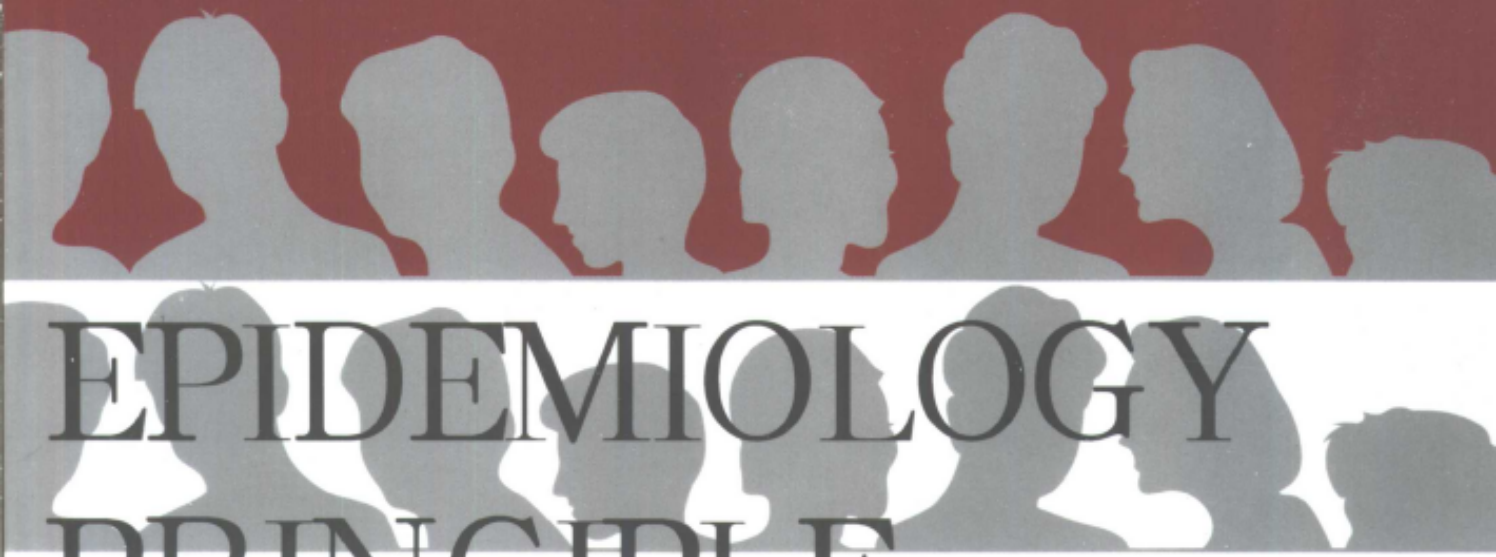


上海研究生教育用书

流行病学 原理与方法

主编 沈福民

A horizontal band across the middle of the cover features a series of grey silhouettes of people's heads and shoulders, facing various directions, creating a sense of a crowd or community.

EPIDEMIOLOGY
PRINCIPLE
AND
METHODS

复旦大学出版社
上海医科大学出版社

责任编辑 宫建平
封面设计 马晓霞

EPIDEMIOLOGY PRINCIPLE AND METHODS

ISBN 7-5627-0660-3



9 787562 706601 >

R·629 定价: 38.00元



□ 上海研究生教育用书

主编 沈福民 副主编 徐忠 徐彪

流行病学原理与方法

编著者 (以章节先后排序)

