

21 世纪高等医药院校教材

(供医学类各专业研究生用)

医学科研设计教程

主编 熊国强

主审 贺石林

编者 (按编写章节顺序)

熊国强 王乐三 胡平成

曾小敏 周 欢

内 容 简 介

本书介绍医学科研设计的基本知识和基本原则,并针对不同学科特点阐述实验设计、调查设计、临床实验设计及干预试验设计的基本原理、具体实施和分析方法。书中各章节举出实例,便于读者理解课程内容。在各章前设置“学习目标”,章后有“结语”、“练习题”、“思考题”,分别提出必须掌握的内容,概括了全章的要点,以帮助读者练习和复习。本书可作为高等医学院校研究生教材,对从事临床医学、预防医学、社会医学、医学管理的卫生工作者也是一本理想的医学科研设计参考书。

图书在版编目(CIP)数据

医学科研设计教程/熊国强主编.-北京:科学出版社,2001.10

21世纪高等医学院校教材

ISBN 7-03-009743-2

I. 医… II. 熊… III. 医学-科学研究-设计-医学院校-教材

IV. R

中国版本图书馆CIP数据核字(2001)第060320号

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

印刷

<http://www.sciencep.com>

科学出版社发行 各地新华书店经销

*

2001年10月第一版 开本:850×1168 1/16

2001年10月第一次印刷 印张:13

印数:1-4000 字数:266000

定价:20.00元

(如有印装质量问题,我社负责调换〈新欣〉)

前 言

医学科研设计是医学生的一门重要的基础课程,也是预防医师、临床医师、基础医学教师和卫生管理工作从事实践工作和科学研究必备的基本知识。医学科研设计方法被医学界比喻为“建筑现代化医学大厦的顶梁柱”。

尽管医药科学分支较多,专业性强,各学科科研都有各自的专业特点,但是无论何种学科的科研,其研究过程、原理、原则和基本方法都有一定的共同性。本教程就是针对医学科研中的共性问题进行论述,并且针对每种设计方法的难点,通过实例深入浅出地介绍基本知识和实际应用原则。试图为医药学各学科研究生、本科生与医师(药师、检验师、护师、教师)涉足科研领域起到抛砖引玉的作用。

医学科研过程包括选题、设计、实施、统计分析与总结成文五个环节,设计思路与统计分析紧密结合将五个环节牢牢地拴在一起,形成一个严密的整体。医学研究的结论往往关系到民众的健康,关系到疾病的预防、诊断、治疗与转归,甚至生死问题,如果设计不当将涉及医学伦理学问题。

21 世纪的医学教育方式将随着电脑、统计软件、多媒体教学和医学新模式的引入而发生深刻的变革,在单位时间内授课的信息量将大幅度增加,这无疑给教师、学生和教材提出更高的要求。本教程在内容的编排上力求顺应这一变化趋势。书中详细介绍了科研设计的基本知识和基本原则,并针对不同学科特点介绍了实验设计、调查设计、临床实验设计及干预试验设计的基本原理和实施与分析方法。教师可根据不同的教学对象灵活选用。该教材的另一特点是各章节设置了学习目标、结语、练习题与思考题,提出了需掌握与熟悉的内容,结语还简明扼要地概括了本章节的基本要点,以便学生顺利完成练习和复习。对从事临床医学、预防医学、基础医学、社会医学和管理科学的卫生工作者也是一本有用的参考书。

本教程的作者都是多年从事医学统计学教学和科研工作的教师,具有丰富的教学经验和扎实的科研功底。总结多年积累的教学经验和科学知识,编写成一本实用性强的教科书,奉献给他们的学生,并用来与同行们交流,这就是主编和作者的初衷。

本教程编写过程中,特邀贺石林教授作为主审,在此衷心感谢贺教授接受我的邀请并为此付出的辛勤劳动。

本教程的编写得到科学出版社和中南大学的大力支持,特别是中南大学教务处刘富友、宋爱丽以及成教学院潘辉英、龚爱云等同志都非常关心这本书的编写,借此机会表示衷心的感谢。湘雅公共卫生学院卫生统计学教研室王一任老师承担了此书的排版打印、图表制作等工作,特此感谢。

在此书的编写过程中,虽然我室同仁全力投入编写并反复修改,但限于时间和我们的水平,难免出现疏误,鉴请读者批评指正,以利再版时更加完善。

熊国强

于 2001 年 7 月

序 言

本人有幸应主编之邀,首先阅读了《医学科研设计教程》一书,感受颇深。医学科研是以科学的观点和方法探索医学及其有关领域的未知事物与规律的认识活动。科研设计就是制定完成选题目标的技术路线和实施方案,它包括专业设计与统计学设计两部分。专业设计的特殊性较强,统计学设计对各专业科研均具有普遍指导意义。本教程从统计学设计出发,讲述了实验设计、调查设计和临床试验设计的基本原理和基本分析方法。全书内容丰富、新颖,实用性强。各章节均有实例、练习题、思考题、学习目标和要点总结,这对促进学生掌握统计学设计的思路、方法和分析问题的能力是十分有益的。本书既可作为医学各专业硕士研究生、本科生及专升本班的教材,又可作为各类医务人员从事科研的参考书。

参加本《教程》的编者都是长期从事医学统计学和医学科研设计方法教学及研究的教师,他们积累了丰富的经验。本书是他们多年共同努力的硕果,读者可从字里行间体会他们坚实的专业功底。祝愿本《教程》在培养学生科研能力中发挥良师益友的作用。

中南大学湘雅医学院 贺石林

2001年8月于长沙

目 录

| | | | |
|----------------------------------|-------|--|-------|
| 第一章 医学科研设计概述 | (1) | 第一节 临床试验基本概念 | (123) |
| 第二章 医学科研的基本知识 | (10) | 第二节 诊断试验研究设计和评价 | (130) |
| 第一节 医学科研假说 | (10) | 第三节 防治效果研究设计与评价 | (141) |
| 第二节 医学科研的基本要素 | (13) | 第四节 疾病预后研究设计与评价 | (149) |
| 第三节 实验设计的基本原则 | (17) | 主要参考文献 | (159) |
| 第四节 数据处理的统计方法选择 | (25) | 附录 统计用表 | (161) |
| 第三章 常用实验设计及其分析方法 | (31) | 附表 1 随机排列表 | (161) |
| 第一节 方差分析 | (31) | 附表 2 随机数字表 | (162) |
| 第二节 完全随机设计与方差分析 | (37) | 附表 3 配对比较(t 检验)所需样本例数... | (163) |
| 第三节 随机区组设计与方差分析 | (39) | 附表 4 两样本均数比较(t 检验)所需样本 例数 | (164) |
| 第四节 拉丁方设计与方差分析 | (43) | 附表 5 两样本率比较时所需样本例数 (单则) | (165) |
| 第五节 多个样本均数间的两两比较 | (47) | 附表 6 两样本率比较时所需样本例数 (双侧) | (166) |
| 第六节 序贯试验设计与分析 | (52) | 附表 7 ψ 值表(多个样本均数比较时所需 样本例数的估计用表) | (167) |
| 第七节 半数有效剂量试验设计与分析 | (56) | 附表 8 λ 值表(多个样本率比较时所需样 本例数的估计用表) | (168) |
| 第四章 多因素设计及其分析方法 | (62) | 附表 9(1) q' 界值表(单侧) | (169) |
| 第一节 2×2 析因试验设计与分析 | (62) | 附表 9(2) q' 界值表(双侧) | (170) |
| 第二节 多因素多水平析因试验设计与分析 | (66) | 附表 10(1) F 界值表(方差分析用表) ... | (171) |
| 第三节 交叉试验设计与分析 | (70) | 附表 10(2) F 界值表(方差分析用表) ... | (172) |
| 第四节 正交试验设计与分析 | (73) | 附表 10(3) F 界值表(方差分析用表) ... | (173) |
| 第五节 重复测量试验设计与分析 | (77) | 附表 10(4) F 界值表(方差分析用表) ... | (174) |
| 第五章 调查设计和分析方法 | (82) | 附表 11 χ^2 界值表 | (175) |
| 第一节 调查设计的基本内容 | (82) | 附表 12 t 界值表 | (176) |
| 第二节 现况调查研究设计与分析 | (87) | 附表 13 q 界值表(Newman-keuls 法用表) | (177) |
| 第三节 回顾性研究的设计与分析 | (96) | 附表 14 基本拉丁方设计表 | (178) |
| 第四节 前瞻性研究的设计与分析 | (107) | 附表 15 $m=2$ 的正交表 | (179) |
| 第六章 临床流行病学基础知识 | (118) | 附表 16 标准正态分布曲线下的面积: $\Phi(-u)$ 值 | (186) |
| 第一节 临床流行病学发展概况 | (118) | 附表 17 F 界值表(方差齐性检验用表)... | (187) |
| 第二节 临床流行病学的特点和任务 | (119) | | |
| 第三节 临床流行病学研究的方法学 | (120) | | |
| 第四节 临床流行病学在临床医学中的 作用 | (122) | | |
| 第七章 临床试验设计 | (123) | | |

| | | | | | |
|-------|----------------------------------|-------|--------------|--|-------|
| 附表 18 | 相关系数 γ 界值表 | (188) | 附表 23 | M 界值表(随机区组比较的 秩和检验用表) | (195) |
| 附表 19 | 百分率的可信区间 | (189) | 附表 24 | D 界值表(各样本例数相等的 Nemenyi 法用表) | (196) |
| 附表 20 | T 界值表(配对比较的符号秩和 检验用表) | (192) | 汉英名词索引 | (197) | |
| 附表 21 | T 界值表(两样本比较的秩和检验 用表) | (193) | 英汉名词索引 | (199) | |
| 附表 22 | H 界值表(三样本比较的秩和检验 用表) | (194) | | | |

