

CLINICAL TRIALS

A Practical Approach

临床治疗科研方法学

(英)S. J. 波可克著

白典 秦绍明译 杨开济 校

天津科技翻译出版公司

临床治疗科研方法学

(英)	S.J.波可克	著
郭曰典	秦绍明	译
杨开济		校

天津科技翻译出版公司



责任编辑 于洪涛

200/17

临床治疗科研方法学

(英) S·J·波可克著

郭曰典 秦绍明 译

天津科技翻译出版公司

(邮政编码: 300192)

新华书店天津发行所发行

天津市宝坻县第二印刷厂印刷

787×1092 1/32 印张11 字数230千字

1991年2月第1版

1991年2月第1次印刷

印数: 1—2800 册

ISBN 7—5433—0159—8/R·23

定价: 6.00元

前 言

对一种新药、一种新疗法，要作出正确而客观的疗效评价，并确定是否宜于推广，决不是轻而易举的事，既要求临床治疗研究者要有严肃认真的科学态度，又要有一套严密而实用的科学方法。由于不少临床工作者对这套方法不熟悉，所以在某些文献报告中的疗效评价，出现不少问题，往往缺乏可信性，甚至以偏概全，夸大疗效，造成不良影响。因此，深感有必要迅速普及这方面的知识。

本书是一本专门介绍临床治疗科研方法的专著，原著作者是英国伦敦大学医学统计学高级讲师，他对临床科研方法的理论和实践，都有很深造诣，撰著本书内容新颖丰富，方法实际。作者用通俗语言，不厌其烦地从实用角度，系统地介绍了临床试验的发展史，以及每一实施过程的实用方法细节，并用大量实例反复说明，以帮助加深理解，每一讨论都有作者自己的鲜明观点。在数据管理中还扼要介绍了电子计算机的具体应用方法。本书在国外是一本不可多得的实用参考书。

目前，由于国内还缺乏此奖专著，因此北京军区后勤卫生部特邀请郭曰典、秦绍明同志抽暇将书译出，并请对医学统计学和临床科研有丰富经验的杨开济教授最后审校，推荐给临床医生和从事临床科研的有关人员。古语说：“他山之

石，可以为错”，相信读者一定可以从本书中受到启发，如能结合我国的实际，有效地加以利用，必将有助于提高我们的临床治疗科研水平，为广大患者造福。

林振堂 1986.10

原著前言

新疗法不断增加，需要用恰当方法确定它们是否对病人真正有益。为获得对疗法效果的可靠评价，以随机对照作为临床试验的主要方法，已经得到广泛应用。本书的目的是想从实用方面阐明临床试验的基本原理，特别强调它的科学性和统计方法的应用，另外对伦理和组织方面的问题，也将作讨论。

本书试图向广大读者介绍一些易于理解的临床试验方法，希望对临床医生及与实施试验有关的其他人员有所裨益，特别希望能促使临床医生更广泛地开展或协作进行经过严格设计的试验，以解决有关的治疗课题。

在临床试验的计划阶段和对试验结果的分析与解释阶段，合理地应用统计方法，具有同样的重要性。鉴于许多临床医生等缺乏专门的数学知识，对理解统计学概念有一定困难。因此，我采用非数学方法叙述有关统计学问题，可能有助于临床试验的实践。本书第13章介绍了统计分析的基本原理，可能对有些读者是有益的。

作为一名医学统计工作者，相信临床试验的实施，需要临床工作者、组织工作者和统计工作者密切地配合，才能获得圆满的结果。

我希望本书能促使其他统计工作者为医学研究多作一些

普及工作，少作一些理论探讨。

本书不是抽象方法的汇集，而是通过实际例子介绍每一个观点，使读者能实际体会临床试验的要旨，作为增加医学知识的实用指南。

S.J. 波可克

目 录

第一章	临床试验的理论基础	(1)
1.1	临床试验的类型	(1)
1.2	临床试验中的对照和科学方法	(5)
1.3	原发性乳腺癌的临床试验	(9)
第二章	临床试验的历史发展	(19)
2.1	1950年以前的临床试验	(19)
2.2	1950年以后的临床试验	(24)
2.3	美国的癌化疗	(28)
2.4	急性心肌梗塞的治疗	(32)
2.5	制药工业	(34)
第三章	组织和计划	(37)
3.1	研究方案	(37)
3.2	管理、工作人员和资金	(42)
3.3	病人的选择	(46)
3.4	治疗计划表	(50)
3.5	病人疗效评价	(54)
第四章	为随机对照试验作辩解	(65)
4.1	无对照试验问题	(66)
4.2	历史对照问题	(70)
4.3	同时非随机对照中的问题	(77)
4.4	随机化是否可行	(81)
第五章	随机化方法	(85)
5.1	病人登记	(85)
5.2	随机化表的制备	(91)

