

精神卫生科研如何严格遵守试验设计四原则之均衡原则

张效嘉¹ 胡良平^{1 2*}

(1. 军事医学科学院生物医学统计学咨询中心, 北京 100850;

2. 世界中医药联合会临床科研统计学专业委员会, 北京 100029

* 通信作者: 胡良平, E-mail: lphu812@sina.com)

【摘要】 明确阐释在进行精神卫生临床试验设计时, 应正确把握“均衡原则”的意义和要领。从基本常识出发, 并基于精神卫生科研的特点, 寻找和发现在此研究领域中, 怎样做才能被称为严格遵守了“均衡原则”。通过结合本专业的特点, 并结合实例, 获得如下的结果, 即在进行精神卫生临床试验设计时, 必须把握好以下三个方面: ①正确领悟均衡的涵义; ②正确把握实现对比组间均衡的方略; ③应清楚在临床试验研究的每一个环节上, 严格落实标准化操作规程和实时精准进行质量控制, 是确保对比组间均衡的重要举措。在如何严格遵守均衡原则问题上, 正确把握好前述提及的三个方面的本质, 就是抓住了问题的本质, 是提高临床试验研究质量的一个重要环节。

【关键词】 精神卫生; 临床试验设计; 均衡原则; 标准操作规程; 质量控制

中图分类号: R195.1

文献标识码: A

doi: 10.11886/j.issn.1007-3256.2016.04.004

Approaches to strictly observe four principles in the clinical trial design of mental health research: the "balance principle"

ZHANG Xiao-jia¹ HU Liang-ping^{1 2*}

(1. Consulting Center of Biomedical Statistics, Academy of Military Medical Sciences, Beijing 100850, China;

2. Specialty Committee of Clinical Scientific Research Statistics of World Federation of Chinese Medicine Societies, Beijing 100029, China

* Corresponding author: HU Liang-ping, E-mail: lphu812@sina.com)

【Abstract】 This paper illustrated that the implication and essential of balance principle should be taken into consideration in the clinical trial design of mental health. By combining common sense with characteristics of mental health research, we explored approaches to strictly observe balance principle in the field. The paper suggested that in mental health clinical trial design, researchers should hold three aspects in terms of balance principle as follows. First, researchers should realize the implication of balance. Second, researchers should master the strategy to realize the equilibrium between groups. Last, researchers should understand that it is significant to strictly implement standardized operation procedures and precisely perform quality control for ensuring equilibrium between groups. The article concluded that in terms of balance principle, the implement of three aspects mentioned above is a significant element to enhance the quality of clinical trial study.

【Key words】 Mental health; Clinical trial design; Balance principle; Standardized operation procedures; Quality control

1 概 述

1.1 何为均衡与均衡原则^[1-6]

1.1.1 问题的提出

进行一项研究, 一般都会涉及两组或两组以上患者的药物疗效或治疗措施效果之间的对比问题, 若各组的患者之间在很多方面几乎都一致, 其区别仅仅在于药物的种类不同或同一种药物的剂量不同或手术的类型不同, 则研究结果就具有较高的说服力。进行科研设计时要求各组患者之间在以下诸多方面应尽可能一致, 这就是“均衡”的真实含义, 即病情构成(轻度、中度、重度的比例), 平均病程及病程分布, 平均年龄及年龄分布, 性别构成(男女比

例) 等任何可能影响疗效评价的各方面。它们在统计学上常被称为重要的非试验因素。

1.1.2 何为均衡

均衡是对一个试验因素的各个水平而言的。在一项试验研究中, 若同时需要考察的试验因素为两个或两个以上时, 谈论“均衡”仍应限定在每个试验因素各水平之间。例如, 在考察三种药物的疗效时, 三个药物组中的患者若用了不同的剂量, 则三个药物组在“剂量”这一重要的非试验因素上是不均衡的。

