仍可收到很好的效果。如吸烟可致肺癌，但吸烟只是肺癌的一个危险因素，病因可能是烟草中的某个成分；尽管如此，控制吸烟仍能有效地预防肺癌。因此，流行病学工作不拘泥于非找到病因不可，若找到一些关键的危险因素，也能在很大程度上解决防病的问题。这是很实际的，是流行病学应用中的一大特点。

流行病学工作常常遇到“未明原因”（指一时原因不明，不意味着原因根本不能查明）疾病的调查。这些疾病是突然暴发或是短时期内多发的，而临床医务人员一时不能作出诊断。以流行病学观点，采取流行病学调查分析的方法，再配合临床检查和检验，从寻找危险因素入手，最终这类暴发大多都能找到原因。这种例子很多。如1957年，某市暴发了预防注射后引起的接种者四肢瘫痪，证明是由鼠脑制作的乙型脑炎疫苗所引起；1958年，新疆的“察布查尔病”证明是肉毒杆菌毒素引起的中毒；1959年，若干地区出现的“烧热病”由长期进食生棉籽油引起；1972年，上海大规模的皮炎流行由桑毛虫引起。有的暴发虽然原因未明，也能得到控制。如1980年，河南、湖北等省中学生中发生的“红斑性肢痛症”，其原因尚未查明，但已得到了控制；20世纪50年代发生的克山病，经数十年的研究，并未最终阐明病因，但疾病已大幅度下降。有关病因与因果推断的内容详见第九章。

### （四）疾病的自然史

该应用是通过流行病学方法研究人类疾病和健康的发展规律，以进一步应用于疾病预防和健康促进。疾病在个体中有一个自然发展过程，如亚临床期、症状早期、症状明显期、症状缓解期、恢复期。在传染病中有潜伏期、前驱期、发病期、恢复期。这是个体的疾病自然史。疾病在人群中也有其自然发生的规律，称为人群的疾病自然史。如对慢性肝炎或迁延性肝炎病人进行定期随访，研究其转归状况和规律，有助于采取有效措施以促进恢复健康。再如，儿童血压轨迹的研究，定期随访儿童血压至成人，考查血压有无轨迹现象以及血压的变化是否受年龄、性别和其他因素的影响。类似的工作还有许多。自然史研究既有理论意义也有实际意义。如通过自然史观察，我们了解到乙型肝炎有很大可能通过孕妇垂直传播给新生儿，故采用接种疫苗来实现早期预防，收到了良好的效果。

### （五）疾病防治的效果评价

这涉及防治疾病效果的最终判断。如观察儿童接种某种疫苗后，是否阻止了相应疾病的发生，可用实验流行病学的方法比较接种儿童和对照儿童的发病情况。又如考察一种新药是否安全有效，除在医院完成三期临床试验并上市后，仍需在大规模的社区人群中长期观察才能作出定论，尤其对药物不良反应的观察，更需要上市后的监测，此即药物流行病学（详见第17章）。在社区中实行大规模干预，如饮水加氟以防龋齿，减少吸烟以降低肺癌等疾病，也需使用流行病学实验方法去评价。类似的评价也用于卫生工作或卫生措施效

## 一、散发

散发 (sporadic) 是指发病率呈历年的一般水平, 各病例间在发病时间和地点上无明显联系, 表现为散在发生。散发一般是对于范围较大的地区而言。确定散发时多与当地近三年该病的发病率进行比较, 如当年发病率未明显超过既往平均水平称为散发。

当疾病预防与控制有效时, 会呈现散发, 常见于如下情况:

1. 病后免疫力持久的疾病, 或因预防接种使人群维持一定免疫水平的疾病常呈散发, 如麻疹。
2. 有些以隐性感染为主的疾病, 常以散发形式存在, 如脊髓灰质炎、乙型脑炎等。
3. 有些传播机制不容易实现的传染病也可出现散发, 如斑疹伤寒、炭疽等。
4. 某些长潜伏期传染病也以散发形式存在, 如麻风。

## 二、暴发

暴发 (outbreak) 是指在一个局部地区或集体单位中, 短时间内突然发生很多症状相同的病人。这些人多有相同的传染源或传播途径。大多数病人常同时出现在该病的最短和最长潜伏期之间。如托幼机构的麻疹、手足口病、腮腺炎、甲型病毒性肝炎等疾病的暴发。

## 三、流行

流行 (epidemic) 是指在某地区某病的发病率显著超过该病历年发病率水平。相对于散发, 流行出现时各病例之间呈现明显的时间和空间联系, 如 2009 年甲型 H1N1 流感的流行表现出明显的人与人间的传播关系和地域间的播散特征。当某地出现某种疾病的流行时, 提示当地可能存在共同的传播因素。

## 四、大流行

某病发病率显著超过该病历年发病率水平, 疾病蔓延迅速, 涉及地区广, 在短期内跨越省界、国界甚至洲界形成世界性流行, 称之为大流行 (pandemic)。疾病世界大流行的危险始终存在, 如流感、霍乱就有过多次世界性大流行。2009 年甲型 H1N1 流感形成世界大流行, 是近三十年来流行规模最大的, 其原因是甲型流感病毒变异。随着世界经济的快速发展, 交通日益便捷, 人群与物资流动的频度和速度是空前的, 病原体和传染源的快速移动会使某种疾病短时间传遍全球, 因而疾病大流行的危险始终存在, 要提高认识, 警钟长鸣。

# 第三节 疾病的分布

由于致病因子、人群特征以及自然、社会环境等多种因素综合作用的影响, 疾病在不同人群、不同地区及不同时间的流行强度不一, 存在状态也不完全相同。疾病的分布既反映了疾病本身的生物学特性, 也集中体现了与疾病有关的各种内外环境因素的效应及其相互作用的特点。疾病的流行特征通过疾病在人群、地区、时间的分布得以表现。对于已知病因的疾病, 流行特征是判断和解释病因的依据。对于病因未明的疾病, 流行特征是病因

