

临床试验所需样本量之管见

张学中, 谭学瑞, 潘红星

张学中, 北京高血压联盟研究所 北京市 100039
谭学瑞, 潘红星, 汕头大学医学院第一附属医院 广东省汕头市 515041

张学中, 1963年毕业于清华大学工程物理系, 工学学士, 研究员, 教授, 博士生导师, 主要从事医学统计学方面的研究。

国家自然科学基金资助, No. 30271158

通讯作者: 谭学瑞, 515041, 广东省汕头市市长平路57号, 汕头大学医学院第一附属医院. tanxuerui@vip.sina.com

电话: 0754-8258290-3455 传真: 0754-8611690

收稿日期: 2006-03-14 接受日期: 2006-04-16

摘要

本讨论了我国有关部门关于临床试验样本量规定中的2个问题, 提出下列看法和理解: 对 I - III 期临床试验的样本量应该规定上、下限, IV 期临床试验的样本量可不加限定。

关键词: 临床试验; 样本数; 样本量; 效应大小

张学中, 谭学瑞, 潘红星. 临床试验所需样本量之管见. 世界华人消化杂志 2006; 14(14): 1339-1340

<http://www.wjgnet.com/1009-3079/14/1339.asp>

0 引言

在我国, 不论1999年发布的《药品临床试验管理规范》, 2002年发布的《药品注册管理办法》(试行), 还是2005年发布的《药品注册管理办法》, 都对临床试验的样本量做了规定, 他们是一致的: (1) 临床试验的样本量应当符合统计学要求和最低样本量要求; (2) 临床试验的最低样本量在试验组的要求是: I 期为20-30例, II 期为100例, III 期为300例, IV 期为2000例. 对于这种规定, 孙瑞元 *et al*^[1]曾介绍过, Hsiao *et al*^[2]也提出了一些观点和意见. 本文的讨论不仅仅局限于统计学问题, 而主要是从临床试验的原则^[3]来综合考虑, 讨论临床试验所需样本量的3个问题。

1 I 期临床试验的样本量界限

我国关于 I 期临床试验的样本量是规定有上、下限的, 且与国外在数值上似乎是一致的. 但对这一数值的理解却有出入(表1)。

我国规定的范围比国外小, 但这是一模糊的最小例数. 也就是说, 该规定容许有更多的例

表 1 I 期临床试验的样本量

制定者	样本量大小	例数的说明
我国	20-30	此为最小例数
欧盟	20-50	建议的例数范围
美国	20-80	建议的例数范围

表 2 II、III 期临床试验的样本量

试验分期	制定者	样本量	例数的说明
II期	我国	>100	此为最小例数
II期	欧盟	100-500	建议的例数范围
II期	美国	100-300	建议的例数范围
III期	我国	>300	此为最小例数
III期	欧盟	500-5000	建议的例数范围
III期	美国	300-3000	建议的例数范围

■背景资料

临床试验样本量的确定是至关重要的, 但国内外规定及文献建议不尽一致, 需要正确理解. 临床试验样本量的规定不只是统计学问题, 更重要的是要符合临床试验的原则. 以往规定或文献强调加大样本可得出显著的统计学结论, 不一定有实际意义。

数. 而且没有上限的限制. 由于 I 期临床试验是在药理实验室的工作和动物实验之后首次过渡到人体的试验, 所以, 一定要精心设计^[4]. 选择适当研究总体, 控制随机误差, 达到一定的预计效应大小等^[5], 使受试对象数的上限不能太大. 如果一种新药只有很小的效益, 仅靠加大样本而得出非常显著的统计学结论, 是没有实际意义的. 根据赫尔辛基宣言精神, 为维护受试者权益, 应规定出受试者的上限. 又因为有些 I 期临床试验是在健康志愿者中进行的, 所以, 例数不宜太多, 最好称为样本数, 而不是病例数。

