

职业流行病学参考资料讲义

上海第一医学院

编者说明

从1977~1982年的五年中，我们曾多次介绍将流行病学的调研方法应用于劳动卫生工作（见王穆兰、顾祖维主编：劳动卫生学进展，第一卷，第一页，人民卫生出版社，1983），并举办了职业流行病学讲习班和进修班。与此同时，职业流行病学也受到国内同道的重视，如华东职业中毒防治协作区于1980年召开了“劳动卫生流行病学调查方法座谈会”。在国内的出版物中，由流行病学和统计学方面的专家，分别介绍了有关方法学的文章，劳动卫生工作中应用流行病学的调查逐渐增多。但到目前为止，尚无一本适当的教材。

本书选择国际上有关职业卫生方面八本著作中职业流行病学的专章八篇，加上自编的有关统计方法和案例各一篇编译而成。为避免一些不必要的重复，有些内容已予删去，也有些内容原文中叙述不清楚，予以必要的注释。最后又加了汉英名词对照。

本书仅作为初学者的参考教材，读通其内容，仅是一个起步，只有通过实践，才能融会贯通。在实践中还应继续读有关文献，进一步深化，提高自己逻辑思维的能力，故选择一些篇章，同时列出原作者介绍的文献，可供进一步自学之用。

职业流行病学

上海第一医学院卫生系

上海第一医学院预防医学研究所

世界卫生组织职业合作中心, 中国上海

(WHO Collaborating Center on Occupational Health, Shanghai, China)

顾学箕 薛寿征 选编

学术秘书 胡景虎

目 录

第一篇	流行病学	J. M. Harrington	1
第二篇	流行病学 I —用途和描述性方法	R. S. F. Shilling, Joan Walford	23
第三篇	流行病学 II —分析性和实验性方法	M. I. Newhouse, R. S. F. Shilling	33
第四篇	职业流行病学	D. H. Wegman, R. Giusti	47
第五篇	职业卫生中的流行病学	S. Hernberg	62
第六篇	职业流行病学方法	J. F. Gamble, M. C. Battigelli	94
第七篇	当前职业流行病学中的一些问题	R. R. Manson	109
第八篇	职业流行病学常用统计方法 (附: 吸烟经历的测量)	张照寰	132
第九篇	流行病学在评价人类遗传危险性方面所起的作用	J. R. Miller	150
第十篇	案例	M. A. Fingerhut	154
汉英名词对照			157

第一篇 流行病学 J. M. Harrington

译自 H. A. Waldron, J. M. Harrington: Occupational Hygiene 第十六章,
Blackwell Scientific Publications, London, 1980

绪 言

【定义】

职业卫生是一门以研究劳动环境为主的学科而不是专门研究劳动者的健康。因为劳动者的健康与他们的工作环境密切相关,卫生学家忽视这点是危险的。

流行病学是研究疾病在人群中发生的频率和分布的规律。每当考虑人群的健康问题时,均必须应用流行病学的原则,卫生工作者不仅要熟悉流行病学的调查方法,而且要能够将流行病学基本原理运用于工作实践。从事职业卫生的医生、护士及卫生工作者都要关切人群的健康,然而,除少数医生外,卫生学家均没有接受过正规的流行病学的训练。本章的目的是为读者提供流行病学的基本概念。

上述定义表明流行病学主要是研究疾病分布及探讨决定疾病分布的因素,卫生工作者主要是关心后者;此外,职业卫生工作搞好协作也是重要的。如果医生怀疑某一健康危害同工作场所有关,对工人的进一步调查可能显示这种危害的分布与年龄、性别、职业和民族的关系。同时要进一步调查与患者有关的职业因素,例如:粉尘浓度、工作周期,疾病的分布。发病日期与病程的关系—所有这些因素都是很重要的。卫生专业知识将大大提高医生认识与患者的疾病过程有密切关系的各种因素的能力。

运用流行病学原理阐述疾病原因,包括3种类型:描述受“危”或接触人群健康状况称描述性流行病学;根据特定的病因学假设设计调查,探索真实的致病原因称分析流行病学;设计和实施调查计划通过改变假定的危险因素,以估计是否导致疾病发病的改变称实验流行病学。由于医德上的原因,实验流行病学研究常常是很难设计的。但归根到底,流行病学是通过证实病因,消除病因而达到预防疾病发生的目的。

【流行病学的发展史】

与其他学科一样,流行病学是在与疫情斗争中发展起来的。以后临床医师提出并确立了流行病学应用的一般原则,因此,正如化学仅用来寻找点金术一样,当时的流行病学仅限于研究传染病。

纵观历史,运用统计学方法研究人群健康是流行病学调查研究的重要工具。1532年中世纪时期,英国已开始运用统计学方法收集死亡率登记资料。直到1563年才开始有正式的死亡报告,资料每周记录。由于这些资料通常是由教区内某些用钱雇用的年老寡妇收集的,准确性较差。但从1665年发表的死亡登记结果可看出主要疾病的组成:

死刑:	21
天花	655
水痘	86
精神刺激	23
窒息和饥饿	45
鼠疫	68,596
火灾	397
坏血病	105

1801年开始人口普查 每十年一次,但收集的资料很有限。1821年增加了年龄一项,因为结婚、出生、死亡登记制度直到1836年才开始实行,故长时间内不能准确地得到年龄资料。在1851年进行的第六次人口普查时增加了职业这一项目,1921年起增补了职业死亡率。William Farr 于1839年被任命为英格兰和威尔士注册登记总部的医学统计医生,为创立现代流行病学奠定了基础。

