

医药高等院校教材  
供临床、药学、检验、影像、护理等专业用

# 医学机能实验设计的 原理和方法

主编 张晓

成都医学院

医药高等院校教材  
供临床、药学、检验、影像、护理等专业用

医学机能实验设计的原理和方法

主 编 张 晓  
副主编 咎加禄

编 委 (按拼音顺序排列)

陈建敏 陈卫中  
梁 楠 刘 静  
荣 成 杨 拯

## 基础医学实验教学中心

### 序 言

二十一世纪高等医学教育的显著变化是强调学科间的交叉融合,而现代医学也已经由经验科学转变为实验科学。因此,医学教育,特别是在基础医学教育阶段,重视实验教学,培养和训练学生的实践能力是全面提高学生综合素质、培养创新意识的重要环节。机能实验学是一门研究生物正常机能、疾病发生机制和药物作用规律的新兴的基础医学实验课程,它继承并发展了生理学、生物化学、药理学和病理生理学实验课程的核心内容,更加强调学科之间的交叉融合,更加重视新技术的应用,更加注重学生创新能力的培养。同时,机能实验技术也广泛应用于生命科学的各个领域和医学实验的研究工作中。因此,机能实验学对于临床医学实践和医学科学研究具有重要的指导意义。

本书由我院青年教师张晓教授牵头,邀请我院生理学、药理学和病理生理学教研室的专家和教授参加编写和审阅。全书系统介绍了机能实验设计的原理和方法,系编者结合设计性实验教学的实践经验而撰写,不仅介绍了实验设计的基本思路,还介绍了实验设计应该细化为专业设计、统计设计和动物设计等方面,具有很好的参考价值。以皮亚杰为代表的建构主义学习理论认为,知识是个体与环境在相互交互作用的过程中逐渐建构的结果。个体在不断与环境的接触中建构知识和行为策略,在这个建构过程中,一方面学生受个人兴趣、需要以及外部环境的推动表现为主动性和选择性;另一方面受本人原有知识经验、思维方式、情感品质等制约,在对信息的内部加工上表现为独立性和创造性。本书的编著者将“知

识建构”和“创新”观念贯穿于教学的各个环节,引导学生通过能动的建构过程学习知识,建构有利于培养学生创新能力和实践能力的新教学模式。

全书深入浅出,循序渐进,内容丰富,图文并茂,章节编排合理,力求使学生在掌握实验技能的基础上,启发式的培养学生的综合分析能力和科研能力,不仅可以作为机能实验学的教学和参考资料,同时也对从事生理学、药理学、病理学和毒理学的技术人员具有一定的参考价值。

本书是在学院党委提出把我院办成“特色鲜明、优势突出”的高等医学院校作为我们的奋斗目标的背景下完成的,对于树立以学生为本,知识传授、能力培养、素质提高协调发展的教育理念,全面提高教学质量具有重要意义。我们要转变实验教学观念,以能力培养作为核心,从传统的显示现象和验证理论向培养高素质和创新型人才转变。该书适应于当前医学教学模式的转变,希望此书对当前医学院校机能实验学的教学改革和发展起到促进和推动作用。

2006年6月于成都医学院

## 目 录

<a href="#">第一章 医学机能实验设计概述... 6</a>	6
<a href="#">第一节 医学科研概论... 6</a>	6
<a href="#">一、医学科研的任务... 6</a>	6
<a href="#">二、医学科研的分类... 6</a>	6
<a href="#">三、医学科研的假说... 7</a>	7
<a href="#">(一) 假说的特征... 7</a>	7
<a href="#">(二) 假说的作用... 8</a>	8
<a href="#">(三) 假说的建立... 8</a>	8
<a href="#">(四) 假说的验证... 9</a>	9
<a href="#">第二节 设计性实验的基本步骤... 9</a>	9
<a href="#">一、设计性实验选题... 9</a>	9
<a href="#">二、实验设计... 11</a>	11
<a href="#">三、实验设计实施... 12</a>	12
<a href="#">四、资料分析... 12</a>	12
<a href="#">五、总结归纳... 13</a>	13
<a href="#">第二章 机能实验专业设计... 14</a>	14
<a href="#">第一节 处理因素... 14</a>	14
<a href="#">一、概述... 14</a>	14
<a href="#">二、确定处理因素要注意以下几个问题... 14</a>	14
<a href="#">三、处理因素确定中常见的问题... 15</a>	15
<a href="#">第二节 受试对象... 15</a>	15
<a href="#">一、概述... 15</a>	15
<a href="#">二、受试对象的确定... 15</a>	15

三、受试对象的条件... 15	
四、受试对象的纯化(Purification)	
五、受试对象的依从性... 16	
六、受试对象选择中常见的问题... 16	
第三节 实验效应... 16	
第三章 机能实验统计设计... 18	
第一节 统计学的基本概念... 18	
第二节 统计设计的基本思想... 21	
一、实验统计设计的特点... 21	
二、正态分布... 22	
三、抽样误差和抽样分布... 26	
四、统计推断... 28	
(一) 计量资料的假设检验... 28	
(二) 计数资料的假设检验... 30	
第三节 实验设计的基本原则... 32	
一、随机化原则... 32	
二、对照的原则... 33	
三、重复的原则... 33	
第四节 常用的实验设计统计方法... 34	
一、完全随机设计... 34	
二、配伍设计... 36	
三、析因设计... 38	
四、拉丁方设计... 38	
五、多个样本均数间的两两比较... 42	
第五节 实验设计的基本步骤... 44	
一、建立研究假设... 44	
二、明确研究范围... 44	
三、确立处理因素... 45	
四、明确观察指标... 45	
五、明确资料类型... 45	
六、选用统计学方法... 46	
第六节 科研实验误差的控制... 49	
一、误差种类... 49	
二、误差的来源... 49	
三、误差的控制... 50	
第七节 医学机能实验统计设计中常见的问题分析... 50	
一、在实验资料的收集和整理方面存在的问题... 51	
二、设立对照方面存在的问题... 51	
三、实验设计重复方面存在的问题... 51	
四、实验分组方面存在的问题... 51	
五、受试对象方面存在的问题... 51	
六、观察指标方面存在的问题... 51	
七、观察或测量方面存在的问题... 52	
八、统计设计类型方面存在的问题：... 52	

第四章	实验动物设计... 54
第一节	选择实验动物应注意的问题... 54
一、	年龄与体重... 54
二、	性别... 54
三、	生理功能状态... 54
四、	健康状况... 54
五、	品系与等级... 55
第五章	实验设计指南... 56
一、	教学目标... 56
二、	主要研究方向和鼓励探索的范围... 56
三、	学习要求... 57
四、	时间安排... 57
五、	工作程序... 57
第六章	设计实验的教学与管理... 58
一、	实验设计大纲... 58
二、	实验报告的撰写... 60
三、	机能设计性实验课的组织与管理... 60
四、	机能设计性实验课考核办法... 60

## 第一章 医学机能实验设计概述

### 第一节 医学科研概论

医学科学研究工作由三个部分组成：科研设计、科研实施、观察与分析结果并得出结论。科研设计是医学科学研究工作中关键的一步，确定一个创意新颖、设计周密、指标合理、科学性强、切实可行的研究方案是取得科研成果的根本保证。

#### 一、医学科研的任务

医学科研工作的基本任务在于了解人体的结构、功能与理化机制，开发促进人民身心健康的科学技术，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，提出有效的防治措施，改善人类生存的环境，提高人民的生命质量。

##### 1.发现生命科学领域中的未知事物及其内在规律

在客观世界中始终存在着大量的未知现象和未知事物的发展过程及规律，医学客观世界是无限的，人们对医学客观世界的认识是有限的。例如，克隆技术，特别是器官克隆，虽然发展迅猛，但是从全局观察与分析，所走过的道路仅为前进中的第一步，胚胎干细胞怎样诱导分化为各种有临床价值的细胞和组织还存在很多问题，还有大量的工作要做。再如，当人们攻克了结核、天花和鼠疫等传染病之后，SARS 和禽流感等又成为新的危害人类健康的疾病，这些问题等待着我们去探索。

##### 2.寻找生命科学领域中已知事物的未知规律

科研的重要任务之一，就是揭示已知事物外在表现的本质与内在的联系，即事物内部的客观规律。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。例如人们已经认识到恶性肿瘤给人类健康造成巨大的损害，但其发病机理至今尚未阐明。虽然恶性肿瘤的发生与癌基因和抑癌基因

被活化有关，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.探索人类生存环境与人类身心健康的联系

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。医学科研工作也朝着这一模式转轨，无论是从宏观观察还是从微观分析都需注重这个目标。现代医学的许多难题和疑点，必须结合生命科学中各学科综合解决，深入探讨生存环境对人类身心健康的影响，全方位寻找保护和增进健康、防治疾病的新途径。

### 4.探索生物技术的开发与应用

医学科研的另一重要任务是探索如何利用已知的科学理论，发展出新的生物技术，应用于防治疾病和增进身心健康。例如应用基因理论和电子工程理论发展出基因芯片技术，用于临床诊断疾病，用基因工程技术改造蛋白质用于疾病的治疗。

## 二、医学科研的分类

医学科研可以分为以下几类：

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和实验场所不同，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、调查性科研。

(1)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有目的的人为的干预，而排除了外界因素的各种干扰，因而可以获得较为可靠的科学数据。此类研究大多在实验室进行，研究对象一般是动物，但有一部分研究对象是社区人群。

(2)临床试验研究：临床试验研究是一种对研究对象的身心健康进行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通常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)。对 clinical trial 应理解为已经通过有关的实验室研究，即将走向人群进行现场的流行病学实验。其研究对象是患者(研究治疗效果)，也可以是“非病人”(研究预防措施)。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。

(3)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，观察其对人群产生的效果。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)。

(4)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。

## 三、医学科研的假说

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做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，并预言了连接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，它是在观察和实验材料不足的情况下，凭借思维活动做出的有待于实践检验的问题，将来有可能被实验证实而进一步发展成为理论，也有可能被证伪而被淘汰。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，才能开阔思路，找出解决问题的突破口。

## (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. 假说是科研工作的方向

科研是一个探索未知事物的发展变化规律的过程，如果没有假说，就会在科研工作中迷失方向。因此，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。这一发现产生了预防白喉免疫的方法，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些因素可能与宫颈癌的发病有关。随后的研究表明:宫颈癌与性交引起的第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。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(validation)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,所以我们将验证的结论绝对化。

## 第二节 设计性实验的基本步骤

设计性实验与基础性和综合性实验有着本质上的区别,基础性和综合性实验



是在前人工作与经验总结的基础上，只要能按照书本上的要求去做就可以达到学习的目的；设计性实验是在借助前人的工作和经验积累的基础上，通过对研究对象进行积极的思考与归纳，对未知因素进行大胆设计，进行设计研究的一种科学实验。设计性实验是一项要求较高、难度较大的实验教学工作，对于学生来说是一种全新的学习方式，对于教师也提出更高的要求。

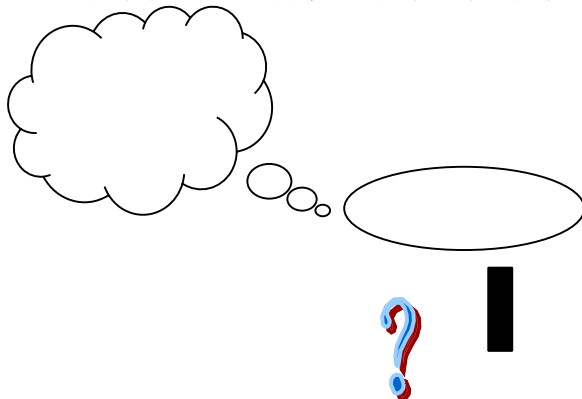
如何开展设计性实验？开展设计性实验的基本思路见图17-2-1，设计性实验首先要从确定探索的对象，提出问题开始。问题从哪里来呢？通过临床观察，课堂学习和阅读文献去发现问题和提出问题。在这个过程中，阅读和学习文献最重要。在阅读文献时，要了解所研究的问题的历史和现状，目前还存在什么问题，确定所提出的问题是否具有研究价值？所提出的问题是否是一个真正的问题？下一步的工作是立题，即根据我们自己掌握的知识，提出解决问题的思路，提出可能的假说，确定研究的内容。立题后要要进行实验设计，包括专业设计、统计设计、动物设计和方法设计等几个方面，实验设计完成后写出实验设计大纲。然后实施实验，观察记录。最后，对得到的数据资料统计分析，得出结论，撰写实验论文。因此，设计实验的基本步骤，丝丝入扣、紧密相关，基本程序包括：①查阅文献、提出问题、拟定研究题目；②实验设计：确立实验所需的观察指标，确立实验所需的统计学方法，确立复制动物疾病模型的方法，制定实验研究方案和实验技术路线，写出实验设计大纲；③认真完成实验；④准确、全面的记录实验结果；收集相关的实验文献资料，分析实验工作，得出结论；⑤撰写论文，报告实验工作。通过设计性实验教学，力求使学生初步掌握医学科学实验的基本程序和方法，培养学生独立进行科学研究的能力。

## 一、设计性实验选题

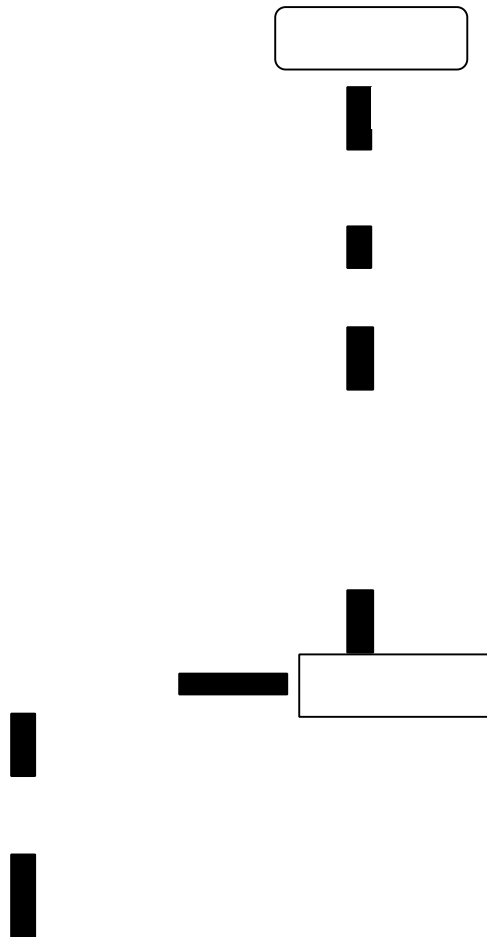
设计性实验的选题或立题就是确定所要研究的内容，选题的关键是发现问题和提出问题。

**1.准备工作：**在设计性实验设计前必须查阅大量的参考文献，认真搜集和阅读与本课题有关的研究资料。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，使您提出的问题得到进一步的深化、系统和完善，让其具有新颖性、科学性和可行性，才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。目前，我们主要通过因特网和光盘检索系统查阅中外文参考文献。在阅读科学文献时我们要把握以下几个问题：本研究领域的历史发展过程，以往解决了哪些问题，还有哪些问题未解决，未来的发展趋势。对所拟研究内容的科学价值、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，要具有创新性、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，而你选择了具有特色的系统性研究。在深思熟虑科学性的同时兼顾创新性、先进性和可行性，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的统一。

### **3.选题范围**

由于科学技术发展速度很快，医学机能实验学科学无论在理论上或实践上都有大量研究课题可供选择。但对在校学生而言，由于各种条件的限制，其选择范围不宜太宽，条件要求不宜太高。主要应该围绕在生理学、病理生理学及药理学

中所学的理论知识和相关文献,按照选题具有科学性、创新性、先进性及可行性的原则进行选题,并在教师的指导下进行。现将其选题参考方向简述如下。

#### (1) 对原有的实验方法进行改进

认真学习生理学、病理生理学及药理学中的实验方法和原理,对以往的实验方法进行改进。设计改进的思路和方法,并在实验中加以证实其可行性、实用性和科学性。例如,在复制高钾血症动物模型时,前人的经验是用3种不同浓度(1.0%、2.0%、3.0%)的氯化钾进行动物高钾血症模型的复制,是否还有另外一种不同于以往浓度的氯化钾溶液呢?如果在实际工作中确认还有另外一种浓度对复制的动物模型效果更好、更可靠,就可以列为研究题目,进行设计性研究。

#### (2) 对动物模型进行设计

对原有的动物模型进行设计时,要注意如下原则:①该动物模型的实验结果好,稳定可靠;②该动物模型的可重复性好;③该动物模型的方法更趋于简单和实用;④该动物模型解决了一些具体的实际问题,而且有推广使用的价值。

#### (3) 对实验原理进行探讨

随着科学的发展,机能实验的很多原理可以进行再认识,再探索。例如,神经递质、生物介质、抗原和抗体等物质在调节机体正常功能及参与疾病演变过程中均有着重要的作用。而新的体液因子又不断被发现,因此,研究这些体液因子的作用机制和生物功能是机能探索实验的重要课题。

#### (4) 对某种药物的体内过程或作用机制进行研究

基础医学研究的根本目的是预防和治疗疾病,保证人类生存的质量和身体的健康。研究药物的作用机制是预防和治疗疾病的重要手段。利用我们在生理学、病理生理学及药理学中学习到的知识和实验技能,对新型药物和改良型药物进行药物的机理和可靠性研究,是机能实验选题的重要内容。因此,从不同的角度研究一种药物的作用是探索实验永不衰竭的课题。

## 二、实验设计

在机能实验过程中,我们遇到的最困难的问题是怎样进行设计性实验?怎样从不确定的偶然事件中发现隐藏在事物现象背后的发生发展规律?怎样从复杂纷繁的事物中搜集数据,怎样将这些资料进行统计分析?怎样通过合理的统计分析得出正确的结论?要解决这些问题必须懂得科研设计。科研设计(research design)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、实验动物学和实验方法学等。就科研程序来说,在确定了题目之后就要进行实验设计,而完善的实验设计包括医学专业设计、统计学设计、实验方法设计和实验动物设计四部分。

**1. 专业设计** 专业设计是从专业角度科学地安排科研,考虑处理因素的来源和施加方法,选择受试对象,确定效应的指标和影响因素,设计记录表格等,其设计水平能反映科研结果的有用性和创造性的大小。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。

**2. 统计设计** 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,怎样从不确定的偶然事件中发现隐藏在事物现象背后的发生或发展的规律性,是科研工作中要考虑的重要问题。医学统计学是解决这个问题的重要方法和手段,应用统计学方法设计实验,通过随机分组,重复观察,描述或推论这

种不确定偶然现象，发现不确定偶然现象背后隐藏事件的发生发展的客观规律。

在统计设计中，要确定资料的类型和资料分析的统计方法、改善实验的有效性和控制实验误差，这是保证专业设计的合理性和实验结论可信性的关键因素。统计学设计应考虑设立何种对照？选择多少受试对象？怎样做到随机化分组？计划搜集哪些资料？如何对原始资料进行整理和分析？保证专业设计的合理性和科研结果的可信性。严格地说，没有统计学设计，搜集的资料就失去了统计分析的意义，这一步工作做的如何，对于能否在较短的时间，消耗较少的人力和物力，获得较多的有用资料以及对科研成果的鉴定和推广都起着举足轻重的作用。为了确保医学科研的科学性，必须做好统计设计。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。