李婉先 沈福民 徐彪 袁鸿昌 俞顺章 边建超
徐茂一 顾惠心 刘佩莉 付朝伟 王曾遂 姜庆五
孟 炜 倪鹏生 金丕焕 刘 勤 徐 彪 赵根明

□ 复旦大学出版社
上海医科大学出版社

图书在版编目(CIP)数据

流行病学原理与方法/沈福民主编. —上海:复旦大学出版社:上海医科大学出版社, 2001. 8

ISBN 7-5627-0660-3

I. 流... II. 沈... III. ①流行病学—理论②流行病学—研究方法 IV. R181

中国版本图书馆 CIP 数据核字(2001)第 035954 号

责任编辑 官建平

责任校对 蒋建安 王汇珊

流行病学原理与方法

主编 沈福民

复旦大学出版社 出版发行
上海医科大学出版社

上海市国权路 579 号

邮政编码 200433

新华书店上海发行所经销

常熟市大宏印刷厂印刷

开本 787×1092 1/16 印张 21.25 字数 517 000

2001 年 8 月第 1 版 2001 年 8 月第 1 次印刷

印数 1—3 000

ISBN 7-5627-0660-3/R·629

定价:38.00 元

敬告读者:奉上级指令,原上海医科大学出版社于 2001 年 2 月 12 日正式与复旦大学出版社合并,组建新的复旦大学出版社。特此告知。

前言

流行病学这一学科在其形成、发展和日趋完善的过程中,始终是预防医学研究的基础和重要手段。流行病学的原理与方法,是医学科学研究及因果推理的重要理论基础和方法学。每个医学科研工作者都有必要掌握流行病学原理与方法。

迄今为止,已经出版了一些有理论、有实践的流行病学参考书。但由于流行病学本身的飞速发展,如从单一病因论发展至多因论,从单因素分析发展至多因素分析,从研究疾病发展至研究疾病和各种健康状态,加之其与不同学科的结合形成了各种分支流行病学,现有的流行病学参考书已越来越不能满足预防医学领域内不同研究的需求。

本书为研究生教材,旨在为研究生撰写课题设计、实施课题研究、进行资料处理及统计学分析提供帮助,同时又可作为预防医学科研工作者和流行病学工作者的参考书。在此次正式出版前,本书已在研究生教学中使用了十余年,颇得好评,曾获原上海医科大学教材一等奖。

本书共含 17 章,内容涉及流行病学的原理(如病因论、偏倚、混杂和效应修饰等)及方法学(如生存分析、Meta 分析、批质量保证性抽样及经济决策分析等)。另外,有些内容已形成独立单元,既有理论,又有实践,如病例对照研究、队列研究、实验流行病学、遗传流行病学、血清流行病学及分子流行病学等。

此次出版,本书在原有的基础上,参考了 Rothman 的第 2 版《现代流行病学》,对其中部分章节作了详细修订和补充,如病因论、病例对照研究、队列研究、筛检、偏倚、混杂和效应修饰、生存分析、Meta 分析、批质量保证性抽样、标

化死亡率比在流行病学中的应用。

本书的编写动员了复旦大学公共卫生学院流行病学教研室的各位老师,汇集了他们在科研及教学过程中积累起来的资料和经验,有的是博士研究生的课题内容或新近从国外进修所获的新信息,因而全书内容新颖、系统、深入、实用。

由于本书是教研室各位老师分头执笔写成,因而文笔、叙述方式等各有不同,虽经努力趋同,仍难免前后的不一致性。此外,本书的内容仍不能包罗万象,难免挂一漏万。笔者衷心希望广大读者能不吝指正,以使本书能得以不断修正、改进和完善,以期成为疾病预防工作者和其他读者案头的一本好参考书。

编者

2001年8月

目 录

第一章 调查设计	1
第一节 概述	1
第二节 人群调查研究类型	3
第三节 观察性研究对象的选择及观察方法	5
第四节 观察性研究的设计类型	8
第五节 观察性研究的设计比较	15
第二章 病因和病因学研究	17
第一节 基本概念	17
第二节 病因的确定	23
第三章 病例对照研究	29
第一节 设计原理	30
第二节 病例与对照的选择	31
第三节 暴露因素的调查	34
第四节 样本大小的估计	34
第五节 病例对照研究中的偏倚	36
第六节 病例对照研究的结果分析	37
第七节 病例对照研究的优缺点	52
第四章 队列研究	54
第一节 概述	54
第二节 研究的实施	58
第三节 资料分析	63
第四节 队列研究的优点和缺点	76
第五章 实验流行病学	78