1.1.3 何为均衡原则

均衡原则就是要求同一个试验因素各水平组之

间除了所考察的因素取不同水平外,在其他试验因素和一切非试验因素方面都应达到均衡一致。例如某研究中,一组为正常人,测定其血压值,另一组是高血压患者,测定治疗前、治疗后的血压值。很多人将这三组定量数据视为来自“单因素三水平设计的定量资料”,这显然是不妥的!为什么?请读者自己独立思考后作出解答。

1.2 均衡原则的作用

确保试验因素各水平组间不受其他试验因素或重要非试验因素的不平衡的干扰和影响,以便使所考察的试验因素取不同水平条件下对观测结果影响的效应大小真实地显露出来。

1.3 一个违背均衡原则的经典案例

【例 1】为了弄清经常吃含赖氨酸的面包是否对少年儿童生长发育有帮助,有人设计了如下的试验:对住校的某班级学生,随机让一半学生每天上午和下午课间各吃一个含赖氨酸的面包,另一半学生课间不吃面包,一年后比较这两组学生的身体发育情况,从而得出“赖氨酸”有利于少年儿童生长发育的结论。试分析,在这个实例中,研究者犯了什么错误?

【解说】因为课间学生肚子会有些饿,即使吃一个不含赖氨酸的面包也会有助于少年儿童生长发育,这两组学生在重要非试验因素(是否吃面包)方面是不均衡的,故其结论是不可信的。

只有随机地决定一半学生吃不含赖氨酸的面包,另一半学生吃含赖氨酸的面包(每个面包的大小、面粉的质量等方面应完全相同),其他(如锻炼身体情况、作息时间等)各方面两组学生应尽可能一致;最好在试验研究期间内,全部学生都生活在学校(因为一旦学生放假回家了,两组学生之间的均衡性就失控了);当然,还有很多需要推敲的细节方面,两组学生都应尽可能一致。一年结束时再比较,才有可能说明“赖氨酸”的价值有多大。

均衡性的道理很简单,即对比组受试对象之间除了所考察的因素取不同水平外,其他一切方面都应尽可能一致,但做到不容易。研究设计者要有很强烈的责任感,要清醒地认识到“违反均衡原则”将会得出错误结论。

2 提高对比组间均衡性的方略^[1-6]

2.1 从人的主观能动性角度考虑

从人的主观能动性角度出发,提高对比组间均

衡性应做到如下三点:①研究者应具有较好的专业知识和统计学知识,对整个试验研究的各个方面有整体的把握,能够制定出较为理想的试验设计方案;②试验设计方案制定出来以后,应该请从事本专业研究的有经验的专家帮助审阅和修改设计方案;③最好再请一两位有丰富试验设计经验的统计学工作者把关,使试验设计方案得到进一步完善。尤其是在三要素、四原则、设计类型和比较类型方面,特别是其中的对照与均衡原则方面,统计学工作者往往可以提出很多富有建设性的建议。

2.2 从研究的整个过程角度考虑

从研究的整个过程角度考虑,提高对比组间均衡性应做到如下六点:①制定合理的纳入和排除受试对象的标准,对参与的研究者进行严格的技术培训;②对试验条件进行标准化管理(如仪器、试剂的校准,制定出科学的诊断、测定和评价标准);③严格按随机(注意按重要的非试验因素进行分层随机)、对照(特别要强调设置合理的对照组)和重复(各试验条件下应有足够的样本量,最好用特定的公式估算出来)原则选取和分配受试对象;④选取合适的设计类型安排多个试验因素;⑤注意整个试验过程中的质量控制;⑥在统计分析时,尽可能利用更多的信息,采取相应的统计分析方法(多因素设计定量资料的协方差分析)消除混杂因素对观测结果的影响。