第一章 医学科研设计概述

学习目标

- (1) 熟悉医学科研工作的基本任务。
- (2) 掌握医学科研工作的基本程序。
- (3) 掌握医学科研选题原则。
- (4) 掌握医学科研三类设计的特点和联系。
- (5) 了解医学统计学在医学科研设计中的地位。

医学科学研究工作由三部分组成:科研设计、科学地实施、观察与分析结果并得出结论。其中科研设计是关键的第一步,确定一个创意新颖、设计周密、指标合理、科学性强又切实可行的实施方案是取得高新成果的根本保证。现代医学是由生物医学模式向生物-心理-社会医学模式转变而发展起来的。医学科研工作也朝向这一模式转轨,无论是从宏观观察还是从微观分析都需注重这个目标。

一、医学科研的任务

医学科研工作的基本任务在于熟悉人体结构、功能与理化机制,认识疾病的分布,掌握疾病的发生、发展过程,揭示疾病与健康的转化规律及其影响因素,探索环境与健康的联系,提出有效的防、诊、治措施,改善人类生存的环境,增进人类的健康,延长人民的寿命,开发促进人民身心健康的科学技术等。

1. 发现医学中的未知事物及其内在规律

尽管现代医学技术相当发达,已从细胞、亚细胞医学水平发展到分子、基因、甚至量子医学水平,但是在医学客观世界中始终存在着大量的未知现象和未知的发展过程及规律。例如,现在发展的基因诊断和基因治疗,特别是基因治疗,虽说发展速度迅猛,但从全局观察与分析,所走过的道路仅为前进的第一步,研究和发展的前景还很广阔。

2. 寻找医学中已知事物的未知规律

科研的重要任务之一,就是揭示已知事物的外在表现(现象)的本质与内在联系(规律性)。因为只有了解事物的本质,掌握了它的规律性,才有可能对它进行利

用、干预或改造。例如恶性肿瘤的死因顺位总是排列在全死因顺位的第 2~ 3 位,给人类健康造成巨大的损害,其发病学至今未阐明;虽然医学家一致认为是抑制基因被抑制与癌基因被活化,但是对这两类基因抑制和活化的规律尚不了解,人类对癌症的防治还未取得根本性突破。

3. 探索生存环境对人类身心健康的联系

现代医学的许多难题和疑点,必须结合生命科学中各层次各专业分支综合解决,深入探讨生存环境对人民身心健康的影响,全方位寻找保护和增进健康、防治疾病的新途径。例如,北方地区的沙尘暴,沿海一带的赤潮,内陆地区江河湖泊的污染,大中城市的空气质量下降,食品和副食品的污染,还有地方病、疫源性等疾病等,都在潜移默化地损害和威胁着民众的健康。从全局来看,这也是医学研究的重要任务之一,而且意义更加深远。

4. 开发医学的应用

医学科研的另一重要任务,是探索如何利用医学事物的内外规律,在防治疾病和增进身心健康方面的应用。如基因工程技术在医学的应用就是典型的例子。

除了上述几方面以外,医学科研还能验证和发展已有的理论与学说,促进医学不断发展和防治水平的不断提高。

二、医学科研的分类

医学科研的分类,按分类方法不同,可分为以下几种。

(1) 按照医学中各专业的属性,医学科研可分为:基础医学科研、临床医学科研、预防医学科研、社会医学科研(包括医学心理学科研)。

(2) 按照研究对象的属性,医学科研可分为:实验性科研、临床试验性科研、社区干预试验性科研。

(3) 按照研究对象的属性和实验场所不同,医学科研可分为:调查性科研、实验性科研、临床试验性科研。

①调查性科研:又称横断面研究、横断面调查或现况研究。是根据研究目的,通过横断面调查,了解某一特定时间断面上特定人群中疾病或卫生服务的现状及其与之相联系的各种因素(如某病的患病率、人群中各种生理、生化指标或病理指标、卫生资源状况等)的分布情况。主要特点是研究过程中没有人为地施加干预措施,而是客观地观察记录某些现象的现状及其相关特征。

②社区干预试验研究:社区干预试验研究是在某一特定人群中通过干扰某些致病因素或施加某些保护性措施,观察对人群产生的效果。例如观察疫苗对预防某传染病发病的效果,观察某些食品对儿童身体发育的作用等。由于社区干预试验难以将研究对象进行随机分组,因此又称为半试验性研究(quasi-experiment)。

③实验性研究:实验性研究是一种在人为严格控制条件下观察和分析客观事物的研究方法。其特点是对研究对象进行干预技术,较好地排除了外界因素的干

扰,因而可以获得较为可靠的科学数据。此类研究场所大多是在实验室进行,研究对象一般是动物,但有一部分研究须走向社区人群。

④临床试验研究:临床试验是一种局限于对研究对象的身心健康无伤害的试验研究方法。观察分析期限可分为短期、中期或远期跟踪。临床试验研究的目的是观察防治措施的效果。由于临床科研的实验条件不易严格控制,影响因素相对较多,因此通常将临床实验研究称为临床试验(clinical trial)。英文中 clinical 常与 laboratory(即实验室工作)相对而言,对 clinical trial 应理解为已经通过有关的实验室研究,即将走向人群(现场)进行流行病学实验。其研究对象是患者(在研究治疗效果时),也可以是“非病人”(在研究预防措施时)。在某些医学科研中,根据选择的课题和研究目的,需将实验室研究与临床试验研究结合起来进行。

三、医学科研的程序

开展一项医学科研工作大体要经过五个基本步骤,各个步骤丝丝入扣、紧密相连。这五个步骤是科研选题、研究设计、实施方法、统计分析和总结归纳。

1. 科研选题

医学科研的选题或立题就是确定所要研究的题目。选题是科研的起点,也是关系到科研成败的关键性问题。

(1) 准备工作:在科研设计前必须查阅大量的参考文献,认真收集、阅读与本课题有关的研究资料,前人和同行的经验与知识都是可借鉴的宝贵财富。科学文献也是人类智慧的结晶,是人们增长知识提高业务水平的重要工具,也是制定科研规划、决定科研题目的重要依据。只有通过全面了解国内外发展趋势,使你提出的问题得到进一步的深化、系统、完善,使其具有新颖性、科学性和可行性,也能有效地防止低水平、重复性选题的弊端。

科学文献的来源:教科书、专著、期刊、会议文件和内部资料。还可以通过光盘检索系统和因特网查阅中外文参考文献。当掌握了有关信息以后,选题将具有初步轮廓;以往解决了的问题,还有哪些问题未解决。对所拟题目的科学价值、社会效益、经济效益、技术性问题、解决的技术路径等都应做到心中有数。

(2) 选题原则:选择确定一个优秀的课题,必须遵循选题原则,即应严格遵循创新性、先进性、科学性和可行性四项基本原则。其中创新性是科研工作最主要的特征。创新是指别人没有研究和涉及的问题,是本学科的空白点,或者是别人虽对此研究过,但你的研究将会在理论或应用上有新的发展和补充。或者是别人过去对这类项目有一些零星报道,而你选择了具特色的系统性研究。在深思熟虑科学性的同时兼顾其他“三性”,所以选题时必须根据具体情况全面分析和考虑,使四项基本原则达到有机统一。

(3) 研究条件和优势:目前科研经费都比较紧张,都希望将有限的经费投入到极有可能获得成功的申请项目上,所以预试验结果及以往的研究基础对中标均有不同程度的影响。选择有研究基础形成优势的领域进行深入探讨,不失为明智之

举。

2. 科研设计(research design)

科研设计是对科学研究具体内容与方法的设想和计划安排。设计的好坏不仅直接关系到科研的创新性、科学性、先进性和可行性,而且还决定了完成课题的速度与经费的开支等问题。要搞好科研设计,设计者不仅应具备丰富的专业知识,还需具备必要的相关知识,如医学统计学、临床流行病学和流行病学知识等。

科研设计由专业设计和统计设计两部分构成。专业设计的第一个问题是课题的选定,即研究什么问题。科研选题非常重要,科研成果的大小取决于选题水平高低;高水平的选题就能获得高水平的科研成果。统计设计就是控制误差、改善实验有效性和确定资料分析的方法,也是保证专业设计的布局合理性和实验结论可信性的关键因素。为了确保医学科研的“四性”,必须做好这两类设计。如果专业设计不科学,统计设计不正确,必然导致错误的结论,结果造成人力、物力和财力的浪费。统计学设计的内容有研究对象数量的确定,对照组的选定、随机化分组原则与统计分析方法的选定。控制误差也是重要内容之一。

总之,只要在科研设计前认真检索、阅读有关文献,全面了解研究专业的过去与现在,确立正确的研究方向,并能选定一个好的题目;在专业设计上做到周密、严谨,在统计设计上做到随机、客观、科学,就能够设计出一个优秀的设计方案,并得到满意、可信的结论。

