

中国医学统计百科全书

医学研究统计设计分册

总主编 徐天和
主 编 徐勇勇
主 审 汤旦林

-61



人民卫生出版社

288120

R195.1-61
X564
2004
C-1

中国医学统计百科全书

医学研究统计设计分册

总主编 徐天和

主 编 徐勇勇

主 审 汤旦林



00132860

人民卫生出版社

图书在版编目(CIP)数据

医学研究统计设计分册/徐勇勇主编. —北京:
人民卫生出版社, 2004. 4

(中国医学统计百科全书)

ISBN 7-117-05953-2

I. 医… II. 徐… III. ①医学统计-中
国-百科全书②医学-科学研究-医学统计
IV. R311-61

中国版本图书馆 CIP 数据核字(2004)第 005841 号

中国医学统计百科全书
医学研究统计设计分册

总 主 编: 徐天和

主 编: 徐勇勇

出版发行: 人民卫生出版社(中继线 67616688)

地 址: (100078)北京市丰台区方庄芳群园 3 区 3 号楼

网 址: <http://www.pmph.com>

E - mail: pmph@pmph.com

印 刷: 北京铭成印刷有限公司

经 销: 新华书店

开 本: 787×1092 1/16 印张: 13

字 数: 306 千字

版 次: 2004 年 5 月第 1 版 2004 年 5 月第 1 版第 1 次印刷

标准书号: ISBN 7-117-05953-2/R·5954

定 价: 26.00 元

著作权所有, 请勿擅自用本书制作各类出版物, 违者必究

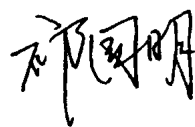
(凡属质量问题请与本社发行部联系退换)

总 主 编	徐天和				
主 编	徐勇勇				
副 主 编	蔡 辉	夏结来			
主 审 委	汤旦林				
编	(以姓氏笔画为序)				
	万崇华	尹全焕	王 玖	刘延龄	孙 高
	张罗漫	李 康	李克均	李炽民	杨振明
	陈平雁	孟 虹	易 东	郑俊池	姜 明
	胡良平	赵清波	夏结来	徐天和	徐勇勇
	高纯书	曹秀堂	黄先敬	蔡 辉	颜 艳
	魏庆琤				
秘 书 长	祁爱琴				
学术秘书	王 玖	石德文			

序

随着医学科学研究的不断深入发展,对科研设计与科研数据的统计处理方法提出了越来越高的要求,现代医学统计科学的新理论、新方法也在不断涌现,但迄今为止国内尚无一部大型医学统计专著。中国医院统计杂志社发起并组织国内百余名从事医学统计研究与教学的专家、学者编撰的这部《中国医学统计百科全书》填补了这一空白。

该书的各位编写者都是工作在医学统计研究与教学第一线的专家、学者,其中归国留学人员占60%,他们为国内带回了最新研究成果。我衷心祝贺本书的出版,并为我国医学统计学界人才辈出、长江后浪推前浪的喜人形势而感欣慰。我相信该书的出版发行一定会大大推动我国医学科研事业的发展。



2003年10月

中国医学统计百科全书

序 言

世界各国尤其是发达国家都非常重视百科全书的编纂工作。它是衡量一个国家科学技术发展水平乃至综合国力的标志之一。随着时代的前进和科技的进步,特别是党的十一届三中全会以来,我国的社会主义现代化建设和文化科学事业蓬勃发展。统计科学和医学统计工作的发展进入了一个新阶段。统计方法作为一种获取信息和科学研究的工具与作出决策的依据,其重要性正被越来越多的人所认识。为总结我国改革开放以来统计科学和统计工作的成果经验,吸收与传播现代医学统计科学的新理论、新方法与新成果,迫切需要编撰一部具有中国特色的医学统计百科全书,以填补国内这一空白。为此,中国医院统计杂志社发起并任总主编单位,组织全国部分高等医学院校及有关医疗卫生机构的百余名统计学教授、专家,经过6年的辛勤劳动,发挥集体智慧,共同编撰了《中国医学统计百科全书》。