3 精神卫生学术论文中均衡性方面存在进一步改进之处

翻阅本刊发表的学术论文,在如何遵守均衡原则这个问题上,绝大多数学术论文通常有如下两种做法。

做法一:基于入组标准和排除标准,呈现纳入研究的受试对象的基线情况。

先给出受试对象的“入组标准”和“排除标准”,然后选取符合入组标准且不符合排除标准的受试对象;接着,以“两组一般资料比较”为统计表的标题^[7-10],罗列出试验组与对照组受试对象在若干个人口学特征(通常为年龄等)方面的平均水平或频数构成情况(通常为性别等)以及在某些定量评价指标上取值的平均水平,并在两组之间进行差异性分析,以显示组间在所考察的项目(包括原因变量与结果变量)上的均衡性好坏。

做法二:基于入组标准和排除标准,呈现接受不同处理后的受试对象的疗效情况。

先给出受试对象的“入组标准”和“排除标准”，然后选取符合入组标准且不符合排除标准的受试对象；接着，用文字描述对比组之间在哪几个人口学特征（年龄、受教育年限等）或病程方面，“两组一般资料比较差异均无统计学意义（ P 均 >0.05 ）”^[11-14]；最后，直接呈现对比组受试对象接受不同处理后的疗效情况。

在以上两种做法的学术论文中，有些作者没有提及是否采取了某种随机化方法分配受试对象，有些虽然提到了“随机分组”，但究竟采取怎样的随机化方法，交代得并不清楚。

以上两种做法，看上去似乎都“无可厚非”。但是，当人们仔细思考之后，可能会提出如下的问题：若在进行疗效评价之前，发现对比组受试对象的某些项目（包括人口学特征和评价指标）在基线期的差异有统计学意义，该怎么办？换句话说，当临床试验研究资料在对比组之间严重违背了“均衡原则”时，接下去关于疗效分析和安全性分析的可信度就很低了。

基于以上在临床科研中普遍存在的问题，建议在临床试验研究过程中，应自始至终牢记“均衡原则”，应将临床试验实施过程划分成多个阶段，并在每个阶段中找出导致对比组之间在基线期可能存在不均衡的“关键点”。从而采取最合适、最有效的措施，将一切“可能导致不均衡性的关键点”消灭在萌芽状态。例如，在制订受试对象质量标准阶段，应将五个标准^[15]制订齐全并尽可能使其达到“无可挑剔”的程度；在实际挑选受试对象阶段，要在标准操作规程指导下，严格按前面提及的“质量标准”纳入受试对象；在将已入选的受试对象进行分组时，应尽可能找准找全对主要评价指标有影响的重要非试验因素，并基于这些因素对受试对象采取“分层随机化分组”；类似的，还有很多阶段或环节，就不一一赘述。总之，“均衡原则”本身并没有具体的“操作内容”，但它要求在临床试验每一个阶段上的“操作”都必须以其为“准绳”，要努力做到“尽善尽美”，甚至是“完美无缺”。

4 对均衡原则方面出现的错误进行辨析与释疑^[16-18]

【例 1】为研究全髋关节置换术中向关节周围局部注射混合镇痛液的疗效。将 76 例行全髋关节置换术的患者分成两组，住院号为单数的患者分入注射组，患者术中关节周围局部注射镇痛液，共 38 例，男性 14 例，女性 24 例，平均年龄 73 岁，手术时间

(88 ± 15) min；住院号为偶数的患者分入对照组，不注射镇痛液，共 38 例，男性 12 例，女性 26 例，平均年龄 75 岁，手术时间(80 ± 20) min。术后采用直观模拟疼痛量表(Visual Analog Scale, VAS) 评估患者疼痛程度；所有患者术后 24 h 内使用患者自控镇痛泵(patient-controlled analgesia, PCA)，于不同时间点分别记录 PCA 的用量。得到的结论是：全髋关节置换术中向关节周围注射混合镇痛液能够显著降低患者术后疼痛并减少镇痛泵的用量。

