

# 医学实验设计分析与 SAS实现

主编 冯国双 刘德平



01 01 01 01 01 01

北京大学医学出版社

# 医学实验设计分析与 SAS 实现

主 编 冯国双 刘德平

副主编 刁玉涛 刘世炜 王俊杰 陈丽娟

编 者 (按姓氏笔画排序)

刁玉涛 王丽娟 王园园 王俊杰

冯国双 刘世炜 刘晓存 刘德平

李秀燕 吴松岭 张 华 陈丽娟

胡跃华 徐丽丽

# YIXUE SHIYAN SHEJI FENXI YU SAS SHIXIAN

## 图书在版编目 (CIP) 数据

医学实验设计分析与 SAS 实现/冯国双, 刘德平主编.  
—北京: 北京大学医学出版社, 2014. 8  
ISBN 978-7-5659-0860-6

I. ①医… II. ①冯… ②刘… III. ①实验医学—医学统计—统计分析—应用软件 IV. ①R-33②R195.1-39

中国版本图书馆 CIP 数据核字 (2014) 第 110051 号

## 医学实验设计分析与 SAS 实现

---

主 编: 冯国双 刘德平

出版发行: 北京大学医学出版社 (电话: 010-82802230)

地 址: (100191) 北京市海淀区学院路 38 号 北京大学医学部院内

网 址: <http://www.pumpress.com.cn>

E - mail: [booksale@bjmu.edu.cn](mailto:booksale@bjmu.edu.cn)

印 刷: 北京佳信达欣艺术印刷有限公司

经 销: 新华书店

责任编辑: 董采萱 责任校对: 金彤文 责任印制: 张京生

开 本: 787mm×1092mm 1/16 印张: 20 字数: 504 千字

版 次: 2014 年 8 月第 1 版 2014 年 8 月第 1 次印刷

书 号: ISBN 978-7-5659-0860-6

定 价: 72.00 元

版权所有, 违者必究

(凡属质量问题请与本社发行部联系退换)

# 序 一

自从 90 年前 Fisher 爵士开创了实验设计的先河以来，实验设计已经在工业、农业、医学等诸多领域得到了广泛应用。在医学领域中，实验设计的重要性也已经得到公认，良好的实验设计已经成为高质量科研工作的必备条件。

尽管如此，国内医学领域对实验设计的重视仍然不够。很多设计方法对医学科研工作者而言仍只是停留在理论层面上，一旦真正需要进行实验设计，设计者便感觉无从下手，存在不少困惑。这导致医学实验设计难以真正推广和应用，阻碍了国内医学科学研究的发展。

因此，目前迫切需要应用性强的医学实验设计书籍，用于指导实验设计的具体操作和实现。由冯国双博士等编写的《医学实验设计分析与 SAS 实现》便是这样一本书。该书最大的特点是实用性和易读性。书中不仅介绍了常用实验设计的原理、思路及分析过程，更重要的是，对所有的设计和分析过程都给出了实现步骤及相应的 SAS 语句。读者只要根据书中提示，仿照书中给出的 SAS 语句编写程序，就很容易得到各种实验的设计过程与分析结果。

本书的写作风格延续了作者以往著作中深入浅出、通俗易懂的特点，避免大而空洞的说教，而是以朴实的语言将各种实验设计的实施思路与分析过程娓娓道来，即便读者的统计学基础薄弱，仍然不影响对本书的理解和应用。

本书的主编冯国双在博士研究生期间便一直协助医院的临床试验设计与分析工作，至今已积累了丰富的实践经验。作者撰写本书时查阅了大量的国内和国外经典文献及著作，并将自己多年来对实验设计及 SAS 分析的理解融入其中，再加上作者本身的实践经验，相信读者一定能从本书中获益良多。

游伟程

## 序 二

我在大量审阅稿件和学位论文中发现，研究设计存在问题的现象屡见不鲜，由此产生了一系列的问题，以致结论的科学性难以令人信服，其造成的损失更难以估计。美国学者 Daniel W. Byrne 对 50 位美国医学学会的审稿人、67 位医学杂志的主编和 25 位诺贝尔奖获得者进行了问卷调查 (Byrne DW. Science Editor, 2000, 23; 39 - 44)。他的研究结果发现，最常见的导致稿件直接被拒绝的单一原因有：科研设计不严谨 (71%)、结果的解释不恰当 (14%)、课题的重要性不足 (14.2%)、结果的表达方式不合适 (0.8%)。可见医学研究设计存在问题的严重性。