多年来,这些全国性的资料,为流行病学家提供了丰富的研究“资源”。在19世纪中期,人们应用流行病学调查方法成功地开展了地方病调查。Victoria 女王的麻醉师及流行病学先驱者 John Snow 肯定了水是1854年伦敦 Soho 地区霍乱暴发传播的媒介物。这比发现微生物所引起的疾病要早得多(流行病历史学家可查阅有关 Snow 访视公寓的经典调查记录,该调查记载了给水站附近水媒疾病的例子)。

流行病学家既要描述疾病与环境的关系,又要收集疾病在自然条件下流行的资料。例如1951年伦敦地区出现了持续5天的“烟雾”致使3,000多人死亡,经流行病学调查推测了空气污染在引起呼吸道疾病死亡原因中所起的作用。

准确地收集死亡资料是流行病学调查的前提。虽然死亡记录不够完整,但现在这些数据仍然是分析疾病的有关致病因素的重要资料。

目前,流行病学家设立群体疫情监测的愿望,同个人不愿提供信息的愿望相抵触。根据作者的经验,假如人们知道了研究工作的价值,他们很少会拒绝向研究者提供健康资料。哈佛大学的MacMahon教授简洁概括了当前有关个人健康资料泄密的争论:“资料大量的保密意味着人们获得的流行病学信息极少。识别新癌症危害极肤浅。……大多数人若能认识到其原因是合理的,他们将会提供甚至是很机密的资料,当数据成为机关化或具有政治色彩时,提供秘密的结果就会变得更加困难。”

在职业卫生中,流行病学有5种主要用途:

- 1)新的危害的识别;
- 2)已知危害的控制;
- 3)建立卫生标准,控制或消除危害因素;
- 4)确立控制各种危害和促进安全生产的重点;
- 5)评价卫生服务计划,保护健康。

资 料 来 源

由于政治上及法律上的原因,首先引用的是人口普查及死亡登记资料,因而,为开展聊

业原因所致疾病的调查提供了取之不尽的资料来源。抚恤金登记、疾病保险补偿计划和治疗计划也同样有用。事实上,即使收集起初目的并不明确的个人健康资料,流行病学家也会加以利用。

【全国性的记录】

在个人生命过程中,重要的事件如出生、结婚、离婚及死亡,在发达国家都是全国性登记的,登记由地方进行,而后存放在中央。

【死亡登记】

在个人健康记录中,死亡是最后的终结。在登记制度完整的国家里,死亡登记是容易实现的,它为人群疾病分布提供了一个准确合理,可用数量表示的衡量标准。

在许多例子中,死亡的原因可无容置疑地确立。但情况并不总是这样的,临床内科医生曾经多次考虑给每个死亡登记下诊断。这可能由于医生对最终引起病人死亡的各种原因混淆不清,或由于对病程的忽视及死前特异性的检查太少。诊断医生可能不是病人的诊治医生,即使同意尸体解剖来搞清死亡的原因,但根据这样的死因调查结果,也难以更改死亡证明书的死因。

临床医生不仅要决定病人最后的死亡原因是什么,而且要决定导致死亡的潜在因素,这更增加了复杂性。这些诊断的准确性随死者的年龄、性别,有关的机体的系统,死亡的环境或地点、甚至随医生签署死亡原因所用的措词而变化。

目前的国际死因分类协定指出,潜在死因应比直接死因优先考虑;虽然近年有些研究人员已设计了些方法,它能使多个死因编码列入死亡率分析中。每十年左右,举行一次国际性的大会以对所有主要的疾病建立编码。这样,死亡资料的流行病学研究就可以在国际范围内加以比较。这项工作很有价值,但不易克服国家(或甚至地区)对疾病命名的差别。多年来,北美医生诊断的肺气肿,若将同一病例介绍给英国医生,会诊断为慢性支气管炎。临床特征并无争论,争论的是那些疾病和体征有关的名称。

可以肯定,死亡资料经流行病学分析后如放宽死因诊断分类,则“真正的”死亡原因有80%以上是正确的。准确性随疾病而变化,死于心血管和呼吸道疾病有时很难鉴别,然而象白血病这样的疾病的确定则是很准确,不过治疗机构首先要有一个精确的病理学诊断。

死亡证明中登记的职业是“最后的”职业,这虽然是有用的,但忽视了那些原先加速病人死亡的职业,会使统计学价值受到限制。例如,“退休”、家庭妇女、公务员这些职业分类在流行病学上是无用的概念。看到尘肺患者死于停车场的例子,这并非是由于汽车经过停车场时所产生的粉尘的作用,事实上,此人可能是一个拿抚恤金的退休煤矿工人。象这样的职业就不准确。有时可以参考工厂工人的抚恤金登记,防止死因诊断的偏差。

【出生登记】

这些记录,主要是应用于计算婴儿疾病发生率的。有关先天畸形,妊娠并发症以及出生体重、妊娠周期的记录为调查职业对母亲及父亲的影响提供了资料。出生登记要同死亡登记一样准确。