该例子中,若随机抽取 1 万人作为调查对象,则按照 30/10 万的现患率估算,调查期望得到的病例数为 3 例。查表 3-2 可知,当期望病例数为 3 时,其 95% 可信限下限为 0.619,上限为 8.77;也就是说,若样本数为 1 万人时,调查结果中可能一个病例也不出现,使调查工作失去了意义。若要使调查结果至少有 1 例或 1 例以上的病例出现,查表 3-2 可知,95% 可信限下限为 1.09 时,期望病例数为 4 例。要达到调查结果中至少有 4 例结肠直肠癌病人出现,则有  $4:X=30:10$  万的比式成立,故  $X=4/30 \times 10$  万 = 13 334 人;即要在 95% 可信限上获得该县结肠直肠癌现患率的样本估计数据,至少应抽样调查 13 334 人。在实际操作时,可适当扩大一些样本量,以免估计的现患率(本例中为 30/10 万)与目标人群的现患率有误差而造成样本量不足。此外,上述方法一般适用于单纯随机抽样的方法,而实际工作中,诸如恶性肿瘤现患率调查等常采用整群抽样的方法,此时,可在上述方法估算的样本量基础上再增加 1/2。

若抽样调查的分析指标为计量资料,则应按计量资料的样本估计公式来计算,公式如下:

$$n = \frac{4S^2}{d^2} \quad (\text{式 3-5})$$

上式中  $n$  为样本量,  $d$  为容许误差,  $S$  为总体标准差的估计值。从式 3-5 可看出,样本量大小与  $S$  的平方成正比,与  $d$  的平方成反比,故在实际应用中,若同时有几个数据可供参考时,  $S$  宜取大一点的值,这样不至于使估计的样本量  $n$  偏小。

2. 抽样方法 抽样可分为非随机抽样和随机抽样,前者如典型调查。随机抽样的样本获得须遵循随机化原则,即保证总体中每一个对象都有已知的、非零的概率被选入作为研究对象,以保证样本的代表性。若样本量足够大、调查数据可靠、分析正确,则可以把调查结果推论到总体。

常见的随机抽样方法有单纯随机抽样、系统抽样、分层抽样、整群抽样和多阶段抽样。

(1) 单纯随机抽样:单纯随机抽样 (simple random sampling) 也称简单随机抽样,是最简单、最基本的抽样方法。从总体  $N$  个对象中,利用抽签或其他随机方法(如随机数字)抽取  $n$  个,构成一个样本。它的重要原则是总体中每个对象被抽到的概率相等(均为  $n/N$ )。

单纯随机抽样的标准误按资料性质根据式 3-6 和式 3-7 计算。

均数的标准误:

$$S_{\bar{x}} = \sqrt{\left(1 - \frac{n}{N}\right) \frac{S^2}{n}} \quad (\text{式 3-6})$$

率的标准误:

$$S_p = \sqrt{\left(1 - \frac{n}{N}\right) \frac{p(1-p)}{n-1}} \quad (\text{式 3-7})$$

式中:  $S$  为样本标准差;  $p$  为样本率;  $N$  为总体含量;  $n$  为样本量;  $n/N$  为抽样比,若小于 5% 可以忽略不计。

在实际工作中,单纯随机抽样往往由于总体数量大,编号、抽样麻烦以及抽到个体分散而导致资料收集困难等原因实际应用的不多,但它是其他各种抽样方法的基础。

# 队列研究

## Chapter 4 Cohort Study

The cohort study is one of the most important analytical epidemiologic methods, which identifies and studies a subset of human subjects from a defined population. These subjects have been unexposed, or exposed, or exposed in varying intensities, to a factor or factors suspected of influencing the probability of occurrence of a disease or other health outcomes such as death. The subjects are then followed up for a sufficiently long period of time to collect data on the outcomes. The incidence rates can then be estimated and compared in groups that differ in exposure status; the results are used to test the relationship between a risk factor and an outcome. As opposed to case-control and cross-sectional studies, a cohort study starts from causes and observes the occurrence of outcome prospectively in time. Thus, cohort studies provide stronger evidence than other observational studies for the temporal order for a cause-effect association and are deemed the most rigorous epidemiological method for studying causes of disease. The cohort study is also used to investigate the natural history of disease, the ability of a diagnostic technique in predicting prognosis, and long-term rare adverse effects of drugs. The cohort study is also commonly known as prospective study, incidence rate study, follow-up study, and longitudinal study.

队列研究 (cohort study) 是分析流行病学 (analytical epidemiology) 研究中两大重要方法之一, 它通过直接观察暴露于某因素不同状况人群的结局来探讨该因素与所观察结局的关系。与之相关的名称还有前瞻性研究 (prospective study)、发生率研究 (incidence study)、随访研究 (follow-up study) 及纵向研究 (longitudinal study) 等。队列研究因其检验病因假设的能力较强而在流行病学病因研究中被广泛应用。

## 第一节 概 述

### 一、概念

队列研究是将人群按是否暴露于某可疑因素及其暴露程度分为不同的亚组, 追踪其各自的结局, 比较不同亚组之间结局频率的差异, 从而判定暴露因子与结局之间有无因果关联及关联大小的一种观察性研究 (observational study) 方法。这里观察的结局主要是与暴露因子可能有关的结局。

等区别。因此,妥善的解决办法是,既按国际或国内统一的标准判断结局,又按自定标准判断,准确记录下其他可疑症状或现象供以后分析时参考。

队列研究的优点之一是一次可以同时收集到多种结局资料,研究一因多果的关系,故在队列研究中除确定主要研究结局外,还可考虑同时收集多种可能与暴露有关的结局,提高一次研究的效率。

由于 PHCC 的病程短,病死率高,诊断明确,很少漏诊,因此,本实例选择的研究结局是 PHCC 死亡。

### 三、确定研究现场与研究人群

#### (一) 研究现场

由于队列研究的随访时间长,因此,队列研究的现场选择除要求有足够数量的符合条件的研究对象外,还要求当地的领导重视,群众理解和支持,最好是当地的文化教育水平较高,医疗卫生条件较好,交通较便利。能选择符合这些条件的现场,将使随访调查更加顺利,所获资料将更加可靠。当然,也要考虑现场的代表性。