随着信息技术的飞速发展，功能强大的统计软件层出不穷，使越来越多计算繁复的统计方法得以推广应用，并使设计方案规范化，推动了医学研究的发展。SAS 软件在医学研究中的广泛应用便是其中典型的一例。

为了在医学研究的设计和分析中更好地推广应用 SAS，将不同的研究设计方法及其分析方法与 SAS 的正确应用结合起来，冯国双博士联系了一批有志为此而努力的中青年医学统计工作者、学者，结合工作实践撰写了本书，这正是当前医疗卫生一线工作者在研究实践中所需要的。

作者们在工作之余为本书的撰写付出了辛勤的劳动。编写过程中，主编冯国双博士带领大家经常交流探讨，以致开展辩论。这种严谨求实、团结共进的优良作风，使曾作为他们老师的我深感欣慰。他们的辛勤劳动将为更多的医务工作者、研究者和研究生提供良好的思路、方法和工具，会有更多的医学研究因此而获益。为此，借本书出版之际，填词一首，以记之。

### 清平乐·咏学人同心著书

夜闌人静，  
月淡寒窗映。  
寻路书山独辟径，  
阅尽青峰名胜。

学人牢结同心，  
挚情共著思深。  
喜鹊晚来鸣柳，  
东风又绿新春。

陈景武

# 前言

中国古人在几千年前便很重视预防的思想，如“圣人不治已病治未病”、“君子以思患而豫（预）防之”等，无不体现了先人的睿智。预防之于治疗，正如实验设计之于统计分析。如果已经病入膏肓，即使扁鹊再世亦是无能为力。同样，对于漏洞百出的数据，即使是伟大的 Fisher 先生重生同样束手无策。所以，实验设计的重要性其实更甚于统计分析。可惜的是，正如人们往往意识不到预防的作用一样，实验设计也很容易被忽视。

实验设计的历史可以追溯到百年前。提到实验设计，就不得不提一下 Fisher 爵士。当我在阅读 Fisher 的事迹时，尽管相隔近百年，仍会有一种莫名的肃然起敬的感觉。我仿佛看到这样一幅情景：一群人围在试验田边上，讨论如何对不同的试验田施以不同的肥料，有的说从南往北，有的说从东往西。Fisher 老先生则一言不发地抽着他的烟斗，直到众人讨论结束，他才笑眯眯地一锤定音：“就用随机的方法吧。”正是这一随机的概念，开创了实验设计的新时代。今天你可能不觉得随机这一概念有多么神奇，因为 Fisher 一旦提出，就变得简单了。但在他提出之前，却无人能想到这一点，你可以想象得到提出这一思想所需要的智慧。

时至今日，实验设计的理念已经渗入到科研领域的每一角落，医学研究领域对实验设计也日渐重视，然而现实的问题是：有不少人很清楚实验设计的重要性，但在实际操作中却无从下手，不知如何实现随机，如何计算样本量。导致这一问题的主要原因就是理论与实际相脱节。关于医学实验设计的书籍并不多，而且基本都是 10 年前的著作。限于当时的知识水平及软件实现能力，对实验设计的解释更多的是停留在理论层面上，而缺乏必要的实施工具，这使得不少医学研究者无法真正实现设计过程。

统计软件的快速发展使许多原本只能停留在理论层面上的方法变得易于实现，这其中也包括实验设计。在众多的统计软件中，SAS 是几乎涵盖所有常见实验设计与分析的权威软件。尽管有人觉得 SAS 需要编程可能较为复杂，但这恰是其灵活的一面，可以将很多不可能变成现实。

正是鉴于以上原因，本书从实际操作入手，介绍常见的医学实验设计与分析过程。本书的主要特色如下：

**(1) 实用性和操作性强。**本书的主要特点是利用 SAS 软件将设计与分析思路转化为实际操作与输出。而且针对实验设计的特点将设计与分析过程模式化，使医学研究者可以在书中程序的基础上只修改简单的数值，便可实现相应的设计与分析。但书中又不是简单地教人模仿，而是对每一设计和分析步骤给出详细解释，不仅授之以鱼，更是授之以渔。相信读者完全可以在本书的基础上做到举一反三，完成自己的实验设计过程。

**(2) 包含 SAS 软件应用指导。**本书涉及诸多 SAS 应用过程，如实验设计过程、样本量计算过程、统计分析过程等。对于所有 SAS 语句，作者在理解的基础上以通俗的语言进行解释，对有些过程语句的解释甚至比 SAS 使用手册更具实际意义。对 SAS 软件的输出结果我们保持了原有的格式，这样便于读者自己实践时进行对照。