【疾病率】

在某些国家,能够提供全国性的职工健康和安全的疾病率记录。这些记录的准确性比死亡率记录为低。从流行病学的目的来考虑,还需要补充特定的材料才能加以利用。在英国,工业事故和必须上报的疾病,要报告给卫生和安全委员会,同时,工伤抚恤金领取和就诊疗疾

病要向卫生和社会安全局报告。诊断的错误，事故，疾病及伤害漏报和法律上不完全包括的范围，都会妨碍这些数据在流行病学上的应用。女工的缺勤记录特别不全，除了按病种粗略地分类外，诊断更不准确。但这种全国范围内的统计资料能够指出长期变化，并可能显示出新的危害。

全国性资料的统计如家庭情况调查和住院病人的询问，虽然对一般流行病学是相当重要的，但对职业卫生来说则不一定适用。

【地方性记录】

这些记录从医院、家庭医生、工厂、学校、抚恤金计划、保险单，职业协会和贸易团体处获得。所有这些记录职业卫生流行病学家已经使用过不止一次。但所有这些资料的收集，都是与流行病学调查的目的无关。所以，它们的准确性，完整性，可比性和适应性都是值得推敲的。

职业组成员应该是毕生从事同一职业的，这种记录才能作为重要的信息来源。本文曾应用这样的记录作为计算完全不同职业的原始资料，如医院病理学家与航海驾驶员。

【特殊记录】

某些大工业组织正试图对“高危”工人进行连续监护，不仅在本公司内调动了工作的工人，当他们离开或退休时还要继续进行观察。这些记录，对评价许多年前曾接触过各种化学物质的工人的健康状况，提供了十分有用的数据。现在，在少数工业企业内仍然提供回顾性的接触数据。

对某种疾病已经建立了登记制度。早期的例子，包括各种传染病如百日咳或麻疹。最近，在许多国家内已建立了临床登记制度。在有些国家，其诊断准确率超过了98%。其他例子包括对某些疾病建立监测制度，如间皮瘤，肝血管肉瘤，尘肺症，某种药物的不良反应，先天畸形和某种损伤，如失明。

尽管这些健康记录很不完整，但流行病学家若经常探索应用各种调查方法，还是能获得他所需要的有关工人健康的一部分资料。为了进一步得到资料，可能还须向工人的家属作调查，甚至有必要收集工人从婴儿时期到死亡总的健康情况。由于工厂工人接触有害因素的资料缺乏，人们仍然不可能精确估计职业危害对工人健康的影响。

接触和反应的衡量

从上述记录所获得的接触危害因素和健康结果的资料，必须显示在人群之间或内部的可比性。这一章节重点在于衡量方法和使用时的基本概念。

【接触】

在将来，这仍然是不够准确的测量方法，在原因和结果的连接上是最弱的环节。大多数流行病学研究调查职业因素有关疾病时，试图对职工接触历史进行准确衡量时，就会犹豫不决。这是卫生学家对将来职业卫生研究能够作出较大贡献的领域。

归根结底，理想的流行病学调查应能显示病因和疾病后果之间的剂量—反应关系，这大大增加了同原因的联联系，并有助于为将来新工人建立安全（或较安全）的工作条件。

现在，这种过去的接触常被分为低、中、高或仅用接触年限来表示，而不考虑接触程度，这是不满意的。理想的是过去的接触资料应该是高质量的，并使用标准的或可信的仪器

准确地测定。这不仅要提供工人潜在的（或实际的）所吸收的毒物浓度，而且也要提供各种准确的数据，如工作周期，每天、每周、每月之内的浓度及其持续时间。另外，必须有其他有关接触的数据，包括工作环境中有毒物质的物理或化学组成的改变以及各种有害因素之间可能存在的相互作用的评价：它们是否是相加作用、增倍作用、或拮抗作用。

目前，很少能够提供这种数据，即使能够提供，总是不完全的，因而这给流行病学家的调查带来了连续不断的误差。

【反应】

在衡量反应方面，分清“比”、构成比和“率”是很重要的。这三个概念是不同的。严格地说，凡是由分子和分母组成的一个分数都可称为比。而一般来说比是指分子和分母是两种不同性质的量，如性别比，人群中的性比例，是男性人数和女性人数之比（有时可乘一常数如100）。构成比是比的一种，其分子包含在分母里面，如在一地区中男性人数占总人数的百分比。率是指一种现象的发生（或改变）的强度。发生这现象的多少依赖于某种基础，如人口。例如发病人数的多少是与人口数的多少有关，是人口数的应变量，因此，发病率是发病人数与当时的人口数之比。

【发生率的衡量方法】

流行病学中最常用的方法是发病率和患病率。普通的表达方式是每“人——期间”（常为人——年）而算出的率。发病率同特定时间内在特定人群中出现的病例数有关；疾病的患病率同某一段时间或某一时刻所存在的病例数有关，因此，严格地说，是在某一特定的时间内，危险人群中所存在的病例数的比值。

对某种疾病来说，患病和发病率是通过疾病周期而相互联系的，这关系可用以下方式来表示：

$$\text{患病率} \propto \text{发病率} \times \text{疾病周期}$$

在实际应用时，对发病不明确或是病程延长的慢性病来说，发病率不是一个合适的指标。然而当疾病是急性的，并能清楚地分出病程时，发病率则是个较好的衡量指标。

【频率的衡量方法】

率有各种各样的，虽然所确定的率使得人群之间具有可比性，但若它们没有建立一个明确的概念，则也会陷入混乱。从本质上看，某一特定的事件（死亡、疾病、意外事故等）发生的频率有三种衡量方法。他们是：