3. 实施方法

按照研究对象的属性和实验场所不同,医学科研实施方法一般可分为调查、实验和观察三大类型。

(1) 调查:调查是在现场进行研究设计的一个重要部分,调查设计区别于实验设计的重要特征是:在调查中,所研究的现象及其相关特征(包括研究因素和非研究因素)是客观存在的,不施加干预措施,不能采用随机分配的方法来平衡或消除非研究因素对研究结果的影响。故调查资料的分析常需借助标准化法、分层分析以及多因素统计分析等方法对混杂因素加以调整。调查设计的基本步骤是:根据研究目的,确定调查对象(目标人群)和对象的选择方法,将研究指标转化为调查项目,进而设计成调查表或问卷,然后确定资料的收集方法、整理的计划和统计分析等。

(2) 实验:实验是在人为地控制一些条件与因素的基础上,对研究对象施加考核因素(防治措施),观察或观测由此引起的结构、功能、生化或疾病过程的变化,这种变化以效应指标表示,通过相应的效应指标揭示事物发生发展的规律性。

实验设计的特点是:①研究者能人为设置处理因素。②研究对象接受何种处理因素或水平是经随机分配而定的。能使对照组与各处理组间具有较好的均衡性,即非处理因素对不同处理组的影响保持均衡,组间具有可比性,可以客观地评价处理因素的作用。③实验设计能使多种处理因素和水平同时安排在较少实验的次数之中,更有效地控制误差,达到高效、精确的目的。

(3) 临床观察:主要是指临床试验,是一种经过精心设计的医学科研方法,常用来评价药物、疗法、诊断技术和预防性干预措施的效果。临床试验的基本原理是:①根据研究目的和客观条件,选择适当的研究对象,然后遵循随机化分组原则将研究对象分为两组或多组。一组不接受处理因素,称为对照组;其他组接受处理因素,称为实验组。②处理因素包括要评价的药物、诊治方法或预防、预后措施等。在有些研究中,研究对象本身的某些特征,如性别、年龄、疾病的类型与病情等,也可成为试验的处理因素。③对不同组的研究对象采用相同的方法,随访观察相同的时间,测量、记录其效应指标,然后比较或鉴别这些指标的差别,以此评价处理因素的效果。

一项成功的临床试验必须具备如下两个基本条件:①对实验组的干预效果及不良反应做出无偏估计。②确定试验结果能够外延推论到总体病人或总体人群,因参与试验的研究对象仅为随机样本。

上述三种设计方法并非彼此完全独立,某些医学科研需要三者或二者互相紧密结合。如开发新药必须经过如下阶段研究:低等动物→高等动物→人体(包括急性、慢性实验、毒理学测试等),如果是预防性药物还需进行人群的预防性干预实验。此例说明设计方法经历了实验设计、临床试验设计和调查设计过程。尽管现场调查研究在流行病学与病因学研究中具有重要的地位,但是调查所得的线索,往往需要实验室与临床试验进一步结合研究,才可能得到比较正确的结论。

4. 统计分析

在生物医学研究领域,实验或观察的结果往往是不确定的偶然现象。描述或推论这种不确定偶然现象,通过重复观察发现不确定偶然现象背后隐藏着事件发生或发展的必然规律性,这就是统计方法应用的显著特征。

医学统计学是将统计学原理和方法应用于生物医学研究的一门学科。医学统计方法在医学研究中的运用主要有三个方面:①以正确的方式收集资料,这种“正确方式”须在科研设计时确定下来。②描述资料的统计特征,如数据归类简化、统计指标的选择与计算、统计结果的表达等。此处的结果仅为局部(样本)的结果。③统计判断并得出正确结论,如根据各种概率分布,对实验和观察结果存在的差异和关联作出统计推断,即由样本推断到总体的结论。

综上所述,正确掌握和运用统计方法是科研中最重要的基本功之一,从科研课题的选题、设计、实施、分析到总结成文的全过程中,统计方法已渗透到各个环节,尤其是设计阶段。有的在设计阶段未考虑统计方法,得到数据以后再咨询统计学专家,这样的例子屡见不鲜,试想这些数据的真实性和可靠性如何?因此没有坚实的专业基础和统计学与流行病学知识,是不可能圆满完成科研课题的。

5. 总结归纳

总结归纳是科研过程中的最后一个步骤,就是根据观察事实与统计分析结果,运用分析、综合、归纳与演绎方法,把感性认识上升为理性概念。在总结概括时应注

意两点:①要根据已有的研究数据来推理。在推理中既要不要违背公理,又要不拘束于传统观念,应当在继承的基础上发展,推陈出新。②按照自己本次研究的范围做出结论,即推断在科研设计时所确定的总体特征,轻易外延推断往往是错误的结论。

总结归纳的基本形式是撰写科研论文。论文写作完成之后应全面衡量论文的学术水平,再将论文投向相应的杂志社发表。在论文发表以后,应注意收集论文被他人引用的情况与评价。

医学科研与其他学科研究的过程相类似,都包括上述五个环节,且环环紧扣,结合严密。其中选题和设计是最重要的两个步骤,能否把握住这两个环节,是科研成败的关键。

虽然医药科学分科较多,专业性强,各学科科研都有自己的专业特点,但是无论何种学科的科研,其研究过程、原理、原则和基本方法均具有一定的共同性。

四、控制科研中的误差

在科研设计、实验过程及研究结果分析等各个阶段中都可能产生误差。如果不加控制,即使出现明显的效果并作了统计处理,也将违背科研的科学性、真实性和可靠性的宗旨。不注意控制误差,研究结果的重复性也往往很差,甚至会得出相反的结论。因此,研究者必须了解和掌握实验中的各种误差和对误差控制的方法,确保研究结果的真实性和可靠性。

误差(error)是指实测值与其真值之差,统计量与参数之差。

1. 误差种类

根据误差产生的来源和性质,可将误差分为系统误差(systematic error)和随机误差(random error)。

(1) 系统误差:在实际观察过程中,由于仪器未经校准、测量者感官的某种偏差、医生掌握的疗效标准偏高或偏低等原因,使观察值不是分散在真值两侧,而是有方向性、系统性或周期性地偏离真值,即测得值普遍偏高或偏低。这类误差可通过科研设计和技术措施来消除或使之减少。观察性研究由于组间无可比性产生的系统误差称为偏倚(bias),如比较两组冠心病患者的死亡率时,其中年龄、伴有糖尿病、伴有脑血管疾病等因素造成两组死亡率的偏倚,失去比较的意义。

(2) 随机误差:又称偶然误差,是指排除了系统误差后尚存的误差。它受多种因素的影响,能使观察值无方向性、无系统性而随机地变化。此类误差变量一般服从正态分布。

随机误差也包括重复测量误差。因为机体是受精神、环境、遗传等多方面因素制约,个体间的变异是客观存在的,因此从同质总体中随机抽样,统计量(样本指标)总是与其参数(总体指标)有一点差异,这种差异用抽样误差指标来描述(标准误)。因此,抽样研究出现抽样误差是必然的。

2. 误差的来源

(1) 生物变异:生物个体之间的差异称为变异,为不明因素所致的随机误差且客观存在。

(2) 判断误差:由于感觉器官(视、听、嗅、触等)判断某项指标时所产生的误差。其中,既有操作者技术熟练程度引起的,也有被试者的主观因素引起的。这也是为何提倡多用客观指标的道理。

(3) 条件误差:由于环境条件不同所产生的误差。此处环境条件系指研究对象所处的“小环境状况”,包括仪器、操作、用具、室温、照明、噪声、季节及动物的饲养条件等。若实验组处于有利条件时,容易出现假阳性结果;若实验组处于不利条件时,容易出现假阴性结果。这种误差既包括实验组与对照组之间条件不同所致的误差,也包括组内各次条件变化所造成的误差。因此,科研设计时应注意这些非处理因素在组间保持均衡,在组内保持稳定。

(4) 分配不均误差:这类误差多发生在调查研究中,由于疾病的分布不均匀,有时呈现一种自然顺序排列,在调查设计时未考虑这些因素的影响往往会产生分配误差或非均匀误差或顺序误差。一般采用分层随机抽样方法消除这类误差。

(5) 过失误差:由于科研设计错误,或实验者的主观片面、粗心大意引起的误差。如不遵守随机化分组原则、主观选择研究对象、记录错误、技术路线不清、计算错误等。

(6) 估计误差:估计误差是研究者对观察结果进行分析或评价时发生的误差。这种误差是因为研究者心理上的偏性和技术上的原因所导致。

3. 误差的控制

误差可产生于科研过程的各个环节,因此应切实注意控制误差,力求使科研得到接近真实与可靠的结果。

(1) 设计阶段:控制误差,应在科研设计的起始阶段就严密防止。事先估计各个环节可能产生的误差,研究者应了如指掌。尽可能控制或减少过失误差和系统误差。其中预试验或预调查可找到产生误差的根源,并对误差大小进行评估,再修订设计方案,对研究总体规定得更明确更具体,或改用能更有效控制实验误差的试验方法,如交叉设计、正交设计或多元分析设计方法。

(2) 分组阶段:遵循随机化原则是控制实验误差的主要方法。尽管随机误差的出现是不可避免的,然而它的大小是预先可以得到控制和估计的。在随机抽样或随机分组的前提下,适当加大观察例数可以缩小误差。小样本研究时,提高样本中个体之间的同质性,使个体之间的变异变小。组间均衡性,是指实验组之间或实验组与对照组间的非处理因素保持一致,目的在于充分突出处理因素的效应结果。

