

前进英才丛书

临床研究

样本含量估算

主编 吴圣贤 王成祥



人民卫生出版社

鲍林春前进英才基金资助出版

临床研究样本含量估算

主 编 吴圣贤 王成祥

副主编 林炳辉 张耀圣

编 委 (以汉语拼音为序)

程志强 方素钦 刘清泉 时 杰 汪 科
王 忠 吴水生 吴松鹰 晏 军 于国泳
张立山 张文生

莫奇量含本料究研宋商

主 编 王成祥 吴圣贤
出版发行 人民卫生出版社 (中组部 010-67818888)
地 址 北京市丰台区万芳园3区3号楼
邮 编 100078
网 址 <http://www.puph.com>
E-mail puph@puph.com
邮 政 热 线 010-67802724 010-65264830
印 刷 商务印书馆
经 销 新华书店
开 本 787×1092 1/16 印张: 10.2
字 数 248千字
版 次 2008年2月第1版 2008年2月第1次印刷
书 号 ISBN 978-7-117-09882-1 R · 9386
定 价 25.00元

人 民 卫 生 出 版 社

图书在版编目(CIP)数据

临床研究样本含量估算 / 吴圣贤等主编. —北京:人民
卫生出版社, 2008. 5

ISBN 978-7-117-09985-1

I. 临… II. 吴… III. 临床医学—研究—医学统计
IV. R4-32

中国版本图书馆 CIP 数据核字 (2008) 第 027345 号

吴圣贤 王成祥 主编

吴圣贤 王成祥 副主编

(本书音译英文) 委 编

王 杰 王 杰 王 杰 王 杰 王 杰 王 杰

王 杰 王 杰 王 杰 王 杰 王 杰 王 杰

王 杰 王 杰 王 杰 王 杰 王 杰 王 杰

临床研究样本含量估算

主 编: 吴圣贤 王成祥

出版发行: 人民卫生出版社 (中继线 010-67616688)

地 址: 北京市丰台区方庄芳群园 3 区 3 号楼

邮 编: 100078

网 址: <http://www.pmph.com>

E - mail: pmph@pmph.com

购书热线: 010-67605754 010-65264830

印 刷: 尚艺印装有限公司

经 销: 新华书店

开 本: 787×1092 1/16 印张: 10.5

字 数: 248 千字

版 次: 2008 年 5 月第 1 版 2008 年 5 月第 1 版第 1 次印刷

标准书号: ISBN 978-7-117-09985-1/R·9986

定 价: 22.00 元

版权所有, 侵权必究, 打击盗版举报电话: 010-87613394

(凡属印装质量问题请与本社销售部联系退换)

序一

临床流行病学是 20 世纪 70 年代后期,在国际医学领域里发展起来的一门新兴临床医学基础学科,它对于改进临床医学的研究、教学和医疗,推动临床医学的发展,具有重要意义。用其原理和方法指导临床医疗科研实践,构建与发展了循证医学(evidence-based medicine, EBM);用以对卫生技术评估(health technology assessment, HTA),促进了健康管理和卫生决策的科学化。

中医药学是科学与人文融合的具有中国特色的生命科学。其产生是我国古代劳动人民长期实践经验的积累,是与疾病作斗争的产物,历数千年而葆其青春,不仅对中华民族的繁衍生息作出了巨大贡献,而且丰富了现代医学的内容。中医中药的现代临床试验需要借鉴当今相关前沿的统计学、信息学、系统科学与复杂科学的临床研究方法学,以使其临床效应得以相对科学与客观地展现。我对 EBM 的主张是:一学、二懂、三要用、四要了解局限性、五要按照中医自身规律加以完善,为中医所用。

样本含量估算是医学临床研究设计中的重要环节,但迄今为止,还没有一本中文专著可供参考。本书内容丰富、纲举目张、方法与实例并重,是一本适合临床研究人员实际应用的参考书,爰为之序。