总之，正确掌握和运用统计方法是设计性实验中最重要基本功之一，从设计性实验的选题、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，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的知识，是不可能圆满完成实验设计的。

**3.实验动物设计** 实验动物设计应该考虑选择什么样的动物，实验动物种类繁多，实验动物的选择是一个内容复杂的问题，不同的实验有不同的目的和要求，而各种动物又有各自的生物学特性和解剖生理特征，因而不能随便选一种动物来进行某项实验研究。事实上，每项科学实验都有其最适合的实验动物，如果选择得当，则可节约人力、物力和时间，以最小的代价最大限度地获得可靠的实验结果。否则，不仅会造成不必要的浪费，而且会影响实验结果的判断。因此，掌握正确的实验动物选择方法非常重要。

### 三、实验设计实施

设计性实验可以对研究对象施加某种处理因素，观察或观测由此引起的有机体的结构、功能、生化和疾病过程的变化，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的原因和发展的规律性。因此，设计性实验的特点是研究者能人为的设置处理因素。在实验实施时，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，实验各组之间具有可比性，可以客观地评价处理因素的作用；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，达到高效和精确的目的。

### 四、资料分析

在机能实验中，实验或观察的结果往往是不确定的偶然现象，怎样从不确定的偶然事件中发现隐藏在资料中的客观规律，是资料分析工作中要考虑的重要问题。再资料分析工作中要注意以下的问题，第一，以正确的方式搜集资料：正确应用统计方法的基本前提是正确搜集资料，这种“正确方式”须在科研统计设计时确定下来。第二，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。第三，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做出统计推断，即得出由样本推断总体的结论。

### 五、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,又要不要拘束于传统观念,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,轻易外延推断往往会导致错误的结论。总结归纳的基本形式是撰写设计性实验的研究论文。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,应注意搜集论文被他人引用的情况与评价。

总之,在设计性实验设计前要认真检索和阅读有关文献,全面了解本专业的过去与现在,确立正确的研究方向,选定一个好的研究题目,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,就能够做出一个优秀的设计方案,并得到满意可信的科学结论。设计实验与其他科学研究的过程相类似,都包括上述几个环节,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性。

## 第二章 机能实验专业设计

探索性实验专业设计要考虑三个基本要素,即处理因素(study factor)、受试对象(study object)和实验效应(experimental effect)。例如,在观察胰岛素对糖尿病患者的疗效研究中,胰岛素为处理因素,糖尿病患者为受试对象,血糖变化为实验效应。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。

### 第一节 处理因素

#### 一、概述

实验研究的特点之一是研究者人为的设置处理因素(study factor)。一般来说,处理因素是由外部施加于受试对象的,它可以是生物的、化学的和物理的,另外,受试对象本身的某些特征(如性别、年龄和遗传特性等)也可作为处理因素。如研究性别和年龄与糖尿病之间的关系,性别和年龄就是处理因素。处理因素为单个时,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,以药物作为处理因素,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:这是机能实验中最常见的实验类型,如观察有机磷酸酯类农药中毒对动物的影响,就属于这类实验。②单因素多水平:这是单因素多组群的实验,如观察比较同一刺激的不同强度对骨骼肌收缩的反应,比较某药物不同剂量的疗效观察,便属于这类实验。③多因素单水平:如比较不同的实验方法,

在某一实验中的作用通常需要采用这类实验。④多因素多水平：例如研究细胞培养的最佳反应条件，探索联合用药的方案等，便是常用的多因素多水平实验类型。

## 二、确定处理因素要注意以下几个问题

**1 抓住实验的主要因素：**任何实验效应往往是多种因素相互作用的结果，我们不可能在一次或几次研究中，得到所有相关因素的作用。例如，寻找神经干细胞的最佳培养方法，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 $10^5=100000$ 次实验，显然是办不到的。因此，一个最优的实验设计，首先要抓住主要的几个处理因素。

**2 确定处理因素的强度：**处理因素的强度是因素的量的大小，如电刺激的强度、药物的剂量等。处理的强度应适当。同一因素有时可以设置几个不同的强度，如一实验药设几个剂量(高、中、低)，但处理因素的水平也不要过多。

**3 确定处理因素与非处理因素：**选择什么作为处理因素，取决于研究目的。例如研究刺激强度与骨骼肌反应的关系，刺激是处理因素；研究利多卡因对某心律失常的疗效，利多卡因是处理因素。除处理因素作用外，同时还有若干其它因素也会影响这些效应和结果。因此，在确定处理因素的同时，还必须明确哪些是非处理因素。对于非处理因素，应当作为误差来源严格加以控制。例如，要进行异博定对大鼠室性心动过速影响的研究，异博定是处理因素，但是血钾的含量、PH值、温度等均可影响异博定的活性。因此，实验组与对照组除处理因素以外，非处理因素都应当相同。非处理因素(干扰因素)会影响实验结果，应加以控制，如离体实验时的恒温、恒压，病人的病种、病情(轻重)、病程(急慢)、年龄、性别等。

**4 处理因素标准化：**处理因素标准化就是指处理因素在整个实验的过程中始终保持一致，包括处理因素的强度、频率、持续时间及施加方法等。通过查阅文献和预备实验找出各自的最适条件，然后制订有关措施。一旦进入正式实验，不允许轻易改变。如处理因素是药物，应当在研究的全过程中使用同一批药物，以防止药物批次间的差异干扰实验效应。

## 三、处理因素确定中常见的问题

在处理因素中常见的问题是只进行单一处理因素的科研验证。在机能实验中，人们习惯于对单一处理因素的定性研究，如某药物作用机理的实验观察，结论只局限于此药有什么作用。为了提高科研实验的层次，应该进行多因素研究和定量研究，如某种药物的不同剂量对不同机能指标的影响。这种设计的对照只有一个，实验组根据处理因素的多少和剂量水平分成几个组，其结果不只局限于此药对机体是否有影响，还可得出某药的剂量多少、疗效的高低等结论。

另外，处理因素作用于受试对象而产生的反应，是以具体的指标来表示的。因此，指标的选定也是科研实验的一个要素。对于一些主观性指标，由于其易受心理状态、暗示程度和感官误差的影响，波动比较大，在机能实验中应尽量少用。

## 第二节 受试对象

### 一、概述

受试对象的选择在探索性实验中十分重要。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，能够为实验成功创造有利条件。遗传学家摩尔根(Morgan)用只有四对染色体，基因结构简单为特点的果蝇，作为动物模型，发

现了基因的连锁与互换规律的遗传特性，在遗传学上取得了举世瞩目的成就，这一点在实验动物学上称为“易化原则”。有时同一药物对不同动物同一器官系统的效应可不同，如吗啡对人、猴、犬、兔的中枢神经系统产生抑制效应，而对虎、猫、小鼠则引起兴奋效应。因此，正确选择受试对象对于实验的成功具有重要的意义。

## 二、受试对象的确定

受试对象的确定取决于研究目的，探索性实验的受试对象绝大多数是动物。根据具体情况可采用整体作为受试对象，在体内(in vivo)进行实验；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)进行实验；还可采用先体内后体外的方式进行实验，这类实验属半体内(ex vivo)实验。一般来说，整体实验反映动物的实际情况，实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实验或者半体外实验。

## 三、受试对象的条件

在机能实验中，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：第一，对处理因素敏感。第二，反应比较稳定。例如，小鼠气管及支气管腺不发达，只有喉部有气管腺，支气管以下无气管腺，不适用于研究体温变化方面的问题，不适宜做慢性支气管炎模型及祛痰平喘药的疗效实验。存在以下情况之一者，不宜作为一般探索性实验的受试对象：(1)存在影响反应结果的病症。(2)机体反应性和致病因素与一般情况不同。

## 四、受试对象的纯化(Purification)

受试对象的动物模型表现具有典型性。非典型的特殊的动物模型不宜作为受试对象，因为特殊动物模型提示动物机体与一般动物存在差异。受试对象的动物模型一般应是标准的，因为，如果动物模型不标准，死亡率很高，不易反映处理因素的疗效，难以区分不同处理因素的优劣。

## 五、受试对象的依从性

受试对象的依从性是指动物按预定计划接受处理因素的合作程度。绝对的依从只有在麻醉动物实验才能实现。在实验过程中由于工作的疏漏等多方面的原因可能出现忘记给药、中途退出实验或出现换组。其次，由于动物在饲养过程中出现病情急剧恶化，必须中途退出实验。这些非依从性的表现必然干扰实验计划的实施。因此，在实验中必须认真、细致，才能提高依从性。

## 六、受试对象选择中常见的问题

机能实验学科研常用的受试对象是动物，在动物实验中，为了减少经费开支，研究者愿意选择价格较低的小白鼠和大白鼠，但这时要考虑动物的生理解剖特点是否适于做此种实验。如大白鼠只在喉部有气管腺，故不宜做支气管炎模型或祛痰平喘药实验；猫减压神经行走于迷走或交感干中，不适于观察减压神经对心脏等的作用；狗不易形成动脉粥样硬化病变，不宜做动脉粥样硬化实验研究。若进行肿瘤的研究，必须了解哪种动物什么肿瘤发生率高？有无自发性肿瘤？另外，考虑到动物的种系差异，有时应重复做几种动物，以保证由动物实验类推到人的安全性。因此，合理地选择受试对象，才能得出正确的科研结论。



### 第三节 实验效应

实验效应是处理因素作用于受试对象所引起的反应，探索性实验成功与否，往往通过具体的实验指标来反映，因此，效应指标的正确选定是非常重要的。

实验指标是指在实验观察中用于反映研究对象中某些可被检测仪器或研究者感知的特征或现象标志，又称观测指标。实验指标选择的基本条件：

#### 1.关联性

选用的指标必须与所研究的题目具有本质性联系，且能确切反映处理因素的效应。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： $L/(min \cdot m^2)$ ]作为指标，否则，不可能得出正确的判断。对于血糖测定，若用班氏实验，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。

#### 2.客观性

实验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。应避免受主观因素干扰造成较大误差，尽可能选用具体数字或图形表达的客观指标，如心电图、脑电图、血压、呼吸等描记和化验检查的客观指标等，而用疼痛、饥饿、疲倦、全身不适、咳嗽等感觉性指标时，由于个体差异的原因，其客观性、准确性则较差，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。

#### 3.特异性

特异性是指实验指标能反映某一特定的现象而不至于与其它现象相混淆。如研究高血压病应该使用动脉压，尤其是舒张压作为实验指标；急性肾炎以尿液和肾功能改变作为实验指标，比用血压作为指标好；胆碱酯酶活性测定对于有机磷毒理学的研究是特异指标。特异性低的指标容易造成“假阳性”，使本不应出现的现象出现。因此，特异性高的指标最能揭示事物的本质，且不易受其它因素的干扰，而非特异的指标极易受其它因素的干扰，使效应结果不准确。

#### 4.灵敏度

实验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处理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，以感应量为0.5g的药物天平即可，若采用感应量为0.1mg的分析天平，显然是不必要的。灵敏度高的指标能使处理因素引起的微小效应显示出来，灵敏度低的指标可使本应出现的变化不出现，造成“假阴性”结果。

#### 5.精确度

精确度包括实验指标的精密性与准确度双重含义。精密性是各次测定值集中的程度，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时，观察值与其均值的接近程度，其差值属于随机误差。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，主要受系统误差的影响。

## 第三章 机能实验统计设计

统计设计(statistical design)是影响科学实验成功与否的关键环节，是提高机能实验教学质量的保证。例如，某科研工作者欲研究神经生长因子是否对胚胎

干细胞的分化具有作用，此时首先应根据研究目的提出一个合理的研究假设：神经生长因子对胚胎干细胞的分化具有作用。为了验证假设，必须在研究的全过程中采取有力的措施，控制各种干扰因素，保证该实验取得较为客观的结果。这是因为影响胚胎干细胞分化的因素，除了研究假设中规定的实验因素(两种神经生长因子)以外，还可能会受到许多非实验因素的影响，如实验动物的来源、实验方法的选择、实验室的环境等等。所以，在进行实验之前，应经过周密的思考，做出一个合理的设计，其中包括采取何种措施控制非实验因素对实验结果的干扰、选用何种统计方法分析实验数据等。此外，还应注意实验统计的各种分析方法都要求资料遵循随机化原则，并要求数据可靠，具有重复性和可比性。只有这样，才能使实验因素(神经生长因子)的效应充分显示出来，保证实验取得预期的结果，回答研究假设提出的问题。一项良好的实验设计能为科研工作节省大量资金和人力、物力资源，达到有效收集数据，少走弯路的目的，取得较为满意的结果。同时，我们在阅读文献和借鉴他人的经验时，也应当运用医学科研设计的基本原理，分析和检验文中结论的正确性。

因此，医学学生掌握实验设计的基本原理和相应的统计分析方法，是我国当今社会对医学各专业学生的要求。从统计研究设计的角度出发，其基本原理是一致的，故本章按实验研究设计的基本原理和步骤为主线，介绍实验设计的基本概念，原理、步骤和方法。

## 第一节 统计学的基本概念

### 1、总体和样本

**总体：**医学实验研究中具有相同性质的全部个体称为总体(population)。医学实验研究的对象可以是一个无限大的抽象的群体，即观察单位数无限，称之为无限总体(infinite population)；也可以是一个有限的含有足够大个体的群体，称之为有限总体(finite population)。譬如，人的血压、呼吸和血液中红细胞数的正常值等等。

**样本：**从总体中抽取的部分个体称为样本(sample)，在一个样本里含有的个体数目称为样本含量(sample size)。例如当我们要把中国成年男性的血压平均值作为总体进行研究时，我们不可能把中国所有健康成年人的血压都进行检查，而只能按照统计学的方法，从总体中抽取一部分人来进行血压检查。然后，根据测量结果来推断中国成年人正常血压的平均值。如何正确地根据样本观察到的情况来推测总体，是医学统计学要解决的重要问题之一。

### 2、抽样

从总体中抽取的部分样本的过程称为抽样(sampling)。从总体中抽取样本应具有代表性(representative)、随机性(randomization)、可靠性(reliability)和可比性(comparable)。

#### (1) 代表性

代表性(representative)要求样本中的每一个个体必须符合总体的规定。这就要求对总体有一个明确的规定。这种规定是根据研究目的而具体确定的。例如，要用标准动物制备糖尿病动物模型，则我们可以规定实验动物来源于某一品系、无急性病、无血液消耗性疾病等，所抽取的样本中的每一个实验动物都必须符合上述规定。

#### (2) 随机性

随机性(randomization)就是使总体中的每一个体都有同等的机会被抽取作为样本。随机化抽样绝不等于随意抽样。例如，将40只小鼠分为两组。如果闭上眼

睛, 随意抓取20只作为第一组, 留下的作为第二组。表面上看来是随机的, 但实际上是不随机的。因为体壮、活泼性强的小鼠难于抓到, 故大部分留在第二组, 两组的活泼性是不同的。为了保证抽样的随机性, 可选用抽签法(sortition)、机械抽样法(systematic sampling)、分层抽样法(stratified sampling)或利用随机数字表(random digits table)抽样等方法。

### (3) 可靠性

可靠性(reliability)是指由抽样得到的实验结果要具有可重复性, 由样本的结果所推测总体的结论有较大的可信度。由于个体之间存在差异, 如果根据少数几例样本就下结论, 则这种结论可靠性差, 体现不了规律性, 因为只有观察一定数量的个体方能体现出其客观规律性。每个样本的含量越多, 可靠性会越大, 但是样本量过大, 会造成人力、物力和时间的浪费, 所以应以“足够”为准。究竟需要多少样本例数, 这与所观察的指标的变异程度有关, 可通过样本例数估计得到。

### (4) 可比性

可比性(comparable)是指实验组与对照组之间, 除处理因素不同外, 其它可能影响实验结果的因素要求基本相同, 即要求比较的样本之间应具有可比性。在基础医学研究的动物实验中可以采用空白对照, 实验组给以某种药物作为处理因素, 对照组使用蒸馏水作为对照。另外还可以采用阳性对照, 通常是采用已知的方法和药物, 用获得确定的实验结果作为对照。

## 3、变量和变量类型

在确定样本开始研究之后, 研究者则应对每个观察单位的特征进行测量和观察, 这种特征称为变量(variable)。变量的测得值称为变量值(value of variable)或观察值(observed value) 亦称为资料。比如某胆管癌的调查资料中, 性别、年龄、发病部位等都是观察单位的特征, 也就是变量。

表14-1-1 某胆管癌调查资料

编号	性别	年龄(岁)	部位	分化程度	分期	肝转移	PCNA 指数	生存时间(月)
1	男	61	上	低分化	I	阳性	52	14
2	女	58	中	高分化	II	阴性	89	20
3	女	63	上	高分化	IV	阴性	93	19
4	女	71	下	中分化	III	阳性	78	5
...	...	...	...	...	...	...	...	...

