

正确把握精神卫生临床试验设计三要素的要领(II) ——影响因素

胡 完,胡良平

(军事医学科学院生物医学统计学咨询中心,北京 100850

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“影响因素”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,怎样做才能被称为正确把握了“影响因素”。通过结合本专业的特点和实例,获得如下的结果,即在进行精神卫生临床试验设计时,必须把握好以下五个方面:①影响因素的分类;②确定因素及其水平的原则;③如何对试验因素进行标准化;④如何找准全重要的非试验因素;⑤如何有效控制重要非试验因素。在如何确定影响因素的问题上,正确把握好前述提及的五个方面,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;影响因素;重要非试验因素;标准化

中图分类号:R195.1

文献标识码:A

doi:10.11886/j.issn.1007-3256.2016.03.003

Correctly grasping the essentials of the "Three Key Points" in the clinical trial design of mental health:Part II the "influential factors"

HU Wan, HU Liang-ping*

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

*Corresponding Author: HU Liang-ping; E-mail: lphu812@sina.com)

【Abstract】 Definitely elaborating the meanings and essentials of grasping the influential factors correctly in the clinical trial design of mental health. Searching and discovering how to hold the influential factors correctly starting from the basic common sense and based on the characteristics of mental health study. The following results are obtained through combining the features with examples in this specialty, that is to say, we must grasp the following five aspects: ①the classification of influential factors; ②the principles of determining factors and their levels; ③how to make the experimental factors standardization; ④how to find non-experimental factors accurately and entirely; ⑤how to control important non-experimental factors effectively. Mastering the five aspects mentioned above correctly equals to gripe the innate of the issue and it is an important link to improve the quality of clinical trial study in answering the question of how to ascertain influential factors.

【Key words】 Mental health; Clinical trial design; Influential factors; Important non-experimental factors; Standardization

1 概 述

1.1 基本概念

人们经常说“吸烟是患肺癌的危险因素”,“饮酒是患肝癌的危险因素”,“盐摄入量过多是患肾脏疾病的危险因素”,“糖摄入量过多是患糖尿病的危险因素”,“有精神疾病家族史是患精神疾病的危险因素”等等。这些说法看起来似乎都“千真万确”,其实,在基本概念上就是错误的。错误的根源在于人们将“因素”与“原因”混为一谈。

若 A 事物的出现会引起 B 事物的发生,可以说 A 事物的出现是导致 B 事物发生的“原因”,而 B 事物的发生则是 A 事物出现所产生的“结果”;若人们关心的是“A 事物是否出现”与“B 事物是否发生”之间的关系时,就要分别考察在“A 事物出现”与“A

事物不出现”两种对立的前提下,“B 事物发生”的概率分别是多少。此时,正确的表述是:“A 事物是否出现”是一个“影响因素”,它有两个水平,一个是“A 事物出现”、另一个是“A 事物不出现”;而“B 事物是否发生”则是一个二值的结果变量,它有两个不同的取值,分别为“B 事物发生”与“B 事物不发生”。也就是说,“原因”与“结果”可以是“一对一”的,而“因素”与“结果”应当是“多对多”的。在统计学上,“因素”的多个“表现”被称为“水平”,而“结果”的多个“表现”被称为“取值”。

由此可知,之所以说前面提及的人们常说的那五句话是错误的,就是因为它们陈述的是“一对一”的因果关系,属于“原因”与“结果”之间的关系,而不是“因素”与“结果”之间的关系。以其中的最后一句为例,给出两种正确的表述如下:第一种,“有

精神疾病家族史是导致患精神疾病的重要原因之一”;第二种,“‘有无精神疾病家族史’是‘是否患精神疾病’的重要影响因素之一”。总之,在统计学上可以被称为“因素”的事物至少有两个水平,而被称为“原因”的事物可以只有一个。