中国工程院院士 王永炎

2008 年 3 月

王永炎 中国社会科学院学部委员 中国工程院

院士 2008

每次举办全国统计学培训班或举办统计学学术讲座,很多听众都会问这样一个问题:我要做一个科研课题,请问我要选择多少病例?作为一名统计学工作者,对诸如此类的问题早已司空见惯了。于是,我会坦率地说,如果您没有提供更多信息的话,这个问题我无从回答。首先,应该肯定,这是一个非常重要的问题,因为它是所有搞科研的人不可回避的一个实际问题;同时,它确实又是一个十分棘手的问题,是一个能让全世界一流统计学家都会难以回答的问题。因为仅有问题,缺少必要的辅助信息,真好比“巧妇难为无米之炊”。

上述问题用统计语言来表述应叫做“样本含量估算问题”,即估计最少做多少病例既能把被研究事物的本质反映出来,又能节省时间、人力、物力和财力。众所周知,根据以往的经验 and 气象资料,人类很早就会预报未来的天气情况;根据以往某些疾病的发病与流行规律,人类很早就会预报未来某些疾病的流行趋势;根据以往某些气候和自然条件,人类很早就会预报未来洪涝灾害发生的可能性大小。这里所说的“根据……”就是“辅助条件”和“先验知识”,否则所有的“估计”或“预报”都是“信口开河”,难以令人置信。估计样本含量更是如此,需要提供必要的“辅助条件”和“先验知识”,才有可能完成此重任。

样本含量估算问题是一个非常复杂的问题,因为往往涉及的“辅助条件”和“先验知识”比较多,有时需要查阅文献资料获得,有时需要通过做预实验获得。从科研课题拟解决的问题分类,可能涉及“参数估计中样本含量问题”、“假设检验中样本含量估算问题”和“多元分析中样本含量估算问题”等;从科研课题拟采取的手段分类,可能涉及“抽样调查的样本含量问题”、“实验(临床试验)的样本含量估算问题”等,每一大类问题又可进一步细分成多个具体问题,情况错综复杂,很难三言两语表达清楚。

值得庆幸的是,本书作者急临床医师之所急,同时也为众多统计学工作者排忧解难,总结了国内外很多估算样本含量的经验,加上作者自己学习和应用的体会,撰写了本书。该书分类清晰、语言表达流畅、内容齐全,可以解决相当大范围内样本含量估算,并以实例为引导,易读、易懂、易学、易用,是同类书籍中写得很成功的一部。本人也从中获益匪浅,并欣然为之作序,愿向广大实践工作者推荐该书。

人无完人,金无足赤,写得再好的书仍会有不足之处,希望广大读者不吝赐教,多提宝贵意见,以便将来再版时使其日臻完善。

中国生物医学统计学会副会长 胡良平

2008年3月

临床研究是医学研究中的一个重要部分。与基础研究比较,临床研究包括临床试验、诊断试验和病因学研究等,通常需要时间更长、耗资更多。通过样本含量估算,不但可以计算出合理的受试者数量以保障研究的成功,而且可以确定研究者所希望的把握度,做到心中有数。因此,样本含量估算在各期临床研究设计中,扮演着重要角色。

在进行样本含量估算之前,研究者必须首先明确试验的主要终点指标、指标数据性质(计量资料或计数资料)、对该指标的统计方法选择、试验总体设计方法(平行对照、交叉对照等)、以及把握度、检验水平、平均数、标准差、有效率、有意义的差值等基础资料。可以说,样本含量估算的多样性和水平直接反映了临床研究设计和统计方法的多样性和水平,因此,本书不仅仅为读者提供了广泛、多样的样本含量估算方法,而且也是对近年来国际上临床研究设计方法、统计分析方法进展的一个全面梳理。