观察上表不难发现, 如性别、部位、分化程度等指标表现为定性的特征, 而年龄、PCNA 指数、生存时间表现为数值、定量的特征, 按资料(变量值)是定量的还是定性的, 可将它们分为以下类型:

1)定量资料(quantitative data) 亦称计量资料、数值变量(numerical variable), 其变量值是定量的, 表现为数值大小, 一般是有单位的, 如岁、月等。

2)分类资料(categorical data) 亦称分类变量(categorical variable)或定性资料(qualitative data), 其观察值是按照某种属性, 将观察单位分为互不相容的类别, 是定性的, 如性别分成男和女, 按部位分成了上、中、下; 此时研究者所关注的是个体属于那个类别, 或者说在一个集合中某个类别有多少个观察单位数, 因此分类资料亦称计数资料。

## 4、误差

在统计学中, 观测值(observed value)与真实值(true value)之差, 样本指标(sample index)与总体指标(population index)之差统称为误差(error)。

## 5、概率

概率(Probability)是描写某一事件发生的可能性大小的一个量度。用  $A$  表示某一事件,  $P$  表示该事件可能发生的概率, 可记为  $P(A)$ 。大小在0到1之间, 即  $0 \leq P \leq 1$ , 常用小数或百分数表示。 $P=1$ , 表示为必然事件;  $P=0$ , 表示不可能事件, 它们是确定性的, 可看成随机事件的特例, 大部分事件的概率都是  $0 \leq P \leq 1$ , 也就是说具有一定的不确定性, 对于某一次观察来讲可能发生, 也可能不发生。其中有一类事件它们发生的可能性很小, 在单次观察或单次实验中可以认为不发生的事件, 称为小概率事件, 习惯上将小概率事件发生的界值定为  $P \leq 0.05$  或  $P \leq 0.01$ 。在统计推断中, 如果小概率事件出现了, 就表明统计数据之间的差别具有统计学意义。但是一定要注意的是小概率事件不等于不可能事件, 小概率事件只是统计学上认为不可能发生, 是可能犯错误的, 而不可能事件是肯定不发生。比如买彩票, 中奖概率是1/10万, 是一个小概率事件, 此时可以认为不会中奖, 但还是有人在中奖。

## 6、平均数和标准差

**平均数:** 平均数(average)是描述一组观察值集中位置或平均水平的统计指标, 它常作为一组数据的代表值, 用于分析观察值。平均数包括多种指标, 常用的有均数、几何均数、中位数等。其中最常见的平均数的直接算法是将所有的观察值  $X_1, X_2, \dots, X_n$  直接相加再除以观察例数。

**变异程度:** 变异程度是用来衡量一批数据的变异大小的指标, 为了说明这一问题, 看下面的例子。

例如对甲乙两名高血压患者连续观察5天, 测得的收缩压分别为:

甲患者(mmHg) 162 145 178 142 186 ( $X_{\text{甲}}=162.6$ )

乙患者(mmHg) 164 160 163 159 166 ( $X_{\text{乙}}=162.4$ )

从列出的数据可以看出, 两人收缩压的均数几乎没有差别, 但是甲患者的血压波动比较大, 而乙患者的血压波动相对比较稳定。由此可以看出一组观察值变量, 虽然平均数相等, 但是, 离散或变异的情况不相等。

怎样衡量均数变异程度大小? 在统计学中是通过计算各观察值偏离平均数的平均差距来表达。为了避免正负抵消, 就将每个观察值与均数之差的绝对值相加,

然后取平均, 称作平均偏差(mean difference), 它可以表示为:

**离均差平方和(sum of squares of mean deviation):** 平均偏差是一个很直观的变异量度, 但由于用了绝对值, 在数学上不便于继续处理, 使它在应用上受到很大的限制, 实际中很少使用。为了克服平均偏差使用绝对值不便于进一步运算的缺点, 可以不通过取绝对值, 而是通过取平方来避免正负抵消, 即使用离均差平方和, 其公式为:

离均差平方和是描述了每个观察值相对于集中位置  $X$  的变异程度。

方差(variance): 如果将离均差平方和再取平均, 其结果称作均方差, 简称方差。需要注意的是, 对于样本资料, 在对离均差平方和取平均时分母用  $n-1$  代替  $n$ , 于是把分母  $n-1$  称为自由度(degree of freedom), 意思是在所有的  $n$  个离均差平方项中, 只有  $n-1$  个是独立的, 即由于样本均数  $\bar{X}$  的限制, 在所有的离均差平方项中只要有  $n-1$  个已知, 最后一个离均差平方项便自动确定。

式中  $s^2$  为样本方差, 方差越大说明数据的变异越大。方差有时也表示为 MS。

标准差(standard deviation): 方差是用取平方后的单位来表示的, 如果原始数据用毫米汞柱表示, 则方差就是毫米汞柱的平方。在统计分析中为了方便, 通常将方差取平方根, 还原成与原始观察值单位相同的变异量度, 计算公式如下:  $s$  称为标准差。显然, 一组观察值的标准差越大说明每一个数据与它们的平均数相减所得差数越大, 各观察值偏离平均数的平均差距越大, 其变异程度越大。

## 第二节 统计设计的基本思想

### 一、实验统计设计的特点

机能实验的特点是研究者可以根据研究目的设置不同的处理因素, 主动加以干预措施, 控制非实验因素的干扰, 观察分析总结, 回答研究假设所提出的问题。

机能实验统计设计的特点是抽样研究。什么是抽样研究? 就是人们对从总体中随机抽取的样本进行研究, 由样本推断总体, 统计学上称之为统计推断 (statistical inference)。机能实验统计设计的主要任务是通过事物的表面现象去揭示隐藏在事物背后的客观规律, 从表现为偶然性的大量数据中, 分析出其中带有必然性的规律。如果仅仅通过观察几个例子来推论, 不可能得出一般规律, 因为观察的例数太少, 容易受到个别偶然因素的影响。从理论上讲, 要了解事物必然性的客观规律, 就要对整个群体中的样本进行研究, 但是, 在实际工作中, 由于受各种条件的限制, 不可能对整个群体中的样本进行研究, 而是通过观察一定数量的样本, 获得带有规律性的结果。因此, 我们通常是从总体中随机抽取一定数量的样本 (见图14-2-1), 对抽取的样本进行统计描述, 然后用样本推断总体, 最后从统计学的角度判断和回答研究假设所提出的问题, 这就是抽样研究。

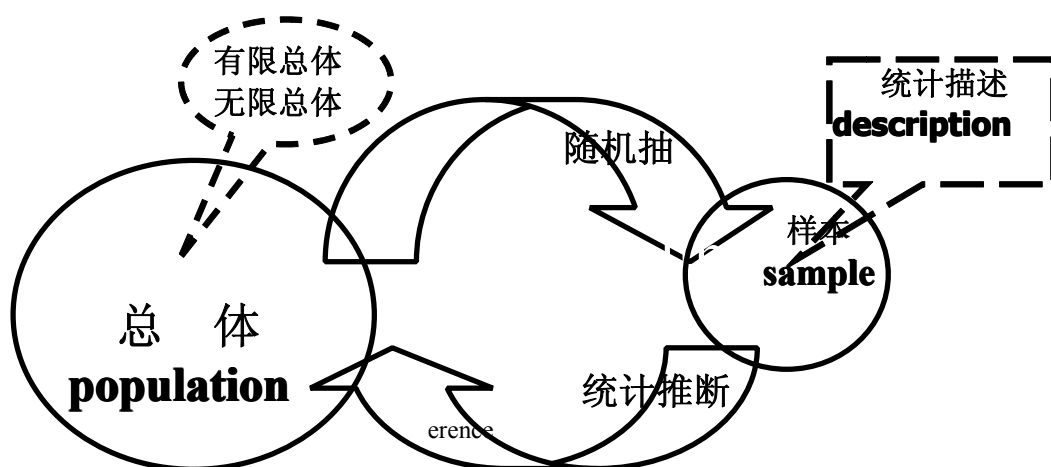


图14-2-1 实验统计分析宏观示意图

抽样研究解决的主要问题是从小数中得到的统计数据之间的差异问题。当我们从总体中抽样时,即使消除了系统误差,并把随机测量误差控制在允许范围内,但是,仍然可能存在误差,这种误差产生的原因一方面可能是由于个体之间存在变异;另一方面也可能是由于抽样时只能抽取总体中的一部分作为样本,样本含量不足,代表性不够,导致样本的数据构成的统计指标(如均数)与总体指标之间有误差。因此,抽样研究要揭示在抽样过程中总体与样本之间是否存在差异?这种差异有多大?这种差异是否具有统计学意义?从大量数据的偶然性中寻找隐藏在数据内部的必然规律性。我们要了解数据内部的必然规律性,首先应该了解正态分布。

## 二、正态分布

医学实验研究中,有很多来自总体的统计数据具有正态分布的特征,例如来自某一年龄段的实验动物的体重、呼吸、血压的平均值和红细胞数等等。正态分布是统计学中最为重要的一种分布。原因有如下几点:第一,在实验科学中,许多现象的频数分布都是服从正态分布或近似正态分布,可以直接应用正态分布的规律对其加以描述,例如动物血压的正常值范围、重量、心率等。第二,根据数理统计中的“中心极限定理”的原理,对于计量资料而言,当样本例数足够大(譬如大于50或大于100)时,无论总体属于何种类型的分布,从总体作随机抽样时得到的样本均数的分布都服从正态分布,这是我们学习许多假设检验方法的理论基础。第三,统计学中的其它一些重要分布,譬如t分布、F分布、 $\chi^2$ 分布等,都是在正态分布的基础上推导出来的;某些分布譬如t分布、二项分布、泊松分布等的极限形式又成了正态分布,在一定条件下,都可按正态分布处理。因此,我们要知道正态分布是最重要的一种分布,是用于了解一组变量值分布的指标,是一种频数分布。下面我们用一个例子来说明正态分布。

例如:对某市某年120名5岁女孩的身高(cm)资料进行统计分析,首先我们要进行统计描述,然后编制频数表,利用频数表描绘正态分布。

表14-2-1 某年某市120名5岁女孩身高(cm)

105.5	118.6	110.5	104.2	110.9	107.9	108.1	99.1	104.8	116.5
110.4	105.7	118.2	117.0	112.3	116.5	113.2	107.9	104.8	109.6
109.1	108.1	109.4	118.2	103.9	116.0	110.1	99.6	109.3	107.5
108.6	100.6	108.8	103.8	95.3	104.4	102.7	101.0	112.1	118.7
100.2	102.1	114.5	110.4	115.0	120.5	115.5	112.7	103.5	114.4
100.7	116.3	105.1	112.8	118.5	113.3	107.9	114.6	121.4	110.7
108.8	114.7	110.6	110.7	116.6	106.9	105.5	107.4	118.4	115.3
119.7	113.9	116.5	112.9	112.9	110.0	99.5	112.7	106.7	119.1
109.6	110.7	102.8	111.3	105.2	117.0	114.9	120.0	103.4	109.3
108.8	105.7	109.0	108.8	108.1	116.4	108.3	111.0	113.0	101.4
108.7	119.1	106.2	115.2	124.0	98.7	106.0	114.7	111.9	107.3
104.1	109.1	108.8	111.0	106.8	120.2	105.8	103.1	105.0	115.0

(数据引自倪宗赞主编的医学统计学)

**1.统计描述 (descriptive statistics):** 当我们获得一组统计数据时, 首先我们要对其进行统计描述, 用于统计描述的两个重要指标是集中趋势(central tendency)和离散趋势(dispersion), 也是描述正态分布的两个重要指标, 描述集中趋势的主要指标是平均数, 描述离散趋势的主要指标是方差或标准差。

均数: 将所有的观察值  $X_1, X_2, \dots, X_n$  直接相加再除以观察例数, 本例为109cm。

标准差: 本例  $s$  为5cm。

## 2.编制频数表

首先应找出该批数据中的最大值和最小值, 用最大值减最小值求出全距, 本例全距 =  $124.0 - 95.3 = 28.7(\text{cm})$ 。其次是确定组数和组距。组距可相等, 也可不等。如果数据的变动比较均匀, 可采用等距分组。等距分组的最大优点是各组的频数不受组距大小的影响, 故一般均取等距分组。然后是划分组段。确定每组的起始点(下限值)和终止点(上限值)。通常以比最小值略小的整数为第一组的起始点, 本例最小值为95.3cm, 可取95作为第一组的起始点。因组距为3, 故第一组的起始点  $95 + 3 = 98(\text{cm})$  即为第二组的起始点, 以下各组依此类推, 直至写出能包括最大值124.0的最后一个组段122~125。本例为连续型的资料, 假如有一数据恰好为113cm, 应分在第六组还是应分在第七组? 为避免两组段界限的互相包含, 除最末一个组段外, 各组段应写为上限开口型, 如表14-2-1中第一组段记为95~, 即凡身高为95至97.99..., 但不足98的都分到第一组。第一组段的下限98既是第二组段的起始点, 同时又是第一组段的上限的近似值。最后是归组, 将原始数据按照各自所属的组段归位, 并计算出频率列表如下。统计表的制作要做到主题鲜明, 层次清楚, 一张表说明一个内容; 在形式上, 统计表的标题概括说明表所表达的内容; 标目清楚, 一般横标目说明分组情况, 纵标目代表观察指标; 线条简洁, 统计表多采用三线表, 一般只有顶线、底线、纵标目下横线, 在合计的上方可以有短横线, 禁止出现竖线、斜线; 表的主体部分要填满, 禁止出现空项, 如果空项用“-”填充, 缺失用“...”填充。

**表14-2-2 某年某市120名5岁女孩身高频数分布**

组段(cm)	频数( $f$ )	频率(%)	累计频数	累计频率(%)
(1)	(2)	(3)	(4)	(5)
95~	1	0.83	1	0.83
98~	7	5.83	8	6.67
101~	10	8.33	18	15.00
104~	18	15.00	36	30.00
107~	25	20.83	61	50.83
110~	21	17.50	82	68.33
113~	15	12.50	97	80.83
116~	15	12.50	112	93.33

119~	7	5.83	119	99.17
122~125	1	0.83	120	100.00
合计	120	100.00	-	-



**3. 利用频数表描绘正态分布：**将表14-2-2的频数( $f$ )分布资料绘制成的直方图（见图14-2-2），在直角坐标系中，以横轴表示身高，纵轴表示频数，以各直方条(宽度为组距)的面积的大小代表各组段的频数的多少。我们通过对学生身高的频数( $f$ )分布进行分析，可以发现频数的分布特征，直观地看到该资料的数据具有越靠近中部，频数越高的特征，这称为集中趋势。此外，还可以直观地看到数据的变异情况，如果数据大部分集中在中部，两端的数据比较少，即绝大部分数据都与109cm差不多，则这样的数据就比较整齐，变异情况小；如果数据的频数分布在横轴上散布得较宽，在横轴上两端的数据比较多，则这样的数据就比较分散，变异情况大，数据的频数由中部向两侧散布的这种趋势或数据的参差不齐的特征就称为离散趋势或变异情况。我们可以设想，当观察人数逐渐增多，组距不断缩小时，图中的直方条将逐渐变窄，其顶端将逐渐接近于一条光滑的曲线。

正态分布(Normal distribution)是高峰位于中央(均数所在处)、两侧逐渐降低，而且左右对称、不与横轴相交的钟型光滑曲线，也叫高斯分布（见图14-2-3），它是许多统计方法的基础。

### 图14-2-3 正态分布图

由于许多现象的频数分布都服从或近似正态分布，故过去有些统计学书籍中又称之为常态分布。其实，正态分布仅仅是各种不同分布类型的一种，只不过就其重要性而言，正态分布是一种最为重要的分布而已。对于连续型的计量资料的频数分布而言，常见的分布类型还有正偏态分布与负偏态分布。正偏态分布（如图14-2-4虚线）与正态分布相比，高峰偏左，曲线右侧拖着一个较长的尾巴；负偏



态分布（如图14-2-4细实线）与正态分布相比，高峰偏右，曲线左侧拖着一个较长的尾巴。由于资料的分布类型不同，所采用的描述指标与统计分析方法也不同，故了解资料的频数分布类型和特点是必要的。

图14-2-4 正、偏态分布图

#### 4.正态曲线下各部分的面积分布规律。

正态分布的正态曲线是一条中间高、两侧完全对称并逐渐下降且永远不与横轴相交的钟型光滑曲线，故该曲线与横轴所围成图形的面积就不象计算长方形、三角形的面积那样简单了，它需要用到高等数学中微积分的知识。

统计学上常用标准差与均数相结合来反映正态分布的分布规律（见图14-2-5）。我们主要了解常用的正态分布区间：① $\mu \pm 1.645\sigma$  区间面积占总面积(或总观察例数)的90%；② $\mu \pm 1.96\sigma$  区间面积占总面积(或总观察例数)的95%，③ $\mu \pm 2.58\sigma$  区间面积占总面积(或总观察例数)的99%，④总面积(或总观察例数)100%。

图14-2-5 正态曲线下常用的面积分布规律图

正态分布的分布规律可以用标准差与均数相结合来反映。假定前面所研究的120名5岁女孩所在的总体的身高平均数为109cm，标准差为5cm，则通过计算可知正态曲线下从  $\mu - 1.96\sigma$  至  $\mu + 1.96\sigma$ ，即99.5cm~119.7cm 之间的数据所围成的面积占曲线下总面积的95%，由正态曲线的对称性可知，99.5 cm 左侧的面积与119.7cm 右侧的面积是相等的，各占总面积的2.5%；从  $\mu - 2.58\sigma$  至  $\mu + 2.58\sigma$ ，即96.1cm~121.9cm 之间的数据所围成的面积占曲线下总面积的99%，96.1cm 左侧的面积与121.9cm 右侧的面积是相等的，各占0.5%。

#### 5.正态分布与概率

概率(probability)是描述某一事件发生的可能性大小的数值，常用  $P$  表示。用  $A$  表示某一事件， $P$  表示该事件可能发生的概率，可以记为  $P(A)$ 。在一定条件下，肯定发生的事件称为必然事件，概率为1；肯定不发生的事件称为不可能事件，概率为0；可能发生也可能不发生的事件称为随机事件或偶然事件，其概率介于0于1之间。

a

b

**图14-2-6 正态曲线下面积分布示意**

在实验设计中常用的是统计概率，即对某一随机现象进行大量观察后得到的一个统计百分数  $f/N$ ，此处  $N$  为观察总数， $f$  为发生数或频数。习惯上把出现在正态分布总面积95%以外的概率记为  $P(0.05)$ ，即出现在  $\mu \pm 1.96\sigma$  区间面积以外的部分，联系图14-2-5和图14-2-6，即是95%以外左右两边的部分，各占2.5%，两侧相加为5%；把出现在正态分布总面积99%以外的概率记为  $P(0.01)$ ，即出现在  $\mu \pm 2.58\sigma$  区间面积以外的部分，联系图14-2-5和图14-2-6，即是99%以外左右两边的部分，各占0.5%，两侧相加为1%。在医学科研实验中，常把  $P \leq 0.05$  作为判断统计数据的差别有统计学意义的指标， $P \leq 0.01$  作为判断统计数据具有显著统计学差别意义的指标。

### 6. 正态分布与小概率事件

小概率事件是指某事件发生的可能性很小，在统计学中，我们习惯上将小概率事件发生的界值定为  $P \leq 0.05$  或  $P \leq 0.01$ 。在单次试验或观察中可以认为是不发生的事件，但是，小概率事件不等于不可能事件。

从数理统计的角度，正态分布的面积密度与概率密度具有相互对应的关系，正态分布的面积（5%）与  $P(0.05)$  相对应，正态分布的面积（1%）与  $P(0.01)$  相对应，并把它们当作是出现小概率事件的判定界线，进行统计推断。例如，在双侧检验时（见图14-2-5）， $\mu \pm 1.96\sigma$  对应0.025（2.5%），两侧相加为5%； $\mu \pm 2.58\sigma$  对应0.005（0.5%），两侧相加为1%；在单侧检验时， $1.96\sigma$  对应  $P(0.05)$ ，一侧为5%； $2.58\sigma$  对应  $P(0.01)$ ，一侧为1%。这时  $1.96\sigma$  和  $2.58\sigma$  就成了两个极重要的特殊数值，用于判断小概率事件发生的可能性。

同时，还应该注意单侧检验和双侧检验的判断标准是不同的，在单侧检验时  $P \leq 0.05$  有统计学意义，在双侧检验时  $P \leq 0.025$  才有统计学意义。例如，在自由度  $n=40$  的单侧  $t$  检验的统计推断中，按照  $\alpha$  为0.05的水准，如果我们得到的统计参数  $t$  值为1.685大于  $1.96\sigma$ ，分布在拒绝域，对应  $P \leq 0.05$ ，我们就认为小概率事件出现了，就表明统计数据之间的误差具有统计学意义；如果是双侧  $t$  检验，按照  $\alpha$  为0.025的水准，则  $t$  值1.685小于  $1.96\sigma$ ，对应  $P > 0.05$ （见  $t$  界值表），不能认为小概率事件出现，表明统计数据之间的误差没有统计学意义。因此，理解正态分布与概率的相互对应关系非常重要，因为它是我们理解统计推断的核心。