第一节	预防接种的效果考核	78
第二节	干预试验	95
第六章	筛检	98
第一节	概述	98
第二节	高危人群的定义和选择原则	102
第三节	筛检效果测定	103
第四节	筛检中的注意事项及偏倚的防止	103
第五节	筛检的记录及数据库的建立	104
第六节	筛检质量控制	105
第七节	筛检的宣传教育	106
第八节	常用筛检程序	106
第七章	血清流行病学	108
第一节	血清流行病学的用途	109
第二节	血清流行病学研究方法	112
第三节	血清标本的收集、运送与保存	114
第四节	血清学试验方法的选择和质量评价	115
第五节	血清流行病学常用的统计学方法	115
第八章	分子流行病学	124
第一节	分子流行病学定义	124
第二节	分子流行病学常用的研究方法	124
第三节	分子流行病学的用途	127
第九章	偏倚、混杂和效应修饰	133
第一节	偏倚概述	135
第二节	选择偏倚	136
第三节	信息偏倚	143
第四节	混杂偏倚	148
第五节	效应修饰	156
第六节	混杂和效应修饰分层分析例示	159
第十章	疾病控制措施经济性和决策分析	167
第一节	干预措施的经济性评估	167
第二节	决策分析	172
第十一章	遗传流行病学	181

第一节 遗传流行病学定义.....	182
第二节 遗传流行病学研究的对象与任务.....	182
第三节 遗传流行病学的研究方法.....	183
附录 遗传流行病学基本研究方法及其常用软件介绍.....	207
第十二章 生存分析	221
第一节 概念.....	221
第二节 生存分析的任务.....	223
第三节 生存分析中常用的估计生存函数的非参数法.....	223
第四节 COX 模型	224
第五节 生存分析 SAS 软件的应用	229
第十三章 疾病频率的测量及标化	250
第一节 常用的计量单位.....	250
第二节 各种率的定义及应用.....	251
第三节 率标准化法.....	256
第四节 标化死亡率比在流行病学的应用.....	259
第十四章 流行病学研究资料分析的策略	265
第一节 资料的整理.....	265
第二节 资料改错后按每项变量作简单的描述性分析.....	266
第三节 变量等级的划分和转换.....	267
第四节 分析指标.....	267
第五节 单因素分析.....	268
第六节 混杂因素与混杂效应.....	269
第七节 交互作用.....	270
第八节 关于配对资料.....	270
第九节 结果的解释.....	270
第十五章 Meta 分析	272
第一节 Meta 分析的历史	272
第二节 文献的收集及其质量评估.....	273
第三节 Meta 分析的模式	274
第四节 几种 Meta 分析常用的方法	281
第五节 Meta 分析今后的发展方向	285
第十六章 批质量保证性抽样	287
第一节 应用扩大免疫规划整群抽样.....	288
第二节 批质量保证性抽样的应用.....	289

第三节	批质量保证性抽样中的二项式	293
第四节	累积概率	297
第五节	抽样样本中的未覆盖数	299
第六节	受方风险固定时的样本大小	301
第七节	确定供方和受方风险	302
第八节	计算 LQAS 数据的覆盖百分率及其可信限	304
第十七章	理论流行病学	307
第一节	数学模型的建立	307
第二节	数学模型的拟合	312
第三节	模型的种类	313
第四节	模型的功用	323
第五节	模型同现实	324
第六节	理论流行病学的发展与前景	324
附录	流行病学专业常用词汇英汉对照	326

第一节 概 述

流行病学是研究疾病的频率与分布及其影响因素,探索病因与流行规律,借以制订相应的防制措施,并对措施的效果进行考核。随着当前危害人群健康的主要疾病谱的变化,流行病学研究的范围也从以传染病为主,发展为包括一切疾病以及健康状况。相应地,疾病的病因不仅可以是生物学因素如细菌、病毒等,亦包括社会因素在内的各种环境因素以及个人行为。流行病学与各临床学科及卫生学科相互渗透,形成了临床流行病学、心血管病流行病学、脑血管病流行病学、精神病流行病学、肿瘤流行病学、环境流行病学、职业病流行病学等分支学科。