《中国医学统计百科全书》是一部大型医学统计参考工具书,主要读者对象是全国医学统计工作者和医疗卫生单位统计信息工作者以及高等医学院校的师生,预计在工作中需要查阅这部百科全书的读者将远远超过这一范围。全书包括描述统计、推断统计、非参数统计、多元统计、统计设计和健康与管理统计等内容,选材着重其在医学上的应用。由于近年来统计理论与方法发展十分迅速,高效、实用的新方法不断出现,像多变量分析方法和非参数统计分析等较高层次的统计方法,在应用上日益普遍,本书以较大的篇幅重点进行了阐述。以便既体现统计理论与方法的完整性,又反映学科的先进水平。全书用条目形式撰写,一个条目介绍一种统计方法,包括问题的背景、方法的理论和直观依据、实施步骤和实例等。由于书的性质,不可能对方法的数学理论根据作严格的推导,但也避免把本书写成一种“菜谱”式的东西。因为,为了用好一种统计方法,对其背景和依据



有些了解,是很重要的。我觉得这是本书写作的一个重要特点,这样使本书能比较全面而确切地介绍了医学统计科学的重要内容与最新研究成果,又不失其以应用为主的特性。另外,本书的编撰方法科学性强,层次分明,结构严谨,既突破了传统的辞典式编撰方法,又汲取了辞典的某些特点。本书强调实用性,深入浅出,具有知识性、可读性、可查性和适用性,它适合于医学各专业、不同层次和不同专业需要的读者阅读。作为医学界、医学统计理论界和医疗卫生统计信息部门的一部大型专业工具书,我相信它会成为这方面专业人士书架上常备之书,对推动我国医学统计科学和统计工作的发展作出积极的贡献,故乐为之序。

陈希孺

2003年7月

编者说明

对于任何一项科学研究,研究设计的好坏直接关系到研究成果的质量,正确的统计设计不仅能帮助研究者节约时间和经费,也是保证统计分析结论令人信服的基本前提。正如现代生物统计学的奠基人之一 R. A. Fisher(1890~1962)所说:“试验完成后再去找统计学家求教,无异于请统计学家为试验结果进行‘尸体解剖’。统计学家或许只能告诉你试验失败的原因。”作为《中国医学统计百科全书》的一个分册,《医学研究统计设计》分册的主要内容分为统计设计概述、实验设计方法、调查设计、抽样设计、临床试验设计、样本含量估计六个方面的内容。各条目的编者,许多是医学院校资深的专家学者,年轻的同志也大多具有博士、硕士学位,并有多年从事医学科研的实践经验。本书将医学研究设计的有关统计学概念、方法以最简洁的形式写成条目,大部分条目不需要读者具备统计学基本知识,不仅可供广大医学科研人员、临床医生、卫生防疫人员作为工具书查阅,也可以作为医学院校学生、在职人员学习医学统计学课程的参考书。

由于本人水平所限,本书一定存在许多不足和缺陷,希望广大读者批评指正。特别是由于医学科学的迅速发展,本书内容不可能涵盖所有医学研究中的统计设计问题,一些重要的内容也难免遗漏,如循证医学中的设计问题。这些遗憾只能等下次本书再版时加以弥补。

徐勇勇

2003年6月

目 录

目 录

医学研究设计	(1)
实验对象与实验单位	(6)
随机数	(7)
随机化	(8)
随机排列表	(9)
干预与对照	(10)
样本大小与检验效能	(11)
实验因素与水平	(12)
实验效率	(13)
实验效应	(14)
数据误差与数据精度	(15)
数据效度	(16)
测量信度	(17)
预试验	(21)
质量控制	(22)
主效应	(24)
交互作用	(25)
关联与因果	(26)
完全随机设计	(27)
配对设计	(29)
配伍组设计	(30)
平衡不完全配伍组设计	(32)
交叉设计	(34)
拉丁方设计	(39)
希腊—拉丁方设计	(42)
析因设计	(44)
正交设计	(47)
均匀设计	(52)
系统分组设计	(54)
裂区设计	(56)
序贯试验设计	(58)