## 三、抽样误差和抽样分布

在抽样研究中，实验研究的结果可能受到以下三个方面因素的影响：①在实验中，采用不同的处理因素而产生的差异；②由于抽样误差而产生的差异；③由于各种系统误差的干扰而产生的误差。因此，用样本信息推断总体的特征时，要了解样本和总体之间是否有差异，这种差异是否具有统计学意义。要理解抽样研究中统计推断的原理，需要了解抽样误差与抽样分布。

1. 抽样误差：抽样误差(sampling error)是在抽样研究中，由于抽样本身的随

机性而造成的样本指标与总体指标之间的误差。抽样时之所以会产生这种误差，是因为每次抽得的样本仅包括总体中的一部分个体，而个体之间又存在着差异，故由样本计算出的统计指标与总体相应的统计指标的大小是不一致的。由于总体的统计指标通常是不知道的，故直接比较样本统计指标与总体统计指标接近程度来表示抽样误差的大小是行不通的，只能用另外的指标来度量抽样误差的大小，这一指标就是统计学中又一重要的概念——标准误。

2. 标准误：从正态分布的总体中进行随机抽样，得到的样本均数与总体均数之间存在差异，这种变异程度可以用标准误(standard error)来衡量。为了有别于一般观察值的标准差(S)，将样本均数的标准差称为标准误。

**表14-2-3 标准差和标准误的区别**

标准差	标准误
-----	-----

**计算公式：**

**含义：**                                   表示观察值的变异程度                                   反映均数抽样误差的大小

**用途：**                                   计算变异系数                                   估计总体均数可信区间  
   确定医学参考值范围                                   进行假设检验  
   计算标准误

3. 数理统计的中心极限原理：从正态分布  $N(\mu, \sigma^2)$  中，以固定  $n$  抽取样本，样本参数的分布仍服从正态分布；即使是从偏态分布总体抽样，只要  $n$  足够大，样本参数的分布也近似正态分布；样本参数的总体均数仍为  $\mu$ ，样本参数的标准差为标准误。

4. **抽样分布：**为了深刻地理解这一概念，让我们先看样本均数抽样分布的情况。根据中心极限定理可知，从正态分布的总体中随机抽样，样本均数的分布也呈正态分布，当  $n$  足够大时，则无论总体呈什么样的分布，样本均数的分布也呈正态分布或近似正态分布。假设从某一正态总体(已知  $\mu$ 、 $\sigma$ )中采用相同的随机抽样方法，抽取了很多个含量相等的样本，分别计算出各个样本的均数，然后将这些样本均数看成是新的观察值，编制成频数分布表就可以发现，这些样本均数的频数分布(如  $t$  分布)也服从正态分布。因此，所谓抽样分布是指所有样本均数的频数分布，该分布的中心即所有样本的均数，其数值等于总体均数  $\mu$  的分布。由图14-2-7可看出，样本均数的均数反映了这些样本均数的平均水平或集中趋势，标准误反映了这些样本均数的变异程度或离散趋势。显然，标准误是衡量样本均数变异程度的指标，它的本质是观察样本均数与总体均数的平均偏差：标准误越大，表明样本均数的分布越分散，用样本均数估计总体均数的误差越大；标准误越小，表明样本均数越接近总体均数，用样本均数估计总体均数的误差越小。

5.  **$t$  分布曲线下面积规律：**用  $t$  分布来说明抽样分布。 $t$  分布曲线下总面积仍为1或100%； $t$  分布曲线下面积以某一值 ( $\mu$ ) 为中心左右对称；由于  $t$  分布是一簇曲线，故  $t$  分布曲线下

$$\sigma_1 > \sigma_2, \mu_3 > \mu_4$$

图14-2-7 t分布曲线图

固定面积(如95%或99%)的界值不是一个常量,而是随自由度的大小而变化。变量值正态分布的形态的“高、矮、胖、瘦”是有很区别的,总体均数 $\mu$ 决定了正态分布在横轴上的位置,称为位置参数;标准差 $\sigma$ 决定了其“胖瘦”,称为变异度参数。当 $\sigma$ 一定后, $\mu$ 增大,曲线沿横轴向右移动;反之 $\mu$ 减小,曲线沿横轴向左移动。 $\sigma$ 越大,曲线的形状越“矮胖”,表示数据分布越分散; $\sigma$ 越小,曲线的形状越“瘦高”,表示数据分布越集中(见图14-2-7)。但是所有正态分布的面积分布规律是相同的,都可以转换成一样“高矮”、一样“胖瘦”的标准正态分布。

$t$ 检验是以原始数据符合正态分布为前提的,这是因为 $t$ 检验所依赖的 $t$ 分布是在正态分布的基础上推导出来的。根据中心极限定理,只要遵循随机抽样原则,而且样本的例数不是很小时,则无论总体的分布为何分布,样本均数的分布均呈正态分布或近似正态分布。因此,在进行假设检验时,由样本均数得出的 $t$ 值的分布呈正态分布。 $t$ 分布与小概率事件的相互关系是我们理解假设检验的核心思想。统计学中的其它一些重要分布,譬如 $t$ 分布、F分布、 $\chi^2$ 分布等,如果符合正态分布,在一定条件下,都遵循这一假设检验的核心思想。

## 四、统计推断

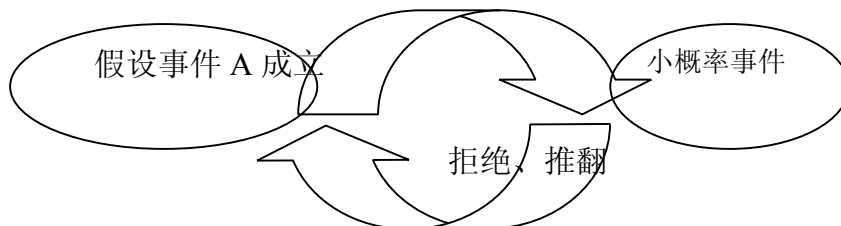
统计推断要解决从抽样中得到的统计数据之间是否具有差异?由样本信息推断总体间是否有差异?因此,要了解总体与样本之间存在的差异是否具有统计学意义就要进行假设检验。

### (一) 计量资料的假设检验

#### 1. 假设检验的基本思想

采用小概率反证法:在假设检验中,假设事件A成立,如果出现了小概率事件,就认为事件A的假设不成立,应予拒绝或否定,并接受它的对立面(见图14-2-8)。

我们用计量资料的假设检验来说明这个问题。例如,当我们在进行总体均数和样本均数的统计推断(双侧检验)时,假设总体均数与样本均数相同,如果得到的总体均数与样本均数的差异( $t$ 值)的分布大于 $\mu \pm 1.96\sigma$ 时,对应 $P$ 值小于0.025,我们就可以把它判断成为一个小概率事件,拒绝假设,接受它的对立面,认为总体均数与样本均数不相同。同时否定了它们之间的差异是由于抽样造成的,肯定了它们之间的差异是由于总体均数与样本均数本身的差异造成的,由此判断总体均数与样本均数之间的差异具有统计学意义,结合专业知识我们可以推断是由于实验中的处理因素的作用而产生的差异。



## 图14-2-8 小概率反证法

假设检验的反证法是基于概率理论的，不同于一般的反证法。在假设成立的情况下，如果出现了小概率事件，说明目前的差异不能单纯用抽样误差解释，差异可以认为是客观存在的；即使出现了小概率事件，也存在犯错误的可能，犯错误的概率一般最大为 $\alpha$ ，通常为0.05（5%）；如果没有出现小概率事件，说明该差异可能是由于抽样误差引起的，但也并不能认为假设一定成立，同样存在犯错误的可能（ $\beta$ ），其大小一般是未知的。

### 2. 计量资料的假设检验

我们通过一个实例来说明计量资料的假设检验的原理。根据大量实验资料，已知健康成年家兔血压的均数为109mmHg，某同学在机能实验中随机抽取了40只健康成年家兔，得到其血压的平均数为98mmHg，标准差为18 mmHg，能否据此认为实验组家兔的血压平均数与正常的血压平均数不同？

在这次实验中，实验组家兔的血压平均数与正常家兔的血压平均数的不同有两种可能性：一是实验组家兔的血压平均数与正常的血压平均数相同，即两总体均数相同，均数之间出现的差异是由于抽样误差造成的；二是除了抽样误差以外，更主要的原因是失血性休克造成实验组家兔的血压平均数与正常的血压平均数不同。因此，我们采用反证法的逻辑推理进行假设检验，来判断109mmHg与98mmHg之间的差别是否具有统计学意义。

在这次抽样研究中，我们先假设事件 A 成立，即实验组家兔的血压平均数与正常家兔的血压平均数相同，均数之间的差异仅仅是由于抽样误差所致。如果所做的这种假设是对的，则根据如前所述的抽样分布的理论可知，由于抽样误差的原因，每次抽样计算得到的样本均数与总体均数的差值（t）有99%的可能性落在 $-t_{0.005, 40} \sim t_{0.005, 40}$ 之间，落在该范围以外的可能性仅有1%或 $P(0.01)$ 。若这次的t值落到了99%范围之内，则表明我们所做的实验组家兔的血压平均数与正常家兔的血压平均数相同的假设是正确的。如果t值落到了1%或 $P(0.01)$ 区域外，则属于小概率事件，即仅由抽样误差的缘故导致这样大差别的可能性是很小的，由于推理过程是严密的，就只能认为事件 A 的假设不成立，应予拒绝或否定，并接受它的对立面，因此推断可能是失血性休克造成实验组大鼠的血压平均数与正常大鼠的血压平均数不同。

当然，做出这种结论是会冒一定风险的，即使出现了小概率事件，也存在犯错误的可能，犯错误的概率一般最大 $\alpha$ 为0.05（5%），即原先所做的假设是对的，现在仅根据t值落到了5%区域而否定所做的假设是否正确，因为在所做的假设成立的前提下，t值落在5%区域内的可能性虽然很小，但毕竟还是有可能的，这种风险率就是人为规定的允许犯错误的概率 $\alpha$ 用统计学术语来说即检验水准或显著性水平。

### 3. 假设检验的程序

明白了以上假设检验的原理，才能使我们对假设检验的结果有较为深刻的理解，才能对检验结果作出合理的解释。至于假设检验的过程，尽管有的简单，有的复杂，但思路大体上是一致的。

1) 建立假设：首先确定检验水准是双侧或单侧，检验水准 $\alpha$ 一般定为0.05。

然后进行假设, 假设有两种, 一种为检验假设, 如假设总体均数与样本均数相等, 因为这种假设设立的目的一般都是为了否定它, 故又可称无效假设, 即所观察到的差异仅仅是由于抽样误差所致。检验假设的符号为  $H_0$ ; 另一种为备择假设, 该假设与检验过程无关, 仅仅是预备着在  $H_0$  被拒绝时才接受的一种假设, 故称为备择假设, 备择假设的符号为  $H_1$ 。

2) 选定检验方法、计算出所需的统计量: 应根据研究设计的不同类型和统计推断的不同目的选用不同的检验方法, 计算出相应的统计量  $F$  值。为了查表的方便, 查表时计算结果一律取绝对值。

### 公式11.1

用公式11.1计算  $t$  值

本例  $t$  值 = 3.865。

3) 确定自由度, 查相应的  $t$  界值表 (见附表4)。表中最上几行表示的是单、双侧概率  $P$  的数值, 第一列为自由度, 表中的数字是自由度为不同数值时, 与某一概率  $P$  的值相对应的  $t$  的界值, 简记为  $t_{\alpha, v}$ , 让我们利用附表4查  $t$  值, 看看各符号的含义。从附表4的第一列查到  $v = 40$  这一行, 再从双侧概率  $P$  的数值这一行中找到双侧概率  $P = 0.05$  这一列, 从行、列交叉点处可查到  $t$  为 2.021, 记作  $t_{0.05, 40} = 2.021$ , 读作自由度为 40 时,  $t$  的 5% 界值等于 2.021, 意思是说从正态的总体做随机抽样  $n = 40$  (即  $v = 40 - 1$ ) 时, 可得到许多样本, 进而可计算得到许多  $t$  值, 在这些  $t$  值中,  $t \leq -2.021$  及  $t \geq 2.021$  的各占 2.5%, 见图 14-2-6 中即 95% 以外部分, 合计为 5%。

4) 做出推断性结论 因为  $t > t_{0.01, 40}$ , 所以  $P < 0.05$ , 所以按  $\alpha$  的水准, 拒绝  $H_0$ , 接受  $H_1$ 。

对推断性结论的理解: 先看假设检验的第一种情况, 即均数之间的差异仅仅是用于抽样误差所致。如果所做的这种假设是对的, 则根据如前所述的抽样分布的理论可知, 由于抽样误差的原因, 每次抽样计算得的  $t$  值有 99% 的可能性落在  $-t_{0.01, 40} \sim t_{0.01, 40}$  之间, 落在该范围以外的可能性仅有 1%。若这次的  $t$  值落到了 99% 范围之内, 则现在这种情况与所做的这种假设是一致的。在我们不知所做的这种假设究竟是对还是错的情况下, 若让我们从中做一抉择的话, 显然我们不能判断说原先所做的假设是不对的。再看假设检验的第二种情况, 如果  $t$  值落到了 1% 区域内, 即分布在  $\mu \pm 2.58\sigma$  之外, 则由抽样误差的缘故导致这样太甚至更大的差别的可能性是很小的, 属于小概率事件, 在无法得知所做的假设是对还是错的情况下, 若必须在所做的假设的对与错之间做出判断的话, 理所当然地应否定发生可能性较小的情况, 即否定这种差异是由于抽样误差所致这种情况。

在上述例子中, 因为得到的  $t$  检验值 3.865 大于  $t_{0.01, 40}$  (2.021),  $t$  检验值分布在拒绝域, 即分布在  $2.58\sigma$  (1%) 之外, 对应于  $P < 0.01$ , 因此, 我们可以判断它是一个小概率事件, 得出拒绝  $H_0$ , 接受  $H_1$  的统计结论。表明在抽样研究中, 假设的总体均数与样本均数不相等, 总体均数与样本均数之间的差异不是由于抽样误差造成的, 而是由于总体均数与样本均数之间本身的差异造成的, 差异具有

统计学意义。

## (二) 计数资料的假设检验

用于分类计数资料的假设检验方法很多，这里仅介绍 $\chi^2$ 检验用于分类计数资料的假设检验方法。在比较按二分类定性变量的两个率的时候，例如阳性率与阴性率、有效率与无效率等，由于有抽样误差存在，我们不能仅凭样本率的值的大小就下结论，而必须经过显著性检验，所用检验方法之一是 $\chi^2$ 检验。检验两个(或多个)率或构成比之间差别是否有显著性，从而推断两个(或多个)总体率或构成比是否相同；检验配对计数资料的差异是否有显著性及两种因素或特征之间有无相关的关系。

**1.  $\chi^2$ 检验的基本思想：** $\chi^2$ 检验的基本思想是检验实际频数(actual frequency)和理论频数(theoretical frequency)的差别是否由抽样误差所引起的，也就是由样本率(或样本构成比)来推断总体率(或总体构成比)。 $\chi^2$ 检验的统计量基本公式为：

公

### 式11.2

这里 R 是 Row(行)的字头，C 是 Column(列)的字头， $A_{RC}$  是位于 R 行 C 列交叉处的实际频数， $T_{RC}$  是位于 R 行 C 列交叉处的理论频数。

$n_R$  是  $A_{RC}$  所在的行的合计， $n_C$  是  $A_{RC}$  所在的列的合计， $n$  是多个样本例数的合计。 $(A_{RC} - T_{RC})^2$  反映实际频数与理论频数的差距；除以  $T_{RC}$  为的是考虑相对差距。由此可看出， $\chi^2$  值反映了实际频数与理论频数的吻合程度。 $\chi^2$  值永远是正值。 $\chi^2$  值大，说明实际频数与理论频数的差距大。但  $\chi^2$  值大于多少才认为差别有统计学意义，要与  $\chi^2$  界限值相比较才能加以判断。一般来说， $\chi^2 < \chi^2_{0.05(v)}$ ，则  $P > 0.05$ ，说明实际频数与理论频数的差别由抽样误差所引起的可能性很大，因而接受无效假设；如  $\chi^2 \geq \chi^2_{0.05(v)}$ ，则  $P < 0.05$ ，说明实际频数与理论频数的差别由抽样误差所引起的可能性很小，因而拒绝无效假设。这里  $v$  是自由度。 $\chi^2$  值的大小除与实际观察频数和理论频数的差值大小有关外，还与它们的组数有关。因此，在考虑  $\chi^2$  值的大小时，要同时考虑组数的多少，也就是要考虑自由度的大小，自由度的大小因具体问题而异。

自由度的计算公式为： $v = (\text{行数} - 1)(\text{列数} - 1)$

在同一自由度下， $\chi^2$  值越大，相应的概率  $P$  值越小； $\chi^2$  值越小，相应的概率  $P$  值大。

**2. 计数资料的假设检验：**我们用一个例子来说明。某药品检验所随机抽取 574 名成年人，研究某抗生素的耐药性。其中 179 例未曾使用该抗生素，其耐药率为 40.78%；而在 395 例曾用过该药的人群中，耐药率为 45.57%，结果见表 14-2-4，试问两种人群的耐药率是否一样？

表14-2-4 某抗生素的人群耐药性情况

用药史	不敏感	敏感	合计	耐药率 (%)
曾服该药	180	215	395	45.57
未服该药	73	106	179	40.78
合计	253	321	574	44.08

本例两个样本是从总体中随机抽取的，其目的是要由两个样本耐药率来推断两种人群对某抗生素的耐药率是否一样。

如果将表14-2-4用符号表示，则变成表14-2-5形式：

**表14-2-5 2×2表格式**

	B1	B2	合计
A1	a	b	a+b
A2	c	d	c+d
合计	a+c	b+d	n=a+b+c+d

在这个表中，a、b、c和d是四个基本的数据，组成二行二列：

a	b
c	d

对其进行 $\chi^2$ 检验，故称为四格表 $\chi^2$ 检验（ $\chi^2$  test for fourfold）。计算四格表 $\chi^2$ 统计量的公式可用式：

1) 建立检验假设并确定检验水准

$H_0$ : 两种人群对该抗生素的耐药率相同，即 $\pi_1=\pi_2$

$H_1$ : 两种人群对该抗生素的耐药率不相同，即 $\pi_1\neq\pi_2$

$\alpha=0.05$

2) 计算检验统计量

无效假设是假定两种总体耐药率相等，即曾用过该抗生素和未用过该抗生素的人群的耐药率是一样的。也就是说，耐药率与是否用过该药无关，那么在 $H_0$ 假设条件下，用公式11.2计算理论值。