本书对于来自医院、研究机构、制药企业、合同研究组织的从事临床研究相关工作的研究者、统计学家、管理者和监察员来讲,都是一本非常实用的参考书。作者撰写该书的主要宗旨是为各种类型临床研究提供样本含量估算方法和实际操作程序。因为本书以方便实际研究设计应用为出发点,所以在结构设置上自始至终体现两个特点:一是每个样本含量估算方法都从临床研究实例入手,力求做到程序清晰、简明易懂;二是对多种样本含量估算方法进行了详细分类,特别重视临床研究设计中的细微差别,力求做到绝大部分临床研究设计都能够直接查阅到实例,而且纲举目张,方便查考。

基于以上考虑,本书不但提供了临床试验的样本含量估算方法,而且提供了诊断试验和病因学研究样本含量估算方法;不但提供了差异性检验的样本含量估算方法,而且提供了非劣/优效性检验、等效性检验的样本含量估算方法;不但提供了两组比较、各组样本量相等情况下的样本含量估算方法,同时提供了单组设计、多组设计、各组样本量不等情况下的样本含量估算方法;不但提供了常用临床研究设计类型的样本含量估算方法(如平行对照设计、交叉设计),而且提供了大量新的临床研究设计类型的样本含量估算方法(如Williams设计、序贯设计、重复测量设计、优化两阶段设计等)。

中国工程院院士、全国人大常委会、中国中医科学院名誉院长王永炎院士为本书欣然作序,谨表诚挚感谢。同时,也非常感谢造诣深厚的统计学家、来自中国军事医学科学院的胡良平教授,他对本书进行了认真审阅并作序。

由于作者学术水平有限,书中难免会出现一些论述不当甚至错误的地方。非常欢迎读者的任何建议和意见,您的反馈将对本书下一版的准备工作起到重要指导作用。

吴圣贤 王成祥

2008年3月20日于北京

第一章 样本含量估算基础	1
第二章 计量资料样本含量估算	3
第一节 单组设计	3
第二节 两组平行对照设计	5
第三节 两组交叉对照设计	9
第四节 多组平行对照设计	12
第五节 多组 Williams 设计	14
第六节 成组序贯设计	16
第七节 剂量反应设计	19
第八节 重复测量设计	23
第九节 直线相关分析	24
第十节 直线回归分析	26
第三章 计数资料样本含量估算	28
第一节 单组设计	28
第二节 两组平行对照设计	30
第三节 两组交叉对照设计	33
第四节 多组平行对照设计	35
第五节 多组 Williams 设计	37
第六节 相对危险度平行设计	39
第七节 相对危险度交叉设计	43
第八节 二项分布检验	45
第九节 Fisher's 精确检验	48
第十节 最优化两阶段设计	49
第十一节 灵活两阶段设计	51
第十二节 最优化三阶段设计	57
第十三节 两组灵活设计	60
第十四节 多组灵活设计	61
第十五节 拟合优度检验	62
第十六节 单层列联数据	63
第十七节 多层列联数据	64
第十八节 麦克尼莫检验	66