重复测量设计	(62)
调查设计	(67)
普查	(73)
典型调查	(74)
横断面调查	(75)
前瞻性调查	(76)
追踪性调查	(77)
回顾性调查	(80)
随机抽样	(84)
抽样总体	(85)
抽样框	(86)
单纯随机抽样	(87)
分层抽样	(88)
系统抽样	(90)
整群抽样	(92)
多阶段抽样	(93)
快速评价抽样	(94)
临床试验分期	(97)
盲法试验	(99)
试验预案	(100)
完全随机分组	(102)
有限随机分组	(103)
区组随机分组	(104)
分层随机分组	(105)
适应性随机分组	(106)
试验监测	(108)
期中分析与成组序贯试验	(109)
生命质量研究的设计	(111)
临床决策分析	(114)
诊断试验	(116)
试验清洗期	(125)
随访与失访	(126)
试验终结	(127)
终结后随访	(128)
实验报告	(129)
纳入标准与剔除标准	(132)
知情同意	(133)
预后因素	(134)
多中心试验	(135)
样本含量估计的条件	(136)



总体均数区间估计样本含量	(137)
总体率区间估计样本含量	(139)
样本均数与总体均数比较样本含量估计	(140)
样本率与总体率比较样本含量估计	(141)
两样本均数比较样本含量估计	(143)
两样本率比较样本含量估计	(144)
多组样本均数比较样本含量估计	(145)
多个样本率比较样本含量估计	(147)
前瞻性(队列研究)样本含量	(148)
回顾性(病例对照)样本含量	(149)
附录一 统计用表	(150)
附表 1 随机数字表	(150)
附表 2 随机排列表($n=20$)	(151)
附表 3 随机排列表($n=10$)	(151)
附表 4 平衡的不完全单位组设计的参数表	(152)
附表 5 平衡的不完全单位组设计表	(153)
附表 6 T 界值表(两样本比较的秩和检验用)	(157)
附表 7 质反应单向序贯试验边界系数表	(158)
附表 8 量反应单向序贯试验边界系数表	(160)
附表 9 小样本翼型设计的边界点坐标	(161)
附表 10 总体率区间估计试验所需样本数表	(163)
附表 11 样本均数与总体均数比较(或配对比较)时所需样本例数表	(166)
附表 12 样本率与总体率比较样本含量估计表(单侧)	(167)
附表 13 样本率与总体率比较样本含量估计表(双侧)	(170)
附表 14 两样本均数比较所需样本例数表	(173)
附表 15 两样本率比较时所需样本例数(单侧)	(174)
附表 16 两样本率比较时所需样本例数(双侧)	(175)
附表 17 Ψ 值表(多个样本均数比较时所需样本例数估计用)	(176)
附表 18 多组样本均数比较时样本含量估计用表	(178)
附表 19 λ 值表(多个样本率比较时所需样本例数估计用)	(180)
附表 20 t 分布界值表	(181)
附录二 英汉医学统计学词汇	(183)
附录三 汉英医学统计学词汇	(186)



医学研究设计

生物的变异性是医学研究中普遍存在的现象。医学研究中的许多事例证明,如果没有统计学的支持,将一次或几次试验的偶然成功看作必然现象,并将其作为研究成果加以推广,不仅使科学研究失去其应有的意义,而且可能会给人类的健康带来灾难性的后果。

1. 医学研究类型与统计学要求

医学研究可分为描述性研究和分析性研究两大类,前者主要以研究者的真实观察结果为依据,描述新发现;后者主要是以研究者的试验为基础,研究和推论试验以外的特定总体是否具有同样的特征。在医学研究中,有时这两类研究并不能严格区分,在有的研究中还会两者兼而有之。以观察对象的选择方法为判断标准,医学研究还可区分为以下几种类型:

(1)单一样本研究:按一定标准选择一个或若干个生物个体或生物材料作为研究对象,主要用于描述性研究。如阐明新的生命现象、形态结构,研究新的检测方法、诊断和治疗技术。在临床上,最多见的单一样本研究是Ⅰ期临床试验,即以志愿者作为观察对象,进行药代动力学和药物安全性的观察。由于单一样本的研究不设立对照组,在基础研究中,一定是新发现的科学事实,在临床研究中,则是Ⅱ期临床试验必须经过的试验阶段,但Ⅰ期临床试验不能作为判定治疗的依据。

(2)观察对比研究:观察对象人为选择,数据资料以观察的方法收集,通过观察结果的前后对照、历史对照或相互对照得出研究结论。常见的观察对比方法有以下三种。

1)无对照研究:通常是回顾性的,也可以是前瞻性的,以治疗前后病情变化判断疗效。多见于临床病例总结。

2)历史对照:通常是前瞻性研究,但由于缺乏足够的病例(如罕见病)或已有公认历史对照,在研究中不设立对照组,以文献报道或以往常规治疗方法的治疗结果作为对照。

3)非随机化同期对照研究:前瞻性的研究,在研究中设立对照组和试验组,但受试者的分组按非随机化的方式进行,如接诊医生根据病情、患者意愿选择治疗方法,或按患者经治医生、所属病区(医院)分组等。在涉及环境等自然因素的较大规模的现场试验中,准试验的分组方法应属于非随机化同期对照研究。出于伦理学的考虑,有些临床研究只能采用观察对比的研究方法,如某病患者与正常人的对比、吸烟者与非吸烟者对比等。在观察对比研究中,通常样本量大(除罕见病),时间跨度长,由于诊断、治愈标准的不同以及对照组和试验组患者的病情程度、年龄、病程、病情、用药史的不同,在进行数据分析时要用统计学方法调整和控制这些因素对观察结果的影响。

(3)随机对照研究:按照实验设计重复、对照和随机化的三个原则进行分析性研究。所谓重复就是同样的试验条件下要有符合统计学要求的观察例数;对照就是在试验内安排同期对照组,而且必须是专业上公认的可靠对照方法;随机化就是保证每个观察对象(生物材料)分到试验组和对照组的机会均等,只有随机对照的研究结果,如Ⅱ期临床试验,才能使研究的推论具有普遍性的科学意义。

(4)调查研究:通过观察部分个体或全体



了解特定总体(人群)的一般特征,如健康状况、生育模式、疾病流行因素、老年人医疗社会保障情况等。调查研究要有明确的抽样框架(“目标人群”)和调查时点,调查部分个体时要采用随机抽样的方法。

2. 干预试验的设计问题

干预试验的主要设计问题是,如何在实验设计和试验过程中遵循重复、对照和随机化的原则。

据文献记载,有控制的一次医学干预试验发生在1747年。1747年5月20日,英国的J. Lind医生将12名病情相类似的患者带到一艘船上。患者的主要症状是牙龈溃烂,皮肤有出血点,双膝无力。Lind将12名患者分为6组,分别给予下列主要干预:

A组:每天饮1夸脱(1.136L)苹果汁;

B组:服25滴硫酸丹剂(elixir vitriol),每天3次;

C组:服2匙醋,每天3次;

D组:每天饮约半品脱海水,服缓和的泻药;

E组:每天食2个桔子,1个柠檬;

F组:每天服由大蒜、芥子等成分组成的干药。

当6月16日船返回英国Plymouth港时,所有患者的病情都有好转,其中E组恢复的最快、最好,其中1人到第6天就可以工作了。B组也有1人比登船时健康。Lind医生的试验并没有得出明确的结论,直到160年后,通过动物实验才真正从科学意义上解释了E组疗效最好的真正原因。

根据现代医学知识, Lind医生描述的病例是坏血病患者的常见症状,但目前已无从考察其中真正坏血病患者的人数。从干预效果看, E组中病情好转最快的患者可能是真正的坏血病患者(坏血病的对症治疗是补充维生素C)。用现代统计学的观点看, Lind医生的干预试验至少有两点符合统计学要求。一是设有对照组, 6个干预组相互