曾用过该药的不敏感例数  $T_{12}=174.10$

曾用过该药的敏感例数  $T_{11}=220.90$

未曾用过该药的不敏感例数  $T_{22}=78.90$

未曾用过该药的敏感例数  $T_{21}=100.10$

计算 $\chi^2$ 值： $\chi^2=1.15$

3) 查表及统计推断

自由度  $v=(2-1)(2-1)=1$

查 $\chi^2$ 界值表（附表8），确定P值，作出统计推断。

$\chi^2_{0.05(1)}=3.84$ ， $\chi^2<\chi^2_{0.05(1)}$ ，所以， $P>0.05$ ，在 $\alpha=0.05$ 的检验水准下，接受 $H_0$ ，认为两种人群对该抗生素耐药性的差异没有统计学意义。

在上述例子中，因为得到的 $\chi^2$ 值是1.15， $\chi^2$ 检验值没有分布在拒绝域， $\chi^2<\chi^2_{0.05(v)}$ ，则 $P>0.05$ ，所以，我们不能判断它是一个小概率事件，不能得出拒绝 $H_0$ 的统计结论。说明实际频数与理论频数的差别由抽样误差所引起的可能性很大，因而接受无效假设；假设的实际频数与理论频数的差别由抽样误差所引起，差异不具有统计学意义。

### 第三节 实验设计的基本原则

一项好的实验设计必须遵循以下3个原则，以便有效地控制非处理因素，保证研究工作的顺利进行。



## 一、随机化原则

**1. 随机抽样(randomization sampling):** 随机抽样是使每一个受试对象有同等机会被抽出来, 并且被分配到各实验组中去, 以抵消处理因素对实验效应的影响, 同时保证实验数据进行统计学分析。实验设计中必须贯彻随机化, 因为在实验过程中许多非实验因素研究者在开始设计时并不完全知道, 随机误差的干扰也在所难免, 有必要采用随机化的办法抵消这些因素的影响, 要做到真正随机化就必须遵循随机化原则, 要使总体中的每一个观察单位都有同等的机会被选入到样本中来。

**2. 随机分配(randomized allocation):** 实验设计中所指的总体已不是泛指无限总体, 而是符合研究假设要求规定的纳入和排除标准的那一部分观察单位, 如根据动物的体重、年龄等所选定的实验对象(即本次实验的有限总体), 再把这些实验对象随机分入实验组和对照组中, 以增强可比性, 称为随机分配。统计学家已经证实, 当总的受试对象数固定时, 各组的受试对象数相等, 实验效率最高。

**3. 随机顺序(randomized order):** 随机顺序在实验设计中十分重要, 随机不等于随便, 必须借助随机分配工具和计算机加以实现, 研究者只有认真贯彻随机化原则, 作到真正随机才能达到预期的目的。

## 二、对照的原则

对照, 就是在实验中设立相互比较的实验组和对照组, 对各组给予相同的处理, 然后观察各组的结果。实验研究的目的是验证研究假设是否正确, 只有经过比较才能鉴别其真伪, 设立对照是比较的基础, 没有对照很难说明研究假设是否正确; 设立对照也是控制实验过程中非实验因素影响的一种有力措施。其正确的方法是将适宜纳入的研究对象随机地分入实验组和对照组, 并保证受试对象间具有可比性, 即对照组中的观察对象除了实验因素不同以外, 实验过程中的实验条件和辅助措施都应相同, 这样有利于反映出所比较的总体之间存在的真实差异, 研究者可以根据上述设立对照的原则和实验的需要, 设立多种对照, 常见的对照方式有:

### 1. 空白对照( blank)

空白对照是不对受试对象作任何处理。例如, 观察盐酸苯肼造成溶血性黄疸的作用, 实验组的动物给一定剂量盐酸苯肼, 对照组的动物给生理盐水, 处理因素完全空白。追踪观察4~5天后, 比较两组动物溶血性黄疸的发生率。

### 2. 假处理对照( dummy treatment )

实验动物经过同样的麻醉、注射、甚至进行假手术、作切开、分离等, 但不用药或不进行关键处理, 假处理所用液体的 PH 和渗透压与处理组相同, 因而可比性好。在做药物实验时, 常将动物做成一定的实验动物模型后用药, 这种对照对于评价药物对这种模型的作用是必需的。

### 3. 标准对照 ( standard control )

用公认的有效标准方法或常规方法作对照, 如在免疫组织化学实验中用已经知道的某种组织和抗体作为标准对照, 又称为阳性对照。当前在新药临床实验中常用已知有效药物作为标准对照。

### 4. 自身对照( own control )

对照和实验措施都在同一实验对象上实施, 但应注意由于实验期间实验对象可能受到诸多干扰因素的影响, 直接影响到实验后的结果, 除了少数目前尚无有效的治疗方法或者尚无理想的对照措施外, 实验前后对照不宜提倡, 而应采用特

殊的自身对照，如交叉实验设计等取代。

### 5. 相互对照( mutual control )

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，将条件基本相同的高血压病人分成三组，分别应用三种方案进行治疗，比较各组的治療结果。

## 三、重复的原则

重复(replication)是指可靠的实验应能在相同条件下重复出来(重现性)，这就要求实验要有一定的例数(重复数)。因此，重复的含义是重现性与重复数。每次实验必须有足够的样本例数，若例数太少，则难以反映出实验因素的真实效应，应当保证有足够的样本含量，否则很可能把某些偶然现象当作真实结论。因此，确定一个合适的受试对象数至关重要。那么，重复数(n)的多少由哪些因素决定呢？一般而言，以下情况例数可以较少：生物差异较小，处理因素的强度较大，实验技术较先进，高效实验设计(如拉丁方设计)，大动物，反之则要较多例数。另外，采用计量指标所需n小于计数指标；确定的显著性水准高要求n大，如P(0.01)高于P(0.05)，则P(0.01)所需n大；把握度高要求n大；单侧检验比双侧检验节省n；两个样本率或两个样本均数相差大，标准差小所需n就小。具体到某一课题，根据不同的设计类型和以上数据，应该由计算和查表求得n的多少。当样本例数较大时，即可认为样本均数的分布呈正态而不必作正态性检验；当样本例数较少时，可应用简易方法迅速做出判断，以便决定是否可以应用这些假设检验方法。重现性可用统计学中显著性检验的值来衡量其是否满意：P≤0.05 差异在统计学上有显著意义，不可重现的概率小于或等于5%，重现性好。P≤0.01 差异在统计学上有非常显著意义，不可重现的概率小于或等于1%，重现性非常好。

机能实验设计中常用的实验动物基本例数如下表。

表14-3-1 常用实验动物的基本例数表

	计量资料	计数资料
小动物(小鼠、大鼠、蛙)	≥50	≥50
中等动物(豚鼠、兔)	≥30	≥30
大动物(猫、猴、犬)	≥20	≥30

## 第四节 常用的实验设计统计方法

常用的实验设计统计方法有以下几种主要类型：完全随机设计、配伍设计、析因设计、拉丁方设计与多个样本均数间的两两比较等，统计数据的资料类型应该与相应的统计分析法相匹配。

### 一、完全随机设计

完全随机设计是把实验动物完全随机地分配到各个处理组及对照组中去，仅涉及一个处理因素，又称单因素设计或完全随机设计(completely random design)。也可分为2组或2组以上，各组例数可相等，也可不等。本法设计及处理简单易行，但只能处理一个因素，效率较低。实验方法采用随机数字表法(随机数字表见附表2)。

**例11.5.1.** 为探讨解磷定在有机磷中毒后对血压的作用，将36只兔子随机等分3组给予不同处理后，测得血压数据见表14-4-2。试问各组血压水平是否不同？

**1. 随机抽样的方法：**将雌兔16只分为2组，2组例数相同，编上动物号(体重

由小至大)。从随机数字表中取一段数字。先以随机数字奇数编为甲组，偶数编为乙组，得甲组9只，乙组7只。因需将甲组1只调为乙组，再取随机数字表中的一个大于9的数字(如76)，将该数字除以9(即甲组9只兔有均等的归入乙组的机会)得余数(4)，故将甲组的第4只归入乙组。

兔号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
随机数字	93	22	53	64	39	07	10	63	76	35	87	03	04	79	88	08
组别	甲	乙	甲	乙	甲	甲	乙	甲	乙	甲	甲	甲	乙	甲	乙	乙
组别调整																乙

如将动物分为3组，过程相似，其中将随机数字被3除，余数为1、2、0者分别归入甲、乙、丙组。

**2. 数据的收集与分析：**数据收集的基本格式如表14-4-1：

**表14-4-1 完全随机设计数据收集的基本格式**

试验序号	观测指标					
	实验因素 A:	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	.....	A <sub>n</sub>
1		x <sub>1</sub>	x <sub>1</sub>	x <sub>1</sub>	.....	x <sub>1</sub>
2		x <sub>2</sub>	x <sub>2</sub>	x <sub>2</sub>	.....	x <sub>2</sub>
3		x <sub>3</sub>	x <sub>3</sub>	x <sub>3</sub>	.....	x <sub>3</sub>
.....		...	...	...	.....	...
n		x <sub>n</sub>	x <sub>n</sub>	x <sub>n</sub>	.....	x <sub>n</sub>

注:表内画"x"处填写定量的实验数据，若因素 A 的某些水平下数据较其他水平下少，则用"."表示缺失数据。

**表14-4-2 兔子在有机磷中毒前后及中毒后解救的血压值 (mmHg)**

序号	正常对照组	有机磷中毒	有机磷中毒后应用	合计
		5分钟组	解磷定组	
1	102	42	104.4	248.4
2	108	44	109.2	261.2
3	103	42.5	103.1	248.6
4	110	48	109.8	267.8
5	100	41.2	103.5	244.7
6	104	47	105.5	256.5
7	105	39	104.9	248.9
8	107	56	107.6	270.6
9	99	45	108.8	252.8
10	115	36	110.2	261.2
11	108	53	98.6	259.6
12	102	41	106.1	249.1
n <sub>i</sub>	12	12	12	36
ΣX	1263	534.7	1271.7	3069.4

**3.完全随机设计数据的分析：**可接单因素方差分析法(F 检验)，如果只有两组

数据，用成组比较 t 检验；如果采用计数资料数据，常采用  $\chi^2$  检验法。

1) 建立检验假设

$H_0$ :  $\mu_1 = \mu_2 = \mu_3$  (各组兔子血压值总体均值相等)

$H_1$ :  $\mu_1, \mu_2, \mu_3$  不等或不全相等,  $\alpha=0.05$ 。

2) 计算统计量

利用统计学专业软件 SPSS 做 One-way ANOVA 分析结论如下:

	F	Sig.	Sum of Squares	df	Mean Square
SGPT	Between GrouPs		29824.044	2	14912.022
687.515	.000				
	Within GrouPs		715.762		33
21.690					
	Total			30539.806	35

3) 查 F 值表，确定 P 值，得出结论。

按  $F_{0.01, 2, 33}$  查附表7，方差分析用 F 界值表，得  $F_{0.01, 2, 33}=5.32 < 687.515$ ,  $P < 0.01$ ，在  $\alpha=0.05$  水准拒绝  $H_0$ ，接受  $H_1$ ，可以认为3组血压水平不同。

## 二、配伍设计

配伍设计(随机区组设计, randomized block design)是配对设计的扩大，每一配伍组的动物数目在3或以上，各配伍组的例数即为组数。本设计涉及2个处理因素，又称为双因素设计。

**例11.5.2.** 为比较不同浓度的肾上腺素对兔子血压的影响，取体重1.5kg~2.0kg 的雌性兔子20只，将年龄相同、体重相近的4只分为一组。每组的5只动物随机分别接受不同浓度的肾上腺素处理后，用生物信号采集系统观察不同浓度的肾上腺素对血压影响是否相同？

**1. 随机抽样的方法:** 先按小鼠体重分成四个配伍组，每个配伍组各有5只体重基本相同的小鼠，依次编号，第一配伍组5只小鼠编为1、2、3、4、5号，第二配伍组编为6、7、8、9、10号，余类推。将已分成4个配伍组的20只动物随机分配到甲、乙、丙、丁、戊五个组。取随机数字，1~5号鼠随机分配到各处理组的方法是：固定 A、B、C、D、E 五种处理的顺序号，在随机数字表(附表2)上任意指定一个起点，假定指定第35行、第1列数字为起点，向右查，依次抄录五个互不相同的两位数，如69、92、06、34、13，排序后得顺序号 R=4、5、1、3、2，则第一个区组内的分配结果是 D、E、A、C、B。同理，再读五个互不相同的两位数的随机数，如59、71、74、17、32，排序后得顺序号 R=3、4、5、1、2，则第二个区组内的分配结果是 C、D、E、A、B。其他配伍组小鼠按同样方法进行分配，结果见下表。

动物编号	1 2 3 4 5	6 7 8 9 10	11 12 13 14 15	16 17 18 19 20
随机数字	69 92 06 34 13	59 71 74 17 32	27 55 10 24 19	23 71 82 13 74
序 号	4 5 1 3 2	3 4 5 1 2	4 5 1 3 2	2 3 5 1 4
组 数	DEACB	CDEAB	DEACB	BCEAD

以上的随机分组结果整理得以下结果：

**表14-4-3 随机分组表**

配伍组	(1)	(2)	(3)	(4)
A 组	3	9	13	19
B 组	5	10	15	16
C 组	4	6	14	17
D 组	1	7	11	20
E 组	2	8	12	18

## 2. 数据的收集与分析：

配伍设计的数据可用双因素方差分析，其基本格式如表14-4-4：

**表14-4-4 随机区组设计数据收集的基本格式**

区组因素	观测指标（单位）					
	实验因素 A:	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	.....	A <sub>n</sub>
1		X <sub>1</sub>	X <sub>1</sub>	X <sub>1</sub>	.....	X <sub>1</sub>
2		X <sub>2</sub>	X <sub>2</sub>	X <sub>2</sub>	.....	X <sub>2</sub>
3		X <sub>3</sub>	X <sub>3</sub>	X <sub>3</sub>	.....	X <sub>3</sub>
...		...	...	...	.....	...
n		X <sub>n</sub>	X <sub>n</sub>	X <sub>n</sub>	.....	X <sub>n</sub>

注：区组因素通常是动物的窝别、体重等重要的非处理因素；当同一行上的数据重复测自同一个受试对象时，称为具有一个重复测量的单因素设计，可近似看作配伍组设计的特例。

**表14-4-5 经不同浓度的肾上腺素处理后兔子血压的变化（mmHg）**

区组号 (因素 b)	不同浓度的肾上腺素（因素 a）				合计
	0.025%	0.05%	0.1%	0.2%	
1	105.6	139.5	175.3	180.7	
2	110.2	146.7	174.6	180.5	
3	98.6	148.2	168.3	193.1	
4	120.5	150.3	164.2	194.6	
5	106.7	120.6	172.1	176.8	
Σx	541.6	705.3	854.5	925.7	3027.1

## 3. 数据的分析

### 1) 建立检验假设

① H<sub>0</sub>: 不同浓度肾上腺素处理兔子后血压的变化总体均数相等，即 = = =

H<sub>1</sub>: 、 、 、 不等或不全相等，α=0.05。

②  $H_0$ : 不同动物的血压值总体均数相等

$H_1$ : 不同动物的血压值总体均数不等或不全相等,  $\alpha=0.05$ 。

2) 计算检验统计量

利用统计学专业软件 SPSS 做 Simple Factorial ANOVA 分析结论如下:

			Sum of Squares	df	Mean
Square	F	Sig.			
SGPT	Main Effects	(Combined)	17784.199	7	2540.600
37.254	.000				
		GROUP	17407.157		3
5802.386	85.084	.000			
		CASE	377.042		4
94.260	1.382	.298			
		Model	17784.200		7
2540.600	37.254	.000			
		Residual			818.350
1268.196					
		Total			18602.550
19979.082					

3) 查 F 值表, 确定 P 值, 得出结论。

本例 按  $F_{0.01(3, 12)}=5.95 < 85.084$ ,  $P < 0.01$ , 在  $\alpha=0.01$  水准处拒绝  $H_0$ , 接受  $H_1$ , 可以认为不同浓度的肾上腺素对血压的影响水平不同或不全相同。

本例 按  $F_{0.01(4, 12)}=5.41 > 1.382$ ,  $P > 0.01$ , 在  $\alpha=0.01$  水准处不拒绝  $H_0$ , 可以认为不同动物的血压值总体均数相等。

结论: 在  $\alpha=0.01$  水准处, 可以认为不同浓度的肾上腺素对血压的影响水平不同或不全相同但不同动物的血压值总体均数相等。

### 三、析因设计

在临床研究中, 许多试验因素之间往往是相互联系、相互制约的, 有时当一种因素的质和量改变时另一种因素现象的质和量也随之改变。例如, 当同时研究两种试验因素(如两种药物)的效果, 每种因素又有两个水平(如用药和不用药)时, 某种药物的水平变化有可能使另一种药物产生的效应也随之发生变化, 此时析因设计(factorial design)是一种十分有用的设计。它不仅可检验两因素各水平之间的差异有无统计学意义, 而且可以检验两因素间的交互作用。若两因素间存在交互作用, 甲因素的水平改变时, 乙因素的效应也相应有所改变; 若无交互作用, 两者是相互独立的。析因设计的优点还在于可以节约样本含量, 若将两种药物分别进行随机对照试验, 析因设计将节约样本含量的1/2, 若用两种药物相互对比的设计, 可节约1/3的样本含量。析因设计(factorial design)中最简单的是两因素方差分析, 其基本格式为:

处理方式 A

		水平1	水平2
处理方式 B	水平1	A <sub>1</sub> B <sub>1</sub>	A <sub>2</sub> B <sub>1</sub>
	水平2	A <sub>1</sub> B <sub>2</sub>	A <sub>2</sub> B <sub>2</sub>

**例11.5.3.** 我们在机能实验中采用肾上腺素和阿托品两种药观察其对兔心率的影响,看是否两种药合用可以显著提高心率?何者为优?它们之间有无交互作用?此时我们可以采用析因设计。现取雄兔40只,随机分配进行以下4种处理方式:①单纯用肾上腺素;②单纯用阿托品;③应用阿托品和肾上腺素;④应用安慰剂作对照。设计如下表:

		肾上腺素		
		不用	用	合计
阿托品	不用	10	10	20
	用	10	10	20
合计		20	20	40

## 四、拉丁方设计

用  $r$  个拉丁字母排成  $r$  行  $r$  列的方阵,使每行、每列中每个字母都只出现一次,这样的方阵叫  $r$  阶拉丁方或  $r \times r$  拉丁方(Latin square)。按拉丁方的字母、行和列安排处理及影响因素的试验称为拉丁方试验。拉丁方设计(Latin square design)是随机单位组设计的进一步扩展,可以是考核3个处理因素,也可用于考核1或2个处理因素,而同时试图控制两种非处理因素,其中处理因素、行单位组因素、列单位组因素的水平数均为  $r$ 。在拉丁方设计中,实验单位按两种属性(或非处理因素)形成单位组,即每个实验单位既属于一个行单位组,又属于一个列单位组,因此,拉丁方设计的基本单位是一个“方格”,有  $r$  行  $r$  列( $r \geq 3$ ),有  $r \times r$  个实验单位,每个单位组内共安排  $r$  个处理,每个处理用拉丁字母表示。例如欲比较 A、B、C 三种饲料对小鼠体重的影响,为了控制小鼠窝别和个体差异的影响,可按下列  $3 \times 3$  拉丁方安排试验,其中,拉丁方的行表示三窝小鼠,拉丁方的列表示各窝小鼠按初始体重排序后序号,拉丁字母表示 A、B、C 三种饲料。