**(3) 内容丰富、涵盖面广。**本书既有对常见设计方法的介绍，如完全随机设计、配对设

计、随机区组设计、拉丁方设计、析因设计、正交设计、交叉设计等，又有对非常见设计的介绍，如尧敦方设计、希腊拉丁方设计、裂区设计、嵌套设计等，同时也有对某些固定领域的设计方法的介绍，如诊断试验设计、重复测量设计、临床试验设计等。不同的读者可以各取所需。

**(4) 思路清晰。**本书各章节的主要结构安排是：首先简要介绍该设计方法；接着对设计中所需的样本量计算给出 SAS 实现过程；然后介绍对该设计的实施思路，并利用 SAS 软件将实施思路转化为实际操作；最后是统计分析，给出了不同类型资料的 SAS 分析过程。

**(5) 案例辨析。**书中部分章节在最后部分给出了错误案例辨析，这些错误应用案例均摘自已经发表的文章。通过对实际研究中常见问题进行分析，避免同类问题再次发生。

**(6) 严谨求实。**作者本着严谨求实的态度来撰写本书，内容都是在参考国外权威专著的基础上，纳入了作者多年来在不同设计方面的思路，具有较强的实际意义。作者对书中每一个概念都多方求证，力求准确。

**(7) 大量案例。**书中列举了大量的分析案例，这些例子大多是真实的，是作者在实际工作中所遇到过的案例。但有的案例只是截取其中部分数据，有的为了使读者更容易理解，对数据做了修改。当然，也有部分例子是我们虚拟的，但这些数据的构造并非空中楼阁，而是以一定的专业知识为基础的。

书中大部分 SAS 程序用的是 SAS 9.3 中文版，输出中文结果，以方便读者阅读。但对有些结果笔者认为中文翻译不妥或意义不明确，故部分结果采用 SAS 9.3 英文版输出英文结果。

为了方便读者，本书中所有案例设计与分析的 SAS 语句均放在北京大学医学出版社的网站上 (<http://pumpress.bjmu.edu.cn>)，有需要的读者可自行下载。

本书所面向的读者主要包括公共卫生领域科研工作者，临床医学工作者，各专业医学生、研究生，医学院校教师，合作研究组织 (CRO) 公司、制药企业等相关职员。

尽管作者尽力完善本书的内容，但由于能力有限，有的理解可能不一定正确，如有错误和不妥之处，敬请读者批评指正。

最后，感谢我的硕士导师陈景武和博士导师游伟程，他们教授我统计分析的技能和科研设计的思路。没有他们当年的指导，就没有今天这本书的面世。

冯国双

2014 年 6 月于北京

# 目 录

第一章 医学实验设计概述	1
第一节 初识实验设计	1
第二节 医学实验设计的基本要素	1
一、处理因素	1
二、受试对象	2
三、实验效应	2
第三节 医学实验误差及其控制	3
一、随机误差	3
二、非随机误差	3
三、误差的控制	3
第四节 医学实验设计的基本原则	4
一、随机化原则	4
二、对照原则	4
三、重复原则	4
第五节 医学实验设计的基本过程	5
一、明确想做什么	5
二、确定处理因素及其水平数	5
三、选择能反映研究目的的指标	5
四、选一个合适的设计类型	5
五、确定研究对象及样本含量	5
六、设计实施	5
七、数据收集与整理	5
八、结果的统计分析	5
第六节 实验数据的整理	6
一、原始数据的整理与录入	6
二、数据核查	6
三、异常值检查和处理方法	7
第七节 实验数据的统计分析	9
第八节 实验设计分析常用统计软件	16
一、Design-Expert 软件	16
二、SAS 软件	17
三、SAS JMP Clinical 软件	17
第二章 完全随机设计及统计分析	20
第一节 完全随机设计简介	20
一、完全随机设计简介	20