5.3	分层随机化	(100)
5.4	不等随机化	(109)
第六章	盲法和安慰剂	(112)
6.1	为双盲试验作辩解	(112)
6.2	双盲试验的实施	(116)
6.3	盲法何时可行	(121)
第七章	伦理问题	(125)
7.1	医学进步和对病人的个别治疗	(125)
7.2	征求病人同意	(132)
第八章	交叉试验	(137)
8.1	“病人内”对比	(137)
8.2	两疗程交叉设计	(139)
8.3	交叉试验的数据分析和解释	(143)
8.4	多疗程交叉设计	(150)
第九章	临床试验的规模	(153)
9.1	确定试验规模的统计方法	(153)
9.2	试验规模的实际评定	(164)
9.3	小规模试验的不适当性	(169)
9.4	多中心试验	(171)
9.5	治疗组数与因子设计	(176)
第十章	试验进程的监督	(181)
10.1	进行监督的理由	(181)
10.2	中期分析	(183)
10.3	屡次的显著性检验——成组的序贯分析	(189)
10.4	连续的序贯设计	(200)
第十一章	表格与数据管理	(207)
11.1	表格的设计	(207)
11.2	数据管理	(216)

11.3	计算机的应用.....	(219)
第十二章	方案偏差	(229)
12.1	不合格的患者.....	(230)
12.2	不服从和未完成评价.....	(233)
12.3	分析中应包括退出病例.....	(238)
第十三章	统计分析的基本原则	(245)
13.1	数据描述.....	(245)
13.2	显著性检验.....	(258)
13.3	估计和置信界限.....	(270)
第十四章	其它方面的数据分析	(277)
14.1	预后因素.....	(277)
14.2	存活数据分析.....	(290)
14.3	数据的多重性.....	(300)
第十五章	结果的发表和解释	(309)
15.1	试验报告及其严格评价.....	(309)
15.2	假阳性过多.....	(316)
15.3	综合证明和整体策略.....	(320)
索引		(330)

第一章 临床试验的理论基础

过去对疾病疗法改进的评价方法一直是不完全和带有偶然性的，直到近些年才逐步得到广泛承认。只有根据科学试验原则组织实施的临床试验，才是提供评价新疗法疗效和安全性的唯一可靠基础。因此，本书的最大目的是阐明在进行有效的和有意义的临床试验时，所必不可少的有关科学和统计学的主要问题。此外，还要讨论一些有关临床试验的伦理和如何组织实施等问题。下文在介绍有关临床试验的历史回顾、现在情况和未来策略时，将对这些问题在方法学方面提供一个轮廓。

1.1 临床试验的类型

首先，我们必须弄清楚“临床试验”的精确含义。简言之，凡是以病人为研究对象，经过设计、探索对符合预定条件的患者，有计划地实施最佳疗法的任何一种类型的试验均称之为临床试验。从本质上也可以认为，临床试验是一个利用有限的病人样本 (Sample) 所得的若干治疗结果，推论如何对未来大量病人总体 (Population) 应用某一治疗方法的问题。

动物试验显然是不符合这个定义的。健康志愿者试验虽

然对患者的疗效只能提供间接证据，但它毕竟是在人体上揭示新疗法作用的重要的第一步，故也可勉强包括在我们的定义之内。

疫苗的现场试验和首次预防试验的受试者，虽处于症状尚未发生以前的状态（如高胆固醇血症），但这和明显患病者治疗时同样也涉及科学和伦理上的问题，所以，在适当时候，也将加以叙述。

个案病例研究，如将一例病人治疗和疗效的典型经验作为新事物报道，是不能算作真正的临床试验的。因为，同样的病人会因生物学变异，几乎可以肯定对某一疗法会出现不同的效应。根据某一个病例的有效经验，推论用同样方法治疗未来病人普遍都有效是不恰当的。因此，真正重要的一个问题，是用足够的病人分成若干不同治疗组，并进行可以置信的治疗对比，才能得出较为可靠的结论。