第十九节	Stuart-Maxwell 检验	67
第二十节	指数模型	68
第二十一节	Cox's 比例风险模型	71
第二十二节	疫苗试验	73
第二十三节	成组序贯设计	74
第四章	诊断试验样本含量估算	77
第五章	抽样调查样本含量估算	79
第六章	病因研究样本含量估算	81
第一节	病例-对照研究	81
第二节	队列研究	84
第七章	样本含量估算软件	86
第一节	nQuery Advisor®	86
第二节	SamplePower	90
第三节	SASA	92
第四节	PASS	126
第五节	SAS	128
附录 1	统计用表	131
附表 1.1	Z 值表(负值)	131
附表 1.2	Z 值表(正值)	132
附表 2	t 值表	133
附表 3	χ^2 值表	135
附表 4.1	F 临界值(方差齐性检验用, 双侧概率为 0.05)	136
附表 4.2	F 临界值(方差分析用, 单侧概率为 0.05)	137
附表 4.3	F 临界值(方差分析用, 单侧概率为 0.01)	138
附表 5.1	两个率差别检验每组样本大小(单侧检验, $\alpha=0.05, \beta=0.10$)	139
附表 5.2	两个率差别检验每组样本大小(双侧检验, $\alpha=0.05, \beta=0.10$)	140
附表 5.3	两个率差别检验每组样本大小(单侧检验, $\alpha=0.01, \beta=0.05$)	140
附表 5.4	两个率差别检验每组样本大小(双侧检验, $\alpha=0.01, \beta=0.05$)	141
附表 6	两个样本均数差别检验时每组样本大小(双侧)	141
附表 7.1	方差分析时每组所需样本大小(3 组, 双侧)	142
附表 7.2	方差分析时每组所需样本大小(4 组, 双侧)	143
附表 7.3	方差分析时每组所需样本大小(5 组, 双侧)	144
附表 7.4	方差分析时每组所需样本大小(6 组, 双侧)	144
附表 7.5	方差分析时每组所需样本大小(7 组, 双侧)	145
附表 7.6	方差分析时每组所需样本大小(8 组, 双侧)	146
附表 7.7	方差分析时每组所需样本大小(9 组, 双侧)	146
附表 8.1	对率作抽样调查时所需样本大小($\alpha=0.05$)	147
附表 8.2	对率作抽样调查时所需样本大小($\alpha=0.01$)	148

附表 9.1	对平均数作抽样调查时所需样本大小($\alpha=0.05$)	149
附表 9.2	对平均数作抽样调查时所需样本大小($\alpha=0.01$)	149
附表 10	直线相关分析时所需样本大小($\alpha=0.05$)	150
附表 11.1	ϕ 值表($\alpha=0.05, \beta=0.1$)	150
附表 11.2	ϕ 值表($\alpha=0.05, \beta=0.2$)	152
附表 12.1	λ 值表($\alpha=0.05$)	154
附表 12.2	λ 值表($\alpha=0.01$)	156
附录 2	希腊字母表	158

本科, 研究对象量级下, 由真量变而致假天量级下因其, 又意对显著效果而底
(1-1 表) 变而量不量含

基础医学与临床

第一章

样本含量估算基础

在临床研究的设计阶段, 下面两个问题是研究者特别感兴趣的: 一是, 为了在希望的把握度(power)下得出具有临床意义的差异, 需要多少样本含量(sample size)? 二是, 如果由于预算问题或某些医学考虑, 只能对一个小样本进行研究, 所得出的结果有多大的把握度? 为了回答上述问题, 基于一项研究的主要目的, 样本含量和把握度的计算是必需的。在临床研究方案设计过程中, 样本含量估算扮演着重要角色。

样本含量估算的重要意义在于, 合适的样本含量有助于我们用最合理的资源发现有意义的临床差异, 如果这个差异确实存在的话。过少的样本含量难以准确回答试验所研究的科学问题, 而过多的样本含量是巨大的资源浪费。

样本含量估算方法必须基于明确设计方法、假设检验和主要目的指标的特性。例如, 临床研究中平行对照设计、交叉对照设计、Williams 设计、序贯设计、连续测量设计等都有不同的样本含量估算方法; 主要目的指标是计数资料或计量资料需要不同的样本含量估算方法; 基于计量资料的 t 检验和 F 检验, 基于计数资料的卡方检验和 Fisher's 精确检验等都有不同的样本含量估算方法; 差异性检验、非劣效性检验和等效性检验也有不同的样本含量估算方法, 等等。

样本含量变化的基本影响因素包括:

1. 检验水平(α) I 型错误表示, 统计推断拒绝了实际上成立的无效假设(H_0)。I 型错误的概率用 α 表示, 故又称 α 错误。在这里, I 型错误指的是, 组间差异实际上不存在, 统计推断的结果, 却错误地承认组间差异的存在, 又称假阳性。进行统计推断时, 研究者需要对容许犯 I 型错误的大小作出规定, 通常是 $\alpha \leq 0.05$, 因为在这一范围内, 若作出拒绝 H_0 的推断, 其所犯的 I 型错误概率已很小, 故承认这一推断的正确性。检验水平 α 值越小, 所需样本含量越大。

2. 把握度(power, $1-\beta$) II 型错误表示, 统计推断的结果不拒绝实际上不成立的 H_0 , 其出现的概率用 β 表示, 故也称为 β 错误。或者说, 当组间的差异确实存在时, 统计推断却不承认该差异的存在, 也称为假阴性。 β 值越小, 所需要的样本含量越大。根据实际情况, 通常规定 $\beta \leq 0.10$, 必要时, 可取 $\beta = 0.20$ 。

把握度(power)也称为检验效能。把握度与 II 型错误互补, 两者的概率之和是 1。若 $\beta = 0.10$, 则把握度 $= 1 - \beta = 1 - 0.10 = 0.90$, 表示若组间确有差别时, 通过 100 次的试验, 有 90 次能够得出组间有差别的结论。把握度越高, 所需要的样本含量越大。足够的把握度能检验出不同治疗水平结果间实际存在的差异。如果某项试验的把握度不够大, 所得

到的结果没有显著性意义,其原因可能是无效假设确实是真的,也可能是把握度低,样本含量不足所致(见表 1-1)。

表 1-1 I 型错误、II 型错误和把握度

统计推断	真实情况	
	两组有差别	两组无差别
两组有差别	判别正确	判别错误 I 型错误,假阳性
两组无差别	判别错误 II 型错误,假阴性	判别正确

3. 变异(σ) 变异指样本中所包含的个体的差异程度。个体之间的差异越大,所需要的样本含量越大,反之,若个体之间的差异越小,则所需要观察的例数越少。例如,对于某一疾病的研究,若包含了病情程度不一的轻、中、重型病例,则对比于仅观察其中单一的某一型所需要的病例数要多。

4. 临床有意义的差值(δ) 临床有意义的差值也可称为组间效应的差异程度,即不同药物可能的疗效差异。组间差异越大,所需要的病例数越少,反之,则所需要观察的病例数较多。假设,甲药对某一疾病的预期总有效率是 80%,乙药是 60%,已知对照药的总有效率是 40%,则当乙药作为试验用药时,要比以甲药作为试验用药时所需要的病例数为多。有时,当缺乏试验用药的疗效数据时,可以用我们希望得到的具有临床意义的差值来推测试验用药的疗效。

5. 单侧检验和双侧检验 差异性检验和等效性检验需要双侧检验,而非劣效性检验和优效性检验需要单侧检验。单侧检验所需样本含量少于双侧检验。

第二章

计量资料样本含量估算

通过观察个体用定量的办法测量某项指标数量大小所得到的资料,称为计量资料。这类资料,一般有度量衡单位,可用数值大小表示。在一定区间内,只要观察个体足够多,其测得数值可以在数轴上连接起来,故也成为连续性资料。如测量病人的身高(cm)、体重(kg)、血压(mmHg)、血液中胆固醇含量(mmol/L)等。对这一类资料的常用描述性指标有平均数、标准差、中位数、最小值、最大值等。推断分析有 t 检验、方差分析等。设计方案有单组设计、两组平行对照设计、两组交叉对照设计、多组设计等。

第一节 单组设计

一、设计方法

在临床试验中,没有对照组的开放设计称为单组设计。单组设计的分析目标有两种,一种是治疗组数据与标准公认值的比较;另一种是基线数据和治疗后数据的分析比较。

例 1 已知北京市某社区 50~70 岁男性的平均收缩压为 158mmHg,用新型降压药物干预 100 位该社区男性 1 个月,平均收缩压降为 148mmHg,问该新型降压药物是否有显著的降压疗效?(测得的具体数据从略)