本实例选择的研究现场是江苏省海门市,这里是 PHCC 的高发区,人口稠密,当地的领导重视,群众支持,文化教育水平较高,医疗卫生条件较好,交通便利,是开展队列研究的理想场所。

#### (二) 研究人群

研究人群包括暴露组 and 对照组,暴露组中有时还有不同暴露水平的亚组。根据研究目的和研究条件的不同,研究人群的选择有不同的方法。

1. 暴露人群 (exposure population) 的选择 暴露人群即暴露于待研究因素的人群。根据研究的方便与可能,通常有下列四种选择。

(1) 职业人群:如果要研究某种可疑的职业暴露因素与疾病或健康的关系,必须选择相关职业人群作为暴露人群。另外,由于职业人群有关暴露与疾病的历史记录往往较为全面、真实和可靠,故如果做历史性队列研究,也常选择职业人群为暴露人群。

(2) 特殊暴露人群:特殊暴露人群是研究某些罕见的特殊暴露的唯一选择,如选择原子弹爆炸的受害者,接受过放射线治疗的人,以研究射线与白血病的关系。

由于对某些职业暴露和某些特殊暴露的危险多半不是一开始就认识到的,一旦认识到了,大多都采取了防护措施以减少暴露,所以一般不宜或不允许进行前瞻性队列研究,而常使用历史性队列研究。

(3) 一般人群:即某行政区域或地理区域范围内的全体人群,选择其中暴露于欲研究因素的人做暴露组。在一般人群中选择暴露组,通常考虑两点:①不打算观察特殊人群发病的情况,而着眼于一般人群及今后在一般人群中的防治,使研究结果具有普遍意义;②所研究的因素和疾病都是一般人群中常见的,不必要选择特殊人群或没有特殊人群可寻,特别是在研究一般人群的生活习惯或环境因素时。美国弗明汉地区的心血管疾病队列研究就是一个很好的例子。

(4) 有组织的人群团体:该类人群可看作是一般人群的特殊形式,如医学会会员,工会会员,机关、社会团体、学校或部队成员等。选择这样的人群的主要目的是利用他们的组织系统,便于有效地收集随访资料。而且他们的职业和经历往往是相同的,可增



出现, 血脂升高或其他生物标志的出现等, 根据研究的要求不同而不同。对观察终点结果的判断应在设计中订出明确的标准, 规定明确的判断方法, 这种规定自始至终不能改变, 即使是实际医疗工作中已有所改变, 但在本研究中也不能改变, 以免造成疾病错分的误差。发现终点的方法要敏感、可靠、简单、易被接受。

4. 观察终止时间 观察终止时间是指整个研究工作截止的时间, 也即预期可以得到结果的时间。终止时间直接决定了观察期的长短, 而观察期长短是以暴露因素作用于人体至产生疾病结局的时间, 即潜伏期(或潜隐期)为依据的, 另外, 还应考虑所需的观察人年数。要在考虑上述两个因素的基础上尽量缩短观察期, 以节约人力、物力, 减少失访。观察时间过短, 可能得不出预期的结果; 但追踪时间越长, 失访率越高, 消耗越大, 结果可能也受影响。

5. 随访问隔 如果观察时间较短, 在观察终止时一次搜集资料即可。但如果观察时间较长, 则需多次随访, 其随访问隔与次数将视研究结局的变化速度、研究的人力、物力等条件而定。一般慢性病的随访问隔期可定为1~2年。如 Framingham 心血管病研究每两年随访一次, 至今已历时63年。

6. 随访者 根据随访内容的不同, 调查员可以是普通的询问调查者, 也可以是实验室的技术人员, 临床医生等, 但随访调查员必须认真进行培训。研究者可以参加随访, 但最好是不亲自参与, 因为研究者随访易于带来主观的偏倚, 而不知情的局外人士反而能够获取更客观的信息。

本实例的所有研究对象(包括 HBsAg 阳性者和阴性者)均在1992年进入队列, 从1993年起到2003年止, 每年随访一次, 观察终点是 PHCC 死亡, PHCC 的诊断依据临床表现、影像学 and 病理学检查结果。

## 六、质量控制

队列研究费时、费力、消耗大, 加强实施过程, 特别是资料收集过程中的质量控制显得特别重要, 一般的质量控制措施包括下列几点:

1. 调查员的选择 调查员应有严谨的工作作风和科学态度, 诚实可靠是调查员应具备的基本品质, 一般应具有高中或大学毕业文化程度, 并具有调查所需的专业知识。另外, 调查员的年龄、性别、种族、语言、社会经济地位等最好与研究对象相匹配, 这样的调查员更具有亲和力, 使调查易于进行。

2. 调查员培训 调查员的工作作风、科学态度、调查技巧与技术, 临床医生和实验技术人员的经验等都将直接影响调查结果的真实性和可靠性。因此, 在资料收集前, 应对所有参加调查者进行严格的培训, 掌握统一的方法和技巧, 并要进行考核。

3. 制定调查员手册 由于队列研究所涉及的调查员多, 跨时长, 因此编一本调查员手册, 内列全部操作程序、注意事项及调查问卷的完整说明等是十分必要的。

4. 监督 常规的监督措施包括: ①由另一名调查员作抽样重复调查; ②人工或用计算机及时进行数值检查或逻辑检错; ③定期观察每个调查员的工作; ④对不同调查员所收集的变量分布进行比较; ⑤对变量的时间趋势进行分析; ⑥在访谈时使用录音机录音等。应注意将监督结果及时反馈给调查员。

2. 用近似法计算暴露人年 如果不知道每个队列成员进入与退出队列的具体时间(精确到天), 就不能用上述方法直接计算暴露人年数; 另外, 如果对暴露人年计算的精确性要求不高时, 也没有必要应用精确法计算。此时, 都可应用近似法计算暴露人年, 即用平均人数乘以观察年数得到总人年数, 平均人数一般取相邻两年的年初人口的平均数或年中人口数。该法计算简单, 但精确性较差。