I	II	III
A	B	C
B	C	A
C	A	B

与随机单位组设计比较,拉丁方设计控制了二个非处理因素,进一步缩小了实验误差,可以得到比随机单位组设计更多一个项目的均衡,因而误差更小,效率更高。拉丁方设计的优点是可以大大减少试验次数,尤其适合于动物实验和实验室研究。缺点是要求处理数必须等于拉丁方的行(列)数,一般的试验不满足此条件,而且数据缺失会增加统计分析的难度。拉丁方设计的设计要求有:①必须是三个因素的试验,且三个因素的水平数相等;②行间、列间、处理间均无交互作用;③各行、列、处理的方差齐。

拉丁方设计的方法是:①根据处理数选定拉丁方(附表列有各种基本型式的拉丁方);②将拉丁方随机化。使用基本型拉丁方时要加以随机化,用列的重排和行的重排来实现,但交换或移动时必须整行或整列进行,不能将列或行拆散;③规定行、列、字母所代表的因素和水平。拉丁方设计,由于各因素、水平是有规

律的安排,就可用方差分析把原来包含在试验误差中的行间和列间的变异分离出来,既减少了试验误差,提高了试验的效率,又可作行间、列间均数的比较。

**例11.5.4.** 为比较五种药物对血压的影响,安排五个兔子在相同日期接受药物并观察血压数,试进行拉丁方设计。

1. 根据处理数在本书后面的附表中选定5×5基本(标准)拉丁方。

2. 将拉丁方随机化: 1) 随机排列拉丁方的行。读取4个两位数的随机数,66、05、32、88,按大小排列,则R=3、1、2、4。即先将3、1行对调,然后将2、4行对调。

2) 随机排列拉丁方的列。读取4个两位数的随机数,53、85、39、06,则R=3、4、2、1,即先将3、4列对调,然后1、2列对调。

3. 随机分配处理。例如,读取5个两位随机数10、28、81、47、20,则R=1、3、5、4、2,于是有A(甲)、B(丙)、C(戊)、D(丁)、E(乙)。将上述最后一个拉丁方的行、列和拉丁字母分别对应于试验日期、受试者和防护服的最终试验方案见下表。

**表14-4-6 拉丁方设计随机分配结果**

试验日期	受 试 者				
	I	II	III	IV	V
1	丁	戊	甲	乙	丙
2	乙	丁	丙	甲	戊
3	丙	甲	戊	戊	乙
4	戊	丙	丁	丁	甲
5	甲	乙	乙	丙	丁

照上述方法我们来看个具体的例子。

**例11.5.5.** 为研究注射甲状腺素对甲状腺体的影响,以豚鼠5个种系,每个种系各5只,分养于5个笼子,每笼内置放各种系豚鼠1只,并以甲状腺素的5个不同剂量分别予以注射。以上剂量、种系和笼子三个因素的分组如下,全部数据列于表14-4-7。试分析不同剂量甲状腺素组均数之间的差别。

**表14-4-7 兔子注射不同剂量肾上腺素后的血压变化**



重量级	笼 号					重量级 小计	重量级 均数
	I	II	III	IV	V		
甲	C65	E85	A57	B49	D79	335	67.0
乙	E82	B63	D77	C70	A46	338	67.6
丙	A73	D68	C51	E76	B52	320	64.0
丁	D92	C67	B63	A41	E68	331	66.2
戊	B81	A56	E99	D75	C66	377	75.4
笼号小计	393	339	347	311	311	1701	
笼号均数	78.6	67.8	69.4	62.2	62.2		
平方和	31303	23443	25549	20383	20041	120719	
剂量	A	B	C	D	E		
剂量小计	273	308	319	391	410		
剂量均数	54.6	61.6	63.8	78.2	82.0		

剂量：A B C D E， 处理数=5

种系：甲 乙 丙 丁 戊， 行数=5

笼子：I II III IV V， 列数=5

(1)建立假设、确定检验水准 $\alpha$

剂量间： $H_0$ ：剂量组间的总体均数相同。

(处理) $H_1$ ：剂量组间的总体均数不相等或不全相等， $\alpha=0.05$ 。

种系间： $H_0$ ：种系间的总体均数相同。

(行间) $H_1$ ：剂量组间的总体均数不相等或不全相等， $\alpha=0.05$ 。

笼子间： $H_0$ ：5个笼子组间的总体均数相同

(列间) $H_1$ ：5个笼子组间的总体均数不相等或不全相等， $\alpha=0.05$ 。

(2)计算检验统计量(F 值)

1)离差平方和的分解：根据变异来源，拉丁方设计资料总的离差平方和( $SS_{总}$ )可分解为  $SS_{处理}$ 、 $SS_{列}$  和  $SS_{行}$  及  $SS_{误差}$  四部分。且

$$SS_{总} = SS_{处理} + SS_{列} + SS_{行} + SS_{误差}$$

2)列方差分析表，计算各离差平方和  $SS$ ，自由度  $v$ ，均方  $MS$  和  $F$  值：方差分析计算参见表14-4-8，其中  $X_k$  为第  $k$  种处理小计， $X_i$  为第  $i$  行小计， $X_j$  为第  $j$  列小计， $X_{ij}$  为第  $i$  行第  $j$  列观察值，校正数  $C = (\sum X_{ij})^2 / r^2$ ， $r$  为拉丁方的阶，即行数、列数或处理数。

各部分离差平方和、自由度、均方、 $F$  值的计算与随机单位组设计的一样，本例：

$$SS_{总} = 120719 - 115736.04 = 4982.96$$

$$SS_{剂量} = (273^2 + 308^2 + 319^2 + 391^2 + 410^2) / 5 - 115736.04 = 2690.96$$

$$SS_{种系} = (335^2 + 338^2 + 320^2 + 331^2 + 377^2) / 5 - 115736.04 = 375.76$$

$$SS_{笼子} = (393^2 + 339^2 + 347^2 + 311^2 + 311^2) / 5 - 115736.04 = 908.16$$

$$SS_{误差} = 4982.96 - 2690.96 - 375.76 - 908.16 = 1008.08$$

列方差分析表，填入离差平方和并计算相应的自由度  $v$ 、均方  $MS$  和  $F$  值，得表14-4-8。

表14-4-8  $r \times r$  拉丁方设计方差分析计算表

---

变异来源	离均差平方和 <b>SS</b>	自由度	均方 <b>MS</b>
	<b>F</b>		

---

(3)查 F 值表, 确定  $P$  值, 下结论。本例  $v_{\text{处理}}=4$ ,  $v_{\text{误差}}=12$ , 查附表7 (F 界值表)得  $F_{0.05(4, 12)}=3.26$ ,  $F_{0.01(4, 12)}=5.67$ 。因  $F_{\text{处理}} > F_{0.01(3, 21)}$ , 故  $P < 0.01$ 。同理, 种系间、笼子间  $P > 0.05$ 。

**表14-4-9 资料方差分析结果表**

变异来源	SS	自由度	MS	F	P
总变异	4982.96	24			
剂量(处理)间	2690.96	4	672.74	8.01	<0.01
种系(行)间	375.76	4	93.94	1.12	>0.05
笼子(列)间	908.16	4	227.04	2.70	>0.05
误差	1008.08	12	84.01		

结论: 对剂量(处理)间差别的检验, 在 $\alpha=0.05$ 水平处, 拒绝  $H_0$ , 接受  $H_1$ ; 对种系(行)间差别、笼子(列)间差别的检验, 在 $\alpha=0.05$ 水平处, 均不拒绝  $H_0$ , 接受  $H_0$ 。故认为不同剂量甲状腺体重量均数不等。不同种系间无差别, 不同笼子间亦无差别。

## 五、多个样本均数间的两两比较

多个样本均数的比较一般分为两种情况: 一种是在设计阶段就根据研究目的或专业知识而决定的某些样本均数间的两两比较, 常见于事先有明确假设的证实性实验研究(confirmatory research), 例如多个处理组与对照组的比较, 处理后不

同时间与处理前的比较等；另一种是在研究设计阶段预先未考虑或未预料到，经数据结果的提示后，才决定的多个均数间的两两比较，常见于探索性分析(exploratory research)，这类情况下往往涉及到每两个均数的比较。多个样本均数间的比较又称多重比较(multiple comparison)，由于涉及的对组数  $k$  大于2，若采用两样本均数差别的  $t$  检验对每两个样本均数间的差别都作  $t$  检验，可作  $k! / 2!(k-2)!$  次比较，每个样本均数都重复比较了  $(k-1)$  次，由于这些比较并非都是独立的，因此会使第一类错误的概率增大，即可能将本来无差别的两个总体均数判为有差别。例如有7个样本均数，可以比较  $C^2_7 = 7! / 2! (7-2)! = 21$  次，即可有21个对比组，若每次比较的检验水准  $\alpha = 0.05$ ，则每次比较不犯第一类错误的概率为  $(1-0.05)$ ，那么21次比较均不犯第一类错误的概率为  $(1-0.05)^{21} = 0.34$ ，这时犯第一类错误的概率，也就是总的检验水准变为  $1 - (1-0.05)^{21} = 0.66$ ，远远超过0.05。故多重比较不宜采用两样本均数比较的  $t$  检验分别作两两比较，这时结合我们的实验可采用几种最常用的多重比较方法，如 LSD- $t$  检验法、Dunnett- $t'$  检验法和 SNK- $q$  检验法。

### 1. LSD— $t$ 检验

LSD- $t$  检验即最小显著差异(least significant difference) $t$  检验。其目的是比较  $k$  组中一对或几对在专业上有特殊意义的均数之间有无差别。检验统计量为  $t$  值。检验步骤为：

1) 建立假设及确定检验水准  $\alpha$

以  $\mu_i$ 、 $\mu_j (i \neq j)$  代表比较的任何两组的总体均数。

$H_0$ :  $\mu_i = \mu_j (i \neq j)$ 。

$H_1$ :  $\mu_i \neq \mu_j$ 。

$\alpha = 0.05$ 。

2) 求检验统计量( $t$  值)

$v = v_{\text{误差}}$

注：完全随机设计资料， $MS_{\text{误差}}$  即是  $MS_{\text{组内}}$ 。

3) 查  $t$  界值表，得  $P$  值，作出结论。

查  $t$  界值表时所用自由度为  $v_{\text{误差}}$ 。

用这种方法可以进一步说明例11.5.2中0.05%和0.1%的肾上腺素对血压的影响是否有差异。方法见上。

### 2. Dunnett— $t'$ 检验

此方法适用于  $k-1$  个实验组与一个对照组均数差别的多重比较。检验统计量为  $t$ 。

检验步骤为：

1) 建立假设及确定检验水准  $\alpha$ ：

以  $\mu_i$  表示第  $i$  个实验组的总体均数， $\mu_0$  为对照组的总体均数。

$H_0$ :  $\mu_i = \mu_0$ 。

$H_1$ :  $\mu_i \neq \mu_0$ 。

$\alpha = 0.05$ 。

2) 计算检验统计量( $t$  值)：

其中

这里,  $x_i$ ,  $n_i$  为第  $i$  个实验组的样本均数及样本例数,  $x_0$ ,  $n_0$  为对照组的样本均数及样本例数。

3) 查  $t$  界值表, 确定  $P$  值, 下结论。查表时所用自由度为  $v_{\text{误差}}$ 。

用这种方法可以说明例11.5.2中0.05%、0.1%和0.2%浓度的肾上腺素对血压的影响与0.025%浓度的肾上腺素对血压的影响是否有显著性的差异。方法见上。

### 3. SNK-q 检验

SNK(Student—Newman-Keuls)检验, 亦称  $q$  检验, 检验统计量为  $q$  值。当方差分析结果为多组间差异具有统计学意义时, 再用 SNK- $q$  检验进行两两之间均数差别的多重比较。在利用 SNK- $q$  检验进行多重比较之前, 须先将要比较的各组均数从小到大依次排列, 检验步骤如下。

1) 建立假设及确定检验水准  $\alpha$ :

以  $\mu_i$  表示第  $i$  个实验组的总体均数,  $\mu_j$  表示第  $j$  组的总体均数。

$H_0$ :  $\mu_i = \mu_j$ 。

$H_1$ :  $\mu_i \neq \mu_j$ 。

$\alpha = 0.05$ 。

2) 计算检验统计量( $q$  值):

其中

这里,  $x_i$ ,  $n_i$  为第  $i$  组的样本均数及样本例数,  $x_j$ ,  $n_j$  为第  $j$  组的样本均数及样本例数。

3) 查  $q$  界值表, 确定  $P$  值, 下结论。

查表时所用自由度为  $v_{\text{误差}}$ , 所用组数  $\alpha$  为两组秩次之差的绝对值加1。

用这种方法可以比较11.5.2中5%、0.05%、0.1%和0.2%浓度的肾上腺素对血压的影响两两间是否有显著性的差异。方法见上。

## 第五节 实验设计的基本步骤

实验设计的基本步骤大体相近, 故这里以实验设计为主线介绍设计中通用的基本步骤, 并说明其中各个步骤自身的特点。

### 一、建立研究假设

在研究题目确定以后, 应当根据研究目的确定研究假设。一般来说, 应表达为问题的形式, 分清研究的主要问题 (primary question) 和辅助问题 (secondary question)。实际上, 主要问题就是本次研究的研究假设。辅助问题是进一步补充说明和完善本次研究的假设。现以研究利多卡因治疗心律失常的疗效举例说明。  
主要问题: 利多卡因治疗心律失常是否有效, 治愈率实验组是否高于对照组?  
辅助问题: 不同性别的动物对利多卡因治疗心律失常的治愈率是否相同? 不同体重的动物对利多卡因治疗心律失常的治愈率是否相同?

由此可见, 主要研究问题是本次实验的研究目的, 实验的结果应对此做出确切的回答。因而在实验设计的各个环节中, 研究者都应围绕此问题进行周密的安排, 采取有效的措施, 控制各种非实验因素的干扰, 保证本次研究取得满意的结果。辅助问题是对主要研究问题的补充说明, 必须紧紧围绕主题进行安排, 不宜

将与此无关的问题作为辅助问题列入该实验的内容，作为本次实验的副产品。

## 二、明确研究范围

根据研究目的建立研究假设之后，应当抓住实验中的处理因素、实验效应和受试对象这3个基本要素。基本要素确定的正确与否，直接影响到实验的结果。首先，应当明确实验对象所组成的研究总体，例如，研究某种药物对糖尿病动物的疗效，应当说，每一个糖尿病动物都应当是实验对象，实际上，在实验中所采用的实验样本往往有限。例如：用四氧嘧啶复制糖尿病模型时，选用的实验动物数量是有限的，不可能去研究一个无限的总体。因此，必须明确规定本次实际研究的总体范围。

根据研究目的，规定排除伴有营养不良和患有感染性疾病的动物，以便集中观察该药对糖尿病的降糖效果。可见，研究者可以通过规定适宜进入实验的对象的标准，简称适宜选入标准(eligibility criteria) 确定研究总体，此标准又可分为纳入标准(inclusion criteria) 和排除标准(exclusion criteria)，并用这些标准选择适宜本次实验的对象。纳入标准主要是使纳入的对象符合研究的目的，如动物实验应按研究目的控制动物的体重、年龄等以符合要求。在确定适宜纳入标准的时候应当注意，在某些实验中应注意纳入那些对处理因素的效应反应灵敏的动物作为研究对象，避免无反应对象的干扰。如研究过敏和抗感染实验，则应选择豚鼠作为实验对象，因为豚鼠容易被抗原物质所致敏，应激反应灵敏。另外，纳入标准和排除标准应当用条文明确规定成为书面形式，让所有参与研究的实验人员都知道，以便认真执行。在总结研究成果时，也应说明本次研究的纳入和排除标准，以便他人引用时参考，并为今后的研究对比提供条件。但有时在实验初期或实验进行过程中，可能会发现某些病例不符合本次实验规定的纳入标准，根据标准应当剔除；在总结资料时往往也会发现某些实验对象并不符合纳入标准，应当及时进行剔除。

## 三、确立处理因素

实验中的处理因素是根据研究目的而施加的特定实验措施，如实验中给予的某种实验药物。为了增强可比性，实验通常设立对照措施，应当强调对照也是一种处理措施。在确定处理因素时应当注意以下几点：

### 1. 分清处理因素和非处理因素

例如，研究神经生长因子是否对胚胎干细胞的分化有效果，处理因素为神经生长因子，但是实验中培养液的PH值、温度、实验环境等非处理因素也可能对实验结果造成影响。研究者应采取各种措施，尽可能使某些非处理因素在所比较的各组中基本相同，以便充分显示处理因素的作用。

### 2. 处理因素应当标准化

应当使处理因素在整个实验过程中始终如一、保持不变，如在实验过程中实验药物的批号、剂量应当一致，实验方法应当保持恒定，否则将会影响实验结果的准确性。

## 四、明确观察指标

实验中的实验效应主要指处理因素作用于实验对象的效应，这种效应将通过实验中的观察指标显示出来，因而指标的选择也是实验设计时应当认真对待的问题，主要应当注意以下几点：

### 1. 选用客观性较强的指标

最好选用易于量化，即经过仪器测量和检验而获得的指标。在一般情况下，

应该选用定量的指标作为判断效果的观察指标更为客观可靠,如果实验难以用定量指标描述而必须选用定性指标时,应当选择易于量化的方法加以表达,如表示免疫组化阳性的指标,应当分级描述免疫组化阳性细胞的程度。

### 2. 选用灵敏度较高的指标

应当选择对处理因素反应较为灵敏的指标,使处理因素的效应能较好的显示出来。例如在盐酸苯肼造成溶血性黄疸的实验中,可以选用临床症状、体征及比色法测定血中黄疸。也可选用分光光度计检测光密度值的变化作为观察指标,前者作为观察指标不够灵敏,只有在黄疸较为明显的情况下才会出现变化,若选用分光光度计检测光密度值的变化作为观察指标,则可敏锐地反映出处理因素的效应。

### 3. 选用精确性较强的指标

指标的精确性应当包括准确度 (accuracy)和精密度 (precision) 两个方面。准确度是观察值与真实值的接近程度。精密度是在重复观察时,观察值与其平均值的接近程度。选择指标时应同时考虑指标的准确度和精密度,两者有着密切的关系,当然准确度更为重要。在实际工作中,应根据研究目的来权衡两者的重要性。总之,所确定的指标应当灵敏、准确地反映处理因素的效应,经过对观察指标的比较分析,能够较为圆满地回答研究假设中提出的问题。观察指标应当精选,与研究目的无关的指标不宜列入,否则将会冲淡主题,影响研究的真实结果。