【对差错的辨析与释疑】该研究资料在对受试对象进行分组时违反了随机原则，“住院号为单数的患者分入注射组，住院号为偶数的患者分入对照组”这种分组方式可能会导致两组患者在某些重要的非试验因素方面相差悬殊，两组间缺乏均衡性，并不是一种随机化的分组方法，同时也违反了均衡原则。这样分组，有可能导致一组患者的病情较轻，另一组患者的病情较重，从而导致两组患者最后测得的疼痛评分和 PCA 的用量不能真实地反应术中局部注射混合镇痛液的疗效。并且该资料也没有对两组患者的病情、病程等基本情况进行详细的描述和相应的统计分析处理，无法确认两组间的一般情况是否均衡，这样的临床试验设计最后得到的结论缺乏可靠性和说服力。

【例 2】有人在研究耐力训练与提高战士体质的关系时，设计了如下的试验：以血乳酸为主要观察指标，采用 20 名连队战士按训练方案进行耐力训练，以机关同龄的 20 名战士为对照，对照组进行日常活动，观察经 4 周训练后两组战士进行一定量运动时血乳酸的变化，结果见表 1。

结论：按此方案训练能够降低运动时血乳酸蓄积，提高战士的耐力。

表 1 两组战士训练前后血乳酸的观测结果

试验分组	血乳酸(mg/L)	
	训练前	训练后
训练组	38.2 ± 3.5	33.1 ± 3.1**△△
对照组	40.0 ± 4.0	38.9 ± 3.5

注：**与训练前比 $P < 0.01$ ；△△与对照组比 $P < 0.01$

【对差错的辨析与释疑】对照组选择不合理，对照组除训练因素外，其他条件应尽可能与训练组一致，而机关兵和连队战士由于工作性质不同，体能的基础可能存在较大差别，一开始两组之间就不具备可比性。

此研究回答了一个问题，即耐力训练提高了战士的耐力，这种结论毫无意义。因人们早就认识到

了训练和不训练不一样,而应研究不同训练方案对提高战士体质之间的差别有无显著性意义,以探索最大限度提高战士体质的最佳方案。另外,分组时连队战士和机关战士都应随机抽取,然后将各自的总样本含量的一半随机分入训练组和对照组。

【例 3】某临床医生收集到临床资料:不稳定性心绞痛患者 60 例,随机分为三组,每组 20 例。三组病例性别、年龄及病程构成差异无统计学意义。布比卡因组(A 组)先胸椎硬膜外穿刺,首次注入 0.5% 布比卡因 5 mL,然后连接电脑控制微量注射泵以 4 mL/h 速度连续给予 0.0625% 布比卡因,连续 9~22 d,平均(15.6±2.8) d;布比卡因吗啡复合液组(B 组)胸椎硬膜外穿刺后,用微量泵以 4 mL/h 速度给予 0.0625% 布比卡因与 10 μg/mL 吗啡复合液,连续 7~16 d,平均(12.6±2.3) d;对照组(C 组)采用内科常规治疗心绞痛的药物如消心痛、倍他乐克等连续治疗 13~27 d,平均(19.8±2.6) d。三组根据病情需要加用静脉滴注硝酸甘油。观察指标为:①心绞痛变化,记录每日发作次数和持续时间,以观察心绞痛症状治疗的效果;②心电图变化,治疗前与治疗后每天做心电图,并记录 ST-T 变化程度,到治疗截止时心电图恢复情况(恢复、部分恢复、无恢复);③观察治疗前后血压和心率的变化。

【对差错的辨析与释疑】原文的目的是比较三种治疗方法的疗效,但在设计时却规定可根据病情需要加用硝酸甘油,我们知道硝酸甘油会改善心绞痛的症状,这就使得各组的效应既包括各药物的效应,又包括硝酸甘油的效应,混杂在一起,分析完毕也不能说清到底是三种药物的作用还是硝酸甘油的作用。另外,三组中各患者的治疗检测时间为 7~27 d 不等,一般说来,治疗的时间越长,其治疗效果就会越好。各组患者的治疗时间不等,对三种药物的疗效进行评价也是缺乏说服力的。进行试验设计时,对混杂因素的处理特别重要。有些方案表面上考虑了一些重要非试验因素的均衡问题,但如果考虑不周,哪怕忽略了一个混杂因素,也会导致试验研究的失败。