一、流行病学研究的目的

1. 描述人群的健康状况,通过疾病发生的数量,了解不同人群发病的相对频数,以及展示疾病的重要趋势,以描述人群的健康状况。
2. 解释疾病的病因,测定引起疾病或某疾病趋势的因素以及发现疾病的传播方式。
3. 预测疾病发生的数量及不同人群的健康状况。
4. 控制人群中的疾病,预防新病例的发生,治疗已有的病例,延长患者生命或增进患者的健康状况。

二、流行病学研究的范围

为达到上述这些目的,流行病学研究可以在两个层次即了解层与干预层上开展。为进一步了解疾病,要进行观察,以积累疾病发生和病因学的知识;为进一步干预疾病,要收集试验信息,为公共卫生决策提供依据。这两个层次都与疾病自然史的4个阶段有联系(图1-1),即:①诱发病因过程,从而产生第1个病因;②诱发病理过程,从而使变化不能逆转,即不可避免地产生了疾病;③临床症状期,出现症状与

体征供临床检测;④疾病转归,包括康复、复发、严重损伤或死亡。

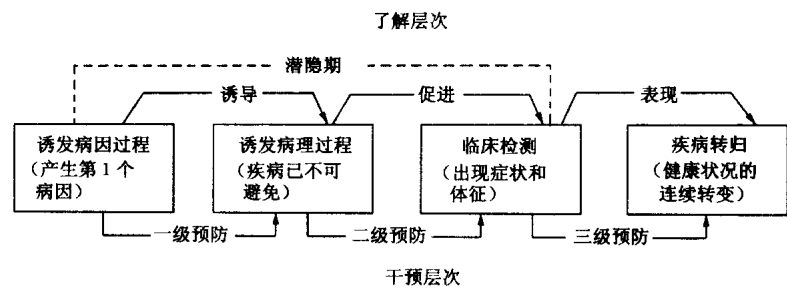


图 1-1 了解、干预与疾病自然史 4 个阶段的关系图解

为了了解疾病,需对疾病的自然史作科学的概括。疾病的自然史也可改成 3 个连贯的时期,即诱导期、促进期和表现期。然而,何时疾病已不可逆转,通常是不知道的,因而在研究中常把诱导和促进两个时期合并起来。此外,对多数疾病而言,第 1 个病因往往在出生时或出生前就形成了。因此,流行病学家用潜隐期来表示从出现病因至检出疾病之间的一段时间。对传染病来说,这段时期即潜伏期。与潜隐期相对应的是表现期,即从检出疾病至结束疾病(康复或死亡),亦称病期。

既然疾病有潜隐期和表现期两个阶段,那么慢性病就可根据在潜隐期或在表现期呈慢性来划分。当然急性与慢性的界限总是人为的,一般假定病程超过 4~12 个月者为慢性(表 1-1)。

表 1-1 急性病和慢性病的举例

表 现 期	潜 隐 期	
	急 性	慢 性
急 性	流行性感胃 肉毒中毒 中毒性休克综合征	胰腺癌 急性淋巴细胞白血病 精神分裂症急性发作
慢 性	梅毒 结核	高血压 老年性痴呆

干预研究的目的是评估预防疾病和促进健康的措施规划和策略,在该层次的研究可将疾病各阶段相连的 3 个连线代表各级预防(图 1-1)。各级预防的干预类型和目的如表 1-2。

表 1-2 预防策略

级 别	干 预 类 型	干 预 目 的
一级预防	改变影响人群中疾病分布的因素	防止或延迟第 1 例新病例的发生
二级预防	早期发现、早期诊断、早期治疗	改善病例的预后,缩短病程或延长生命
三级预防	治疗与康复医疗	减少或防止病后伤残,延长生命,改善疾病的转归