(3) 抽样或实验阶段:重复试验、平行试验或抽样研究,适当加大样本含量是增加样本可靠性的又一有效手段。

概括起来,采用随机抽样或随机分组方法,定期、经常校准仪器,处理因素标准

化,实验过程统一编号、统一的记录表格,完整的实验记录,完善操作人员的技术水平考核、提高专业素质都是减少实验误差的重要措施。

结 语

(1) 医学科研设计是确定一个创意新颖、设计周密、指标合理、科学性强又切实可行的实施方案和工作计划,也是科研工作中第一个关键性的步骤。

(2) 医学科研工作的基本任务是熟悉人体结构、功能与理化机制,认识疾病的分布,掌握疾病的发生、发展过程,揭示疾病与健康的转化规律及其影响因素,探索环境与健康的联系,提出有效的防、诊、治措施,改善人类生存的环境,增进人类的健康,延长人民的寿命,开发促进人民身心健康的科学技术等。

(3) 医学科研的分类大致可分为调查性研究、社区干预试验性研究、实验性研究和临床试验性研究。

(4) 医学科研工作的基本程序是科研选题、科研设计、实施方法、统计分析和总结归纳。其中科研选题和科研设计是关系到科研成败的关键性步骤。

(5) 科研选题原则是选择具有创新性、先进性、科学性和可行性的课题。其中创新性就是敢于标新立异,勇于坚持和发展真理,不拘泥于旧概念和旧观点,善于根据自己的研究结果,正确运用唯物辩证法,通过科学的理论思维,在学术上有所创新。重点考虑创新性的同时,兼顾其他“三性”。

(6) 科研设计是科研的具体内容与方法的设想和计划安排。科研设计由专业设计和统计学设计两部分构成。专业设计是针对专业性课题进行的研究设计,其内容有选题、形成假说、干预措施、研究对象、实验方法和技术路线等。统计设计是针对统计数据收集进行的设计。其内容有样本来源、样本含量、随机化分组、统计设计类型、选择指标和误差的控制等。统计设计是统计分析的基础。任何设计上的缺陷,都不能寄希望于统计分析阶段弥补和纠正。

(7) 按照研究对象的属性和实验场所不同,医学科研设计一般可分为调查设计、实验设计和临床试验设计三种类型。调查设计的特点是所研究的现象及其相关特征是客观存在的,不施加干预措施,也不能采用随机分配的方法来平衡或消除非处理因素对研究结果的影响,只能借助统计方法加以调整。实验设计的特点是可以人为设置处理因素,并能随机化分组,保持了组间的均衡性和可比性,有效地控制实验误差,因此能达到高效精确的目的。临床试验设计的特点是可将研究对象随机化分组,同时施加防治措施,设置对照组的意义更大,但不容许设置空白对照组,须注意医学伦理学问题。混杂因素较多,需采取方法消除或缩小,尽量保持效应结果的真实性和可靠性。

(8) 统计分析的任务是统计描述和统计推断。统计描述是描述样本信息的某些特征,统计推断是根据统计量来推断参数的特征。统计方法贯穿于医学科研的各个环节,尤其是科研设计阶段,因此在设计时须考虑采用何种统计方法去验证假说,怎样收集资料以及资料的类型等。

(9) 误差在科研过程中都可能存在,只是大小不同而已,但研究者须尽可能设

法消除系统误差,控制和缩小随机误差,让处理因素产生的效应突显出来。

思考题

- (1) 结合自己的专业能否解释预防性干预试验和治疗性干预试验?
- (2) 科研选题的原则是什么? 其重要性是什么?
- (3) 科研设计中的专业设计和统计设计的基本内容是什么?
- (4) 统计学方法在科研设计中的地位如何?
- (5) 医学科研的基本程序中的关键步骤是什么?

(熊国强)

第二章 医学科研的基本知识

学习目标

- (1) 了解科研假说的概念、作用和建立假说的方法。
- (2) 掌握医学科研的三大基本要素。
- (3) 掌握实验设计的四大基本原则。
- (4) 熟悉数据处理的统计方法的选择原则。

第一节 医学科研假说

在医学科研中,人们通过实验和观察积累了一定的实验资料之后,依据已有的理论知识对研究的问题的某些现象和规律作出假定性的说明和推断。这种根据已知的科学事实和科学理论,对研究的问题提出的假定性说明和推测就是假说(hypothesis)。

一、假说的特征

1. 科学性(characteristic of science)

假说的科学性表现在,它来源于已有的事实材料,是以一定的科学理论为依据,绝不是随意的幻想和毫无根据的猜测。例如哈维(Harvey)提出的血液循环假说,是根据他多年的实验结果,即发现在半小时内通过心脏的血液,已经达到人体血液的全部。他认为人体无论如何不可能在半小时内产生这么多血液,血液是沿着一条封闭的管道不断地在全身循环。因此,他提出了血液循环假说,并预言了联结动脉和静脉的毛细血管的存在。

在这里也可看出,假说的两个条件——事实和科学理论中,事实更为重要。科学理论虽然也是假说的科学依据,但科学理论不是真理,它只是相对完成的认识,需要随着新事实的发现而完善其内容,理论要服从事实,假说必须能解释事实。

2. 假定性(characteristic of suppose)

假说虽然有一定的科学依据,但毕竟是对未知的研究问题及其规律的猜测和

推断;也就是说,它是在观察和实验材料不足的情况下,凭思维活动做出的,尚有待于实践的检验,将来既有可能被确证而发展为理论,也有可能被证伪而淘汰。假说的假定性,意味着科学探索活动需要发挥人的主观能动性,尤其是当旧的科学理论与新的科学事实之间发生矛盾时,作为科研工作者既要实事求是,更要大胆假设,只有这样,才能开扩思路,找出解决矛盾的突破口。

二、假说的作用

1. 假说是科研创新的起点

科学研究的目的是创新,就是形成新思想、新理论,实际上它们都是科研假说被验证后的具体体现形式。假说本身就是人类创造性的高度表现,它显示的是一种伟大的洞察力。因此,提出假说的能力,往往被认为是科学创造性的重要标志;也就是说,假说的提出是创新性工作的开始。

2. 假说为科研活动提供了进一步的研究方向

科研是一个探索未知事物和规律的过程,如果没有假说,便无法着手解决这一问题。因此,在进一步的研究中,我们必须依靠假说的引导,避免盲目的、无方向的科研活动。也就是说,假说能够为科研设计提供清晰的思路,进行有目的、有计划的研究。

科学领域中的许多重大发现都是在不同假说的引导下获得的。例如,预防白喉免疫法的发明,就很好地说明了这一问题。德国药物学家贝林(Behring)在白喉的早期研究中发现,实验动物注射白喉杆菌死亡时,细菌仍留在注射点周围。那么动物是什么原因引起死亡的呢?他提出细菌毒素是其原因的假说。根据这一假说,他做了大量的实验。最后,他不但查明了白喉杆菌的毒素是实验动物死亡的原因,而且发现细菌培养液中的毒素与细菌的培养时间有直接关系,培养时间不够,产生毒素就不足,增加培养时间,则能制成毒性很大的滤液。这一发现导致了预防白喉免疫法的发明,并使抗血清用于治疗。可见,如果没有细菌毒素致使动物死亡的假说,就不会有关于细菌毒素的进一步研究,也就不可能发明预防白喉的免疫方法。

3. 不同假说的争论有利于科学的发展

假说不仅能够指导人们去进行各种科学实验和观察,而且能够引起不同学说的争论,有利于学术的繁荣和科学的发展。对于同一问题的解决,需要多途径多方法的探索,每个途径甚至方法都可能提出一个假说。假说的相互争论,有利于揭露各种假说中存在的问题,取长补短,相互补充,乃至发展出新的假说,促使科研向更广、更深的方向发展,从而推动科学技术的进步。

三、假说的建立

建立假说,在医学上常见的方法有以下几种。

1. 差异法(method of difference)

差异法是根据观察到的事物的差异提出假说。例如,宫颈癌是妇科常见癌症,但是在尼姑和修女以及终生未婚妇女中极少发生宫颈癌,她们与普通妇女的差异在于性生活。所以怀疑性生活中的某因素可能与宫颈癌的发病有关。随后的研究表明:宫颈癌可能与性交引起的第Ⅱ型疱疹病毒感染有关。

2. 共变法(method of concomitant variation)

共变法是根据某种现象往往总与另一现象伴随发生,提出两者间可能有因果关系的假说。例如,根据大量流行病学调查,肺癌患者往往大量吸烟,因而提出吸烟可能是肺癌的危险因素的假说。

3. 类推法(method of analogize)

类推法是根据已知事实或事物规律推论未知事物的方法。例如,阿托品具有扩张血管的作用,可以改善微循环,并已知阿托品(atropine)属于M胆碱能受体阻滞剂。由中药藏茄分离出来的山莨菪碱也属于M受体阻滞剂,因此推断它也具有改善微循环的作用。这种作用已被大量实验与临床试验所证实。

4. 类同法(method of agreement)

类同法是根据事物发生的一致性提出假说。例如,大量的流行病学调查发现冠心病多有高血压,因而提出冠心病与高血压具有共同的危险因素(高血脂、吸烟、肥胖)的假说。

5. 剩余法(method of residues)