(1)粗率； (2)校正率； (3)标准化率

为了具体说明目的，我们集中说明死亡频率的衡量方法，其他指标也可以如此衡量。

$$\text{粗死亡率} = \frac{\text{受危人群中的全部死亡数}}{\text{受危人群的人数}}$$

这样的死亡率离开理想的死亡率太远了，因为死亡同性别，年龄有关，而粗死亡率的比较会导致错误的结论。例如：

A镇（一个工业中心）的粗死亡率是12/10⁵/年

B镇（一个海滨胜地）的粗死亡率是15/10⁵/年

结论是B镇的健康状况比A镇差。这是因为人群中未能作出年龄分类的缘故。A镇人群的年龄别死亡率高于B镇，而人群的年龄构成较年轻，而B镇人群的年龄构成较年老。解决这个

难题的方法是计算校正死亡率。在这个例子中，通过使用另一个以年龄为权重，校正了粗率内部潜在的误差。若对A镇和B镇都作了以上校正，两者之间的对照显示出A镇的年龄别死亡率较B镇为高。

在职业卫生中，人们常希望将某一职业人群组（所谓指征人群）与其它一些人群（标准人群）的死亡进行比较，这个过程产生了一种简单的统计方法，称为标（准）化死亡比（SMR）。将标准人群的死亡率作为1（或100%），所调查的职业组的死亡率就可以简单地用标准死亡率的百分数来表示，并能应用合适的技术方法检验它的统计学的显著意义。在表1中列出一例。

SMR虽已广泛使用，但仍须谨慎，粗心大意将会产生错误的解释。首先，由于此方法的年龄校正是间接的，故使用一个标准人口死亡率计算期望死亡数以计算二个或以上人群的标化死亡比，它们之间是不能相互比较的。

表1 1949—1953英国农业职工标化死亡比计算

年龄	观察死亡数 (1949—1953)	男性农业 人口数 (1951)	全国死亡率 (1/1000)	期望死亡数*
20—	541	83,400	1.4	116.76
25—	963	133,300	1.6	213.28
35—	1,500	131,600	2.9	381.64
45—	3,422	117,200	8.2	961.64
55—64	7,527	90,600	23.0	2033.52
20—64	13,953	556100		3756.52

$$SMR = \frac{13953}{5 \times 3756.52} \times 100 = 74$$

* 期望死亡数计算： $\frac{\text{第4栏} \times \text{第3栏}}{1000}$

第二点是SMR的大小依赖于标准人群的选择。由于标准人口通常是全国性人口，因此它的健康状况不如工作（职业）人群。这是因为全国性人口中包括了病人和丧失劳动能力者。因此，如果研究人群并不接触严重的职业危害，用全国性人口作为标准计算该职业人群的SMR必低于100%。事实上，不接触职业危害的人群与全国人口比较时，SMR约为80~90%。

第三个因素是总SMR可掩盖职业人群中年龄别死亡规律之间的差别。并不是所有的工人都同等地接触危害因素，可能年青人接触危害更为严重，或由于危害因素的累积作用，老年人更值得引起注意。解决这一问题的办法是计算年龄别SMR。

第四个有关的因素是社会经济状况。由于社会经济状况不同造成SMR不同。两个比较的人群中不同经济状况所占的比例可能不同。所以这个问题也需要考虑在内。

【时间、地点和人群的关系】

流行病学家处理数据应该按照影响结果的因素不断调整。然后对结果作出判断。

流行病学家 Donald Reid 教授曾经说过，人群中出现异常疾病，在进一步采取措施前需要询问以下 6 个问题：

(1) 什么人？(2) 什么地方？(3) 什么时间？(4) 被什么所致？(5) 怎样产生？(6) 结果怎样？

1859 年在 Soho 发生霍乱流行，John Snow 在调查了人群发病地点和时间后，发现流行仅累及那些饮用 Brad Street 泵站自来水的居民。1951 年 11 月伦敦烟雾引起伦敦市中心老人和儿童死亡率显著升高。主要死因为肺心病、支气管炎和肿瘤（图 1）。