---

二、完全随机设计的要点 .....	20
第二节 完全随机设计的样本含量估计及 SAS 实现 .....	20
一、两组均数比较的样本含量估计 .....	24
二、两组率比较的样本含量估计 .....	25
三、多组均数比较的样本含量估计 .....	26
四、多组率比较的样本含量估计 .....	27
五、两组生存资料比较的样本含量估计 .....	28
第三节 完全随机设计的实施及 SAS 实现 .....	29
一、两组完全随机设计的实施 .....	30
二、多组完全随机设计的实施 .....	31
第四节 完全随机设计的统计分析及 SAS 实现 .....	33
一、两组正态资料比较的统计分析 .....	35
二、两组非正态资料比较的统计分析 .....	37
三、多组正态资料比较的统计分析 .....	38
四、多组非正态资料比较的统计分析 .....	42
五、两组率比较的统计分析 .....	43
六、多组率比较的统计分析 .....	46
七、两组有序资料比较的统计分析 .....	48
第五节 案例辨析 .....	49
第三章 配对设计与分析 .....	51
第一节 配对设计简介 .....	51
一、配对设计简介 .....	51
二、配对设计的要点 .....	51
第二节 配对设计的样本含量估计及 SAS 实现 .....	51
一、定量资料的样本含量估计 .....	53
二、分类资料的样本含量估计 .....	54
第三节 配对设计的实施 .....	54
第四节 配对设计资料的统计分析及 SAS 实现 .....	56
一、定量资料的统计分析 .....	57
二、二分类资料的统计分析 .....	60
三、多分类资料的统计分析 .....	61
第五节 案例辨析 .....	64
第四章 随机区组设计与分析 .....	67
第一节 随机区组设计简介 .....	67
一、完全随机区组设计 .....	67
二、平衡不完全随机区组设计 .....	67
第二节 随机区组设计的样本含量估计及 SAS 实现 .....	69
第三节 随机区组设计实施及 SAS 实现 .....	70
一、完全随机区组设计的实施 .....	71
二、平衡不完全区组设计的实施 .....	72

第四节 随机区组设计的统计分析及 SAS 实现	74
一、完全随机区组设计的统计分析	76
二、平衡不完全区组设计的统计分析	82
第五节 案例辨析	87
第五章 拉丁方设计与分析	89
第一节 拉丁方设计简介	89
一、拉丁方设计	89
二、重复拉丁方设计	90
三、尧敦方设计	90
四、希腊拉丁方设计	91
第二节 拉丁方设计的实施及 SAS 实现	92
一、拉丁方、尧敦方设计的实施	94
二、希腊拉丁方设计的实施	96
第三节 拉丁方设计的统计分析及 SAS 实现	97
一、拉丁方设计的统计分析	98
二、尧敦方设计的统计分析	101
三、希腊拉丁方设计的统计分析	102
四、重复拉丁方设计的统计分析	103
第六章 析因设计与分析	107
第一节 析因设计简介	107
一、基本概念	107
二、析因设计中的效应	108
三、析因设计的主要用途	109
第二节 析因设计的样本含量估计及 SAS 实现	109
第三节 析因设计的实施及 SAS 实现	111
一、齐水平析因设计的实施	112
二、混水平析因设计的实施	114
第四节 析因设计的统计分析及 SAS 实现	116
一、定量资料的析因设计分析	117
二、分类资料的析因设计分析	121
第五节 案例辨析	123
第七章 正交设计与分析	124
第一节 正交设计简介	124
一、正交设计简介	124
二、正交表与交互表	124
三、正交设计的步骤	132
四、正交设计主要用途	134
第二节 正交设计的实施及 SAS 实现	134
一、齐水平正交设计的实施	135
二、混水平正交设计的实施	139

第三节 正交设计的统计分析及 SAS 实现 .....	143
一、定量资料的正交设计分析 .....	143
二、分类资料的正交设计分析 .....	148
第四节 案例辨析 .....	150
<b>第八章 交叉设计与分析</b> .....	152
第一节 交叉设计简介 .....	152
一、交叉设计简介 .....	152
二、交叉设计的特点 .....	152
第二节 交叉设计的样本含量估计及 SAS 实现 .....	153
第三节 交叉设计的实施及 SAS 实现 .....	154
第四节 交叉设计的统计分析及 SAS 实现 .....	155
一、定量资料的交叉设计分析 .....	157
二、分类资料的交叉设计分析 .....	168
<b>第九章 诊断试验设计与分析</b> .....	171
第一节 诊断试验设计简介 .....	171
第二节 诊断试验的样本量估计及 SAS 实现 .....	172
第三节 诊断试验评价的常用方法和指标 .....	175
一、分类资料的评价方法 .....	175
二、连续资料的评价方法 .....	178
三、诊断试验评价的 SAS 实现 .....	178
第四节 诊断试验的评价分析及 SAS 应用 .....	180
一、分类资料的诊断试验分析 .....	180
二、定量资料的诊断试验分析 .....	189
第五节 案例辨析 .....	194
<b>第十章 重复测量资料的设计与分析</b> .....	195
第一节 重复测量设计与数据结构 .....	195
第二节 重复测量设计的样本含量估计 .....	196
第三节 重复测量设计的常用统计分析方法 .....	198
一、重复测量方差分析 .....	199
二、广义估计方程 .....	201
第四节 重复测量资料的统计分析及 SAS 实现 .....	202
一、定量资料的重复测量分析 .....	202
二、分类资料的重复测量分析 .....	211
<b>第十一章 嵌套设计与裂区设计</b> .....	217
第一节 嵌套设计与分析 .....	217
一、嵌套设计简介 .....	217
二、嵌套设计的实施 .....	218
三、嵌套设计的统计分析 .....	219
第二节 裂区设计与分析 .....	224
一、裂区设计简介 .....	224