另一个问题是关于回顾性调查，这是对过去用各种不同方法治疗病人所得结果的检验。在这类无计划的观察研究中，包含有严重的潜在偏倚（如疗效相当高的疗法，用于预后相当差的病人，可能人为地降低了疗效），所以对不同疗法进行相互比较，很少有助于作出可信的评价。因此，作为非随机化的不适当案例除在第4章将提到以外，对这类研究不考虑列入临床试验和作更多的讨论。

在早期就考虑到从不同途径对临床试验进行分类是有用的。首先是治疗的类型。大多数临床试验都涉及药物疗效评价，而且其中多数是制药公司感兴趣并有财务支援的。但是，临床试验也包括其它形式的治疗。例如：外科操作方法，癌症的放疗，不同形式的医学咨询（如心脏病发作后的

饮食、运动指导），以及对病人护理方法的选择（如腹股沟肠疝手术后的家庭护理或医院护理），都应看作是不同方式的治疗，都可以用临床试验进行评价。可惜，对这些非药物治疗和保健方法，一般都未能用经过妥善设计的临床试验进行评价。下文将对这一题目进行讨论。

在制药公司内部的药物试验，常常将试验分为四期。这只是如何开展某药治某病的临床试验的一般性准则，不能看作是不许打破的硬性规定。

I 期试验：临床药理和毒性作用

在人体上首先要作的试验，是观察药物的安全性，而不是药效。所以经常在志愿者身上，而且往往是在制药厂职员身上实施。第一个目标是，确定可以接受的一次服药剂量（即不致引起副作用的剂量）。这经常是采用志愿者根据预先规定的程序，逐步增加服药剂量的方法，它也涉及药物代谢和生物利用率问题。稍后，将进行多次服药剂量的研究，以确定适合Ⅱ期试验要用的剂量程序。在正常志愿者研究完成以后，在第Ⅰ期中也可以让患者进行初步试验。按标准，第Ⅰ期研究需要志愿者和患者的总数约为20~80例。

Ⅱ期试验：临床治疗效果的初步探讨

本期是对药物的疗效和安全性进行较小规模的试验研究，但要求对每一个病人都进行严密的观察。第Ⅱ期试验有时也用作筛选手段，从大量无效或毒性大的药物中选出很少、但真正有效的药物。这样，才能着手将选出的药物作Ⅲ期试验。对一种药物的Ⅱ期试验，很少需要患者数超过100~

200例。

Ⅲ期试验：治疗的全面评价

在某药初步显示有较好的疗效以后，必须用相当大量病情相同的患者，与现行的标准疗法进行大规模的对比试验。对某些人说，临床试验这一名词与“全面的Ⅲ期试验”是同义语，是临床研究某一新疗法最严密、最广泛的形式。因此，本书将重点叙述实施Ⅲ期试验的一些原则。

Ⅳ期试验：销售后的监视

经过以上研究程序，使药品得到批准销售后，还需要进行更重要的调查，监视有无副作用，並同时进行大规模的长期致病率和病死率的研究。“Ⅳ期试验”这一名称有时也被用于宣传新药，目的是使更多的临床医师注意这一新药。后一类型调查的科学价值有限，不应算作临床试验研究的组成部分。

这种由制药公司主办的药物试验分期方法，不可避免地会使真正的临床药物研究过程过于简化。但是，早期（第Ⅰ、Ⅱ期）的人体研究，有其特殊的作用，它可以解决Ⅲ期试验进行前需要完成的许多工作，如组织实施办法、伦理学和科学性等方面的问题。在美国食品和药物管理处（1977年）发布的药物开发程序准则中，包括有对15种特定疾病如何建立药物的第Ⅰ至Ⅲ期试验的有关内容。

应该记住，每一个制药厂都有一个与临床试验同等重要的临床前研究程序，如新药合成、各种动物实验、测定药物代谢以及稍后的药物效力鉴定，尤其是毒力试验等。自从发

生了反应停一种有使胎儿畸形的药物事件，促使许多国家制定了相应的药品管理法规。此后，动物试验的规模愈来愈大，试验的科学性也有了较大的提高，並规定任何药物都必须先经过严格的动物安全试验之后，才允许用于临床试验。

I 至 III 期的分期系统，並不只限于制药工业，也可以作为临床试验的一般性准则。例如，癌症的化疗和放疗研究计划，是美国国立卫生研究所投入大部分研究经费的重点项目；也按第 I 至 III 期的顺序，分期组织实施了临床试验。因为这时所用的抗癌药物和放疗都有很大的副作用，所以 I 期试验是在患者身上，而不是在正常志愿者身上进行的。