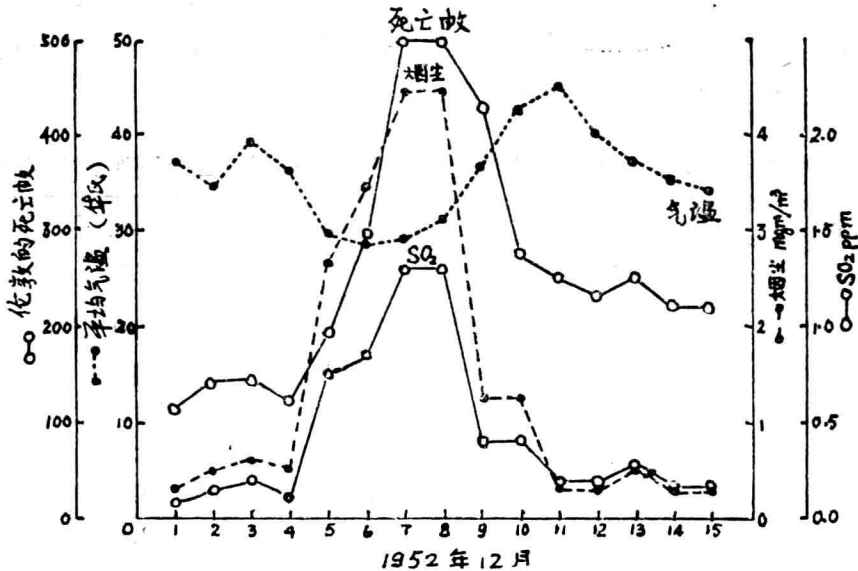


图 1 空气污染（SO₂和烟雾）、气温和死亡数曲线图，伦敦，1952年11月1日~15日

影响这些因素的主要特点是：

人群——	地点——	时期——
年龄	自然地理分界	天
性别	行政区域	月
种族	城乡	季
社会等级	工厂中工作地点	年
职业	环境	} 长期或周期
婚姻状况	气候	
家庭	移民情况	
遗传		

病因需要探索，病因的作用方式需要阐明。当解决了上述问题时，另一个严肃的问题

是——结果如何？这不是一个儿戏的问题，它应有助于确信进行调查是为了获得建设性意见，如果事实表明无重要危险，目前环境可不予改变。

【危险度测量】

危险度测量主要是资料分析。通常有二种方法：

(1) 相对危险度

(2) 归因危险度

相对危险度是接触人群发病率除以非接触人群发病率的比值。

归因危险度是接触人群发病率和非接触人群发病率的差值。

表 2 提供了一个通过纵向调查得出的危险度计算的例子。计算注意事项如下：

(1) 相对危险度的大小是危险因素和疾病联系强度的计量值。统计意义大小与此强度无关。

(2) 病例对照研究通常无法获得相对危险度和归因危险度的数值。因为这种调查方法通常无法计算发病率。而仅测量危险因素接触的频度。

表 2 1950年10月到1966年9月广岛居民白血病（资料来自原子弹伤亡委员会）

辐射剂量 (rads)	接触人群每年 粗发病率(1/10万)	相对危险度	归因危险度 (1/10万)*	%**
< 5	3.0	1.0	0.0	0
5—10	5.1	1.7	2.1	41
20—49	20.9	7.0	17.9	86
50—99	18.3	6.1	15.3	84
100—199	41.5	13.8	38.5	93
200—299	55.6	18.5	52.6	95
300以上	150.5	66.5	147.5	98
	7.2	2.4	4.2	58
合计				

* 非接触人群每年发病率为 3/10万

**归因危险度/接触人群每年粗发病率

在考虑流行病学调查前，需要强调一个重要的概念：危险因素与疾病在统计学上联系不一定存在因果关系。这种联系可能是虚假的也可能是由于其它已知或未知的间接因素造成的。流行病学技术决不能证明 A 引起 B，但它能对病因假设提供肯定或否定的支持。

这种支持将在下面 9 条中得到论证。（对这 9 个因素的详尽说明参考 Bradford Hill 的文章——环境和疾病，联系或因果？（1965 in Proc. Roy. Soc. Med, 58, 295—8.）。

(1) 联系的强度——疾病在特定的工人中更常见吗？如果是，有多少？

- (2) 一致性——是否已有多个调查者用不同的研究方法描述过这种联系？
- (3) 特异性——疾病是局限在一定的人群和场所吗？
- (4) 时间——可疑的病因总是位于疾病之前吗？有固定的时间间隔吗？
- (5) 生物学梯度——有肯定的剂量反应关系吗？
- (6) 生物学合理性——联系是合理的还是荒谬的呢？
- (7) 连贯性——因果关系符合逻辑和可行吗？
- (8) 实验证据——因果关系能被实验证实吗？
- (9) 类推——对有关的原因或结果，类似的可疑原因是否已被证实？

完全符合上述 9 点论证某一假设的情况是罕见的。它们也不具有相同的权数。但是在假设的证实中出现越多，联系就越强，因果关系就越可能存在。

研 究 设 计

流行病学研究设计首先要有计划。一成不变的研究设计通常是不可能的，因为实际上需要不断修改原来的计划。

研究设计分为二部分。

- (1) 目的 (图 2)
- (2) 方法选择 (图 3)