假设某队列研究从2006年8月10日开始, 开始观察人数为30 054人, 至2010年2月10日结束, 结束时的人数为30 023人, 资料详列于表4-7, 以说明其计算方法。表中35岁年龄组的人年数 = (9836 + 9847)/2 + (9847 + 9857)/2 + (9857 + 9843)/2 + (6/12) × (9843 + 9840)/2 = 34 464(人年), 余类推。

表 4-7 近似人年算法

| 年龄<br>(岁) | 观察人数       |            |            |            |            | 观察人<br>年数 |
|-----------|------------|------------|------------|------------|------------|-----------|
|           | 2006-08-10 | 2007-08-10 | 2008-08-10 | 2009-08-10 | 2010-02-10 |           |
| 35 ~      | 9836       | 9847       | 9857       | 9843       | 9840       | 34 464    |
| 45 ~      | 12 593     | 12 568     | 12 578     | 12 566     | 12 562     | 44 008    |
| 55 ~ 64   | 7625       | 7636       | 7643       | 7623       | 7621       | 26 714    |
| 合计        | 30 054     | 30 051     | 30 078     | 30 032     | 30 023     | 105 186   |

3. 用寿命表法计算人年 利用简易寿命表方法也可以计算人年。该法计算简单, 并有一定的精确度。常用的计算方法是规定观察当年内进入队列的个人均作 1/2 人年计算, 失访或出现终点结局的个人也作 1/2 人年计算。其观察人年数计算式如下:

$$L_x = I_x + \frac{1}{2}(N_x - D_x - W_x) \quad (\text{式 4-2})$$

$$I_{x+1} = I_x + N_x - D_x - W_x \quad (\text{式 4-3})$$

式中  $L_x$  为  $x$  时间内的暴露人年数,  $I_x$  为  $x$  时间开始时的观察人数,  $N_x$  为  $x$  时间内进入队列的人数,  $D_x$  为  $x$  时间内出现终点结局的人数,  $W_x$  为  $x$  时间内失访的人数。试以表 4-8 的资料为例, 说明其计算方法。

表 4-8 寿命表法计算人年实例

| 观察时间<br>(第 $x$ 年) | 年初人数<br>( $I_x$ ) | 年内进入人数<br>( $N_x$ ) | 年内发病人数<br>( $D_x$ ) | 年内失访人数<br>( $W_x$ ) | 暴露人年数<br>( $L_x$ ) |
|-------------------|-------------------|---------------------|---------------------|---------------------|--------------------|
| 1                 | 2903              | 123                 | 8                   | 64                  | 2928.5             |
| 2                 | 2954              | 115                 | 6                   | 76                  | 2970.5             |
| 3                 | 2987              | 42                  | 9                   | 18                  | 2994.5             |
| 4                 | 3002              | 35                  | 8                   | 34                  | 2998.5             |
| 5                 | 2995              | 0                   | 7                   | 18                  | 2982.5             |
| 合计                |                   |                     | 38                  |                     | 14 874.5           |

第一年的暴露人年数为



## 第二节 研究类型

### 一、病例与对照不匹配

在设计所规定的病例和对照人群中，分别抽取一定量的研究对象，一般对照数目应等于或多于病例人数。此外没有其他任何限制与规定。

### 二、病例与对照匹配

匹配 (matching) 或称配比，即要求对照在某些因素或特征上与病例保持一致，目的是对两组进行比较时排除匹配因素的干扰。如以年龄作匹配因素，使两组在年龄构成上类似或一样，在分析比较两组资料时，可避免由于两组年龄构成的差别对疾病和因素关系的影响，从而更正确地说明所研究因素与疾病的关系。匹配分为频数匹配与个体匹配。

1. 频数匹配 (frequency matching) 频数匹配首先应当知道或估计出匹配变量每一层的病例数，例如做年龄匹配，应当知道 20~24 岁组、25~29 岁组等各组的病例数，然后从备选对照中选择对照，直至达到每层所要求的数目，不一定要求绝对数相等，重要的是比例相同。例如，病例组中男、女各半，则对照组中也应一样。

2. 个体匹配 (individual matching) 以病例和对照个体为单位进行匹配叫个体匹配。1:1 匹配，即为每一个病例配一名对照，又称配对 (pair matching)，1:2、1:3、…、1:R 匹配时，直接称为匹配。

定量指标一般要求在一定范畴内匹配。例如年龄匹配，病例为 50~59 岁组，则对照亦应为 50~59 岁组。或者要求对照在  $\pm 2$  岁、 $\pm 3$  岁或  $\pm 5$  岁等范围内匹配，如要求对照与病例的年龄之差在  $\pm 3$  岁之内，则一个 39 岁的病例，其对照的年龄应当在 36~42 岁之间。匹配指标范围的大小应当根据可行性而定，在预实验 (pilot study) 中可以从较窄的范围开始，探求多大的范围最合适。很显然，范围越宽，两组的可比性就会越差，会造成较大的残余混杂 (residual confounding) 而达不到匹配的目的。

在病例对照研究中采用匹配的目的，首先在于提高研究效率 (study efficiency)。其次在于控制混杂因素的作用。所以匹配的特征或变量必须是已知的混杂因子，或有充分的理由怀疑为混杂因子，否则不应匹配。

匹配同时也增加了选择对照的难度。而且一旦对某个因素作了匹配，我们将不能再分析该因素与疾病的关系，也不能充分分析它与其他因素的交互作用。把不必要的项目列入匹配，企图使病例与对照尽量一致，就可能徒然丢失信息，增加工作难度，结果反而降低了研究效率。这种情况称为匹配过头 (over-matching)，应当注意避免。有两种情况不应使用匹配，否则会造成匹配过头。一是疾病因果链上的中间变量不应匹配。例如，吸烟对血脂有影响，而血脂与心血管疾病有病因关系，在研究吸烟与心血管病关系的病例对照研究中，按血脂水平对病例和对照进行匹配，则吸烟与疾病的关联可能消失。另一种是只与可疑病因有关而与疾病无关的因素不应匹配。例如，避孕药的使用与宗教信仰有关，但宗教信仰与研究的疾病并无关系，因此不应将宗教信仰作为匹配因素。换句话说，上述两种情况中提到的因素都不符合混杂因素的特征，所以不应用来匹配。