应控制非试验因素的干扰,各组患者禁止加用硝酸甘油,如果出于治疗的需要必须加用,要注明加用的次数和剂量,以便后期统计分析时加以处理。各组患者的治疗时间应根据专业知识确定统一的观察时间或多个重复测量的时间点,以使各组具有较好的可比性。

参考文献

- [1] 胡良平. 临床科研中如何正确把握均衡原则 [N]. 中国医学论坛报, 2006-04-13(14).
- [2] 胡良平, 贾元杰, 鲍晓蕾. 试验设计中均衡原则的把握与常见错误的辨析 [J]. 药学服务与研究, 2011, 11(1): 11-14.
- [3] 李子建, 刘惠刚, 胡良平. 医学论文中统计分析错误辨析与释疑——实验设计原则的正确把握(14) [J]. 中华医学杂志, 2004, 84(14): 1228-1230.
- [4] 胡良平, 李子建, 刘惠刚. 医学论文中统计分析错误辨析与释疑——实验设计原则的正确把握(15) [J]. 中华医学杂志, 2004, 84(15): 1315-1317.
- [5] 胡良平, 刘惠刚, 李子建. 医学论文中统计分析错误辨析与释疑——实验设计原则的正确把握(16) [J]. 中华医学杂志, 2004, 84(16): 1406-1408.
- [6] 胡良平, 李子建, 刘惠刚. 医学论文中统计分析错误辨析与释疑——实验设计原则的正确把握(17) [J]. 中华医学杂志, 2004, 84(17): 1481-1484.
- [7] 高帅, 张瑜良, 蔡丽莉. 阿立哌唑联合奥氮平对精神分裂症患者体质量、糖脂代谢影响的研究 [J]. 四川精神卫生, 2015, 28(2): 127-130.
- [8] 范洪峰. 帕利哌酮缓释片和齐拉西酮对精神分裂症患者认知功能和生活质量影响的对照研究 [J]. 四川精神卫生, 2015, 28(2): 134-137.
- [9] 范妮, 孙彬, 何红波. C-反应蛋白在精神分裂症患者中的阳性率及其与精神症状及住院时长的相关性分析 [J]. 四川精神卫生, 2015, 28(3): 208-210.
- [10] 陈景旭, 梁雪梅, 卞清涛, 等. 阿立哌唑治疗利培酮所致高催乳素血症效果的性别差异 [J]. 四川精神卫生, 2015, 28(3): 215-218.
- [11] 刘芙蓉, 周平, 刘联琦, 等. 伴有抑郁症状的精神分裂症患者的血清皮质醇及甲状腺功能特征 [J]. 四川精神卫生, 2015, 28(3): 224-226.
- [12] 熊洪莉, 陈君辉, 王雪. 心理护理对介入治疗围手术期肝癌患者焦虑抑郁水平影响的临床对照研究 [J]. 四川精神卫生, 2015, 28(3): 227-229.
- [13] 衣磊, 于慧, 崔维珍. 老年抑郁症患者认知功能损害的随访研究 [J]. 四川精神卫生, 2015, 28(3): 230-233.
- [14] 平军辉, 仲照希, 王东平. 齐拉西酮注射液治疗儿童精神分裂症急性激越症状 35 例 [J]. 四川精神卫生, 2015, 28(4): 314-316.
- [15] 郭春雪, 胡良平. 正确把握精神卫生临床试验设计三要素的要领(1)——受试对象 [J]. 四川精神卫生, 2016, 29(3): 197-201.
- [16] 胡良平, 李子建. 医学统计学基础与典型错误辨析 [M]. 北京: 军事医学科学出版社, 2003: 242-259.
- [17] 胡良平. 科研课题的研究设计与统计分析——提高学位论文统计学质量的对策 [M]. 北京: 军事医学科学出版社, 2010: 121-350.
- [18] 胡良平. 科研设计与统计分析 [M]. 北京: 军事医学科学出版社: 129-227.

(收稿日期: 2016-07-26)

(本文编辑: 陈霞)