剩余法是在逐一排除可能影响某结果的各因素后,剩下的不能排除的因素就是可能的原因,从而提出假说。例如,弗莱明(Fleming)发现青霉素的过程,运用的就是剩余法。弗莱明长期从事溶菌酶的研究,曾发表多篇有关溶菌酶的文章,葡萄球菌不易发生溶菌现象。但有一次,他的培养皿里发生了葡萄球菌被溶解的现象;培养皿的营养成分和各种实验条件与前几次完全一样,不同因素就是培养皿中发现了青霉菌,因此,他提出青霉菌很有可能是这次葡萄球菌发生溶菌原因的假说。

四、假说的验证

假说提出后,医学上最常用的验证(test and verify)方法是通过实验或观察。即根据研究目的,经过周密的研究设计,收集有关信息与数据,对所获得的数据进行统计学分析,结合专业知识,对假说的真实性进行验证。

需要指出的是,科学技术在不断的发展,事物在不断的变化,人们的认识水平也在不断提高,今天验证的结论,可能明天就需要修改与补充,所以我们不能将验证的结论绝对化。

第二节 医学科研的基本要素

医学科研由三个基本部分组成,即被试因素(study factor)、受试对象(study object)和实验效应(experimental effect)。例如,观察巯甲丙脯酸对肾性高血压患者的疗效。其中,巯甲丙脯酸为被试因素,肾性高血压患者为受试对象,动脉压变化为试验效应。通常将这三个组成部分称为医学科研的基本要素。如何正确选择三个要素,是科研设计的关键问题。

一、被试因素

1. 概述

一般来说,被试因素是由外部施加于受试对象的,它可以是生物的、化学的、物理的。然而受试对象本身的某些特征(如性别、年龄、遗传特性、心理因素等)也可作为被试因素。如研究性别和年龄与心肌梗死患者死亡率的关系,性别和年龄就是被试因素。被试因素为单个时,称单因素。单因素在量上可有不同,这种量的不同就称为水平。例如,以药物作为被试因素,不同剂量就是不同水平。依照研究因素与水平的不同,可产生四类实验:①单因素单水平:这是科研中最常见的实验类型,如夏枯草提取物对原发性高血压患者降压作用的观察,就属于这类试验。②单因素多水平:这是单因素多组群的实验,如比较同一刺激不同强度的反应,比较某药物不同剂量对某病的疗效观察,便属于这类试验,特别是珍贵药物,毒性较大的药物或新药剂量的最佳选择,往往需要采用这类实验。③多因素单水平:比较不同药物、不同疗法、不同复方或者不同因素在某一疾病中的作用通常需要采用这类试验。④多因素多水平:例如研究酶学试验的最佳反应条件、探索联合用药方案、研究中药复方等。多因素多水平便是常用的实验类型。

2. 确定被试因素要注意以下几个问题

(1) 确定试验中的重要因素:事物之间是普遍联系的,任何试验效应往往是多种因素作用的结果。我们不可能在一次或几次研究中,获知一切有关因素的作用。例如,寻找某真菌的最佳培养方法,与其有关的因素很多,如温度、pH值、营养、气体、湿度等,其中每个因素又可分为若干个水平。若选定这5个因素,各取10个水平,将要做 $10^5 = 100000$ 次试验,显然是办不到的。因此,一个最优的试验设计,首先要抓住主要的几个被试因素。

(2) 确定被试因素与非被试因素:选择什么作为被试因素,取决于研究目的。例如研究刺激与平滑肌机械反应的关系,无疑刺激是被试因素;研究药物对某疾病的疗效,显然药物是被试因素。但是对于反应、疗效、病情和预后而言,除被试因素作用外,同时还有若干其他因素也会影响这些效应和结果。因此,在确定被试因素的同时,还必须明确哪些是非被试因素。对于非被试因素,应当作为误差来源严格

加以控制。例如,开展异博定(isoptin)对大鼠肺动脉内皮细胞黄嘌呤氧化酶活性的影响的研究,异博定是被试因素,但是细胞数量、细胞外钙浓度、氧分压、黄嘌呤含量、pH 值、温度等均可影响黄嘌呤氧化酶的活性。因此,试验组与对照组除被试因素不同以外,所有这些非被试因素都应当相同。

(3) 被试因素必须标准化:被试因素标准化就是指被试因素在整个试验的过程中始终保持一致,包括被试因素的强度、频率、持续时间及施加方法等,都应通过查阅文献和预备试验找出各自的最适条件,然后制订有关措施,以保证相对固定,一旦进入正式实验,不允许轻易改变。如被试因素是药物,应当在研究的全过程使用同一批药物,以防批间差异干扰试验效应。

二、受试对象

1. 概述

受试对象的选择在医学科研中十分重要。研究课题不同,对受试对象的要求也不一样,受试对象选择得当,能够为试验成功创造有利条件。遗传学家摩尔根(Morgan)用果蝇作为动物模型,发现了基因的连锁与互换规律,在遗传学上取得了举世瞩目的成就。正是因为果蝇结构简单(仅四对染色体),而又能反映遗传特性,这一点在实验动物学上称为“易化原则”。由此可以看出,正确选择受试对象同等重要。

2. 受试对象的确定

受试对象的确定取决于研究目的。医学科研的受试对象绝大多数是人或动物,但在药物研究中也可以是植物。根据具体情况可采用整体作为受试对象,在体内(in vivo)进行试验;也可采用器官、组织、细胞、亚细胞或分子作为受试对象,在体外(in vitro)进行试验;还可采用先体内后体外的方式进行试验,这类试验属半体内(ex vivo)试验。一般来说,整体试验反映人体或动物的实际情况,试验结果对临床医学参考意义较大。然而体内影响因素十分复杂,为了深入探讨作用机制,往往需要配合适当的体外试验或半体外试验。以人体作为试验对象,譬如研究参考值(正常值)范围,应当选择健康人或“相对健康者”(指未患有影响该指标的疾病和未患有重要脏器疾病的人)。假如研究病因学,往往需要将暴露于某危险因素与未暴露于某危险因素的人进行比较。倘若研究某病诊断方法,通常需要选择确诊本病的患者与未患本病的人,以便了解方法的敏感度与特异度。如果观察某药与某疗法对某种疾病或病证的疗效,需选择确诊为本病或本病证的患者。例如进行计划生育的研究,则应以身体健康具有较强生育能力的青壮年作为受试对象。倘若观察药物的毒性与不良反应及药代动力学过程,则可采用志愿者(volunteer);但是志愿者由于心理与经济等多种因素的影响,可能影响研究结果,甚至易于出现假阳性,因此采用志愿者应持慎重态度。

3. 受试对象的条件

在医学科研中,作为受试对象的前提是所选择的对象必须同时满足两个基本条件:①对被试因素敏感。②反应比较稳定。例如,临床上观察某种药物对高血压病的疗效,Ⅲ期高血压患者对药物不敏感,Ⅰ期患者本身血压波动范围较大,因此一般应选择Ⅱ期高血压病患者作为受试对象。其次,受试对象必须能与实验人员很好地配合,不仅易于取样,安全性好,而且应当以用于病人诊治为原则。此外,在临床疗效科研中,应当选择具有客观指标明显变化的患者作为受试对象。例如,冠心病不宜选择仅有心前区不适症状的患者,应选择有缺血性心电图改变的患者。

存在以下情况之一者,不宜作为一般临床科研的受试对象:①存在影响反应结果的并发症。②危重状态。③多种疗法无效(机体反应性和(或)致病因素与一般病例不同)。④不能配合者。但是,若专门研究合并症、危重病症或顽固性(难治性)病症,理所当然应以这类患者作为受试对象。

4. 受试对象的纯化(purification)

受试对象的疾病应是确诊的,且表现具有典型性。非典型的特殊病例不宜作为受试对象,因为特殊病例提示机体或致病因素与一般病例存在差异。受试对象的病情一般应是中等的,因为病情过重,死亡率很高,不易反映被试因素的疗效;若病情过轻,稍加处理即可痊愈,难以区分不同被试因素的优劣。因此,只有通过病情中等病例得到肯定结论后,才能扩大观察轻、中、重三类不同病情患者。

5. 受试对象的依从性

受试对象的依从性指他们按预定计划接受被试因素的合作程度。绝对的依从只有在麻醉动物实验才能实现。患者由于心理、社会、经济等多方面原因可能出现忘记服药、中途退出实验或出现换组。其次,由于病情急剧恶化或存在难以忍受的不良反应,必须中途退出实验。这些不依从性表现必然干扰实验计划的实施。因此,必须充分关心体贴病人,做好思想工作,使病人建立充分信任并服从医务人员的心理状态,从而提高依从性。同时要注意控制试验时间,试验时间过长,受试者的依从性往往降低。此外,在科研设计时应制定对不依从情况出现时的补救措施。

6. 受试对象影响因素的控制

不同性别与年龄的人,他们的激素、代谢水平和器官功能均有一定的差异,这些对许多疾病的疗效都有一定的影响。因此,实验组与对照组的性别和年龄要均衡。一般试验,首先选择中青年作为受试对象,只有肯定疗效后,才能扩大到儿童与老年。除非专门研究妊娠有关课题,一般不应选择孕妇作为受试对象。

生活习惯与嗜好可影响实验结果。首先,吸烟可使PGI₂合成减少,明显影响治疗动脉粥样硬化与慢性阻塞性肺疾病患者的效果。其次,居住条件与家庭经济状况等因素也会影响疗效,尤其是发病与精神因素有关的疾病更是如此。此外,饮食因素