助于控制早期环境影响和遗传因素的混杂作用；配偶对照则主要考虑控制环境的影响。

为了保证研究的真实性，在选择对照时必须考虑对照的代表性、对照与病例的可比性，以及可能出现的选择偏倚等问题。就对照而言，没有哪一种对照一定优于其他对照，不同的对照要解决不同的问题。病例与对照和研究人群严格可比时，不存在选择偏倚。当对照来自产生病例的源人群，而且是一个有代表性样本时，即可做到这一点。这时研究人群不需要代表任何特定的人群。

### 3. 比较以社区为基础的和以医院为基础的病例对照研究（表 5-1）

表 5-1 以社区为基础的和以医院为基础的病例对照研究各自的相对优点

| 以社区为基础的病例对照研究             | 以医院为基础的病例对照研究            |
|---------------------------|--------------------------|
| (1) 可以较好地确定源人群            | (1) 研究对象的可及性好            |
| (2) 容易保证病例和对照来自于同一源人群     | (2) 研究对象更易合作             |
| (3) 对照的暴露史更可能反映病例源人群的暴露情况 | (3) 比较容易从医疗记录和生物标本收集暴露信息 |

## 四、确定样本量

### 1. 影响样本量的因素

(1) 研究因素在对照组中的暴露率  $p_0$ 。

(2) 预期的该因素的效应强度，即相对危险度 RR 或暴露的比值比 OR（其含义详见后文）。

(3) 希望达到的检验的显著性水平，即假设检验第 I 类错误的概率  $\alpha$ 。

(4) 希望达到的检验功效也即把握度  $(1 - \beta)$ ， $\beta$  为统计学假设检验第 II 类错误的概率。

2. 估计方法 不同匹配方式的样本量计算方法不同，除了利用公式计算外，还有现成的表可查。需要注意的是：①所估计的样本量并非绝对精确的数值，因为样本量的估计是有条件的，而这些条件并非是一成不变的。②样本量越大，结果的精确度越好，但是样本量过大，常会影响调查工作的质量，增加负担和费用。实际工作中应当权衡利弊。③总样本量相同的情况下，病例组和对照组样本量相等时统计学效率最高。

(1) 非匹配设计且病例数与对照数相等

$$n = 2\bar{p}\bar{q}(z_\alpha + z_\beta)^2 / (p_1 - p_0)^2 \quad (\text{式 5-1})$$

其中：

$$p_1 = \frac{p_0 RR}{1 + p_0 (RR - 1)}$$

$$\bar{p} = 0.5 \times (p_1 + p_0)$$

$$\bar{q} = 1 - \bar{p}$$

式中  $p_1$  为病例组的暴露率， $p_0$  为对照组的暴露率， $z_\alpha$  和  $z_\beta$  可查表 5-2 得到，也可直接查表 5-3 得到  $n$ 。

表 5-2 标准正态分布的分位数表

| $\alpha$ 或 $\beta$ | $z_\alpha$ (单侧检验) | $z_\alpha$ (双侧检验) |
|--------------------|-------------------|-------------------|
|                    | $z_p$ (单侧和双侧)     |                   |
| 0.001              | 3.09              | 3.29              |
| 0.005              | 2.58              | 2.81              |
| 0.010              | 2.33              | 2.58              |
| 0.025              | 1.96              | 2.24              |
| 0.050              | 1.64              | 1.96              |
| 0.100              | 1.28              | 1.64              |
| 0.200              | 0.84              | 1.28              |
| 0.300              | 0.52              | 1.04              |

表 5-3 病例对照研究样本量 (非匹配、两组人数相等)  
 $[\alpha=0.05$  (双侧),  $\beta=0.10]$

| RR   | $P_0$ |      |      |      |      |      |      |
|------|-------|------|------|------|------|------|------|
|      | 0.01  | 0.10 | 0.20 | 0.40 | 0.60 | 0.80 | 0.90 |
| 0.1  | 1420  | 137  | 66   | 31   | 20   | 18   | 23   |
| 0.5  | 6323  | 658  | 347  | 203  | 176  | 229  | 378  |
| 2.0  | 3206  | 378  | 229  | 176  | 203  | 347  | 658  |
| 3.0  | 1074  | 133  | 85   | 71   | 89   | 163  | 319  |
| 4.0  | 599   | 77   | 51   | 46   | 61   | 117  | 232  |
| 5.0  | 406   | 54   | 37   | 35   | 48   | 96   | 194  |
| 10.0 | 150   | 23   | 18   | 20   | 31   | 66   | 137  |
| 20.0 | 66    | 12   | 11   | 14   | 24   | 54   | 115  |

(Schlesselman, 1982)

例如, 拟进行一项病例对照研究, 研究吸烟与肺癌的关系。预期吸烟者发生肺癌的相对危险度为 2.0, 人群吸烟率 (此处以对照组的吸烟率代表) 为 20%, 设  $\alpha=0.05$  (双侧),  $\beta=0.10$ , 估计样本含量  $n$ 。

先求得  $p_1$ :

$$p_1 = (0.2 \times 2) / (1 + 0.2 \times 1) = 0.333$$

$$\bar{p} = (0.2 + 0.333) / 2 = 0.267$$

$$\bar{q} = 1 - 0.267 = 0.733$$

再用公式 5-1 求  $n$ :

$$n = 2 \times 0.267 \times 0.733 \times (1.96 + 1.282)^2 / (0.333 - 0.2)^2 = 232$$