了解和干预这两类调查研究有不同的出发点和不同的研究策略及假设。对疾病过程的了解对于控制该病具有重要意义,但对规划一个成功的干预还不一定足够。事实上,不同领

域的学者还常有争论。在解释他们的研究结果时,科学家们往往注重对假设疑点的解释,而制订决策者则往往衡量此结果能否实际推广及其重要性,以便考虑是否采用及如何采用。

三、流行病学研究的方法

成功的流行病学研究,常常借助于若干其他学科或综合其他学科的原理与方法,诸如生物医学(病理学、生理学、微生物学、病毒学、免疫学及临床医学)、社会科学(心理学、社会学、人类学、政治经济学)和定量学科(数学、统计学、人口学、运筹学)等。当然,每个流行病学家对上述各学科的应用程度是不同的。由于知识面和各人所用方法的不同,流行病学研究可以分为下列3种类型。

1. 实验室研究:将基础学科的发展用于促进检测、控制疾病及其影响因素。例如制造疫苗或建立一种能证实某抗原的方法。
2. 流行或暴发的调查:找出病因、传播方式,并提出相应的控制措施。
3. 人群调查研究:研究影响疾病的生物学、环境和行为的因素以及预防措施。本章主要讨论此类研究。

第二节 人群调查研究的类型

根据是否由观察者将所研究的因素加之于观察对象,以及在安排此研究因素给观察对象时是否采用了随机方法,可以将人群调查研究分为实验研究、准实验研究和观察性研究3类。

一、实验研究

随机地将实验研究(experimental study)对象分成实验组与对照组。实验组给予所研究的因素或新疗法,对照组给予安慰剂或传统疗法;观察一段时间,测量和记录两组的结果并作比较。为防止观察者及研究对象对实验结果主观因素的影响,可采用盲法。

1. 实验室实验:为期最短,只需数小时至数天。实验室实验是用于检验病因假设并估计即时的生物学和行为上的作用。实验对象往往有高度选择性,实验室的结果很少能代表一般人群。如检验某特殊的急性精神紧张因素、不能解答的难题等是否会使健康人的儿茶酚胺增高,如获阳性结果,可能解释情绪紧张与冠状动脉硬化性心脏病(简称为“冠心病”)的联系。

2. 临床试验:需时稍长,一般数天至数年。往往限于高度选择的人群,如筛选对象、确诊病例或其他志愿者。这种实验在流行病学中已有悠久的历史。临床试验的主要目的是检验治疗或预防措施的效果。例如,检验某新的化学疗法是否能延长急性淋巴细胞白血病患儿的生命。故临床试验可用于检验病因假设,也可估计研究因素的长期作用,检验干预措施是否能改变健康状况以及提示此措施是否适宜在人群中采用。

3. 社区干预:历时较长,一般需进行6个月以上。它是在某一自然人群(社区)中开展针对某危险因素的一级预防。常用以测定改变某种个人行为、生物学特征或某环境条件后可能获得的收益。社区干预的结果可供卫生规划及决策参考。

实验的条件是最易受到控制的。然而即使在设计很好的实验中,仅靠随机分组来控制可能的混杂因素,也往往不是很容易。研究对象可能有选择性,因而观察到的结果不一定能

在人群中出现。有时医生和病人都认为某疗法有效,因为随机分组而不给某些人此疗法时,也存在着伦理学上的问题。有时盲法亦不易办到。在这种情况下,可采取准实验(quasi experiments)。

二、准实验

由观察者将研究因素安排给实验对象,但不经随机分组。可以进行一组内部或多组组间比较。一组比较系将措施前后的情况相比;多组比较时,各组可根据不同地区或根据受试对象的志愿行为划分。根据如何进行干预,可将准实验分为以下两类。

1. 临床或实验室研究:实验对象自愿参加某实验组,例如接受饮食指导而改变饮食习惯,以控制轻度高血压,并与未接受饮食指导(未改变饮食习惯)组相比较。

2. 规划或策略研究:相当于没有随机分组的社区试验,研究开始时干预已由他人设计并实施了。它可评价干预的效果,以决定是否需作相应的规划或策略的变更。还可发现开展措施后没有预见到的问题、措施的失败或成功的原因,并比较干预的成本效益。