【目的】

流行病学调查研究的基本目的是获得关于研究对象的正确资料，但实际上调查前后和调查期间出现的各种问题往往使获得最正确的描述性资料受限制。

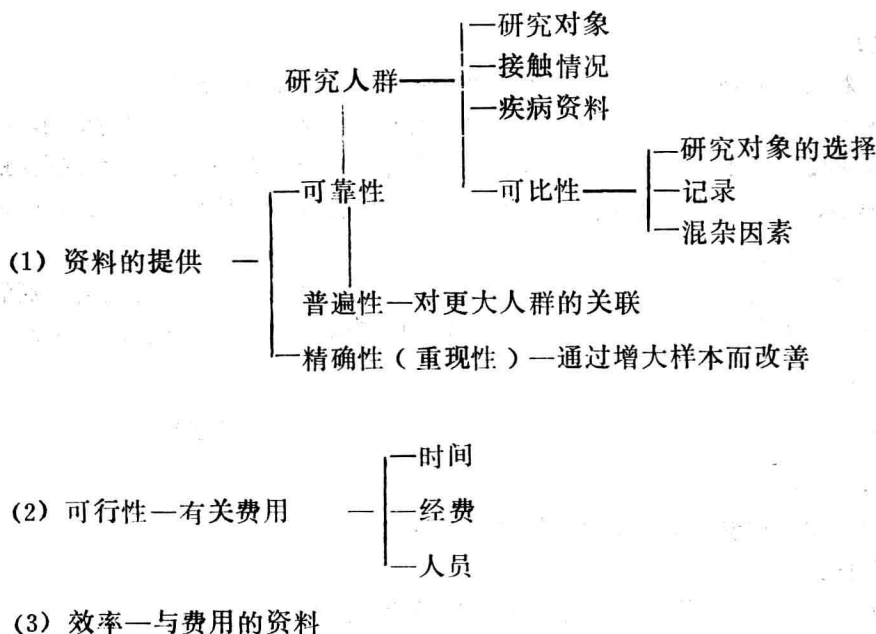


图 2 研究设计—目的

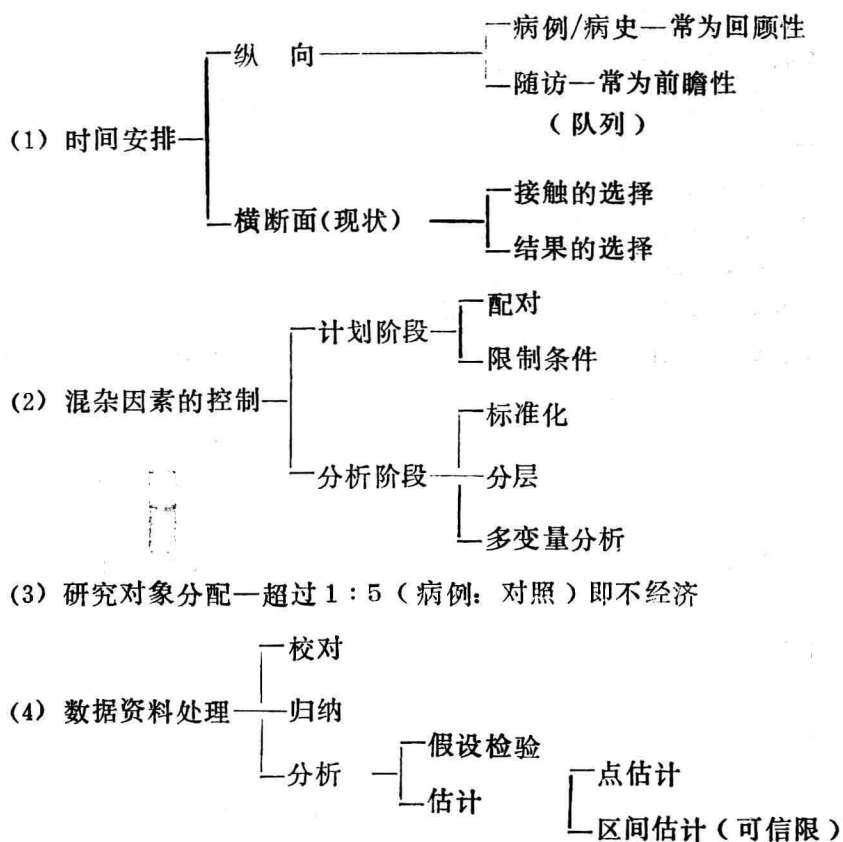


图 3 实验设计—选择性

【资料的提供】

衡量所收集资料的正确性包括两部分，无偏畸性或可靠性，重现性或精确性。

可靠性与普遍性和特殊性有关。一般认为它关系到研究结果怎样外延。例如，调查指出，某地区农民腿部骨折高于办公室人员，这个结果能外推到全部农民吗？可能在研究地区存在特殊情况（地形、气候、使用拖拉机等），因此，农民骨折并不具有普遍性。（或许对照人群不恰当），不具有普遍性。

从特殊性意义说，研究人群可能有偏倚性。包括选择研究对象，收集接触资料，疾病的定义及其诊断特点的偏差。举一个简单的例子，两个人同测量一组工人血压值时，可产生随机误差和系统误差。列表如下：

误差原因	误差类型
(1) 病人血压随时在变化	随机
(2) 特殊情况，如害怕或忧虑能使血压增高，休息能使血压降低。	系统
(3) 观察者爱读血压(毫米汞柱)的整数值——尾数化为整数	系统
(4) 每个测量者由于疲劳、积极性、视力、经验等原因读数不同	系统

(5) 一个测量者的读数经常会比别人偏低(或偏高)

系统

(6) 一个病人两次测得的血压各不相同。

随机

(7) 血压计不准确

系统

还要考虑精确性方面问题。如果某一个特定人群的研究计划被重复地进行若干次,各变量值将会接近于均值。偏离均值的程度将表示随机误差的大小。随机误差小,表明精确性高。增加调查研究的精确性以通过增大样本或增加测量次数来达到。