与许多疾病的发生发展有关。如刺激性强的食物易使胃部疾病复发;食用大量黑木耳可能导致出血倾向。最新的研究表明,许多药物效应存在种族差异。因此,选择受试对象时必须控制影响受试对象试验效应的非被试因素。

三、试验效应

1. 概述

被试因素作用于受试对象所引起的试验效应或反应,往往是通过具体实验指标来反映的,因此效应指标的正确选定是非常重要的。

2. 指标的关联性

选用的指标必须与所研究的题目具有本质性联系,且能确切反映被试因素的效应。一般来说,功能性指标应与所反映的功能存在本质联系。比如以心电图作为心脏收缩力的指标,这显然是不正确的。若要了解心脏泵血功能,就应选择心输出量(单位:L/min)或心脏指数[单位:L/(min·m²)]作为指标,否则,不可能作出正确的判断。对于血糖测定,若用班氏试验,易受体内还原物质影响,读数往往偏高,应该用特异性高的葡萄糖氧化酶方法。必须明确,科学技术是不断发展的,科研工作者应当及时了解最新信息,以供自己的科研工作应用的指标更好地具有高度关联性。

3. 指标的客观化

指标分为主观指标和客观指标,这是由指标数据来源决定的。指标数据由观察者或受试对象根据主观感受程度判定的,称为主观指标。指标数据由仪表指示的,称为客观指标。如心电图、血细胞自动计数等就属于客观指标。主观指标易受心理状态及暗示程度的影响,并且感觉器官的感受往往由于背景条件与对比诱导可发生较大的差异,因此在科研中应尽量少用。客观指标较少受心理状态等的影响,应用价值较大。

4. 指标的灵敏度

指标的灵敏度通常是由该指标所能正确反映的最小数量级或水平来确定。在形态学方面,光学显微镜可判断组织和细胞水平的改变,电子显微镜则可判断亚细胞超微结构的改变。作为科研指标,要求其灵敏度能正确反映被试因素对受试对象所引起的反应,并不是灵敏度越高越好,如测定小鼠体重,以感量为0.5g的药物天平即可,若采用感量为0.1mg的分析天平,显然是不必要的。

5. 指标的特异性

在选用指标时,必须选用能准确反映被试因素的效应本质且特异性高的指标。特异性高的指标最能揭示事物的本质,且不易受其他因素的干扰;而非特异的指标

极易受其他因素的干扰,使效应结果不准确。例如胆碱酯酶活性测定对于有机磷毒理学的研究是特异指标。

6. 指标的精确性

所谓精确性包括指标的精密度与准确度双重含义。准确度是测定值与真实值接近的程度,也就是说,准确度是测定正确性的量度。精密度是各次测定值集中的程度,平时强调实验结果的可重复性,就是在相同条件下多次取样测定结果的精密度。如果一个指标有几个测定方法,在设计时应首选既准确而又精密的方法,准确度差则是不可取的。

7. 指标的观察

试验效应的观察应避免带有偏性(bias)。如研究者和资料的分析者的心理常偏于阳性结果,医生常偏于新疗法组,病人常对新疗法持怀疑态度等。这些都有可能致试验效应指标测定时带有偏性,从而影响结果的比较和分析。为了消除或最大限度减少这种测量偏差,在设计时,常采用盲法(blind method)。如使受试对象不知道设计者的方案,何组是试验组,何组是对照组,预期结果是什么,称为单盲法(single blindness);若受试对象及试验执行者均不知道设计方案,称为双盲法(double blindness);如果受试对象、试验执行者和资料分析者均不知道设计方案,称为三盲法(triple-blind trial)。

第三节 实验设计的基本原则

实验设计的意义在于用较少的人力、物力和时间,获得较为可靠的结果,最大限度地减少误差,以达到高效、快速和经济的目的。为此,实验设计必须遵循以下四个基本原则。

一、设立对照的原则

1. 对照的意义

对照,就是在实验中设立相互比较的实验组和对照组,对各组给予不同的处理,然后观察各组的结果。对照的基本要求是,除被试因素作有计划的变化外,实验组和对照组的非被试因素要相同或相近,从而抵消或减少由此引起的实验误差,以便分离出被试因素的效应。例如,临床上的许多疾病(像感冒、流行性腮腺炎、早期高血压等)不用药物治疗,经过一定时期,它本身可以自愈。再者,影响疾病的因素是复杂的,除治疗因素外,气候、营养、休息、精神状态等也对疾病产生影响。因此,要做到正确的鉴别,就必须设立对照组。

2. 对照的形式

对照的形式有多种,可根据研究的目的和内容加以选择。常用的有以下几种。

(1) 空白对照: 对照组不施加任何被试因素。例如,观察维生素 A 的防癌作用,实验组的石棉矿工每天口服一定剂量的维生素 A,对照组的石棉矿工不服维生素 A,被试因素完全是空白的。追踪观察相等的一段时间后,比较两组癌症的发病率。

空白对照简单易行,但容易引起实验组与对照组在心理上的差异,从而影响试验效应的测定。临床上不宜用空白对照,即使对无肯定疗效的疾病,医生也不应放弃具有一丝希望的疗法。

(2) 安慰剂对照: 对照组使用一种无药理作用的假药,药物的剂型、外观、包装或处理上不能为受试者识别,称为安慰剂(placebo)。使用安慰剂有助于避免对照组病人产生与实验组病人不同的心理作用,但在使用时一定要慎重,做好保密工作。例如,观察某种药物对轻度精神忧郁症的疗效,让实验组病人服用某种药物,对照组病人服用安慰剂,然后观察两组病人的疗效。一般说来,这种方法只适用于小规模的研究。

(3) 实验对照: 对照组不施加被试因素,但施加某些与被试因素有关的实验因素。例如,在动物实验中,如果研究切除甲状腺对白细胞生成的影响,对照组也必须进行麻醉、切开、暴露甲状腺、缝合等手术操作,只是不切除甲状腺。

(4) 标准对照: 用公认的标准疗法或常规疗法作对照。这种对照在临床试验中用得较多,因为在很多情况下不给病人任何治疗是不符合医德的。例如,观察人参皂苷的抗心肌缺血作用,对照组可用已知的有效药物心得安。

(5) 历史对照: 以过去的研究结果作为对照。例如,某种疾病的治愈率为 0,现用一种新药连续治愈 2 例。尽管没有设立对照,仍可认为是一种很好的药,因为过去大量从未治愈的该病病人就是对照。历史对照形式,除公认的难治性疾病如恶性肿瘤、狂犬病等以外,一般不宜使用。因为随着时间的推移,医学高速发展,历史资料往往不具有可比性。

(6) 自身对照: 对照与实验在同一受试对象进行。例如,以病人用药前后的血压值作对比,这种对照形式不需要另设对照组,这样可节省样本量,并避免个体差异所引起的误差。由于简单易行,使用广泛,在实验中往往将处理前作为对照;严格地说,它不是随机分配的,有一定的局限性。需要注意的是,对有短期自愈倾向(如感冒)或有周期性发作倾向的疾病,以及实验前后某些环境因素或自身因素发生了改变,并且会影响实验结果时,则不宜用自身对照。

(7) 相互对照: 不设立对照组,而是几种处理(或水平)互为对照。这种对照要做到同期平行进行。例如,将条件基本相同的贫血病人分成三组,分别应用三种方案进行治疗,比较各组的治理结果。

二、随机化的原则

1. 随机化的意义

在实验设计中,受试对象的分组以及施于受试对象的实验顺序等必须遵循随机化(randomization)的原则。随机化就是使每一个受试对象有同等机会被抽出并分配到各实验组中去,以抵消被试因素对试验效应的影响,同时保证了实验数据可进行统计学分析。因为随机化是数据统计的先决条件,倘若分组不是随机的,即使统计分析结果有差异,也是毫无意义的。

值得注意的是,随机不是随便或随意。在实际工作中常见的不随机举例如下:

(1) 动物分组:先抓到的小鼠放入 A 组,后抓到的小鼠放入 B 组。看似随机,其实进行了挑选。因为先抓到的多为不活泼的或雌性的,后抓到的多为活泼的或雄性的,而这些因素有可能影响实验结果。

(2) 病人分组:将单日门诊病人放入 A 组,双日门诊病人放入 B 组。这就存在研究者的主观意识,况且病人往往对门诊日是有选择的,因而是非随机的。

2. 随机化的方法

随机化的方法有多种,最简单的如抽签。这种方法简便易行,但不适用于样本含量大的多组分配。在实际设计中,广泛应用的随机化方法有随机数字表(random number table)和随机排列表(random permutation table)。随机数字表,常用于抽样研究及对病人、标本、实验动物等的分组随机化。表内数字互相独立,无论从横行、纵列或斜向等各种顺序均是随机状态。使用时可从任何一个数字开始,可单行、单列或双行、双列,亦可多行、多列,按任何一个顺序录用。随机排列表适用于分组、排列等的随机化,但不适用于随机抽样研究。随机排列表有 $n=10$ 、 $n=20$ 、 $n=30$ 等多种。表中 r_s 是该行与自然序列 1、2、3... n 的 Kellidall 等级相关系数。 r_s 的绝对值愈小,表示随机化愈好。随机排列表的用法比随机数字表更为简便,使用时可取任一行的随机排列数字,但不能按列查,它可以简便地将受试对象随机分配到实验所要求的各组中去,且各组不用调整。也可以利用比较高级的科学型计算器,如 Casio fx-180p 或 fx-360p 等,这些计算器都能产生随机数字。还可以利用各种统计软件如 SAS 等进行随机化分配。