开展准实验所遇到困难较少,较易开展,费用亦较低,便于大规模进行。然而因实验对象并非随机分组,不易控制可能存在的混杂因素的影响。例如某医师欣赏某新疗法,在考核此疗法效果时,他可能选择一些预后较好的患者作试验对象,因而使结果受到影响。

某些自然试验本质上虽是准实验,实验对象未经过随机分组,但其性质却类似于实验研究。例如,1885年 John Snow 在伦敦观察的霍乱流行。流行不久,有一家自来水公司改变了其进水口的位置,水质明显改善。由这家公司所供水的居民,虽然与其他居民居住地区、经济状况等都类同,居民自己也不知道饮用何家公司的水,但饮用这家公司水的居民的霍乱死亡率要低得多。

三、观察性研究

观察性研究(observational studies)的研究因素不经人为安排,可有多种研究形式。最简单的设计类似实验研究或准实验。例如可以在基线检查时明确一组老年高血压患者及一组老年非高血压患者,对他们随访若干年。如果老年性痴呆在前组发生较多,我们可以认为高血压是老年性痴呆的一种危险因素。此外,亦可观察不经干预的血压变化与继发老年性痴呆的关系。

观察性研究为流行病学研究的中心内容。根据以往对疾病的了解情况,可将观察性研究分为以下两类。

1. 描述性研究:对某病的发生、自然史及其决定因素了解甚少时,通常先开展描述性研究。描述性研究的目的在于估计某人群中该病发生的频率和时间趋势,以形成该病的病因假设。例如,对老年人精神卫生了解甚少时,可以计划一个社区研究,估计发生某些精神异常的相对频率。如果发现某种忧郁性精神异常在某年龄组较一般人群更为常见时,可以进一步研究,探求其原因,并寻找有关的社会因素与保健因素。

2. 分析性(或病因性)研究:对某病已有一定了解时,可通过分析性研究来检验其病因假设。分析性研究的目的在于证实此病的各种危险因素,估计它们对疾病作用的大小,并提出可能的干预策略。

观察性研究往往是最现实、最便于开展的科学研究,一般是在较接近于自然条件下进行

的,研究人群较能代表靶人群。卫生规划或策略制订者往往要根据这些研究的结果,作出决策。观察性研究的主要缺点是不能控制研究条件,有时研究结果可能受到影响。

第三节 观察性研究对象的选择及观察方法

在调查设计中,研究对象的选择和观察方法是两个极为重要的问题。

一、研究对象的选择

对研究对象加以限制、随机抽样和分层抽样是选择研究对象的3个关键。

1. 限制研究对象:为防止其他外界危险因素的影响,要剔除具有所研究疾病或健康状态的另一已知危险因素的对象。为防止因诊断标准不同而影响研究结果,也应剔除通过某种特殊医学检验而查出原先不能检出的对象。存在与研究因素有联系的其他疾病或健康状态时,亦应予以剔除,以免其他疾病或健康状态混淆研究因素与所研究疾病的关系。

为研究方便,可以将研究对象限制在居住于一定地区或在某一时期内的确诊者或某年龄段者,这样选出的对象观察较为方便,且较省经费。

还可将研究对象限于有一定程度的研究因素者,以便明确地观察这种因素与某疾病的联系。例如在研究职业暴露于某化学药物的后果时,对照比较重度暴露组与未暴露组,可能有助于找出研究因素与其后果的联系。但是这样做也可能丢弃许多具有此因素的对象,而不利于研究。

(1) 全面选择:指经上述限制后,凡符合选择条件的对象,全部选为研究对象。例如规定40~50岁的汉族妇女作为观察对象,比较肥胖与发生高血压的关系。如全面选择,则当地40~50岁的汉族妇女均为研究对象。