比较;二是进行了质量控制,所有患者在同一条船上,便于监督服药和观察病情。关于无对照和质量控制不严而导致研究结果“失真”的例子,在近代医学的研究中也没有完全避免。

例1 门腔分流手术治疗肝硬化门静脉高压症,前后有51次研究(表1)。无对照研究支持门腔分流手术的比例是24/32(75%),而按统计学原则进行的4次随机对照研究,认为手术几乎没有价值。无对照的研究夸大了门腔分流手术的作用。

表1 关于门腔分流手术的51次研究结果

对照方式	门腔分流手术的价值			小计
	非常支持	支持	不支持	
无对照	24	7	1	32
非随机对照	10	3	2	15
随机对照	0	1	3	4
合计	34	11	6	51

Journal of Gastroenterology. Vol 50. 1966. 646-691

例2 一项随机对照的双盲试验,受试者为8341名有心脏病的中年男子,干预是可能防止心脏病突发的5种药物和安慰剂(相同的药物胶囊装入乳糖),用5年追踪观察的死亡率评价6种干预的效果。服用安妥明(Adherence)的患者中有357人没有达到应该服用药量的80%,其死亡率达24.6%,而坚持服药者的死亡率为15.0%,所以不坚持服药是增加死亡率的主要原因,不能归究于安妥明的预防效果。再看表2中的失访人数(尽管在慢性病研究中失访不可避免),合计死亡率18.2%,是假定失访的38人(失访率3.6%)全是病情轻的患者,且无一例死亡。如果失访者全是重症患者,或失访率达到20%、30%或更多,你还能信任表2中的合计死亡率吗?

受试者是否遵从医嘱,并使干预措施得到充分落实的程度,称为患者的依从性(compliance)。依从性和失访是干预效果评价的两个“陷阱”。



表 2 服用安妥明的死亡率

分 组	人 数	死亡人数	死亡率/%
服药 \geq 80%	708	106	15.0
服药 $<$ 80%	357	88	24.6
失 访	38	—	—
合 计	1065	194	18.2

如果今天重复 Lind 医生的干预试验,在统计设计和统计分析上还应做到:

(1)所有受试者均为经临床和实验室检查确诊的坏血病患者。

(2)有足够多的试验病例。样本估计方法详见样本含量估计有关条目。

(3)采用随机分组方法,保证各组干预效果的可比性。Lind 医生的试验将病情最重的 2 名患者分到 D 组(泻药是当时流行的治病方法),D 组的干预效果不及 E 组和 B 组也可能是病情所致,未必是干预的真实效果。

(4)用定量测量数据表示干预效果,如血清(血浆)的抗坏血酸的含量、尿中的抗坏血酸排出量等。Lind 医生的干预效果评价用的是“病情是否好转”,属定性观察。定性观察所用的试验样本通常远远多于定量观察。

(5)用假设检验的方法比较两组均数(多组均数)间的差别,以推论本次试验的结果是否对所有坏血病患者都有效。

到 19 世纪,现代科学思想和科学方法在医学研究中的应用初见端倪。代表性的人物是法国医生 P.C.A. Louis。1835 年,对当时流行的“放血”疗法治疗肺炎的效果进行了比较,发现“放血”的疗效不像预期的那么好,而且早期“放血”和晚期“放血”组比较,患者的诊断、病情、病程、年龄等方面的差异很大,比较平均治愈时间意义不大。因为晚期“放血”组的平均治愈时间长,但该组患者病情重、病程长、年龄大。Louis 的研究虽然没有彻底否定“放血”疗法,使在西方延续了 2000 余年的“放血”疗法一直持续到 20 世纪初,但他对医学研究的方法学作出了很大贡献:用数字的方法表示不同疾病患者的特征和预后;提出

了临床疗效对比的前瞻性原则;提出了临床研究的抽样误差和混杂(confounding)概念。其中前瞻性研究和减少混杂,分别成为当今临床试验和流行病学研究的基本研究方法。