开发外科的新技术操作，也可以广泛地遵循相似的计划进行，並认为 I 期试验是开发外科技术的基本方法。但是，在外科，也有少数试验采用了试验设计周到的 III 期试验。

1.2 临床试验中的对照和科学方法

本节将集中叙述 III 期试验问题，並探讨实施中的科学基础。当然，临床研究中最重要的是提出好的改进治疗设想。只有当临床研究者有见识、並有想象力，能提出对病人有益的革新疗法时，研究工作才能有所进展。自然，当临床前研究和早期试验表明有希望时，任何新疗法的提出者，都容易对新疗法的潜力感到兴趣。特别是制药公司在全面试验未实施以前，都对他的产品满怀希望。遗憾的是，有许多新疗法原来並不象希望的那样有效，一旦接受正确设计的 III 期试验的严格检验，原来的期望就可能会落空。这可参照 Gilbert 等（1977 年）有关外科和麻醉药的例子。

Ⅲ期试验的基本原则就是对比，也就是需要用一组新疗法的病人，和一组用标准疗法的病人，进行对比。如果没有真正有意义的标准疗法的话，也可以用不治疗的病人作对照组。为获得新疗法的无偏评价，通常需要随机地把每个病人分配到新疗法或标准疗法组。（见第4、5章细节）。一般承认随机对照试验是进行临床研究最可靠的方法。

在此，用几个例子说明，随机对照试验的意义。表1.1列举了6个试验例。

（1）细菌性脑膜炎试验，情况是简单的。以新疗法（先锋霉素）与标准疗法（氨苄青霉素和氯霉素合用）比较，观察前者是否对杀灭细菌更为有效。

表1.1 随机对照试验示例

文 献	疾病种类	试 验 组	对 照 组
瑞典研究组（1982年）	细菌性脑膜炎	先 锋 霉 素	氨苄青霉素+氯霉素
苯磺唑酮预防再梗塞试验（1980年）	急性心肌梗塞	苯 磺 唑 酮	安 慰 剂
医学研究委员会工作组（1977年）	轻症高血压	苯氯噻嗪、心得安	安 慰 剂
Douglass等（1978年）	进展期结肠直肠癌	甲基-CCNU、 5-FU+TG、 5-FU+CTX	5-FU
Garraway等（1980年）	卒中	专门卒中护理小组	一般内科病房护理
Langlands等（1980年）	可手术乳腺癌	根治切除术	单纯切除+放疗

* 5-FU=5-氟尿嘧啶，TG=6-硫代鸟嘌呤，CTX=环磷酰胺

（2）苯磺唑酮试验，反映另一种常见情况。以新疗法（苯磺唑酮）与安慰剂（无活性的口服药片）进行对比。对

照组的心肌梗塞病人未接受任何有效的治疗，目的是观察在发生梗塞后的第1年，苯磺唑酮治疗组的病死率是否有所减少。

(3) 轻度高血压试验，用两种有效疗法与安慰剂比较，观察哪一种能减少心血管和肾脏疾病的发病率和病死率。

(4) 进展期结肠直肠癌试验，采用了不常用的3种新疗法与标准药5-FU(氟脲嘧啶)进行对比。在大多数试验中只用两组(新法和标准法)对比，用多组比较时，一般应慎重考虑，因为这给收集治疗所需的足够病例数增加了困难。

(5) 最后两种试验是为了提醒人们，临床试验也可用于评价非药物治疗的效果。卒中试验是对病人护理方法进行比较，即观察卒中病人是否由专门小组护理比一般内科护理更能使病人恢复得好些。

(6) 乳腺癌试验，是一种不常用的方法。是对两种不同手术处理方法(根治乳腺切除和单纯乳腺切除加放疗)进行比较。两种疗法中的任一种方法，在不同医院均可能作为标准疗法，即一种疗法对另一疗法都是对照组。这种试验对解决治疗上长期存在着的争议，是极为重要的。在此以前还未见到用随机对照试验加以解决的先例。

现在我想探讨：如果按科学原则要求，临床试验应如何进行的问题。图1.1表示试验过程的一般顺序。从一种疗法可能有改进的最初设想出发，第一步我们需要作有疗效和安全性的特殊假设，对试验目标作更为精确的规定。具体说，我们必须对病人的类型，供疗效对比的不同疗法，以及对每