(2) 部分选择:又称配比。要求对象都具有某些条件,将病例组与一个或更多的非病例组比较。例如选择了一组病例组,按非病例组的年龄分布配比。事实上是每个病例配一个或几个作比较的对照,使对照组的年龄分布同病例组。个例配比可分为人为与自然两类。人为配比指由观察者根据自己的知识,从可比人群中找对象。自然配比指自然存在的配比对照。例如用双胞胎患病的一方或用未患病的同胞作对照。人为与自然配比的主要区别在于,人为配比的条件是讲明的,而自然配比的条件是内涵的。例如双胞胎可能在若干基因和环境上都类似,它们都可能与所研究的疾病有关。

2. 随机抽样:指靶人群中每个人都有已知的一个概率被选为研究对象。当被选中的概率大家都等同时,称为单纯随机抽样。这样,抽出的研究人群的各种因素可以代表靶人群。

随机抽样与实验中随机安排对象接受不同措施或不同治疗。实验中的随机分配是为了使其他各种危险因素在两组中均匀分布,使两组有可比性。随机抽样是为了能有代表性,因而观察研究中的随机选择不能取代实验中的随机分配。

3. 分层抽样:指在靶人群的不同亚组或不同层次中分别抽样,通常各层中再用随机抽样即分层随机抽样,这样,不同层次的成员可有不同的被选择的机会。层次可根据下列一个或几个变量的不同等级或程度来确定。

(1) 已知或疑似某病的危险因素:根据一个危险因素的程​​度分层,各层内以不同概率的分层随机抽样,使每层都有足够多的病例。例如研究冠心病时可多选择一些年轻人,因为他

们较少患冠心病。但如果研究的目的是比较研究因素对不同民族或种族疾病的影响,则尽管他们在靶人群中的人数不相同,仍将在各民族或种族中选择相等的人数。

(2) 便于抽样的因素:分层抽样可使研究工作易于实施。例如可从不同医院中抽样,即使各医院疾病患病率差别不大。

(3) 研究因素:当研究因素的某些层次人数很少时,按研究因素来分层抽样可能较方便。例如比较输精管结扎与未结扎者。因为人群中输精管结扎者人数不多,到输精管结扎诊所去收集这些对象较为方便。等数量的两组比较,其统计效率亦较高。

(4) 根据所研究的疾病分层再选择对象:主要用于处理少见病,如同对较少暴露的研究因素分层。

分层抽样的目的是使分析结果更为有效或资料更易收集,因而在观察性研究中,其重要性仅次于限制研究对象及随机抽样。

二、观察方法

1. 人群类型:研究人群的类型与调查的性质有关。横断面调查只对人群观察 1 次,而纵向调查则至少观察 2 次,并有随访。

(1) 固定人群:在某一时间明确一组对象,随访一段时间,以检测新病例。“固定人群”是指在开始观察后,虽因不参加、迁移、死亡等原因人群不断有所减少,但不允许其他人员进入。一般说来,固定人群的平均年龄依随访年数的增加而上升。

(2) 动态人群:在随访过程中允许退出,也允许进入。例如,随访某县居民或某厂工人,观察 1985~1990 年发生黑色素瘤的新病例。如果动态人群的数量及年龄分布不变时,称为稳定。因而稳定的动态人群的平均年龄一般不变。

有时亦可将动态人群看作固定人群。将参加研究的时间作为开始暴露时间。例如,观察冠心病患者中吸烟与冠心病发作的关系,可以收集某医院 3 年内所有第 1 次发生心肌梗死的病例作为研究对象,收集他们的吸烟史并随访一时期,观察再次发作的情况。虽然设计中用的是动态人群,分析时可作为固定人群。不管他们进入研究的确切日期,将第 1 次心肌梗死作为观察的起点,开始随访。但这里对各人随访的时间是长短不一的。

2. 疾病的含义:流行病学研究中,疾病可以指一时序期后健康状况的改变或某种健康状况。疾病的变量可以是连续性(定量)的,如血压值,或分类(定性)的,如有无高血压。最常用的是分成有无疾病两大类,但有时也不易区分。例如诊断成人糖尿病,医院一般根据症状和体征及实验室检查包括葡萄糖耐量试验的结果来确定。然而,对于 65 岁以上的人群,不能认为糖耐量试验的某水平可以区分病例与非病例。描述疾病常用的率有:患病率、发病率、总发病率(total incidence)、死亡率、病死率等。