20 世纪 20 年代,英国统计学家 R.A. Fisher 爵士(1890~1962)在伦敦附近的 Rothamsted 农业实验站,创立了现代实验设计方法和统计分析技术,如完全随机分组设计、随机区组设计、拉丁方设计、析因设计等,奠定了生物统计实验设计的基础。1948 年,英国发表了评价链霉素治疗肺结核疗效的临床试验报告,第一次采用随机对照的设计方法进行临床干预试验。目前,大样本随机对照设计是国际公认的临床疗效评价的“金标准”,并且发展为多个大样本随机对照研究结果的系统评价(systematic review, SR)或 Meta 分析,以此作为证明某种疗法的有效性和安全性的最可靠依据之一。

3. 观察对比的设计问题

观察是在没有干预或没有控制的情况下了解事物的自然变化规律。观察对比的设计问题是,如何获得有代表性的观察样本,如何尽可能地保证各对比组均衡可比。

早在 1839 年,英国就有了覆盖全人口的死亡登记制度。通过对 1848~1849 年霍乱流行时期伦敦地区的死亡人数分析,在死亡登记处工作的统计学家 W. Farr 发现,霍乱的死亡率与地势高度呈负相关。当时对霍乱的流行原因主要有两种观点,一种是 Farr 的有统计结果支持的“瘴气”学说(混浊空气致病),另一种是 J. Snow 提出的“水污染所致的传染病”的假说(比 Kock 和 Pasteur 的微生物学早 20 年)。当 1853~1854 年霍乱再次流行时,Farr 和 Snow 一起分析资料。Snow 对比了伦敦自来水厂供水范围的霍乱死亡人数(表 3)。Lambeth 公司在污染轻的泰晤河上游取水, Southwark 和 Vauxhall 在污染重的泰晤河下游取水。显然,表 3 的统计结果支持 Snow 水污染的假说。1883 年,



R. Koch 在水中发现霍乱弧菌,最终证实了 Snow 的假说,否定了“瘴气”学说。

表 3 1853~1854 年伦敦霍乱死亡率(1/万户)

水 源	用户 数	死亡 人数	死亡 率
重污染(Southwark 和 Vauxhall 公司)	40 046	1 263	315.4
轻污染(Lambeth 公司)	26 107	98	37.5
伦敦其它地区	256 423	1 422	55.5
合 计	322 576	2 783	86.3

观察呈现出的统计规律是许多医学研究形成假说的基础,但不是支持假说的唯一证据,更不能代替其它科学方法进行的观察和实验。伦敦霍乱流行时,“瘴气”学说和“水污染”假说都有统计结果的支持,用 Koch 在水中发现霍乱弧菌的“金标准”评价,前者是在解释上出现了偏倚(bias),或者说海拔高度是水污染严重程度的一个混杂因素。事实上, Snow 对表 3 观察结果的解释十分谨慎。他认为只有在实验的条件下,即保证各自来

水用户除水的来源不同外,其它各方面均相同的条件下,对表 3 的统计分析才有意义。

英国的死亡登记制度,后来发展为包括生命统计、生育统计、儿童生长发育统计、传染病统计、肿瘤统计为主要内容的卫生统计。生命统计虽然是全人口的资料,但统计分析也不能简单地比较数值的大小,否则也会得出荒谬的结果(例 3)。对于难以依靠全人口的登记制度获得数据,如许多疾病的发病率和患病率的统计,必须依赖抽样调查。通过调查设计和相应的随机抽样方法,使抽样调查的结果能代表全人口的情况。

例 3 根据英国某年全人口的统计资料,比较男性某病的发病率(表 4)。简单地比较发病率。英格兰和威尔士的发病率为 132.1/10 万,移民的发病率为 65.7/10 万,结论是英国本土居民比移民更容易得病(2 倍的相对危险度)。如果将人口按年龄分组,就明显看出上述结论是错误的,因为每个年龄组都是移民的发病率高于本土居民。

表 4 英格兰和威尔士男性与移民男性的发病率(1/10 万)