---

二、裂区设计的实施	225
三、裂区设计的统计分析	227
第三节 案例辨析	235
<b>第十二章 序贯设计与分析</b>	236
第一节 序贯设计简介	236
第二节 成组序贯设计的实施与 SAS 实现	236
一、定量反应资料的实施与样本量计算	239
二、分类反应资料的实施与样本量计算	242
第三节 成组序贯设计的统计分析与 SAS 实现	245
一、定量反应资料的统计分析	246
二、分类反应资料的统计分析	256
<b>第十三章 临床试验设计与分析</b>	265
第一节 临床试验简介	265
第二节 临床试验的设计	265
一、临床试验的设计方法	265
二、临床试验的随机化方法	266
三、临床试验的盲法	269
四、临床试验的对照法	270
第三节 临床试验的分类	271
一、非劣效性试验	271
二、等效性试验	272
三、优效性试验	273
第四节 临床试验中的样本含量估计及 SAS 实现	273
一、均值的非劣效、等效和优效性检验样本量估计	276
二、率的非劣效、等效和优效性检验样本量估算	282
三、生存资料的非劣效性检验样本量估算	286
第五节 临床试验中的统计分析及 SAS 实现	287
一、均值的非劣效、等效和优效性检验	290
二、率的非劣效、等效和优效性检验	296
三、生存资料的非劣效性检验	300
<b>参考文献</b>	303

# 第一章 医学实验设计概述

## 第一节 初识实验设计

实验 (experiment) 是为了证实或验证某种假说而开展的系统研究的过程。一个好的实验可以帮助你回答很多问题, 例如:

(1) 一种新研发的降血压药物对高血压患者的血压降低是否有效?

(2) 中药的熬制需要考虑多个条件, 如时间、温度、剂量等, 在多长时间、多高温度、多少剂量的条件下效果最优?

(3) 两种药物或多种药物联合应用治疗某病, 是否一定优于单种药物的治疗?

(4) 如果减少机动车辆的出行, 空气质量是否会有所好转?

通过良好的实验设计, 不仅可以回答上述问题, 而且还可以最大限度地节省成本。著名统计学家 G. E. P. Box 说过, 假如有 10% 的工程师使用各种实验设计方法, 产品的质量与数量都会得到很大提升。这对于医学工作者进行实验设计工作同样具有重要的意义。

本书重点介绍医学方面的实验设计实施与分析过程。依据研究对象不同, 常将实验分为三类: 动物实验、临床试验和社区干预试验。

动物实验 (animal experiment) 是以动物为研究对象进行的实验。动物实验分为急性、亚急性及慢性动物实验三种, 其中前两种多用。例如, 药理实验、病理变化实验、毒物的致癌实验等。

临床试验 (clinical trial) 是以患者或健康人为研究对象, 以临床为基础对治疗效果作出评价的实验。可以是短期、中期或远期追踪观察。例如, 对某种新型化疗方法治疗淋巴细胞白血病效果的研究。

社区干预试验 (community intervention study) 是以社区人群为研究对象的干预试验, 常在某一地区的人群中进行, 持续的时间一般较长。试验目的常是观察某项保护措施对抑制危险因素致病的效果。例如, 某地区饮水中加氟预防龋齿的试验。

## 第二节 医学实验设计的基本要素

医学实验包括三个基本组成部分: 处理因素 (treatment)、受试对象 (study subjects) 和实验效应 (experimental effect)。例如, 用两种降眼压药治疗青光眼患者, 观察比较应用两种不同降眼压药的两组患者眼压值的下降情况, 这里所用的降眼压药就是处理因素, 青光眼患者为受试对象, 被测的眼压值则为实验效应。构成医学实验的这三个基本组成部分称为医学实验的三个基本要素。任何一项实验研究在进行设计时, 必须首先明确这三个要素, 并以其为重点制订详细、严密的研究计划。