例 2.1 试将 12 头动物随机分配到甲、乙两组。

(1) 抽签法:先将动物按体重(或性别、月龄等)顺序从小到大编号为 1、2、3...12,制作与动物编号相对应的数字标签,充分混合后,按预先规定,先抽取的 6 个数字标签为甲组,余下的数字标签为乙组。

(2) 随机数字表法:先将动物按体重编号为 1、2、3...12,再从附表 1 随机数字表中任意指定第 11 行第 11 列数字“2”起,向下查取 12 个随机数字,依次抄录于“动物编号”下。预先规定:遇单数,该动物分入甲组,遇双数分入乙组,得结果如下表:

| | | | | | | | | | | | | |
|------|---|---|---|---|---|---|---|---|---|----|----|----|
| 动物编号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 随机数字 | 2 | 2 | 0 | 2 | 5 | 3 | 4 | 6 | 9 | 9 | 0 | 5 |
| 所属组别 | 乙 | 乙 | 乙 | 乙 | 甲 | 甲 | 乙 | 乙 | 甲 | 甲 | 乙 | 甲 |

即将 5、6、9、10、12 五头动物分入甲组,1、2、3、4、7、8、11 七头动物分入乙组。如要求两组动物数相等,应从乙组七头动物中按随机化原则拨出一头动物归入甲组。可由上述随机数字表继续向下查取 7 以内的数字,先遇到哪个数字,即将该序次的动物拨归甲组。本例先遇到 1,因此,将乙组第 1 头动物(1 号)拨归甲组。

(3) 随机排列表法:先将动物按体重编号为 1、2、3 … 12,再从附表 2 随机排列表中任意指定第 8 行,舍去 12~ 19,将 0~ 11 的数字依次抄录于“动物编号”下。按预先规定,将随机数字为单数的分入甲组,双数的分入乙组,结果如下表:

| | | | | | | | | | | | | |
|------|---|---|---|---|---|---|---|----|---|----|----|----|
| 动物编号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 随机数字 | 3 | 2 | 6 | 1 | 8 | 0 | 9 | 11 | 5 | 4 | 10 | 7 |
| 所属组别 | 甲 | 乙 | 乙 | 甲 | 甲 | 甲 | 乙 | 乙 | 乙 | 甲 | 甲 | 乙 |

分入甲组的动物编号为 1、4、5、6、10、11,分入乙组的动物编号为 2、3、7、8、9、12。

(4) 应用计算器所产生随机数法分组:先将动物按体重编号为 1、2、3 … 12,凡按 INV RAN 键显示随机数 ≥ 0.500 者分入甲组, < 0.500 者分入乙组。若其中一组已达到 6 时则停按,余下的动物分在另一组,结果如下表:

| | | | | | | | | | | | | |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----|
| 动物编号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 随机数字 | 0.450 | 0.945 | 0.553 | 0.156 | 0.742 | 0.339 | 0.281 | 0.246 | 0.998 | 0.997 | 0.518 | |
| 所属组别 | 乙 | 甲 | 甲 | 乙 | 甲 | 乙 | 乙 | 乙 | 甲 | 甲 | 甲 | 乙 |

结果为 2、3、5、9、10、11 号动物分入甲组,1、4、6、7、8、12 号动物分入乙组。

(5) 应用 SAS 统计软件法:先将动物按体重编号为 1、2、3 … 12。

SAS程序

| 行号 | 程序 | 行号 | 程序 |
|----|----------------------|-----|-----------------------|
| 10 | DATA AAA; | 80 | FACTORS UNIT= 12; |
| 20 | DO UNIT= 1 TO 12; | 90 | OUTPUT OUT= BBB |
| 30 | IFUNIT \leq 6 THEN | | DATA= AAA; |
| | TRTMENT= 'A'; | 100 | PROC PRINT DATA= BBB; |
| 40 | ELSE TRTMENT = 'B'; | 110 | PROC SORT OUT= CCC |
| 50 | OUTPUT; | | DATA= BBB; |
| 60 | END; | 120 | BY UNIT; |
| 70 | PROC PLAN | 130 | PROC PRINT DATA= CCC; |
| | SEED= 26853; | 140 | RUN; |

SAS 程序解释:10~ 60 号语句生成数据集 AAA,内含 12 个实验对象,分 A、B

两组,每组 6 个。70~ 100 号语句请求 PLAN 过程,FACTORS 是对 12 个实验对象随机排列,将排列的前 6 个分入 A 组,后 6 个分入 B 组,分配结果输出到 BBB 文件中并打印结果。110~ 130 号语句是按实验对象号重新整理,输给 CCC 文件并打印结果。

三、重复的原则

1. 重复的意义

重复是指在相同实验条件下进行多次观察。即要求各处理组的实验单位都要有一定的数量,要考虑样本含量的大小。随机化能在很大程度上避免非被试因素所产生的影响。但是,如果观察数量太小,就有可能把个别情况误认为普遍现象,把偶然性或巧合的现象当作必然的规律性,以至将实验结果错误地推广到群体;如果观察数量太多又会增加严格控制实验条件的困难,造成不必要的浪费。在实验研究中对样本含量的要求是,既要保证实验结果可靠,又要避免不必要的浪费。这就要求受试对象具有较好的同质性,精选的小样本优于庞杂的大样本。设计样本含量时,各组例数最好相等。

2. 影响样本含量的因素

(1) 第一类错误的概率 α :即检验水准。 α 越小所需样本含量越多。

(2) 检验效能 $(1 - \beta)$:即两总体确有差别时,按 α 水准发现它们有差别的能力。 $(1 - \beta)$ 越大,所需样本含量越多。

(3) 最小差值 δ :即在 α 及 β 水准下期待检出的最小差异或客观上可能存在的差异大小。 δ 越小,所需样本含量越多。

(4) 总体标准 σ : σ 大,所需样本含量大。实际工作中,可凭经验或通过预实验取得的样本标准差 s 来点值估计 σ 。

(5) 资料的性质:一般来说,在同等情况下,数值变量所需样本含量少于分类变量。

(6) 实验设计的类型:从常用的实验设计来看,完全随机设计所需样本量多,配对设计与随机区组设计所需样本量较少,拉丁方、正交、析因等设计所需样本量更少。与常用实验设计相比,序贯实验设计所需样本量可少 30% ~ 50%。

3. 估计样本含量的方法

(1) 凭经验:例如在动物实验中,分为大动物(犬、猪)、中等动物(兔、豚鼠)和小动物(大鼠、小鼠),每组分别为 5~ 15 头、10~ 20 头、15~ 30 头。难治的疾病,疗效显著时(如癌症)5~ 10 例即可,急重病死亡率高的(如休克、心衰、肾衰、呼吸衰竭)需 30~ 50 例,一般病和慢性病 100~ 300~ 500 例;确定正常值范围每组多于 100 例;恶性肿瘤的流行病学调查至少 10 万以上人口;多因素分析中样本含量数至少是研究因素个数的 5~ 10 倍以上等。

(2) 算法:① 参数估计时的样本含量估计(参见第五章)。

② 假设检验时的样本含量估计。

- 样本均数与总体均数比较(或配对比较),按式 2.1 计算。

$$n = \left[\frac{(u_{\alpha} + u_{\beta})s}{\delta} \right]^2 \quad (2.1)$$

式中 n 为所需样本含量; s 为总体标准差 σ 的估计值; δ 为容许误差, u_{α} 和 u_{β} 由 t 界值表($v = \infty$)查得, u_{α} 有单侧值和双侧值之分, u_{β} 只取单侧值。

例 2.2 用某药治疗矽肺患者,估计可增加尿矽排出量,其标准差为 89.0 mmol/L,若要求 $\alpha = 0.05$, $\beta = 0.10$ 的概率,能辨别出尿矽排出量平均增加 35.6 mmol/L,试问需用多少矽肺病人做试验?

本例 $\delta = 35.6$, $s = 89.0$, 单侧 $\alpha = 0.05$, $u_{0.05} = 1.645$, $\beta = 0.1$, $u_{0.1} = 1.282$, 代入式 2.1 得:

$$n = \left[\frac{(1.645 + 1.282) \times 89}{35.6} \right]^2 = 53.5, \text{取 } 54$$

故可认为需治疗 54 个矽肺病人。即以 54 例进行试验,如该药能增加尿矽排出量,则有 90% (即 $1 - \beta = 0.9$) 的把握可得出有差别的结论。

查附表 3, 单侧 $\alpha = 0.05$, $\beta = 0.1$, $\delta/\sigma = 35.6/89.0 = 0.4$, 得 $n = 55$, 与上述计算结果相近。

- 两样本均数比较,按式 2.2 计算。

$$n_1 = n_2 = 2 \left[\frac{(u_{\alpha} + u_{\beta})s}{\delta} \right]^2 \quad (2.2)$$

式中 n_1 和 n_2 分别为两样本所需例数,一般假设其相等; s 为两总体标准差 σ 的估计值,一般假设其相等; δ 为两均数差值; u_{α} 和 u_{β} 的意义同前。

例 2.3 比较黄芪与生血散对粒细胞减少症的疗效。据以往经验,黄芪可增加粒细胞 1000 个/ μL ,生血散可增加 2000 个/ μL ,两组合并标准差估计为 1800 个/ μL ,拟取双侧 $\alpha = 0.05$, $\beta = 0.10$,试问每组需观察多少病例?