1.2 与“因素”有关的名词概念

1.2.1 影响因素

在各种试验研究中,所有可能影响治疗效果或安全性的因素被统称为影响因素,常简称为“因素”或“因子”,它包括试验因素和非试验因素。影响因素通常可分为以下三大类:第一类,因素的水平为性质对立的两种,即“有”与“无”,例如,是否用 A 药(“不用”与“使用”是其两个水平);第二类,因素的水平为“数量型”或“有序型”,例如,温度的水平分为“20℃、40℃、60℃”,风速的水平分为“小、中、大”;第三类,因素的水平为“并列型”,性别的水平分为“男性、女性”,职业的水平分为“工人、农民、商人、公务员、军人”,药物种类分为“A 药、B 药、C 药”。

1.2.2 试验因素

试验因素是研究者希望通过试验考察其对试验结果是否有影响的性质对立(前面提及的第一类因素)或性质相同(前面提及的第二类或第三类因素)的不同试验条件的总称,包括物理的(如针刺、射线、理疗等)、化学的(如药物、毒物等)、生物的(如细菌、病毒等)以及社会的因素等;受试对象自身具有的、可影响其发展过程的某些因素(如性别、年龄等)也可作为试验因素。同一种因素还可能有不同水平,如药物的浓度高低、时间的长短、温度的高低等。

不同因素或同一因素的水平不同,就造成试验条件的多样性,包括:①单因素单水平(本质上应属于单因素两水平),这是科研中较常见的一种试验设计类型,如某降压药对原发性高血压患者降压作用的观察,比较时常以“正常血压”为基准(它是此项试验研究中降压药种类这个因素的理论水平),就属于这类试验;②单因素多水平,比较某药物不同剂量对某种疾病的疗效观察,便属于这类试验,特别是珍贵药物、毒性较大的药物或最佳新药剂量的选择,往往需要采用这类试验;③多因素多水平,例如探索联合用药方案、研究中药复方等,生物体的复杂

性决定了多因素多水平研究成为最常用的试验研究类型。

在精神卫生科研中,研究者关心的试验因素通常有如下几类:第一类,两种治疗某种精神疾病的药物,即试验因素为“药物种类”^[1-4];第二类,在基础治疗基础上是否使用非药物手段治疗,即试验因素为“是否用非药物手段”^[5-8];第三类,来源于社会、家庭和心里的因素,即试验因素为多个社会或心理因素^[9-12]。

1.2.3 非试验因素

所有与试验因素同时存在、能使受试对象产生效应的其他因素被称为非试验因素。最常见的非试验因素有如下几类:①受试对象自身具有的、可影响疾病发生和发展过程的某些属性(如性别、年龄、血型等)或状态(如病型、病情、患病时间、生理或心理所处的阶段、有无某些疾病的家族史等);②试验者的技术水平和情绪等是否稳定;③药物或试剂的质量和性能是否稳定;④试验时的环境和条件是否稳定等。

非试验因素虽然不是试验因素,但由于其中有些会严重影响试验结果,一旦处置不当,它们就变成了混杂因素而产生混杂效应,导致结论出错。在进行试验设计时,应明确哪些是非试验因素,应有意识地控制和消除其干扰。

1.2.4 重要的非试验因素

因考察的效应指标不同,有些非试验因素对试验结果的影响是不容忽视的,必须在试验设计阶段或在统计分析阶段加以控制,这样的非试验因素被称为重要的非试验因素,常简称为区组因素。

在众多的非试验因素中,应结合具体问题,选定其中对观测结果具有最大影响的一个或若干个,将其定为“重要的非试验因素”。在试验设计时,设法对它们加以控制,这就是所谓“含区组因素”的试验设计类型,如配对设计、随机区组设计、平衡不完全随机区组设计、含区组因素的析因设计等;更多的时候是将受试对象按重要非试验因素进行分层,然后将每层中的受试对象完全随机地均分入对比组(通常为试验组与对照组)中去,此为“分层随机化”;若重要的非试验因素是定量的,则应从试验开始就记录每个受试对象的定量影响因素的数值。在