3. 方向性(directionality):是观察性研究设计中的重要问题。

(1) 前向设计(forward design):从研究因素开始随访对象,检测新病例或其他健康状况的改变。所有的实验研究及准实验均属此类设计,而大多数观察研究却不是。

(2) 后溯设计(backward design):调查者从疾病状态开始(病例及非病例),取得个人的有关研究因素的历史,例如以往暴露史、事件或特征等假设的危险因素。

不论前向设计或后溯设计,调查者都能明确研究因素是否真正出现在疾病发生之前。

(3) 不定向设计(nondirectional design):调查者同时观察研究因素与疾病,不能分清先

后。研究可以是一个靶人群样本的横断面调查,亦可对人群进行随访。此期间,研究因素与疾病都有变化。例如,测定 1960~1970 年血压变化与冠心病变化的关系,这里并不能肯定假定的病因是否较疾病先发生。

方向性与流行病学结果的解释有关系,左右着研究者能否证实因先于果的能力。

4. 时序(timing):指从开始研究或从开始收集最新资料至出现所研究的现象,即研究因素和疾病的时间先后关系。

(1) 前瞻性研究:在一个全部前瞻性的研究中,研究者在开始研究后直接观察研究因素与疾病情况,因而能按研究计划作记录。它可采用前向设计或不定向设计,但不会采用后溯设计。

(2) 回顾性研究:在完全回顾性研究中,研究因素与疾病均在研究开始前已经发生。研究者从各种记录或从对象及其亲属的回忆中获得资料。它可采用任何方向,但很少为不定向设计。

(3) 双瞻设计(ambispective):它结合了前瞻性及回顾性设计。有一个变量在前瞻中测量,另一个变量却是回顾的。例如儿童发育对今后肿瘤发生的关系研究,可找一群中年人询问他们儿童期的经历和社会关系(回顾),并随访观察其肿瘤发生情况(前瞻)。

某原始变量既可在回顾中测量,又可在前瞻中测量。例如,可询问这群中年人在研究开始前的肿瘤史。即肿瘤发病率既可用回顾方式获取,亦可用前瞻方式取得。

时序并不直接影响我们判别前因后果的能力,然而时序使我们便于估计资料的质量。一般说来,回顾性资料不论来自记录或回忆都较前瞻性资料易包含测量误差。

5. 观察单位与分析单位:观察单位指能直接测量某变量的人的某集聚水平。此单位可能是个人或一群人,如家庭、社区或县等。流行病学多数的资料是以个人为观察单位收集的。

分析单位指各有关变量可以进行分析的共同水平。流行病学中分析单位往往是个体(个体分析)或组(生态或集聚分析),都可从前向性或不定向性观察中获得。后溯研究一般为个体分析。个体分析与生态分析的区别是:在生态分析中每组兼有研究因素及疾病的个体数是不知道的,而在个体分析中是明确的。

如某变量的观察单位与分析单位相同,此变量就称为原始变量。例如,比较心血管病的死亡率与研究人群中每人所饮水的平均水质硬度。如作个体分析,则这两个变量都是原始的。当这两个变量不是从个人经历中提取时,分析单位可以是组。如观察不同地区水质硬度与心血管病死亡率的关系,水质硬度是从各地给水中获得的,心血管病死亡率也是地区性的,这里的研究因素仍是原始变量。

如果观察单位和分析单位不同,此变量称为衍生变量。例如,可将当地水质硬度值赋予该地的每个人,作个体分析水质硬度与心血管病死亡率的联系。这里的研究因素是衍生的,因为观察单位是地区,而分析单位是个体。若将平均个人饮水硬度作为地区的水质硬度进行生态分析,研究因素是衍生的。这里的疾病变量也是衍生的,因为它用个人疾病的合计情况来估计各地区的率。

因而,可以将某分析中的各变量归入下列 4 类中。

(1) 原始变量:作个体分析。

(2) 原始变量:作生态分析。

(3) 衍生变量:作个体分析。