### 一、处理因素

处理因素是研究者为实现研究目的而主动安排的, 通常在设置时应注意以下几点。

1. 明确实验中的主要因素 实验中的主要因素必须根据研究目的来确定,可适当地确定几个能回答研究目的的因素作为处理因素。处理因素不宜设置太多,否则会使分组增多,受试对象的总例数相应增多,使实际工作的难度也增加。但如果处理因素过少,则常难以增加实验的广度和深度。

2. 找出可能影响实验结果的非处理因素 非处理因素不是我们要研究的因素,但其中有些可能会影响实验结果,其效应混杂于处理因素的效应中,有时这些会影响实验结果的非处理因素又称为混杂因素。例如,观察两种降眼压药治疗青光眼患者的效果,非处理因素可能有病程长短、病情轻重等,若两种降压药组患者的病程、病情构成不同,则很可能对两种药物的疗效比较造成混杂。故实验中必须设法消除这些非处理因素的干扰作用。

3. 处理因素要标准化 处理因素在整个实验过程中必须标准化,即保持稳定不变。如实验的处理因素是药物,那么药物的质量(成分、厂家、出厂批号等)必须保持不变,否则会影响对实验结果的评价,或产生混杂作用。

4. 处理因素应安排合适的水平 所谓水平(level),通俗来说就是处理因素的类别。例如,研究两种药物治疗青光眼的效果,药物就是处理因素,两种药物就是两个水平。再如,观察某药物三种剂量治疗某病的疗效,该药物就是处理因素,药物的三个剂量就是三个水平。处理因素的水平数通常需要根据研究目的来确定。如果是初筛实验,可以先设定水平数少一些;如果是验证性实验,可以设置得细一些,以便找出最优组合。

## 二、受试对象

医学实验中的受试对象主要是动物或人。

1. 受试动物的选择 研究课题、研究目的不同,对受试动物的要求也往往不同。选择受试动物时应注意种类、品系、年龄、性别、体重、窝别和营养状况等。

2. 受试人的选择 受试人可分为健康人或患者。选择受试病例时应当有明确的诊断,并且受试者依从性好,能真实、可靠地反映主观感觉。依从性是指受试者在实验过程中对处理因素的服从程度,如配合护理、按时服药等。另外,还要求受试病例退出实验的可能性小,以免失访率过高(如大于20%)而影响研究结果的评定。预防医学中的人群实验常以健康人为受试对象,如小学生中预防红眼病的实验。

## 三、实验效应

受试对象对处理因素的反应称为实验效应。实验效应是通过观察指标来表达的。因此,选好观察指标并正确地进行观察是至关重要的。

1. 指标的选择 对实验效应的观察指标有多方面的要求,重要的是要客观性强、灵敏度高、精确性好。客观指标是测量和检验的结果,比主观指标更可靠。指标的灵敏度高,就能使处理因素的效应更有效地显示出来。指标的精确性包括准确度和精密度两方面。准确度(accuracy)是观察值与真值的接近程度,主要受系统误差的影响。精密度(precision)是重复观察时,观察值与其均数的接近程度,其差值属随机误差。

2. 指标的观察 观察实验效应时一定要避免带有偏性(bias)。如医生常会偏重于新疗法组,而患者常对新疗法持怀疑态度等。这些都可能导致测定实验效应指标时带有偏性,从而影响实验结果的比较和分析。为避免或减少带有偏性,设计时常采用盲法(blind method)。例如,使受试对象、实验的执行者(医生等)甚至研究者等均不清楚哪组是实验组,哪组是对照组,预期结果是什么。

### 第三节 医学实验误差及其控制

误差 (error) 是实测值与真值之差, 按产生的原因及其性质可分为两类: 随机误差和非随机误差。

#### 一、随机误差

随机误差 (random error) 是由一系列偶然因素引起的不易控制的误差, 该误差值具有统计规律性, 如重复误差。在测量过程中, 存在的变异成分越多, 则测定值越分散, 随机误差就越大。例如, 同一名医生多次测量同一患者的眼压值所产生的随机误差, 要小于多名医生同时测量同一患者眼压值所产生的随机误差。随机误差虽然不能避免, 但是却有规律可循。大量的观察说明, 随机误差的分布是以 0 为中心的正态分布。