解答:在本例中只测定了一组(100 例)服用新型降压药物后的血压数值,属于计量资料,对高血压患者未按任何因素进行分组,此为“单组设计”,需给定标准值或理论值(本例为某社区 50~70 岁男性的平均收缩压为 158mmHg)方可进行假设检验,这是单组设计治疗组数据与标准公认值比较的典型设计方法。

例 2 用新型降压药物治疗 100 位男性高血压患者 1 个月,治疗前的平均收缩压为 158mmHg,治疗后的平均收缩压降为 148mmHg,问该新型降压药物是否有显著的降压疗效?(测得的具体数据从略)

解答:在本例中只测定了一组(100 例)服用新型降压药物前后的血压数值,属于计量资料,对高血压患者未按任何因素进行分组,此为“单组设计”,需对治疗前后的血压差值进行比较分析,这是单组设计基线数据和治疗后数据分析比较的典型设计方法。

二、样本含量估算方法

1. 与标准公认值比较,双侧检验的公式

$$n = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2 \sigma^2}{\delta^2}$$

公式中:

- (1) n 代表样本含量
- (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
- (3) σ 代表总体标准公认数据的标准差
- (4) δ 代表具有临床意义的两组差值

2. 与标准公认值比较, 单侧检验的公式

$$n = \frac{2(Z_{1-\alpha} + Z_{1-\beta})^2 \sigma^2}{\delta^2}$$

公式中:

- (1) n 代表样本含量
- (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
- (3) σ 代表总体标准公认数据的标准差
- (4) δ 代表具有临床意义的两组差值

3. 与基线比较, 双侧检验的公式

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \sigma^2}{\delta^2}$$

公式中:

- (1) n 代表样本含量
- (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
- (3) σ 代表治疗前后数据差值的标准差
- (4) δ 代表治疗前后均数差值

4. 与基线比较, 单侧检验的公式

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \sigma^2}{\delta^2}$$

公式中:

- (1) n 代表样本含量
- (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
- (3) σ 代表治疗前后数据差值的标准差
- (4) δ 代表治疗前后均数差值

三、举例

例 1 已知北京某社区 50~70 岁男性的平均收缩压为 158mmHg, 标准差为 18mmHg, 用药物 AAA 干预, 如果平均收缩压下降 10mmHg, 则认为有临床意义, 取 $\alpha = 0.05$, $\text{power} = 90\%$, 双侧检验, 需要多少病例?

$$n = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2 \sigma^2}{\delta^2} = \frac{2 \times (1.96 + 1.28)^2 \times 18^2}{10^2} \approx 35$$

该研究需要至少 35 例患者。

例 2 用药物 AAA 治疗老年期收缩压, 治疗前平均收缩压为 158mmHg, 希望治疗后平均降低收缩压 10mmHg, 估计差值的标准差为 15mmHg, 取 $\alpha=0.05$, $\text{power}=90\%$, 双侧检验, 需要多少病例?

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \sigma^2}{\delta^2} = \frac{(1.96 + 1.28)^2 \times 15^2}{10^2} \approx 24$$

该研究需要至少 24 例患者。

第二节 两组平行对照设计

一、设计方法

在临床试验中, 研究者通常希望通过治疗药物与安慰剂或阳性对照药物比较, 来评价治疗药物的有效性和安全性。其典型的研究目的通常包括: ①治疗药物的疗效与安慰剂比较是否存在有意义的临床差异; ②治疗药物的疗效是否不低于现有的公认有效药物; ③治疗药物的疗效是否优于现有的公认有效药物; ④治疗药物是否与现有公认有效药物有同等疗效。

两组平行对照设计常用以下两种设计方案:

1. 完全随机成组设计 也称为完全随机单因素两水平设计, 即将合格受试对象完全随机分为 A、B 两组。

例 一位研究者欲研究新型降压药品 AAA 治疗高血压的作用, 和安慰剂进行对照, 该如何设计?