## 五、明确资料类型

医学统计资料一般可分为计量资料 (measurement data) 和计数资料 (enumeration data)两大类。

**1. 计量资料** 计量资料是对每个观察对象的观察指标用定量方法测定该项指标的数值大小所得的资料,一般用度量衡单位表示。如身高(cm)、体重(kg)、浓度(mg/L)、脉搏(次/分)、血磷(mg%)、血红蛋白(g%)等。

**2. 计数资料** 计数资料是先将观察对象的观察指标按性质或类别进行分组,然后计数各组该观察指标的数目所得的资料。比如,用“生脉散”治疗脑血管病,其疗效可分为治愈、显效、有效、无效,那么其对应的人数将是计数资料。又如,调查某人群的血型分布,按照 A、B、AB、O 四型分组。计数所得该人群的各血型组的人数也是计数资料。

等级分组资料是指在医学实践中,有些资料具有计数资料的特性,同时又兼有半定量的性质,被称为按等级分组资料。如临床化验中,将化验结果按 -、+、++、+++等级分组,计数得每组病人数,就是等级分组资料。

根据分析的需要,计量资料、计数资料和等级分组资料可以互相转化。例如每个人的血红蛋白,原属计量资料;若按血红蛋白正常与异常分为两组,得出各组的人数,是计数资料;若按血红蛋白含量的多少分为五个等级: <6(g%)(重度贫血)、6(g%)中度贫血、9(g%)轻度贫血、12.5(g%)血红蛋白正常、>16(g%)血红蛋白增高,计算各等级人数,就是等级分组资料。

**3. 考察数据** 检查数据是否有错误。在统计描述和统计推断时,首先要找出过大或过小的数据,他们可能是数据错误。数据错误可能是由于错误而造成的误差,如抄错的数据或仪器失灵记录的数据,也可能是奇异数据(outlier)。比如利用探索性分析(explore),要求其列出五个最大和最小的数据,看是否在可能的范围内。其次,要分析原因。最后,要决定是否从分析中剔除:对错误数据,要坚决剔除;对奇异数据,要根据统计学方法决定是否保留。考察数据的分布特征:许多统计分析方法对数据的分布有一定的要求,例如配对设计的两个小样本比较

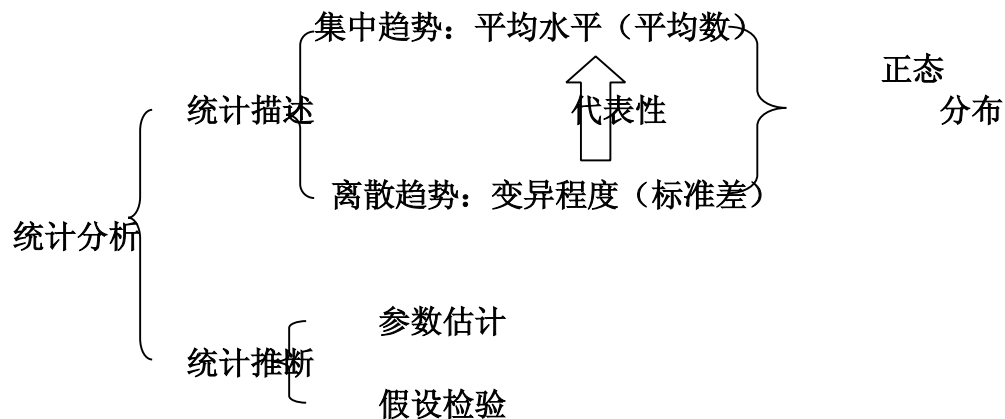
的 t 检验，就要求样本来自正态总体，要求两样本的总体方差相等。为此，要考察数据是否近似服从某种分布，如正态分布、二项分布或指数分布等。对两组或多组间数据的差异性做分析时，要考察其总体方差是否相等。观察数据的规律，通过初步观察尽可能找出内在的一些规律。例如两个变量间是否存在线性相关的关系。在机能实验设计中，计量资料的统计方法主要涉及计量资料的描述，两组样本均数实验统计方法的选择，多组样本均数实验统计方法的选择。

## 六、选用统计学方法

不同的统计资料应采用不同的统计分析方法。统计方法的选用要考虑实验的目的、实验设计类型、资料类型、实验因素与水平数、数据分布特征和样本量大小等，同时，应当根据专业知识与资料实际情况，结合统计学原则，灵活地选择统计分析方法。统计设计的核心问题不在于如何进行数据处理和计算，而在于如何选择合理的统计方法。

### 1. 计量资料统计方法的选择

在机能实验设计中，计量资料的统计方法主要涉及计量资料的描述，两组样本均数实验统计方法的选择，多组样本均数实验统计方法的选择。



1) 统计描述：统计描述是初步的统计方法，了解变量值的平均水平或集中趋势。采用的指标和方法：描述集中趋势的统计量主要有算术均数、几何均数和中位数。应用：选择何种统计量来描述其平均水平(即集中趋势)，取决于数据的分布类型，判断其是否更适用，要看计算出来的统计量是不是更靠近其集中位置的变量值，即某组数据最中间的变量值。了解变量值的离散趋势。采用的指标和方法：描述离散趋势的统计量主要有极差、四分位数间距、标准差、方差、变异系数以及标准误。应用：统计量的选择主要与资料的分布类型和研究目的有关。一般来讲，要全面描述一组数据的分布特征，既要描述其集中趋势，又要描述其离散趋势。如果我们选择均数作为描述某组数据集中趋势的统计量，那么一般可以选择标准差或方差作为描述其离散趋势的统计量；如果选择中位数作为描述集中趋势的统计量，那么一般选择四分位数间距作为描述其离散趋势的统计量；如果选择几何均数作为描述集中趋势的统计量，那么一般选择对数标准差(变量值取对数后再计算标准差)作为描述其离散趋势的统计量。

2) 统计推断：统计推断是进一步的统计方法，属于估计性统计。估计总体中个数的频数分布，用正态分布法；估计总体均数的大小，用点值估计、可信区间估计。

3) 统计比较：统计比较属于推断性统计，用于样本均数的比较。采用的指标和方法：①单组实验统计检验方法的选择：单组实验主要包括自身前后配对设

计和自身左右配对设计。自身前后配对设计是以受试对象在被试因素前的变量值作为对照值,将被试因素作用后的变量值作为实验值,然后对被试因素作用前后效应指标的差值进行统计检验。如果这些差值来自正态分布,我们就选择配对 t 检验,否则,就选择符号秩和检验(用于资料配对设计计量差值的比较和单一样本与总体中位数的比较)。自身左右配对设计时,一般选择方差分析。自身左右配对设计,是考虑到绝大多数情况下同一个体左右两侧器官或组织是对称的,它们的效应也是相近的。如四肢、眼、耳、脑、肾、肺等,对一侧施加被试因素,另一侧给予对照物,然后比较两侧的效应。② 两组样本均数实验统计检验方法的选择:计量资料的异体配对设计是将受试对象按照一定的条件(专业知识),将条件相同的个体配成对子,然后在对子内部按照随机方法,将一个分配到实验组,另一个分配到对照组,最后将其结果以配对分析的统计方法加以处理。两组样本均数完全随机设计,是将受试对象完全按随机原则分配到实验组和对照组,然后分别给予被试因素和对照物,对它们的效应进行同期平行观察,最后对实验结果做出组统计分析。其统计检验方法的选择,主要取决于资料类型、数据分布类型、方差是否齐性和样本含量大小。如果样本来自正态分布的总体,两组的总体方差相等,一般选择配对 t 检验,否则,选择符号秩和检验。③ 多组样本均数统计检验方法的选择:不同设计的多组实验统计检验方法是不完全相同的。完全随机实验统计检验方法的选择,基本上都是采用方差分析,只是变异来源有所不同而已,它是将总变异按设计的类型和需要分为两个或多个组成部分,每个部分的变异可由某因素的作用来解释,通过比较可能由某因素所致的变异与随机误差,即可了解该因素对测定结果有无影响。

**表14-5-1 计量资料统计分析方法**

统计分析的类型	分析目的	可采用的指标和方法
描述的方法(初步统计分析)	了解变量值的平均水平或集中趋势	均数、几何均数、中位数、众数
	了解变量值的变异情况或离散趋势	极差、四分位数间距、标准差、方差、变异系数
	反映某种现象的动态变化过程	发展速度、增长速度
	综合分析	平均分析法
估计的方法(进一步的统计分析,属推断性统计)	估计总体中个数的频数分布	正态分布法、百分位数法
	估计总体均数的大小	点值估计、可信区间估计
	由过去与现在的情况预测未来	动态数列分析
比较的方法(进一步的统计分析,属推断性统计)	样本均数与总体均数的比较	t 检验、u 检验
	两个样本均数的比较	t 检验、u 检验(秩和检验)
	配对样本均数的比较	配对 t 检验(秩和检验)
	两个以上样本均数的比较	F 检验(秩和检验)
	两个以上样本均数间的两两比较	q 检验(秩和检验)
关联的方法(更深层的描述)	了解客观事物或现象间相互关系的密切程度与方向	简单相关分析 多元相关分析
	了解某一变量随着其它变量的变化而变化的数量关系	简单回归分析 多元回归分析

## 2. 计数资料统计方法的选择

1) 统计描述:初步的统计方法,了解变量值的特征。采用的指标和方法:率,反映某种随机事件发生的频繁程度。



2) 统计推断: 进一步的统计方法, 由样本率推断总体率。采用的指标和方法: 由样本率推断总体率的主要方法有点值估计、期间估计和动态数列分析。

3) 统计比较: 属于推断性统计, 用于样本率和总体率的比较。采用的指标和方法: 计数资料的配对设计原理与计量资料的配对设计原理是一致的, 只不过其结果为计数数据。样本率与总体率的比较: 选用  $u$  检验、二项分布和泊松分布。两个样本率比较: 选用  $u$  检验、四格表  $\chi^2$  检验; 多个样本率或构成比的比较: 选用行  $\times$  列表的  $\chi^2$  检验; 配对样本的比较, 选用配对  $\chi^2$  检验。

**表14-5-2 计数资料统计分析方法**

统计分析的类型	分析目的	可采用的指标和方法
描述的方法 (初步统计分析)	反映某种随机事件发生的频率程度	率
	反映某一指标是另一指标的多少倍或百分之几	相对比
	反映某一事物内部各组成部分所占的比重或分布	构成比
	反映某现象的动态的变化过程	发展速度与增长速度
	反映人口的生存与死亡的动态过程	现时寿命表、定群寿命表
估计的方法 (进一步的统计分析, 属推断性统计)	由样本率推断总体率	正态分布法、百分位数法
	由过去与现在的情况估计 (预测) 未来	动态数列分析
比较的方法 (进一步的统计分析, 属推断性统计)	样本率与总体率的比较	$u$ 检验, 二项分布或泊松分布的直接概率法
	两个样本率的比较	$u$ 检验、四格表 $\chi^2$ 检验
	多个样本率或构成比的比较	行 $\times$ 列表的 $\chi^2$ 检验
	配对样本的比较	配对 $\chi^2$ 检验
	内部构成不同的两个或多个样本率的比较	标准化率的假设检验
关联的方法 (更深层的描述)	了解分类变量间有无联系	四格表或行 $\times$ 列表的 $\chi^2$ 检验
	了解分类变量间关联程度	列联系数 C

## 第六节 科研实验误差的控制

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。如果不加以控制, 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)是指实测值与其真实值之差, 统计量与参数之差。

### 一、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:** 在实验和收集资料的过程中, 由于仪器初始状态未调整到零、标准试剂未经校正、实验动物的疗效标准偏高或偏低等原因, 使观察结果造成倾向性的偏大或偏小, 这叫系统误差(system error)。系统误差影响原始资料的准确性, 必须克服, 予以校正。

**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。在收集原

始资料过程中,即使仪器初始状态及标准试剂已经校正,但是,由于各种偶然因素的影响也会造成同一对象多次测定的结果不完全一致。譬如,实验操作员操作技术不稳定,不同实验操作员之间的操作差异,电压不稳及环境温度差异等因素造成测量结果的误差。这种误差往往没有固定的倾向,有时高有时低,被称为随机测量误差(random measurement error)。对于这种误差应采取措施,尽最大可能来控制,至少应控制在一定的允许范围内。一般可以用技术培训,指定固定实验操作员,加强责任感教育及购置一定精度的稳压器、恒温装置等措施,从而达到控制的目的。

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的,个体间的变异是客观存在的,因此从同质总体中随机抽样,统计量的样本指标总是与其总体指标之间有一定的差异,这种差异用抽样误差指标标准误来描述。因此,抽样研究出现抽样误差是必然的。

**3. 抽样误差:**从总体中抽样时,由于抽样引起的总体参数与样本统计量之间的差异叫做抽样误差(sampling error),对它要用统计方法进行正确分析。例如从一个计量资料的总体中进行抽样时,从总体中得到的样本均数与总体均数之间会存在误差,这种误差是抽样误差,它可以是由于抽样时偶然的、次要的因素产生的误差造成的,也可以是由于总体均数与样本均数之间存在的真实误差造成的。一般说来,样本含量越大,则抽样误差越小,样本的观察指标与总体的该指标越接近,样本越能说明总体的规律。反之,样本含量越小,则抽样误差相应地越大,因此,我们不能仅仅凭观察指标的大小进行简单判断,而应该使用概率与数理统计方法来辨别哪些实验研究的结果是有统计学意义的,哪些可能是由抽样误差所造成的,从而得出正确的结论。常用的统计方法是假设检验(hypothesis test),也称显著性检验(significant test)。

## 二、误差的来源

**1. 生物变异(biomutation):**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(judgment error):**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(conditional error):**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(malapporportioned error):**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(gross error):**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(estimation error):**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。

### 三、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，使处理因素标准化，实验过程统一编号，统一记录表格，完整实验记录，并完善操作人员的技术水平考核，提高专业素质等都是减少实验误差的重要措施。

## 第七节 医学机能实验统计设计中常见的问题分析

在医学实验设计过程中，我们经常会遇见下面的问题，面对辛辛苦苦搜集到的一堆数据不知道怎样处理，怎样将这些资料进行统计分析？通过分析得出合理的结论？因此，周密地统计设计是科研的良好开端，可以有效地避免搜集资料时的盲目性和分析资料时的困惑感。

### 一、在实验资料的收集和整理方面存在的问题

由于缺乏完善的统计研究设计方案的指导，人们常不能正确地收集和整理观测或实验数据，在整理实验数据时，随意将那些自认为“过大或过小”的数据舍弃掉，这不仅使实验研究的真实性受到了破坏，有时还容易失去发现奇迹（如基因的突变）的机会。

### 二、设立对照方面存在的问题

设立对照的不合理，对空白对照，实验对照和标准对照的含义理解不深，相互混用。空白对照是在不给任何措施的情况下观察反应自发变化的规律；实验对照指的是采用与实验组相同操作条件的对照，以排除实验中环境的温度，溶液的PH、抗体浓度等对结果的影响，在机能实验中常用。标准对照是以公认的有效的药物或实验方法作为对照，在药物实验中常用。

### 三、实验设计重复方面存在的问题

确定的重复数不恰当。由于生物存在着个体差异，同一处理因素施加于不同的受试对象会发生不同的效应。若受试对象太少，效应的稳定性差，做统计推断的可靠性当然就差；而受试对象太多，则科研过程中会消耗大量的人力、物力和财力，同时也使搜集的资料比较粗糙，带来较大的测量误差。因此，确定一个合适的受试对象数至关重要。那么，重复数（n）的多少由哪些因素决定呢？概括

地说,采用计量指标所需  $n$  小于计数指标;确定的显著性水准高所需  $n$  大;把握度高要求  $n$  大;单侧检验比双侧检验节省  $n$ ;两个样本率或两个样本均数相差大,标准差小所需  $n$  就小。具体到某一课题,根据不同的设计类型,应由计算和查表求得  $n$  的多少。

## 四、实验分组方面存在的问题

分组的随机性较差,相当多的人习惯于将实验组分配较多的受试对象,认为以此能提高实验效率。其实,当总的受试对象数固定时,各组的受试对象数相等,实验效率最高。如何在分组时做到随机化?最好的方法则是利用随机数字表和随机排列表来分组,避免以受试对象的自愿选择和科研人员的主观意愿分组。

## 五、受试对象方面存在的问题

受试对象的选择偏颇。医学机能实验常用的受试对象是动物。在动物实验中,为了减少经费开支,研究者愿意选择价格较低的小白鼠和大白鼠,但这时要考虑动物的生理解剖特点是否适于做此种。如大白鼠只在喉部有气管腺,故不宜做支气管炎模型或祛痰平喘药实验。若作肿瘤的研究课题,必须了解哪种动物什么肿瘤发生率多?有无自发性肿瘤?另外,考虑到动物的种系差异,有时应重复做几种动物,以保证由动物实验类推到人的安全性。

## 六、观察指标方面存在的问题

主观定性的指标多。处理因素作用于受试对象而产生的反应,是以具体的指标来表示的。因此,指标的选定也是科研的一个要素。医学机能实验经常采用一些主观性指标,如腹痛、唾液、肌肉颤抖等等。统计分析时,这些指标并非不能用,但由于在观察中波动比较大,在科研中应尽量少用。对于定性结果的判定,多数人习惯用定性的(如有效和无效)和半定量(如尿蛋白-、 $\pm$ 、+、++、+++ )的指标,这样的指标在具体应用过程中容易带来主观和测量误差而且所需受试对象较多。若采用定量指标,不仅可减少受试对象数,而且能提高结果的可靠性。可选用灵敏度和准确性高的化验或测量指标,如溶血性黄疸的检测可以采用血清总胆红素作为观察指标,对肾炎模型的效用24小时尿蛋白总量等指标。

## 七、观察或测量方面存在的问题

在实验过程中由于研究者的主客观原因,经常可能产生观察或测量而造成的偏差。因而在设计时必须考虑到实验过程中可能引起偏差的某些因素并加以防范,在研究总结中也应进行分析和说明,便于读者正确估计科研结论的可靠性。在实验研究中常见的因素有:(1)污染:对照组的实验对象接受实验组的处理措施,提高了对照组的阳性率,其结果是造成实验组和对照组间差异缩小的假象。(2)干扰:实验组从实验外接受了对实验因素有效的药物或措施(非处理措施),提高了实验组的阳性率,其结果是扩大实验组和对照组之间的差异。(3)观察记录的失误:在实验观察中往往出现一些记录的失误,如记录不完整、记录失真等。如不经纠正,输入计算机后将会造成偏倚。设计时应明确规定记录的方式、时间及记录的人员,并规定经常定期检查记录的制度。

## 八、统计设计类型方面存在的问题:

在选用统计分析方法方面存在的问题:将  $t$  检验和  $\chi^2$  检验分别视为处理定量资料和定性资料的"万能工具";不考虑资料所具备的前提条件,盲目套用各种统计分析方法;将在专业上毫无关联的变量强硬地放在一起进行相关与回归分析;用直线回归取代曲线回归。