随机抽样产生的抽样误差服从随机误差分布的规律。抽样误差越大, 对实验结果进行统计推断的困难也就越大。实验设计的目的之一, 就是要控制和减小随机误差。

#### 二、非随机误差

非随机误差 (nonrandom error) 又称系统误差 (systematic error) 或偏倚 (bias), 是由于对受试对象、实验因素或条件等控制不严密而产生的一种误差, 有一定来源, 并使实验结果有倾向性地偏离真值。根据来源不同, 系统误差可分为以下几种:

1. 选择偏倚 选择偏倚 (selection bias) 是由于对受试对象选择或分组不当所产生的偏倚。例如, 在选择病例时未严格执行入选标准和剔除标准, 或某组患者的依从性不好等, 都可能使实验产生选择偏倚。

2. 信息偏倚 信息偏倚 (information bias) 又称测量偏倚, 是由于测量仪器、操作方法、实验条件以及主观因素等方面的原因造成的偏倚, 如数据记录错误等。

3. 混杂偏倚 混杂偏倚 (confounding bias) 是由影响实验结果的非处理因素在各对比组中不平衡所引起的偏倚。如疾病的疗效除了药物的治疗作用外, 还与患者的性别、病情、心理状态等多方面的因素有关, 若忽略了这些有关因素在各对比组的均衡性, 就会发生混杂偏倚。例如, 比较 A、B 两种药物对青光眼的治疗效果, 如果服 A 药的患者大多数为重型, 而服 B 药的患者多数为轻型, 此时若结果显示 B 药疗效优于 A 药, 则很难说清到底是因为 B 药确实优于 A 药, 还是因为服 B 药的患者中轻型者较多, 比较容易取得好的疗效, 以致造成虚假结果。

#### 三、误差的控制

实验误差在实验设计、实施和结果分析的各个阶段都存在, 因此, 在各个阶段都要针对误差产生的原因采取相应的控制措施, 力求使研究得到接近真实的可靠结果。

例如, 对研究总体的规定明确具体, 正确拟定观察对象的纳入标准和排除标准, 确定合适的样本含量, 随机抽样、随机分组, 保证样本有好的代表性等, 可有效防止选择偏倚。

又如, 实验因素标准化, 实验过程中的统一、规范、完整记录, 定期或经常校准仪器, 对参与研究的实验人员进行技术培训, 盲法的应用等, 可控制、防止信息偏倚。

再如, 设计时明确处理因素和对实验效应有影响的非处理因素, 以选用合适的实验方法, 均衡非处理因素的影响; 在实验阶段始终贯彻标准化的实验条件; 在统计分析阶段采用合适的统计方法, 以排除混杂因素的干扰等, 皆可防止混杂偏倚的发生或减少其影响。

## 第四节 医学实验设计的基本原则

如前所述, 实验设计最重要的作用就是减小误差、提高效率。为达到这一目的, 在设计时必须遵循三项基本原则, 即随机化原则、对照原则、重复原则。

### 一、随机化原则

随机化 (randomization) 是实验设计的基石, 几乎所有的实验都需要随机化, 其目的是使各对比组间不可控的非处理因素的分布尽量一致。在实验对象的抽样、分组及实施过程中都要随机化。如用随机的方式抽取样本, 使符合条件的实验对象参加实验的机会相同; 将处理因素随机地分配给样本中的每一个受试单位, 使每个受试单位被分到实验组与对照组的機會均等。

随机化可以将不可控的非处理因素“平均掉”, 使非处理因素在实验组与对照组中的影响较为一致, 是保证比较组间有良好均衡性的重要手段之一。例如前述的比较 A、B 两种药物治疗青光眼的效果, 如果采用随机分配, 通常可以避免两组的病情差别过大, 保证两组轻型和重型病例的比例基本一致。

随机不等于随便, 既不是由研究者主观决定的, 也不是由受试者随意选择的, 而是通过随机化的分组程序实现。目前多数统计软件均可通过产生随机数的方式来实现随机化, 这是目前应用最广泛的方法。例如, 将 8 只小鼠随机分为两组, 首先将小鼠按体重大小编号为 1~8, 然后利用计算机产生 8 个随机数字, 假定分别为 0.984、0.837、0.414、0.548、0.122、0.970、0.054、0.193, 按从小到大的顺序分别编号为 8、6、4、5、2、7、1、3。规定前 4 个较小的随机数分配到 A 组, 后 4 个较大的随机数分配到 B 组, 则 8 只小鼠的排列为: B、B、A、B、A、B、A、A。