设 $\delta = 2000 - 1000 = 1000$, $s = 1800$, 双侧 $\alpha = 0.05$, $u_{0.05} = 1.96$, $\beta = 0.1$, $u_{0.1} = 1.282$, 代入式 2.2 得:

$$n_1 = n_2 = 2 \left[\frac{(1.96 + 1.28) \times 1800}{1000} \right]^2 = 68.10, \text{取 } 69.$$

查附表 4, 双侧 $\alpha = 0.05$, $\beta = 0.1$, $\delta/\sigma = 1000/1800 = 0.56$, 利用内插法求出:

$$n = 71 - \frac{0.56 - 0.55}{0.60 - 0.55} \times (71 - 60) = 69$$

- 两样本率比较,按式 2.3 计算。

$$n = \frac{(u_{\alpha} + u_{\beta})^2 2P(1 - P)}{(P_1 - P_2)^2} \quad (2.3)$$

式中 n 为两样本分别所需例数, P_1 和 P_2 分别为两总体率的估计值, P 为两样本合并率, $P = (P_1 + P_2)/2$, u_{α} 和 u_{β} 为 $v = \infty$ 时,由 t 界值表中查得的 t 值。

例 2.4 比较两种中药使肝炎表面抗原(HBsAg)阴转的疗效,预试验甲药阴转

率为 30%, 乙药为 50%, 正式实验拟取 $\alpha = 0.05$, $\beta = 0.10$, 试问各组需要多少病例?

$$\text{设 } P_1 = 0.30, P_2 = 0.50, \text{ 则 } P = \frac{0.30 + 0.50}{2} = 0.40$$

双侧 $u_{0.05} = 1.96$, 单侧 $\beta = 0.10$, $u_{0.10} = 1.28$ 。代入式 2.3

$$n = \frac{(1.96 + 1.28)^2 \times 2 \times 0.4(1 - 0.4)}{(0.30 - 0.50)^2} = 125.97, \text{ 取 } 126$$

查附表 6, $\alpha = 0.05$, $\beta = 0.1$, $1 - \beta = 0.90$, 较小率 $P_1 = 30\%$, $\delta = P_2 - P_1 = 50\% - 30\% = 20\%$, 得 $n = 125$ 。

● 多个样本均数比较, 按式 2.4 计算。

$$n = \psi^2 \left[\sum_{i=1}^k \frac{s_i^2}{k} \right] / \left[\sum_{i=1}^k (\bar{X}_i - \bar{X})^2 / (k - 1) \right] \quad (2.4)$$

式中 n 为各样本所需例数, \bar{X}_i 和 s_i 分别为第 i 个样本的均数和标准差的初估值, $\bar{X} = \frac{\sum_{i=1}^k \bar{X}_i}{k}$, k 为组数。 ψ 值由附表 7 查得。若 α , β , $v_1 = k - 1$, $v_2 = \infty$, 查得 ψ 值, 代入式 2.4 中求得 $n_{(1)}$, 第二次由 $v_1 = k - 1$, $v_2 = k(n_{(1)} - 1)$ 查 ψ 值, 代入式中求 $n_{(2)}$, 仿此进行, 直至前后两次求得结果趋于稳定为止, 即所求样本例数。以上为完全随机设计时 ψ 值的求法, 当随机区组设计时, 第一次 $v_2 = (k - 1)(n - 1)$, 式 2.4 中 $\sum_{i=1}^k \frac{s_i^2}{k}$ 用误差均方代替。

例 2.5 拟用 A、B、C 三种方案治疗贫血患者, 估计治疗后血红蛋白含量(g/L)增加的均数分别为 17.2、11.8、15.6, 标准差分别为 8.0、7.1、7.9, 设 $\alpha = 0.05$, $\beta = 0.1$, 若要得出有差别的结论, 每组需观察多少例?

本例 $\bar{X}_A = 17.2$, $\bar{X}_B = 11.8$, $\bar{X}_C = 15.6$; $s_A = 8.0$, $s_B = 7.1$, $s_C = 7.9$; $\alpha = 0.05$, $\beta = 0.1$ 。

$$\bar{X} = (17.2 + 11.8 + 15.6) / 3 = 14.87$$

$$\sum_{i=1}^k (\bar{X}_i - \bar{X})^2 = (17.2 - 14.87)^2 + (11.8 - 14.87)^2 + (15.6 - 14.87)^2 = 15.39$$

$$\sum_{i=1}^k s_i^2 = 8.0^2 + 7.1^2 + 7.9^2 = 176.82$$

以 $\alpha = 0.05$, $\beta = 0.1$, $v_1 = k - 1 = 3 - 1 = 2$, $v_2 = \infty$, 查附表 7, 得 $\psi_{0.05, 0.1, 2, \infty} = 2.52$, 代入式 2.4,

$$n_{(1)} = 2.52^2 (176.82 / 3) / [15.39 / (3 - 1)] = 48.6 \approx 49$$

若 $\alpha = 0.05$, $\beta = 0.1$, $v_1 = k - 1 = 3 - 1 = 2$, $v_2 = k(n_{(1)} - 1) = 3(49 - 1) = 144$, 查附表 7, 表中 v_2 无 144 值, 故取近似值 $v_2 = 120$, 得 $\psi_{0.05, 0.1, 2, 120} = 2.55$, 代入式 2.4

$$n_{(2)} = 2.55^2 (176.82 / 3) / [15.39 / (3 - 1)] = 49.8 \approx 50$$

两次结果相近, 故可认为每组需要观察 50 例。

● 配对分类资料, 按式 2.5 计算。

$$n = \left[u_\alpha \sqrt{\frac{1}{2p}} + u_\beta \sqrt{\frac{2(p_1 - p)(p_2 - p)}{p}} \right]^2 / (p_1 - p_2)^2 \quad (2.5)$$

式中 n 为所需观察的对子数, p_1 和 p_2 为两总体阳性率的估计值, p 为两种处理结果一致的阳性率, $p = (p_1 + p_2 - 2p) / 2$, u_α 和 u_β 意义同前。

例 2.6 拟比较甲、乙两种血清学方法检查鼻咽癌患者的情况。初估计甲种血清学检查的阳性率为 90%，乙种血清学检查的阳性率为 71%，两种血清学方法检查均为阳性率者 64%，试问至少需取多少份鼻咽癌患者的血清做检查？

本例 $p_1 = 0.90$, $p_2 = 0.71$, $p = 0.64$, $p = (0.90 + 0.71 - 2 \times 0.64) / 2 = 0.165$ ；
 设 $\alpha = 0.05$, $u_{0.05} = 1.960$, $\beta = 0.1$, $u_{0.1} = 1.282$ ；代入式 2.5

$$n = \left[1.96 \sqrt{2(0.165)} + 1.282 \sqrt{2(0.90 - 0.64)(0.71 - 0.64) / 0.165} \right]^2 / (0.90 - 0.64)^2 = 44.2 \approx 45$$

故认为共需 45 份鼻咽癌患者的血清。

● 多个样本率比较,按式 2.6 计算。

$$n = 2N \left[2 \sin^{-1} \sqrt{P_{\max}} - 2 \sin^{-1} \sqrt{P_{\min}} \right]^2 \quad (2.6)$$

式中 n 为每个样本所需观察例数, P_{\max} 和 P_{\min} 分别为最大率和最小率,当仅知最大率和最小率差值 P_d 时,则取 $P_{\max} = 0.5 + P_d / 2$, $P_{\min} = 0.5 - P_d / 2$ 。 λ 是以 α, β , 自由度 $\nu = k - 1$, 查附表 8 λ 值表而得, k 为组数。

例 2.7 某省观察三个地区的花生被黄曲霉毒素 B₁ 污染的情况,初估甲地区污染率为 79.3%,乙地区的污染率为 31.3%,丙地区的污染率为 27.6%,试问各地区至少需观察多少个样品？

本例 $P_{\max} = 0.7930$, $P_{\min} = 0.2760$, 以 $\alpha = 0.05$, $\beta = 0.1$, $\nu = k - 1 = 3 - 1 = 2$, 查附表 8, $\lambda_{0.05, 0.1, 2} = 12.65$, 代入式 2.6, 得

$$n = 2(12.65) / 2 \left[\sin^{-1} \sqrt{0.7930} - 2 \sin^{-1} \sqrt{0.2760} \right]^2 = 30.1 \approx 31$$

即每组需 31 个样品,三组共 $31 \times 3 = 93$ 个样品。

(3) 查表法:统计学家在计算法的基础上,编制出各种情况下所需样本含量的表格,然后根据研究目的选择不同类型的表格,再根据 α, β, δ 等参数值查阅出所需样本含量。具体的查表方法在计算法中已经介绍。

四、均衡的原则

1. 均衡的意义

实验组和对照组必须遵守均衡的分组原则,即在实验中,受试对象除了观察的某种被试因素不同外,其他一切条件应尽可能与对照组相同或相近,以消除非被试因素对试验效应的影响。

2. 均衡的方法

为达到均衡的目的,首先要了解抽样总体的大致情况,找出影响实验结果的非被试因素,将非被试因素中的重要影响因素分层,再在分层的基础上,层内随机抽样进行实验对象分配。至于次要影响因素,按随机化原则处理。例如,比较两种药物对高血压的疗效,因为高血压有轻度、中度和重度之分,如果随意将一些高血压病人分组治疗,即使完全随机也不合适,会导致某一组重型者多,另一组轻型者多。