统计学上,常将“重要的定量非试验因素”改称为“协变量”,将来进行数据处理时,采用类似“协方差分析”等技术将其影响降至最低程度。

1.2.5 复合型区组因素

区组因素常指一个非试验因素,也可以是由多个非试验因素复合而成的。例如,某病患者的年龄分为老年、中年、青年,性别分为男性、女性,职业分为工人、农民、军人和商人四种,假定在某项试验研究中,这三个因素都属于重要非试验因素,由它们可以形成 $3 \times 2 \times 4 = 24$ 种不同的水平组合,我们就说有一个复合型区组因素,它有 24 个水平。

在科研设计时,如何真正发挥“复合型区组因素”的作用呢?首先,将其每个水平视为“一个层”,将全部符合纳入标准且不符合排除标准的受试对象分配到相应的各层中去;然后再将各层中的受试对象完全随机地均分到各对比组中。在统计分析时,如何真正发挥“复合型区组因素”的作用呢?先将构成复合型区组因素的所有非试验因素与全部试验因素同等看待,确定其对应的是什么样的多因素试验设计类型,选择相应的统计分析方法处理数据;然后,根据初步分析的结果修改统计模型,直至找到最恰当的统计模型为止。

1.2.6 次要的非试验因素

有些非试验因素是可以忽略不计的(被称为次要非试验因素)。事实上,在任何一个试验研究中,次要的非试验因素往往很多。在试验设计阶段,研究者通过制订严格的受试对象质量标准和采取分层随机化等措施选取和分配受试对象,以使一切次要的非试验因素在对比组(通常由所有试验因素不同水平组合而形成)之间尽可能达到动态平衡。

2 因素及其水平的选定

2.1 选定试验因素与重要非试验因素的原则

任何试验效应都是多种因素综合作用的结果,我们不可能也没有必要在一次或几次试验中,同时研究所能想到的一切有关因素,只能在设法严格控制某些重要非试验因素的前提下,抓住对试验结果可能有较大影响的重要试验因素加以研究。通常将重要试验因素简称为试验因素,将重要非试验因素

及其组合简称为区组因素。

在一项试验研究中,拟选定哪些试验因素(如药物种类、剂量大小、给药途径、作用时间等)取决于研究者的研究目的和可以投入试验研究的人力、物力、时间和经费。试验研究的目的不同,对试验的要求也不同。若在整个试验过程中影响观察结果的因素很多,就必须结合专业知识,对众多的因素作全面梳理,必要时需做一些预试验,区分哪些是重要的试验因素;而考察哪些重要非试验因素(如病型、病情、发病时间等)取决于基本常识、专业知识和试验分组后组间的均衡性的高低、研究结果的适用范围的大小和所得结论可信度的高低。

2.2 确定试验因素与重要非试验因素水平的原则

一般来说,定性因素的水平是相对固定的。例如,患者性别通常只考虑男性和女性两种性别,尽管有极少数性别畸形的人(可以通过制订相应的质量标准将畸形性别的拟参选受试对象排除掉);患者 ABO 血型系统通常只有 A、B、AB 和 O 型四种。然而,定量因素的水平是有很多的,有时其水平数有无穷多个,如某种药物的剂量、作用时间、患者年龄等。

值得注意的是,因素的水平数一般不应定得太多。例如,研究某种精神疾病的治疗方法,与其有关的因素有很多,如年龄、病程、疾病类型等,其中每个因素又分若干水平。如果选定其中的两个因素各有 10 个水平,就有 $10^2 = 100$ 种试验条件,即 100 个试验组;若选定 4 个因素,各取 10 个水平,那就有 10 000 种试验条件,即 10 000 个试验组。每种试验条件(或试验组)下至少要做 2 次独立重复试验(具体应重复几次,应有根据地估算出来),因此,一个好的试验设计,首先要求在众多因素与水平中抓住主要的几个,水平数不能定得太多。