解答: 本例是一个典型的安慰剂对照试验, 属于随机单因素两水平设计。将合格受试对象完全随机分为两组, 一组使用新型降压药品 AAA 治疗, 另一组使用安慰剂治疗即可, 设计模型见图 2-2-1。



图 2-2-1 随机单因素两水平设计模型图

2. 随机区组设计 它是将全部受试对象按某个或某些重要的属性(即区组因素, 如多中心临床试验中的中心, 血压的轻度、中度和重度等)分为若干个区组, 在每个区组中随机分配两个受试因素。

例 一位研究者欲研究新型降压药品 AAA 治疗高血压的作用, 和临床有效药品洛丁新进行对照, 欲考察新型降压药品的临床疗效是否不劣于洛丁新, 该研究将在四个中心进行, 应如何设计?

解答: 本例是一个典型的多中心随机对照试验, 属于随机区组设计。需要将合格受试对象按中心分为 4 个区组, 然后把每个中心中的受试对象随机分为两组, 一组使用新型降压药品 AAA 治疗, 另一组使用洛丁新治疗, 设计模型见图 2-2-2。

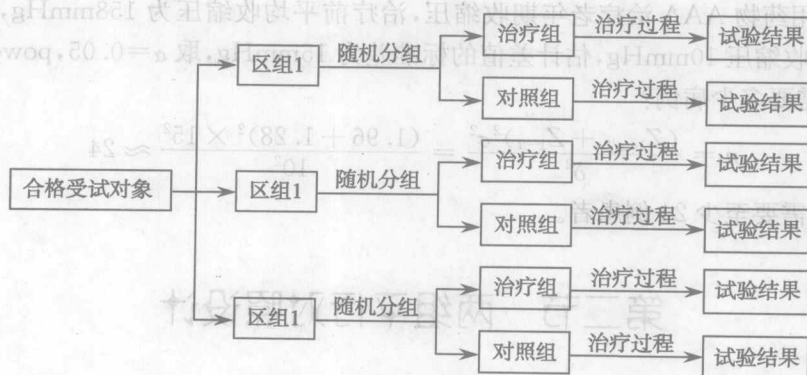


图 2-2-2 随机区组设计模型图

二、样本含量估算方法

1. 差异性检验, 两组病例数相等, 双侧检验的公式

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times (\sigma_1^2 + \sigma_2^2)}{\delta^2}$$

公式中:

- (1) n 代表每组样本含量
- (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
- (3) σ_1 代表第一组的标准差
- (4) σ_2 代表第二组的标准差
- (5) δ 代表具有临床意义的两组差值

2. 差异性检验, 两组病例数相等, 单侧检验的公式

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \times (\sigma_1^2 + \sigma_2^2)}{\delta^2}$$

公式中:

- (1) n 代表每组样本含量
- (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
- (3) σ_1 代表第一组的标准差
- (4) σ_2 代表第二组的标准差
- (5) δ 代表具有临床意义的两组差值

3. 差异性检验, 两组病例数不等, 双侧检验的公式

$$n_1 = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times (\sigma_1^2 + \sigma_2^2) \times (1 + 1/K)}{\delta^2}$$

$$n_2 = K \times n_1$$

公式中:

- (1) n_1 代表第一组样本含量
- (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
- (3) σ_1 代表第一组的标准差

- (4) σ_2 代表第二组的标准差
 (5) δ 代表具有临床意义的两组差值
 (6) K 代表两组病例数的比值
 (7) n_2 代表第二组样本含量

4. 差异性检验, 两组病例数不等, 单侧检验的公式

$$n_1 = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \times (\sigma_1^2 + \sigma_2^2) \times (1 + 1/K)}{\delta^2}$$

$$n_2 = K \times n_1$$

公式中:

- (1) n_1 代表第一组样本含量
 (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
 (3) σ_1 代表第一组的标准差
 (4) σ_2 代表第二组的标准差
 (5) δ 代表具有临床意义的两组差值
 (6) K 代表两组病例数的比值
 (7) n_2 代表第二组样本含量

5. 非劣效性检验或优效性检验, 两组病例数相等的公式

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \times (\sigma_1^2 + \sigma_2^2)}{(\epsilon - \delta)^2}$$

公式中:

- (1) n 代表每组样本含量
 (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
 (3) σ_1 代表第一组的标准差
 (4) σ_2 代表第二组的标准差
 (5) δ 代表具有临床意义的低限或高限
 (6) ϵ 代表两组的实际差值

6. 非劣效性检验或优效性检验, 两组病例数不等的公式

$$n_1 = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \times (\sigma_1^2 + \sigma_2^2) \times (1 + 1/K)}{(\epsilon - \delta)^2}$$

$$n_2 = K \times n_1$$

公式中:

- (1) n_1 代表第一组样本含量
 (2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表
 (3) σ_1 代表第一组的标准差
 (4) σ_2 代表第二组的标准差
 (5) δ 代表具有临床意义的低限或高限
 (6) ϵ 代表两组的实际差值
 (7) K 代表两组病例数的比值

(8) n_2 代表第二组样本含量

7. 等效性检验, 两组病例数相等的公式

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta/2})^2 \times (\sigma_1^2 + \sigma_2^2)}{(\delta - |\epsilon|)^2}$$

公式中:

(1) n 代表每组样本含量

(2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表

(3) σ_1 代表第一组的标准差

(4) σ_2 代表第二组的标准差

(5) δ 代表具有临床意义的低限或高限(低限和高限的绝对值相同)

(6) ϵ 代表两组的实际差值

8. 等效性检验, 两组病例数不等的公式

$$n_1 = \frac{(Z_{1-\alpha} + Z_{1-\beta/2})^2 \times (\sigma_1^2 + \sigma_2^2) \times (1 + 1/K)}{(\delta - |\epsilon|)^2}$$

$$n_2 = K \times n_1$$

公式中:

(1) n_1 代表第一组样本含量

(2) $Z_{1-\alpha}$ 和 $Z_{1-\beta}$ 需要查阅 Z 值表

(3) σ_1 代表第一组的标准差

(4) σ_2 代表第二组的标准差

(5) δ 代表具有临床意义的低限或高限(低限和高限的绝对值相同)

(6) ϵ 代表两组的实际差值

(7) K 代表两组病例数的比值

(8) n_2 代表第二组样本含量

三、举例

例 1 一项临床试验研究新型降压药品 AAA 治疗高血压的作用, 和安慰剂进行对照。如果新型降压药品 AAA 能够比安慰剂平均多降低收缩压 10mmHg, 则认为有推广价值。已知安慰剂降低收缩压的标准差为 8mmHg, 估计新型降压药品 AAA 降低收缩压的标准差为 15mmHg。1:1 平行对照设计, 选用 $\alpha=0.05$, $\text{power}=90\%$, 双侧检验, 需要多少样本?

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times (\sigma_1^2 + \sigma_2^2)}{\delta^2} = \frac{(1.96 + 1.28)^2 \times (8^2 + 15^2)}{10^2} \approx 31$$

该研究每组需要至少 31 例患者。

例 2 一项临床试验研究新型降压药品 AAA 治疗高血压的作用, 和安慰剂进行对照。如果新型降压药品 AAA 能够比安慰剂平均多降低收缩压 10mmHg, 则认为有推广价值。已知安慰剂降低收缩压的标准差为 8mmHg, 估计新型降压药品 AAA 降低收缩压的标准差为 15mmHg。3:1 平行对照设计, 选用 $\alpha=0.05$, $\text{power}=90\%$, 双侧检验, 需要