2 II、III 期临床试验应有上限

我国关于 II 期和 III 期临床试验的样本量的规定是没有上限的, 与国外在数值上的比较见表2. II 期和 III 期试验的目的不单是要确认哪个药或疗法有效, 而主要的是要求他的有效性应达到一定程度^[6], 即在设计样本量的大小时, 要让临床试验的处理效应足够大, 从而限制了样本量的大小, 因而样本要有上限. 如果只是靠把样本加大到“足够大”才能得到统计学上的显著差别, 这种药物及其临床试验是不应被批准的. 另一个制定临床试验上限的原因是, 从保护受试者权益出发, 在批准上市前, 对参与试验的样本量应有限制, 使受影响的人数尽量少。

■创新盘点

对 I - III 期临床试验的样本量应该规定上、下限, IV 期临床试验样本量可不加限定。

■名词解释

I型错误:即某药物或疗法事实上无效而试验后判为有效的错误。

为了科学估算样本量,需要对I型错误的概率和把握度以及试验的效应大小作出假设,这些往往是因问题而异的。文献中的建议也各有所异^[7-12],例如,按预试验或已往的数据,由极差D,估计标准差SD,有人用 $SD = 1/4D$,有人用 $SD = 1/6D$,因而得出不同估计结果。面对诸多不确定的选择,规定适当范围是合理的。

3 IV期临床试验的样本量可不作规定

IV期临床试验是在药品上市后进行的,这类试验因目的不同,所需样本量也会有很大差异。如果是为了了解可能会出现罕见的副作用,样本量可能会相当大。如果与某种标准疗法比较,要达到一定把握度,所需样本量也会大些。这时候应该认真估算,算出样本量可能达数万。但也有例外的情况,在特定地区、对特定疾病的疗效做简单对比的试验,对于全是已上市的药物,不要求较大样本也是容许的。国外IV期临床试验的批准权限和赞助者都与前三期不同。我国规定IV期临床试验样本量最小大于2000,似乎没考虑到他与前三期的不同之处。既然是已经

上市的药品,其包括样本量的各种规定就可放宽,甚至对样本量不作规定,如国外那样。

4 参考文献

- 1 孙瑞元,王学融,李玉红. 新药临床研究的例数估算问题. 中国临床药理学杂志 1999; 8: 559-561
- 2 Hsiao CF, Xu JZ, Liu JP. A two-stage design for bridging studies. *J Biopharm Stat* 2005; 15: 75-83
- 3 谭学瑞,张学中. 临床试验刍议. 世界华人消化杂志 2005; 13: 2173-2178
- 4 Ensign LG, Gehan EA, Kamen DS, Thall PF. An optimal three-stage design for phase II clinical trials. *Stat Med* 1994; 13: 1727-1736
- 5 Logan BR. Optimal two-stage randomized phase II clinical trials. *Clin Trials* 2005; 2: 5-12
- 6 Stallard N. Approximately optimal designs for phase II clinical studies. *J Biopharm Stat* 1998; 8: 469-487
- 7 郑青山,孙瑞元,陈志扬. 临床两组药物等效判断方法及其辨析. 中国临床药理学与治疗学 2002; 7: 559-561
- 8 刘玉秀,姚晨,陈峰,陈启光,苏炳华. 非劣性/等效性试验的样本含量估计及把握度分析. 中国卫生统计 2004; 21: 31-35
- 9 赵国龙,王曙炎. 多中心两组临床试验所需样本量的测定. 中国医院统计 2003; 10: 27-29
- 10 李大庆,李颢,王志萍,李会庆. 基因与疾病研究的样本量计算方法. 中国公共卫生 2004; 20: 886-887
- 11 谷红梅,李康. 诊断试验ROC参数估计双正态样本量估计方法探讨. 中国卫生统计 2006; 23: 2-4

电编 李琪 编辑 潘伯荣

ISSN 1009-3079 CN 14-1260/R 2006年版权归世界胃肠病学杂志社

• 消息 •

2006广州肝病防治研究新进展学术报告会

本刊讯 2006广州肝病防治研究新进展学术报告会将于2006-07-01/05在广州举行,现将征文通知公布如下:

1 稿件要求

要求摘要和正文各一份,并附软盘或寄到电子信箱。

2 联系方式

广州市广州大道北1838号南文医科大学南方医院消化病研究所(510515),白岚 教授 收. 电话: 020-61641658;

E-mail: gdsgzbxh@163.com.