对于定量因素,通常有两种确定其水平的方法。其一,依据基本常识、专业知识和实际情况,取有限个有实用价值的水平。如患者年龄,可以取老、中、青三个水平(假定纳入标准中要求年龄为 18~60 岁);药物剂量,可以取 2 mg、6 mg、10 mg(依据预试验结果来确定作为剂量的起始值、变化范围和间隔大小;选取几个水平,与试验的工作量有关,水平定得越多,工作量将会越大,特别是在多因素试验中,因素的水平略有增加,因素不同水平组合数将大幅度增加)。其二,

定量因素的水平可以取任意多个,即不限制其仅取哪几个具体值。如患者年龄,是多大年龄就记录多大年龄;在观测药物作用时间的结果时,药物进入生物体内有多长时间就记录多长时间。

3 试验因素与重要非试验因素的处置

3.1 处置因素的关键技术

3.1.1 试验因素要标准化

试验因素标准化就是要保证试验因素在整个试验的过程中始终如一,保持不变,包括处理因素的施加方法、强度、频率和持续时间等。例如试验因素是药品,除应确定药品的名称、性质、成分、作用及用法外,还应明确生产厂家、药品批号、出厂日期及保存方法等;若试验因素是手术方法,操作者的熟练程度应当自始至终保持恒定,否则将会影响结果的稳定性;若试验因素水平的定义与实际情况不符,所得结果当然就不同了,此时很容易得出错误的结论,因为可能会张冠李戴。

同样,重要非试验因素也必须遵循标准化原则,否则它们或多或少会影响试验结果,导致结论出现较大偏差。因此,在试验设计时应制订或摸索出标准化的具体措施和方法,使试验因素和重要的非试验因素真正达到标准化,并尽可能使重要的非试验因素始终处于可控状态,使次要非试验因素在对比组之间处于平衡状态。

3.1.2 注意因素之间的交互作用

多种因素之间往往存在交互作用。最易于理解的是两药联用可发生药效的协同作用或拮抗作用。事实上,不仅两种药物同时使用可能会产生交互作用,非药物的试验因素之间也常可产生交互作用。所谓交互作用,就是一个因素不同水平对观测结果的影响情况会随着另一个因素水平的改变而改变。例如,A、B 两药物的疗效谁好谁差不能一概而论,因为用于某疾病男性患者时,A 药的疗效优于 B 药的疗效,而用于该病女性患者时,A 药的疗效劣于 B 药的疗效。此时,就可以说,药物种类与性别之间存在交互作用。因此在试验设计时除了要考虑各因素单独施加于受试对象的效应,还应特别关注因素配合施加于受试对象后产生的效应。在试验设计时,如何将不可忽视的交互作用项考虑进去,这个问题

比较复杂,可参阅有关文献^[13],此处从略。

3.1.3 基于因素及其水平构造特定的设计类型

一般来说,当因素(包括试验因素与重要非试验因素)为定性的且取有限的几个水平(通常为 2~5 个水平)时,由所有因素及其水平构造出一种架构,再结合专业知识(因素对观测指标的影响有无主次之分、因素之间在属性上是否有嵌套关系)和具体操作的特点(如因素施加有无先后顺序)等,给“架构”取一个设计名称,即由定性因素决定设计类型;而那些定量因素对形成什么样的设计类型不起作用,它们只在统计分析时发挥其应有的作用。在统计学上称其为“协变量”,在对受试对象进行分组时,定量因素或称为自变量被视而不见,仅在进行统计分析时才“激活”它们。

在构造设计类型时,定性试验因素与定性重要非试验因素所起的作用是有区别的。一般来说,由定性试验因素及其水平的全面组合构成设计的“框架”,而由定性重要非试验因素对“框架”进行分层。这些内容比较难懂,可参阅有关文献^[14-15],此处从略。