在处理多因素多指标统计资料方面存在的问题：用单因素分析方法取代多因素分析方法，用一元分析取代多元分析，这样无法揭示多因素之间的相互关系和多指标之间的内在联系，容易得出片面的、甚至歪曲事实的结论。

在应用计量资料进行显著性检验时存在的问题：首先是多组间均数的比较采用多次 t 检验，而不用方差分析。多组间均数比较时，如果资料呈正态分布，且方差呈齐性时，应该用方差分析（也叫 ANOVA 分析或 F 检验）。方差分析只能告诉我们多组间均数是否全部相等，即便  $F < F_{\alpha}$ ，也只能说明在  $\alpha$  水准上至少有两组均数差异有显著性，并不能告诉我们到底哪两组均数间有差异。要研究某两个或某几个总体均数是否相等，还要在方差分析的基础上，进一步作两两比较的 q 检验。当多个观察组与一个对照组进行均数间比较时，应作 Dunnett-t 检验。其次是配对 t 检验与完全随机 t 检验的相互误用。配对 t 检验与完全随机 t 检验主要有以下不同：①两种检验方法适用的定型分组设计不同，完全随机 t 检验用于完全随机设计的资料，配对差值的 t 检验适用于随机配对设计的资料，包括自身对照和配对对照等设计；②由于资料的分组设计不同，因此观察的内容也不同，对于完全随机设计的资料，研究者需要观察的是两样本的均数标准差，并采用完全随机 t 检验进行统计推断，以比较两样本均数有无差异。自身对照设计是在同一个体上进行观察，研究者需要观察的是每一个实验单位处理前后（或两种处理方法）两次结果差值的均数和差值的标准差，采用配对差值 t 检验，目的是推断同一样本处理前后（或两种处理方法）有无差异；③配对差值的 t 检验适用于随机配对设计的资料，而随机配对设计的资料最大限度地减少了个体差异（生物变异）对实验结果的影响，因而减少了实验误差，提高了实验精确度，效率较高，用较小样本可得出较多的信息和较大的精确度，故对自身对照设计的资料应用配对差值的 t 检验。

就某一课题而言，到底选择何种设计类型？要根据实验研究的目的、实验组数和受试对象的具体情况来确定。对于两组实验，如果处理因素施加前后其它影响效应的因素无变化，可选用自身前后配对设计，如有机磷中毒前后血压值的变化。对于三组及多组的实验，可用完全随机设计。但若受试对象中影响效应的因素不整齐，如动物的体重相差多，则将体重相近的受试对象分成一个区组，然后按随机区组设计。但若受试对象较少，处理因素施加顺序对结果的影响大，而各处理因素之间又无交互作用时，也可选用拉丁方设计。如果分析药物方剂中各成份作用大小，有无交互作用时，可采用析因设计。因此，全面地了解常用实验类型的适用条件和优缺点，对于正确选择设计方案，有效地利用资金，合理安排实验有重要意义。机能实验设计是在统计学的理论指导下，结合研究的具体情况而产生的。在运用的过程中，可能会发现新的问题，需要研究人员在实践中不断摸索和探讨解决问题的新方法、新思路。使科研统计设计在机能实验研究中完善和发展，真正发挥它应有的作用。

## 第四章 实验动物设计

### 第一节 选择实验动物应注意的问题

实验动物对外界刺激的反应存在着个体差异，为了减少实验误差，在实验动物的选择上应注意动物的年龄、体重、性别、生理状态、健康状况及动物的品系和等级等因素。

#### 一、年龄与体重

动物的年龄是一个重要的生物量，实验动物的机体反应性随着年龄的增长而有明显的变化。不同年龄的动物，对药物的毒性反应亦有差异。大多数机能实验都应该选择成年实验动物。因为幼年动物敏感性高，其实验结果常与成年动物不一致，而老年动物，则因代谢功能降低，反应不灵敏，如非特殊需要亦不宜选用。因此，急性毒性试验规定小鼠体重在18~22g，大鼠体重在200~280g，其目的在于减少由体重、年龄差距造成的影响。急性毒性实验选择的动物种类，最好与药效学、药物动力学以及长期毒性试验一致。但有例外，如慢性实验或长期毒性试验，由于实验周期长，大多选择幼年和体重较小的实验动物，而老年医学研究，则多选用老年实验动物。

动物的体重，在同一实验中应尽可能一致，若相差悬殊，易增加动物反应的个体差异，影响实验结果的准确性。

#### 二、性别

实验动物对药物的毒性反应有性别差异,并主要表现在成年动物。一般而言,雌性动物

对药物的敏感性高于雄性,但因受试物的不同而有差异。如雌性小鼠对巴比妥酸盐类、四氧嘧啶和土的宁等药物的敏感性高于雄性小鼠;但雄性小鼠对肾上腺素、乙醇、氨基比林和烟碱等药物的敏感性则明显高于雌性小鼠。一般认为,动物对激素的依存性,是其药物毒性反应性别差异的主要原因。研究发现小鼠对氨基比林和氯仿,大鼠对巴比妥类,犬对地高辛等药物的毒性反应的性别差异与性激素有关。

在实验动物设计时,通常选择雌雄各半的实验动物。若发现有明显的性别差异时,则应

分别测定不同性别实验动物的LD<sub>50</sub>值。

### 三、生理功能状态

在实验动物设计时,要考虑实验动物不同的生理功能状态对实验的影响。例如,在怀孕和哺乳时,实验动物的体重及某些生理生化指标将有所不同,许多酶的活性降低,肝微粒体单胺氧化酶对某些外来物的代谢作用减弱。此时,它们对外界环境因素作用的反应与非怀孕的实验动物有着较大差异。因此,一般不宜采用怀孕和哺乳的实验动物进行实验研究。另外,实验动物的功能状态不同,对药物的反应也有影响。如:体温升高时对解热药较敏感,体温不高时就对解热药不敏感;血压高时对降压药较敏感,血压较低时,对降压药就不敏感。

### 四、健康状况

动物的健康状况直接影响实验结果的正确与否,患病动物对药物的耐受性较健康动物

小,因而患病动物容易在实验过程中中毒死亡。兔食量不足,体重下降10%后,麻醉时间显著延长,可以因为饥饿和创伤等原因而导致休克。因此,为确保实验动物健康,要做好实验前的动物健康检查和病原学检查工作。

### 五、品系与等级

在实验动物设计时,要选用相匹配的品系和等级。品系表示实验动物按遗传学控制的条件进行标准化处理的程度。一般情况下,近交系动物的生物反应稳定性和实验重复性都较封闭群好;F1代动物生命活力强,带有两个亲代品系的特性,个体间的遗传型和表型都是一致的,应用时能获得正确的结论。封闭群动物和杂种动物在实验的重现性上有较大的误差。

等级是指实验动物的微生物标准化控制程度。各级动物具有不同特点,分别适用于不同的研究目的。无细菌动物是一种超常生态模型,既能排除微生物对实验结果的干扰,减少免疫功能的影响。无特殊病原体动物是正常的健康无病模型,能排除疾病或病原的背景性干扰。普通动物具有价廉、易获得、饲养设施简便、容易管理等特点,但选用时应考虑微生物对实验结果有无影响,同时特别注意已标准化的等级动物由于各种原因重新污染。

## 第五章 实验设计指南

机能学实验教学是医学院校基础医学教育中极其重要的部分，其教学内容主要包括以下三方面：(1) 基本知识与基本技能的训练：主要向学生介绍机能实验的基础知识、实验动物的特点，进行一些基本操作技能训练（如动物的抓取、给药，血管的分离、插管等操作），以及 BL-410、BL-420 等生物机能实验系统和常用手术器械及仪器的正确使用。(2) 基础与综合性机能实验：在保留了一些生理学、病理生理学、药理学三门学科经典传统实验基础上，进一步将三学科实验方法相似、理论知识相关联的实验有机地结合起来，通过正常动物整体实验、建立实验动物疾病模型以及用药物进行实验性治疗等方法，观察和探讨疾病的发生以及药物对机体各系统功能的影响。(3) 机能学实验设计：由教师介绍实验设计的目的、意义及如何选题、进行设计等方面知识，然后在教师的指导下，学生自己查文献、选课题，写出实验设计方案，并通过开题答辩，确定实验的可行性后，先进行预实验，再做正式实验，最后通过分析、总结，写出论文。

为了满足当今培养高级医学人才的需求，把学生培养成为应用型、综合型人才，我们在做好前两个教学内容的基础上，把重点放在了设计性实验和探索性实验的教学上，并根据我校机能实验室具体情况，制定了机能学实验设计的学习指南。

### 一、教学目标

#### 1、总体目标

提高学生进行实验设计的兴趣，激发和培养学生的创新思维，不断提高教学质量，充分体现自己的特色，达到本科教学的标准。

#### 2、具体目标

通过机能实验学设计课程，使学生掌握进行医学科学研究的基本程序和方法，提高学生实验设计的能力，使学生掌握实验设计方面的相关知识；通过实验操作及论文的写作，使学生掌握实验论文的写作方式，并在论文的写作过程中，提高学生对实验结果进行分析的能力。

### 二、主要研究方向和鼓励探索的范围

#### 1、研究方向

- (1) 感染性休克的机理和解救的基础研究
- (2) 弥散性血管内凝血的机理和解救的基础研究



- (3) 糖尿病机理的基础研究
- (4) 不同类型黄疸模型的机理和解救的基础研究
- (5) 急性中毒性肝炎的机理和解救的基础研究
- (6) 肝坏死的机理和解救的基础研究
- (7) 肝硬化的机理和解救的基础研究

## 2、鼓励探索范围

- (1) 神经系统发育的基础研究
- (2) 脊髓损伤修复的基础研究
- (3) 神经生长因子的基础研究
- (4) 干细胞的相关研究基础研究
- (5) 生物信息学中新理论、新技术和新方法研究

## 三、学习要求

- 1、凡是必修机能学实验课的本科生，都要参加实验设计的教学活动。
- 2、设计实验的内容由学生自由选择。
- 3、实验设计是以小组为单位进行的，参加实验设计的每组成员不得超过五个；每个组员都要积极、主动，配合小组完成好实验设计。
- 4、每位学生都要写出一份机能实验设计大纲和相关方面的综述材料；在实验结束后，以组为单位写一篇实验论文。

## 四、时间安排

从机能学实验课的第二次课开始，最后一次课结束。

## 五、工作程序

首先由教师在第一次课时以专题讲座的形式讲述有关实验设计的相关内容，如介绍实验设计的目的与意义、如何选题、设计的步骤、注意事项、教研室现有的仪器设备、实验设计书的书写格式及如何进行课堂答辩等。其次，学生查阅文献，拟订立题报告，写出机能学实验设计大纲（内容包括：）——经教师初步审核，修改设计大纲——以组为单位，随机抽一位同学进行答辩，筛选实验，确定实验内容——进行预实验，根据预试结果，调整或修改设计方案，正式进行实验——收集、整理实验资料并进行统计分析和结果分析——总结和完成论文

## 第六章 设计实验的教学与管理

设计性实验是指采用科学的逻辑思维，配合机能实验学的方法与技术，对拟定研究的目标或问题，进行有明确目的的探索性研究。

# 一、实验设计大纲

实验设计大纲要求包括以下内容：

## 1. 立题

简练、扼要地阐明实验设计的目的、意义，国内外研究现状，目前研究中存在的问题，拟解决问题的思路。根据同学们在生理学、病理生理学和药理学中所学的知识自由选题，选题时应注意科学性、先进性、可行性和实用性。

## 2. 专业设计

从专业的角度科学地考虑处理因素的来源和施加方法，选择受试对象，确定效应的指标和影响因素。

## 3. 统计设计

统计设计的主要任务是要解决数据的收集、整理和分析的问题。实验数据的完整性和准确性是对实验研究的最根本要求，恰当的、充分的和可靠的数据分析则是建立在完整的、准确的实验数据基础之上的。因此，只有高质量的数据，才谈得上高质量的实验研究。数据收集和步骤：

(1) 确定数据的类型。

(2) 确定数据所采用的统计方法，确定样本含量的大小及抽样误差排除的方法。

(3) 遵循对照、随机和重复三个基本原则。在实验过程中，为了避免和减少实验误差，取得可靠的实验结论，必须始终遵循这些原则。

随机：实验动物按照随机的原则分组，根据研究的目的和处理的方式进行分组：如果是单因素设计，只有一种因素而无非处理因素者用完全随机设计分组法；如果有一种处理因素和一种非处理因素，则采用配伍处理分组；如果一种处理因素和两种非处理因素，则用拉丁方设计；在多因素设计时，实验安排两种以上处理因素时，采用析因的设计方法。

重复：实验动物按照重复的原则，要确定一定的数量。

对照：设计对照组。

均衡：按照均衡的原则，还要考虑实验动物的品种、规格和来源。根据与人类疾病尽可能相似的原则，选择经济实用的动物。

(4) 明确观察指标：必须明确观察指标，而且是最能说明所要研究问题的指标。指标测定的具体步骤，包括样本采集的时间、样本量、样本处理和测定方法。注意指标的特异性、客观性、重现性和灵敏性。

(5) 设计数据收集的表格：实验原始记录表格包括两部分，一是常规数据，二是来源于观察指标的实验原始数据。常规数据应包括实验日期、室温、动物体重、性别、编号、麻药（麻醉药品）名称、浓度与剂量、实验参加者等。实验原始数据即所有一级实验观察指标，应该以表格形式表示。

## 4. 实验动物设计

(1) 实验动物的来源，是否选用标准动物，标准动物的微生物等级，实验动物的种类。

(2) 动物处理：包括疾病模型复制和实验治疗两部分，注意模型相似性、重复性与实用性原则。须写明动物麻醉、固定、疾病模型复制、给药治疗(时间、剂量、途径)等具体的方法。

**5. 实验方法：**简要阐述所采用的实验步骤和操作方法，实验技术路线和方法，可行性分析，实验设计工作时间安排。

**6. 完成实验的条件：**简要阐述实验中所采用的主要仪器设备的情况，注明所

用仪器的型号、生产地、生产厂家，如果有必要可以附仪器装置简图。药品试剂的来源情况：药品和试剂：药品和试剂的批号、规格及来源（生产厂家）；实验用药品和试剂的配制方法；实验用药品的数量。

## 机能学实验设计大纲

实验名称：

课题来源：

设计班级：

设计人员：

设计日期：

指导教师：

成都医学院实验技术教研室

2006年制

一、实验设计目的意义：实验设计的目的意义、国内外研究现状、存在的问题、解决问题的思路（附参考文献目录5-10篇）：

1. 实验设计的目的意义
2. 国内外研究现状
3. 存在的问题
4. 解决问题的思路

二、实验设计方案

1. 实验设计目标、拟解决的主要问题
2. 实验统计设计
3. 实验动物设计
4. 实验技术路线和方法
5. 可行性分析
6. 实验设计工作时间安排
7. 预期结果

三、完成实验的条件

1. 主要仪器设备的情况
2. 实验动物的情况
3. 药品试剂的来源情况

四、教研室评审意见

1. 实验设计是否合理，可否达到预期目标
2. 实验设计和技术路线安排是否恰当
3. 实验条件是否具备

教研室审查意见

年 月

日

## 二、实验报告的撰写

设计实验报告的撰写的基本结构通常包括4大部分，一般称之为“四段式”结构，即“前言”、“材料与方法”、“结果”、“讨论与结论”。这几部分并不是绝对的、不变的，可以根据不同的情况或分或合，灵活运用，但这几部分所包含的内容是不可缺少的。按照编排格式的标准化和规范化要求，设计性实验报告撰写的内容包括：文题、作者、作者单位、摘要、主题词、正文、参考文献。

实验报告大致包括以下内容：

1. **实验报告的题目**：包括实验名称、姓名和单位的一般情况说明。
2. **摘要和关键词**。
3. **前言**：主要说明实验设计的目的意义，国内外研究现状，目前研究中存在的问题，拟解决问题的思路。
4. **实验方法和步骤**：说明所采用的实验技术路线和方法。
5. **实验结果**：得出的实验结果。
6. **实验结果分析与讨论**：对实验结果进行分析与讨论。
7. **结论**：得出什么结论。
8. **参考文献**：列出所采用的参考文献。

### 三、机能设计性实验课的组织与管理

设计性实验由4~5名学生组成实验小组，经过查阅文献资料，调查研究，选择实验项目，写出实验设计方案并在小组会上进行开题论证，其方案经指导教师审查同意后进行预实验，继而转入正式实验，实验结束后写出总结论文并以教学班为单位组织论文总结与答辩。时间分配：调研选题、实验设计和开题论证共约2学时；预实验约3学时；正式实验约4学时；论文总结答辩约2学时。

### 四、机能设计性实验课考核办法

机能实验学与机能综合实验设计课是机能三学科实验教学整合、重组所形成的实践课。该门课程考核对于客观、公正地评价学生的实践能力，引导学生重视实践，从而保证学生实践能力得到有效的锻炼，具有很重要的意义。该课程的考核内容、要求及标准如下：

**1. 实验设计质量（30%）**：按照实验设计大纲的要求进行设计。实验设计体现个人查阅文献、发现问题、提出问题的能力，是个人科学思维、创新思维的完整体现。根据设计是否合理、规范，是否具有创新性进行综合评分。查阅文献量大、内容新颖，立题依据明确、可靠，设计思路清晰、合理，记15分；设计规范、完整，有一定的创新意义，记15分。

**2. 实验设计讨论（10%）**：设计者向同学陈述自己设计的理论依据、构思、存在的问题并回答同学的质疑。语言流利、表达清楚、回答问题正确者，记5分。在听取其他同学的设计时，勤于动脑、发言积极者，加5分。

**3. 实验设计的实施（20%）**：根据实验操作是否熟练规范，实验结果是否准确可靠，实验技术难度的大小进行综合评定。在实施过程中，组内人员团结一致、群策群力，既分工明确，又互相帮助，在规定的时间内完成实验内容，实验结果准确、可靠者，组内集体得分，满分20分。实验时，操作积极、规范，特别突出者，加5分。

**4. 实验论文的质量（20%）**：论文书写格式正确、重点突出、语言流畅、讨论科学者，组内集体得分，满分15分。作为负责人的实验设计者，能组织、管理

组内人员顺利完成实验，加5分。

**5. 答辩表现（10%）：**由指导教师或其他组学生对其报告的论文提出问题，被评组的学生均应能对所提问题进行答辩。提问内容包括：文献准备与背景知识，设计思路与技术手段，操作环节与实验结果，分析讨论与存在问题等方面。数据处理、资料汇总，对论文书写起决定作用者，加5分。参加论文答辩，报告效果好，集体加3分，报告者个人另加2分。

**6. 小组互评（10%）：**各小组成员在其他组进行论文报告和答辩时对其实验设计质量、实验结果、论文质量及论文报告答辩情况予以综合评定，现场按照优、良、中3个等级分别给出10、6、3分。各组评分的平均值即为被评小组学生的共同得分。

**7. 在实验设计和论文中的排名得分：**学生在实验设计和论文中排名由小组根据个人贡献大小(内容包括文献调研、实验选题、方案设计、资料整理、结果分析及论文撰写等方面)民主评议确定并上报指导教师审定。