即每组需要调查 232 人。

如查表 5-3, 得  $n=229$ 。

(2) 非匹配设计且病例数与对照数不等

设: 病例数: 对照数 = 1:  $c$ , 则需要的病例数

$$n = (1 + 1/c) \bar{p} \bar{q} (z_\alpha + z_\beta)^2 / (p_1 - p_0)^2 \quad (\text{式 5-2})$$

表 5-5 不同患病率时 OR 与 RR 的差异 (%)

| 非暴露组患病率 (%) | RR  |      |      |      |      |
|-------------|-----|------|------|------|------|
|             | 1.5 | 2.0  | 3.0  | 4.0  | 5.0  |
| 0.1         | 0.1 | 0.1  | 0.2  | 0.3  | 0.4  |
| 0.5         | 0.3 | 0.5  | 1.0  | 1.5  | 2.1  |
| 1.0         | 0.5 | 1.0  | 2.1  | 3.1  | 4.2  |
| 5.0         | 2.7 | 5.6  | 11.8 | 18.8 | 26.7 |
| 10.0        | 5.9 | 12.5 | 28.6 | 50.0 | 80.0 |

表 5-6 不同发病率和 OR 时的 RR

| OR | 非暴露组发病率 (I <sub>0</sub> ) |      |      |      |
|----|---------------------------|------|------|------|
|    | 0.20                      | 0.10 | 0.05 | 0.01 |
| 2  | 1.7                       | 1.8  | 1.9  | 2.0  |
| 3  | 2.1                       | 2.5  | 2.7  | 2.9  |
| 4  | 2.5                       | 3.1  | 3.5  | 3.0  |
| 5  | 2.8                       | 3.6  | 4.2  | 4.8  |
| 6  | 3.0                       | 4.0  | 4.8  | 5.7  |
| 7  | 3.2                       | 4.4  | 5.4  | 6.6  |
| 8  | 3.3                       | 4.7  | 5.9  | 7.5  |
| 9  | 3.5                       | 5.0  | 6.4  | 8.3  |
| 10 | 3.6                       | 5.3  | 6.9  | 9.2  |

1. 不匹配不分层资料 这是病例对照研究资料分析的基本形式。

(1) 每个暴露因素可整理成表 5-4 的四格表形式

例如，一项关于口服避孕药与心肌梗死的病例对照研究，结果如表 5-7：

表 5-7 口服避孕药 (OC) 与心肌梗死 (MI) 关系的病例对照研究结果

|       | 病例  | 对照  | 合计  |
|-------|-----|-----|-----|
| 服 OC  | 39  | 24  | 63  |
| 未服 OC | 114 | 154 | 268 |
| 合计    | 153 | 178 | 331 |

(2) 利用  $\chi^2$  (卡方) 检验，检验病例组与对照组两组的暴露率有无统计学显著性的差异

$$\chi^2 = \frac{(ad - bc)^2 n}{(a + b)(c + d)(a + c)(b + d)} = 7.70$$

的发病占全部发病的比例。换句话说就是，假如消除该暴露后，人群中发病降低的比例。

归因分值是具有公共卫生学意义的指标，它还反映了人群中随机抽取一个病例，可能因该暴露引起的概率。

### 三、功效

研究功效 (power) 也叫做把握度，可以解释为拒绝无效假设的能力，即当无效假设不成立时，该假设被拒绝的概率。以 1:1 匹配病例对照研究资料的功效估计为例。

例如，假定人群中暴露于所研究的危险因素的比例  $p_0 = 0.30$ ，统计学双侧检验的显著性水平  $\alpha = 0.05$ ，病例与对照各 50 例。计算该研究有多大的功效发现  $OR = 2$ 。

首先计算  $z$  值

$$z_\beta = \sqrt{\frac{n(p_1 - p_0)^2}{2\bar{p}\bar{q}}} - z_\alpha \quad (\text{式 5-21})$$

功效  $= 1 - \beta = P\{z \leq z_\beta\}$ 。

计算出  $z_\beta$  之后，根据标准正态分布查出小于  $z_\beta$  时的概率， $p_1$  与计算样本量时相同。

$$p_1 = \frac{p_0 RR}{1 + p_0(RR - 1)} = \frac{0.3 \times 2}{1 + 0.3 \times (2 - 1)} = 0.4615$$

本例中， $n = 50$ ， $p_0 = 0.30$ ， $\alpha = 0.05$  (双侧检验)， $z_\alpha = 1.96$ ， $OR = 2$ ，则：

$$\bar{p} = (p_0 + p_1) / 2 = (0.3 + 0.4615) / 2 = 0.3808$$

$$z_\beta = \sqrt{\frac{50 \times (0.4615 - 0.3)^2}{2 \times 0.3808 \times 0.6192}} - 1.96 = -0.297 \approx -0.30$$

查正态分布表，当  $z_\beta = -0.30$  时， $\beta = 0.62$ ，功效  $P = 1 - \beta = 38\%$ 。

结论：如果该研究选用 50 个病例和 50 个对照，在给定的条件下，该研究能检出  $OR$  显著地不等于 1 的概率为 38%，如果  $OR$  确实  $>1$  且  $\leq 2$ ，则该研究成功的希望不大。一般认为一项研究的检验功效应在 80% 以上。

以上计算方法没有考虑控制混杂因素 (如分层) 和评价交互作用的情况，因此所计算的研究功效只是一个粗略的估计，计算的结果可供设计阶段参考。

如果以不一致对子数 ( $n$ ) 来表达时，研究功效可用下式计算：

$$z_\beta = \frac{\left(n - \frac{1}{2}\right)\sqrt{n} - \frac{z_\alpha}{2}}{\sqrt{\pi(1 - \pi)}} \quad (\text{式 5-22})$$