### 二、对照原则

对照原则指在确立接受处理因素的实验组时, 同时设立不接受处理因素的对照组 (control)。实验组与对照组除了处理因素不同外, 其他重要的、可控的非处理因素的分布尽量保持一致, 即要求均衡 (balance)。这样才能平衡非处理因素对实验结果的影响, 把处理因素的效应充分显示出来。所以, 实验组与对照组间均衡一致是正确显示实验效应的前提。例如, 动物实验的不同对比组常在年龄、体重、健康状况等构成上保持一致, 临床试验常考虑病情、病程、以往接受治疗情况等的一致。对照组设立后, 还应对各对比组的基线情况 (baseline) 进行比较, 检验组间的均衡性。

注意, 对照组是专为相应的实验组设立的, 不可借用以往研究结果或其他研究资料作为现研究的对照, 因为当时的条件跟现在的条件并不同, 这种对照已经失去了原有意义。

### 三、重复原则

重复 (replication) 是指在相同的实验条件下, 进行多次研究或多次观测。由于存在变异和实验误差, 即使同一实验条件下的实验单位也可能产生不同的观测结果。为了保证结果可靠, 通常需要在相同实验条件下, 进行一定数量的重复观察, 这里的“一定数量”就是指样本含量。一般说, 样本含量大, 抽样误差相对就小, 但样本含量过大时, 会加大实验的规模, 延长实验时间, 造成人力、物力的浪费, 还增加了系统误差出现的可能性。所以, 正确估计一个实验的观察单位数也是实验设计的重要内容。



注意这里的重复与重复测量 (repeated measurement) 是有区别的。重复测量是对同一受试对象的多次重复观测, 通常是为了保证结果的精度, 或为了观察某指标随时间变化的情况。

## 第五节 医学实验设计的基本过程

### 一、明确想做什么

确定研究目的是第一步要做的。如果不明白自己要做什么, 那后面的所有步骤就都不需要了。所以如果要开展一项实验, 首先要明确自己想研究的内容是什么, 而且一定要具体、明确, 不能是空中楼阁。例如, 要研究药物对眼病的治疗效果, 这就很容易让人摸不着头脑, 但是如果改为研究某滴眼液对结膜炎的治疗效果, 就明了得多。

### 二、确定处理因素及其水平数

处理因素是研究目的的直接体现, 需要紧密结合研究目的恰当选择处理因素及其水平。如果想观察某滴眼液对结膜炎的治疗效果, 则处理因素就是滴眼液, 这时需要考虑对照药物选哪一种, 选择几种对照药物, 用不用考虑不同剂量的效果等。

### 三、选择能反映研究目的的指标

选择研究指标就是为了把研究目的转换为看得见、摸得着的具体指标, 比如研究目的是想看一种新药对降低眼压是否有效, 要想真正达到这一目的, 就必须通过具体指标来反映, 如眼压降低值, 这是比较实在的、能具体看到的指标。

### 四、选一个合适的设计类型

一旦明确了目的, 便可以根据目的、处理因素及其水平来选择一个合适的实验设计类型。如果只想考查一个处理因素, 可以用完全随机设计; 如果考虑到还有非处理因素的影响, 可以用随机区组设计、拉丁方设计等; 如果考查多个因素, 可选用析因设计、正交设计等。

### 五、确定研究对象及样本含量

前面这些都确定后, 基本就可以进入实施阶段了。当然首先得选择研究对象, 所以需要考虑下面的这些问题: 选择哪些研究对象, 是动物实验还是人群研究? 选择多少例研究对象? 如果是临床试验, 需要确定哪些入选和排除标准?

### 六、设计实施

选定研究对象后, 就可以实施处理因素了。不同实验设计类型有不同的实施方式, 如随机化分组、安排处理因素等, 详见以后各章。

### 七、数据收集与整理

实验结束后, 数据也就相应收集到了。对于较正式的实验, 最好是有专门的数据库用于录入数据, 如 Access、Epidata 等。当然, 如果是简单的动物实验, 也可以直接录入在 Excel 中。总之, 关键就是保证数据不要录入错误。

### 八、结果的统计分析

数据整理好后, 就可以进入统计分析步骤了。尽管国内不少医生都是全才, 能够胜任从设计到统计分析的所有工作, 也没有找人做统计分析的习惯, 但还是建议能够跟生物统计学家合作, 毕竟他们更为专业一些。