【可行性】

各种调查方法均会限制完善调研计划的实施,因为涉及到调查研究的人力、物力和财力方面问题。在工厂的调查,不易找到合适的调查对象或生产环境。

总之,提供资料和可行性必须一起考虑。在两种相同资料的实验设计中,必须选择可行的。如果调查是同样地可行的话,必须选择能提供更多资料的调查。

【选择性】

合理的调查设计要求按照调查实施的方法作出选择。

【时间安排】

头等重要的选择是调查的时间安排。图4列举了一个工厂人群按日程(时间/年龄—时间安排)的生活历史。本图用于阐明时间安排上的选择。当然现实生活中的例子要比这复杂得多,但原则仍是相同的。

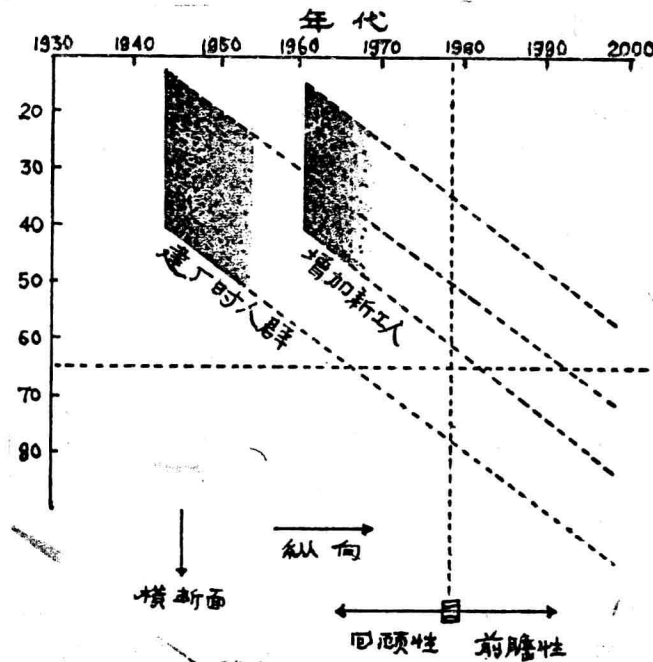


图4 调查研究设计中的年龄时间因素

此厂建于1945年,工人年龄在18~40岁之间。随着时间的推移(横坐标)人群年龄老化(纵坐标)。假设人群不变则人群年龄沿斜线方向向前推移。1960年实施扩展计划,增加了

以青年人为主的劳动力。新的队列年龄与原有的人群以同一速度向前推移。(1960年年青工人进入队列后,使1945年进厂的工人队列年龄模式发生改变)调查可以是横断面的(垂直定向)也可以是纵向(水平定向)的调查。

横断面调查节省时间和费用。缺点是在有限的时间范围内估计受危人群造成的影响有一定局限性。观察时间短,调查者看不出接触和结果间的关系。横断面调查研究有选择结果或选择接触的倾向。估计接触或结果发生率是可行的。

纵向调查研究费时多、费用大,由于调查是在一段时期而不是瞬时,为观察接触和结果间的时间联系提供了良好机会。常用的两种纵向调查:病史调查(病例/对照调查)和随访调查(包括队列调查)。

病史调查倾向于回顾性的。从确定病例组开始,并将病例组和非病例组(对照组)与过去接触情况联系起来。在职业性流行病学中,病史调查的主要缺陷是难以确切查明长达40年的接触史。

如接触因素已明确,则随访调查没有以上的缺点。随访某一个接触组(或非接触组)可估计接触的最终结果。这种调查通常是定向前瞻性的。队列分析是随访调查的一种特殊类型,事先明确接触特点,随访一段时期后计算结果。设计随访调查以观察其发病率,应持续观察一段时期,超过接触因素的最长诱导期,以得出推断性的结果。

【控制混杂因素】

混杂因素有二种属性,一种与接触有关,一种与疾病有关。典型的例证是年龄。工人年龄越大,接触职业危害的可能性越多,患病的可能越大。(通常说来,疾病与年龄有关)关于混杂因素其它注意点是:

(1) 混杂因素与接触和疾病间不是简单的因果关系。

(2) 混杂因素与疾病的联系具有预期性的,是危险度的指标(虽然反过来不一定是正确的)。

(3) 危险度特征与接触情况无关。

(4) 混杂因素与疾病联系必须出现在调查人群中,而不适合于一般的人群。(反过来也不一定是正确的)。

(5) 必须同时考虑各种混杂因素,以减少个体的权重。

在设计阶段或数据分析时可控制混杂因素。在设计时,病例配对减少或消除混杂因素的干扰。例如,年龄配对可消除把老人与年青人同样看待的问题。挑选配对的原则是病例组和对照组分层,可以有效地排除干扰。在数据资料分析中可以自由分层或采用标准化程序。多变量分析能把复杂的相互作用的干扰因素减到最低限度。

【研究对象分配】

对照人群常超过病例人数两倍或更多。统计理论说明这样能增强资料的比较效果。病例/对照比例超过1:3时,意义逐步减少,由于经济原因,不必超过1:5。

【数据处理】

数据处理包括三个主要程序。整理收集资料,使资料随时都能加以利用,归纳资料以便处理,最后统计分析。分析过程目的在于检验在调查开始时所提出的假设,或用于估计各种参数,更常用于对算出的估计值确定可信限。

调查研究类型

主要介绍调查研究类型和举例说明实际中如何应用。

【纵向调查】

多年来,随访调查被认为是典型的流行病学调查方法,此方法的准确性在于内在的无偏倚的设计——在疾病出现前就按照研究特征选择一组或多组人群。然后随访全过程,观察疾病发生频率。这种调查时间长,代价昂贵且常相当复杂。因此目前有些流行病学家认为用详细的病史很好地设计病例对照调查,可能是既精确省时而又经济。