公式中的  $\pi = OR / (1 + OR)$ 。该公式可检验已完成研究的研究功效。因为这里已经有了不一致对子数  $n$ 。

## 第五节 常见偏倚及其控制

病例对照研究是一种回顾性观察研究，比较容易产生偏倚。这些偏倚可以通过严谨的设计和细致的分析加以识别、减少和控制。常见的偏倚有选择偏倚、信息偏倚和混杂偏倚。

# 实验流行病学

## Chapter 6 Experimental Epidemiology

Experimental epidemiology is one of most important epidemiologic methods. It is an experiment in which the intervention is applied to individuals (patients or healthy people) and its effects are evaluated. The essential features of experimental epidemiology are prospective, in which the intervention is compared with a control and the allocation to treatment or control is randomized. Experimental studies include clinical trials, field trials and community trials. A randomized clinical trial is to evaluate the potential efficacy of a treatment on disease, in which the patients with interested disease are the subjects. Field trials are useful in the evaluation of preventive measures and usually organized in such as the communities and schools. The subjects in the field trials are often free of interested clinical outcome but at risk for that one. In the community trial, groups of people rather than individuals are allocated to receive an intervention or not. Well-conducted randomized clinical trial is the most robust and reliable experimental methodology for comparing preventive and therapeutic clinical interventions. However, ethical consideration is important.

实验流行病学是在 20 世纪发展起来的, 早期的实验流行病学主要是用动物在实验室模拟传染病的流行规律。随着医学科学研究的快速发展, 特别是在第二次世界大战后, 在英国的流行病学和统计学家 Bradford Hill 等学者的大力推动下, 实验性研究逐渐在流行病学研究中应用。这期间实验流行病学研究逐渐走出实验室, 以人群为对象, 以工厂、学校、医院或社区作为研究现场开展不同内容的研究工作。1938 年, 美国耶鲁大学的 Paul 教授第一次提出了“临床流行病学”这个术语, 建议把流行病学的思维应用到临床实践和研究中去。Bradford Hill 在 1951 年发表了题为“The clinical trial”的论文, 较系统地论述了临床试验的设计和分析问题。1962 年, 他又出版了第一部有关临床试验的著作《*Statistical methods in clinical and preventive medicine*》。此后, 很多学者将流行病学方法与临床实践紧密结合相继开展了多次临床试验研究, 并最终形成了临床流行病学这一分支学科。目前, 临床试验已成为实验流行病学的一个重要类别。

### 第一节 概 述

#### 一、定义

实验流行病学 (experimental epidemiology) 亦称流行病学实验 (epidemiological experi-



果、受试者可能被分配到不同的组别及可能发生的问题。受试者同意并在知情同意书上签字后，方可参加试验，不能强迫受试者。

#### （四）要科学评价临床疗效

研究者对临床试验的效果要进行实事求是的评价，科学评价应包括试验的真实性、重复性及实用性三个方面。

### 四、设计和实施

临床试验，尤其是随机对照试验（randomized controlled trial, RCT），强调将研究人群随机分为试验组与对照组，将研究者所控制的措施施加给试验人群后，随访观察并比较两组人群的结局，以判断干预措施的效果。因此，其结果真实可靠，但设计和实施复杂。试验步骤如下：

#### （一）制订试验计划

随机对照试验在实施前必须制定科学、严谨的研究计划。研究计划应包括以下几方面内容：

1. 明确试验的目的，即本研究要解决的问题是什么。
2. 明确试验对象的具体要求和来源。选择研究对象时，要使用统一的疾病诊断标准，并严格遵照执行；要注意研究对象的代表性，即研究对象的年龄、性别、病情、病变部位、病变的范围等。
3. 明确规定研究因素，如药物的剂量、用药的时间或疗程、给药途径等。
4. 确定观察指标，即反映干预措施效果的指标，通常确定1~2个主要结局指标，以及几个或十几个次要结局指标。选择研究指标时，最好选择客观的或定量的指标。
5. 确定随访观察时间及资料的收集方法。通常是根据药物的治疗效果、疾病严重程度和结局来确定药物治疗后随访观察的时间。
6. 资料收集后要进行整理和分析，设计时应说明要采用的统计分析方法。

#### （二）确定研究人群

在制订试验计划时，应首先考虑选择哪些人做研究人群。所谓研究人群是指符合研究对象入选标准的人群，包括试验组和对照组。选择研究对象时应注意以下几点：

1. 必须使用统一的人选和排除标准，以确保试验组和对照组的可比性。
2. 入选的研究对象应能从试验中受益，即参加研究，当试验结束时病人的疾病得到了有效的治疗或症状得到缓解。例如，以泌尿系统有感染的病人作为研究对象，首先要求对试验用药的作用机制有清楚的认识，即应该知道哪种抗生素是通过泌尿系统消灭细菌的。
3. 尽可能选择已确诊的或症状和体征明显的病人作研究对象。如研究治疗心肌梗死药物的药效时，试验对象最好是近期心绞痛频繁发作的病人，而不是病情稳定的病人，这样才能容易获得要观察的试验效应。
4. 尽可能不用孕妇作为研究对象。临床试验中，任何药物都有可能产生不良反应，有些药物的不良反应对正常人虽然很小，但对孕妇或对胎儿可能影响很大，所以试验前要充分估计不良反应的程度，衡量其得失后再选择恰当的对象。能不用孕妇作研究对象的，要尽量避免。

异和非特异效应，为了排除非特异效应的干扰，常用安慰剂作对照。安慰剂是指没有任何药理作用的物质，常用的物质有淀粉、乳糖、生理盐水等。使用安慰剂对照时，要注意两点。第一，要求安慰剂的剂型和外观尽量与试验药物相同，而且对人体无害，以便于进行盲法处理；第二，要掌握安慰剂的使用指征，应限在目前尚无有效药物治疗的疾病研究中，或虽然使用安慰剂，但对病人的病情和预后基本没有不良影响，否则不要应用安慰剂对照。

(3) 交叉对照 (crossover control): 这是一种特殊的对照方式，即按随机方法将研究对象分为甲、乙两组，甲组先用试验药，乙组先用对照药。一个疗程结束后，间隔一段时间以消除治疗药物的滞留影响，然后甲组再用对照药，乙组用试验药，最后分析和比较疗效。这样既能自身前、后对比，又可消除试验顺序带来的偏倚。两次治疗的间隔时间因疾病的症状或药物残留的时间长短而不同。设立此种对照的目的：①研究药物应用先后顺序对治疗结果的影响；②研究药物最佳配伍方案。