年龄分组 /岁	英格兰和威尔士			移 民		
	人口/千人	发病数	发病率	人口/千人	发病数	发病率
0~	1 900	1 406	74.0	26	21	80.8
5~	3 100	186	6.0	30	2	6.7
15~	9 400	1 786	19.0	127	27	21.3
45~	4 900	7 350	150.0	25	42	168.0
65~	2 000	17 400	870.0	5	48	960.0
合 计	21 300	28 128	132.1	213	140	65.7

用观察对比的方法对非干预致病因子的研究,如病例对照研究(case-control study)、队列研究(cohort study)等,是流行病学调查最常用的方法。1964 年 R. Doll 和 A. B. Hill 将接受调查的 4 万名英国注册医生分为吸烟和不吸烟两组,通过以后的肿瘤统计发现,吸烟组和不吸烟组肺癌的年发病率分别为 1.66% 和 0.07%,吸烟者患肺癌的相对危险度为 $1.66/0.07 = 23.7$,强烈提示吸烟的致癌作用。在吸烟致癌的假设下进行的动物

实验也证实了这一点。尽管 Doll 和 Hill 无法按随机的原则对两组医生进行合理地分组,对吸烟者或不吸烟者的可比性也存在种种质疑,如吸烟者的年龄大、善饮酒、工作压力大、经济状况差等等,但如此大的相对危险度足以抵消组间可比性不好所带来的偏倚。

值得提醒的是,像吸烟与肺癌这样的强关联观察对比结果,在现在慢性病流行病学研究中已比较少见。对于多个致病因子综合作用引起的慢性病,某个致病因子的单独作



用可能很微弱(相对危险度约 1~2),则组间可比性所带来的偏倚不能忽视,或者说组间的可比性决定了相对危险度的大小。不幸的是,现在对与日常生活有关的致病因子的研究,如洗发水与淋巴癌、咖啡与心脏病、电磁场与白血病,常会出现这种微弱联系,并导致同一类问题得不同研究结果。尽管组间可比性差的资料,在统计分析时可以用复杂的统计方法控制和调整组间的可比性,但影响组

间可比性的因素如果根本不知道或者没有记录,再复杂的统计分析方法也无能为力。因此,研究“弱关联”的现象,在研究设计阶段应该尽可能合理选择样本,收集数据要标准化和数量化,研究设计可以多个层次,如分子—细胞—个体—家庭—社区—社会—环境—生态,以便采用更精细的研究设计方法。

(徐勇勇)



实验对象与实验单位

医学实验对象主要是人体和动物,实验单位则是接受处理的基本单位。实验对象与实验单位有时是一致的,如临床试验,每个患者接受一种治疗,实验对象(患者)就是实验单位。但在许多情况下,实验对象与实验单位是不一致的。例如用动物作实验,根据处理的方式,实验单位可以是动物个体,也可以是动物的某一肢体或组织。实验单位亦可划分为若干级别。如用家兔作实验,观察指标是家兔眼房水中某种物质的含量,如处理是给家兔全身注射药物,则家兔是一级实验单位,若分别给家兔的两眼造成不同程度的局部损伤,则兔眼为二级实验单位。如果在试验中既给家兔注射药物,又给兔眼造成局部损伤,则在该试验中既有一级实验单位,又有二级实验单位。在比较药物作用时,将家兔

看作实验单位;在比较局部损伤作用时,将兔眼看作实验单位。也就是说,药物的作用在一级实验单位中比较,局部损伤的作用在二级实验单位中比较。在以上实验中,实验单位是兔眼,而不是家兔个体。同理,在有些医学实验中,尽管观察单位很小,处理间的差异也应在实验单位中比较。如用毫米波辐射小鼠后观察小鼠肝细胞超微结构,尽管每个小鼠的肝脏可制成很多切片进行观察,但接受处理的单位仍是小鼠,不是小鼠的肝脏切片,所以,实验单位数仍是小鼠个数。

准确地说,实验设计的样本大小估计是对实验单位数的估计,而不是实验对象的估计。

(徐勇勇)