【随访调查】

进行队列调查不仅是由于方法本身的需要,而且与调查人群的属性有关,主要选择标准列举于下:

1. 特殊接触:(1)高浓度;(2)特异性;(3)新出现的可疑因素。
2. 容易追踪:(1)医院病人,包括产科病房母亲或婴儿和接受特殊治疗的病人,如放射治疗;(2)工人;(3)享受保险者;(4)享受养老金者;(5)专业团体。
3. 按地理或种族分组——如移民调查。

队列调查的实例,如离子辐射作用的调查。调查对象集中在高浓度接触人群——日本原子弹爆炸地区的幸存者,接受放射治疗的病人和放射部门工作人员。偶然发生灾害事故,一组人群可幸运地也许是灾难性地接触于某一种因素之下。例如,1976年7月10日发生在意大利的Seveso灾害,这样一个事件需要进行彻底调查,因为它提供了一个追踪一组严重接触二恶烷人群的机会。调查者不仅调查了妊娠结局和对动植物的危害,还调查了人群的远期发病率和死亡率。

近年来在动物实验的研究中,发现有机氯杀虫剂氯丹可能起引癌症,虽然对人体的研究几乎还没有开展。1976年3月,浓缩的氯丹进入美国田纳西州Chattanooga县的市政供水系统,使得这个地区的自来水中氯丹的浓度超过0.1%。通过对供水人群的调查,评价了其急性效应,还可以进行远期效应的调查。

一般说来,职业性接触比环境性危害更适合于进行队列调查。对于职业危害通常需要考虑接触在前和疾病在后的现象。大多数职业流行病学需要进行纵向调查,因队列的确定使随访研究能够实行,虽然能够收集到现在和将来的资料,但过去的接触资料往往残缺不全,调查者为了得到最后的结果不得不等待时间。流行病学家很少愿意选择一些经过研究而不能发表的课题,这是病例对照调查比队列调查更吸引人的原因。如果一个职业人群的资料,能及时地追溯到,那么,调查这个人群中患有某种疾病是可行的。如果所调查的接触人群生活在同一个地方或个人持有养老金或保险单记录的专业团体,对于队列调查是有利的。

前不久Trinity House的海港领港员要求调查领港员死于心脏病的危险性。他们认为,领港员死于心脏病的危险性比非领港员高。这些人从35岁左右开始担任领港员,直到退休或死亡才离开此职。对领港员的调查几乎无一人漏失,而且那些已故领港员的遗孀领到抚恤金的收据就是死亡证书的副本。因此,Trinity House不仅能提供这些领港员死于心脏病的危险性的详细资料,而且也提供了领港员的健康状况的资料,死亡者附有死因证明书。

这个调查采用队列“寿命表”法,归纳如下:

(1) 规定调查的领港员在1920年以前出生, 1956年1月1日起工作—393名领港员符合此标准。

(2) 追踪观察这些领港员到1968年12月31日。在这期间, 有60名死亡。Trinity House提供了其中44名领港员的死亡证明书的副本, 调查者又发现两名领港员的死因, 其余14名死因不明。

(3) 从出生、结婚、死亡登记中心追踪到其中8个人的死因, 余下的6个人经进一步调查得知死因在国外, 随后也得到了死因证明书。

(4) 设计调查表, 确定5个年龄组的领港员在5年期间从事于领港工作条件下的人年数。

(5) 总结以上资料, 并分析领港员的死亡年龄和年份, 就能推算出死亡率, 然后与同一地区, 同一经济状况的其他人群的死亡率进行比较。

调查结果见表3。

表3 英吉利海峡领港员冠心病死亡实际数和期望数

年龄	受危 (人年)	1956~1960			受危 (人年)	1961~1965		
		死亡率 (‰)	期望死 亡数	实际 死亡数		死亡率 (‰)	期望 死亡数	实际 死亡数
35~49	825	0.75	0.62	0	896	0.94	0.84	1
50~64	672	5.00	3.36	2	683	5.72	3.91	4
65~79	183	21.22	3.08	1	274	21.85	5.99	4
			—	—			—	—
			7.06	3			10.74	9
		1966~1968				1956~1968		
	379	1.02	0.39	6			1.85	7
	422	5.78	2.44	3			9.71	9
	223	21.01	4.69	5			13.76	10
			—	—			—	—
			7.52	14			25.32	26

如同这个调查那样追踪一组人群是队列调查的关键问题。队列调查的质量和完整性将大大提高研究结果的正确性。由于队列调查需要在许多年中追踪一组人群, 因此总是存在一部分人“失访”这个严重的问题。结婚, 迁移, 工作调动, 拒绝协作都能减少调查人群, 使调查陷于困境。因为不回答和“中途失访者”在一些重要方面是不同于回答者的, 前二者的缺失在结果分析时可使回答者存在不均衡性, 从而产生偏倚, 回答率至少要达到85%, 除非调查者能够保证回答者与不回答之间的差别无显著意义。这一点可通过其它途径得到的不回答者的有关资料(年龄、性别、工种、居住条件而估计, 努力证实在这些方面或在其它方面不回答者相同于回答者。