(4) 互相对照 (mutual control): 如果同时研究几种药物或治疗方法时，可以不设专门的对照，分析结果时，各组之间互为对照，从中选出疗效最好的药物或疗法。

(5) 自身对照 (self control): 试验中研究对象不分组，在同一研究对象中进行试验和对照，如比较用药前、后体内某些指标的变化情况，以判断药物的疗效；或研究皮肤科用药时使用左、右肢体作试验和对照，分析何种药物疗效更好。

### (五) 随机分组

研究对象和使用的对照方式确定之后，下一步的工作就是随机将研究对象分配到试验组与对照组。

1. 原则 随机化是为了使对照组与试验组具有可比性，提高研究结果的真实性，减少偏倚。每位研究对象被分配到试验组或对照组的机会相等，而不受研究者或受试者主观愿望或客观因素所影响。

2. 方法 在临床试验中常用的随机分组方法有三种：

(1) 简单随机法 (simple randomization): 最常用的方法是利用随机数字表或随机排列表，也可用抽签或抛硬币等方法。简单随机法的应用和操作简单，但分配到各组的样本量可能不等，如果组间样本量差异较大时有必要再次经随机化原则调整。多中心试验研究时，不适合用单纯随机法，因为可使各医院两组人数比例不等，甚至可相差悬殊，从而产生偏倚。

(2) 区组随机法 (block randomization): 当研究对象人数较少，而影响试验结果的因素又较多，简单随机化不易使两组具有较好的可比性时，可以采用区组随机化法进行分组。其基本方法是将条件（如年龄、性别、病情）相近的一组受试对象作为一个区组，每一区组内的研究对象（通常4~6例）数量相等，然后应用简单机分配方法将每个区组内的研究对象进行分组。该法的优点是在分组过程中，任何时刻试验组与对照组病例数保持相对一致，并可根据试验要求设计不同的区组。

(3) 分层随机法 (stratified randomization): 按照对治疗效果影响较大的特征（如年龄、性别、病情、临床分型等）进行分层，再运用简单随机化方法将每层内的研究对象分到治疗组和对照组。当影响因素较多时，如果不使用分层随机法，两组患者都包含很多影响疗效的因素，试验对象的个体差异较大，需要对较多的患者进行对比才能显示两组的差

别。而使用分层随机化方法能将影响疗效的因素按影响程度的大小依次分层加以考虑,就使得两组的临床特征比较相近,增加组间可比性,结论更可靠。分层随机法的优点是所用样本量小,效率高;缺点是分层不可太多,分层越多选择可比性的研究对象越难,需要的样本量越大。

### (六) 应用盲法

临床试验中的研究对象一般都是人,为了去除人(包括研究对象、观察者及资料整理和分析者)的主观心理因素对研究结果产生的某些干扰作用,观察结果时最好使用盲法。盲法分单盲法、双盲法和三盲法。

1. 单盲(single blind) 单盲是指研究对象不知道自己所在分组和所接受的处理,但观察者和资料收集分析者知道。单盲方法简便,容易进行,观察者知道受试者分组的情况,对受试者的健康和安全性有利。观察者可根据病情变化,采取方案允许的相应处理措施,做到心中有数。单盲法可以减少来自研究对象的偏倚,但不能防止来自观察者的观察性偏倚,即不能避免观察者主观因素引起的偏倚。若观察者认为研究对象的主观回答和客观反应与研究前的假说不一致,就可能下意识地在研究过程中给研究对象以暗示或引导他们按照观察者的意图回答,或是有目的增加某些辅助处理以得到阳性或阴性结果,由此可影响结果的可靠性和研究结论的正确性。

另外研究人员可能通过许多方面去影响病人的治疗,如医生对接受新疗法的病人观察特别细致,护士对新疗法组病人的关心和热情,这些都可能影响或暗示他们产生不同的应答反应。

2. 双盲(double blind) 研究对象和观察者都不知道分组情况,也不知道研究对象接受的处理措施,称为双盲。病人与医生只知道药物的编号,如A和B。待试验结束和资料分析后才宣布A药和B药何为试验药,何为对照药。双盲法较复杂,执行起来也比较困难,应用时必须考虑其可行性,在执行中要有严格的管理制度和办法。在试验过程中,双盲状态可因种种原因遭到破坏,因此在使用时应注意一些问题:①试验药的制剂应防止破盲,试验药和安慰剂两种制剂的颜色、气味、大小、外形要相同,甚至容器和外包装也要一样,一般常用胶囊制剂;②保证试验对象的安全,在双盲试验中,当医生发现病人出现了严重的不良反应、治疗无效或病情加重时,不应单纯为追求完整的资料而继续试验,必须从医德和伦理的观点出发,对该病人应立即停止盲法治疗,并公开该病人所用的真实药物。因此,在盲法试验之前,应由设计者预先制定出停止盲法的指标和条件,以利观察者执行,避免给病人带来不良影响或严重后果;③不适用于危重病人,双盲法的缺点是在管理上缺乏灵活性,因而不适于危重病人的抢救。此外,有特殊副作用的药物容易破盲。

双盲法虽然可收集到病人用药后的客观真实反应结果,但在资料整理和分析时,资料的整理和分析者为了得出某种结果,不按统计学原则处理资料,如去掉试验组的极小值,去掉对照组的极大值,人为导致两组资料的差异。

3. 三盲(triple blind) 三盲法是研究对象、观察者和资料整理分析者均不知道研究对象分组和处理情况,只有研究者委托的人员或是药物的制造者知道病人分组和用药情况,直到试验结束时才公布分组和处理情况。这种方法在理论上可减少资料分析时产生的偏倚,使研究结果更符合客观情况。该法减弱了对整个科研工作的监督作用,使科研的安全性得不到保证,应用并不普遍。在实际应用中通常用双盲随机对照试验。