3.1.4 处置重要的非试验因素的关键技术

在试验设计阶段,首先,要找全一切可能的非试验因素。其次,结合基本常识、专业知识和工作条件,将可有可无的非试验因素淘汰掉,保留那些最重要的非试验因素。这个过程可以被称之为“找全找准重要的非试验因素”。

若试验中存在重要的非试验因素,设计时应有意识地控制和消除其干扰和影响。若非试验因素是定性的,如性别、窝别等,可以采用分层随机化的方法对其加以控制,常见的区组随机化和配对随机化可看作是分层随机化的实际应用。此类试验设计包括配对设计、随机区组设计、平衡不完全随机区组设计、含区组因素的析因设计等。有些非试验因素虽然是定量的,但可以通过分组将它们看作定性的因素,如年龄是定量的,若按照年龄大小分为儿童、少年、青年、中年、老年,就转变成定性的因素了。有时,若认为将定量因素定性化会丢失一些信息,也可以按下述方法处置定量的非试验因素。

若重要的非试验因素是定量的又不便将其定性化,则应从试验开始之时,就记录每个受试对象的定量影响因素的数值,在进行数据处理时采用“协方差

分析”等技术将其影响降至最低程度。例如在评价不同药物或治疗方案对精神疾病的疗效与安全性时,若不同患者从发病到接受治疗的“时间间隔”相差悬殊,就应该准确记录每位患者的“时间间隔”,在进行统计分析时,将其视为“协变量”构建统计模型。

4 某些实际工作者在确定试验因素方面常犯的错误

有些人喜欢用“组别”这个词作为全部试验分组的总称,无论试验中涉及多少个组,他们总把“组别”理解为“一个试验因素”,很自然地将其下的各组理解为该因素的各个水平。还有人毫无根据地选择很多试验因素,使试验研究变得十分复杂,往往因为试验条件、人力和时间都达不到起码的要求,使试验研究半途而废或因安排不当,组间夹杂着某些混杂因素的影响,导致结论可信度低。

由于研究者试验设计水平贫乏,不会科学地利用多因素试验设计技术,只能将其他试验因素视而不见或固定在各自的特定水平上,每次只让自己关心的某个试验因素取不同的水平,当试验因素之间不独立时,基于这样的试验方法获得的结果进行分析,常得出错误结论。

参考文献

- [1] 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.
- [2] 陶正宇,马妍芹,常馨. 不同剂量文拉法辛与帕罗西汀治疗抑郁症的效果及安全性研究[J]. 四川精神卫生,2015,28(4):324-326.

- [3] 吴胜,唐丽,罗捷. 氟伏沙明与氯米帕明治疗躯体化障碍的随机对照研究[J]. 四川精神卫生,2015,28(5):417-419.
- [4] 陈思宇,张辉,易春燕,等. 喹硫平与利培酮对老年人脑血管疾病所致精神障碍的疗效观察[J]. 四川精神卫生,2015,28(6):526-528.
- [5] 周平,张瑶,谭庆荣. 电针联合舍曲林治疗创伤后应激障碍的效果观察[J]. 四川精神卫生,2015,28(6):504-506.
- [6] 郭新宇,杨媛,田丽. 草酸艾司西酞普兰联合重复经颅磁刺激对改善难治性抑郁症患者执行功能的疗效研究[J]. 四川精神卫生,2016,29(1):26-30.
- [7] 刘晓,马元业,李捷,等. 低频重复经颅磁刺激联合草酸艾司西酞普兰治疗广泛性焦虑障碍的临床对照研究[J]. 四川精神卫生,2016,29(2):146-149.
- [8] 牛亚杰,杨少杰,杨可冰,等. 西酞普兰合并改良电休克治疗对抑郁发作的疗效及认知功能的影响[J]. 四川精神卫生,2015,28(2):123-126.
- [9] 杨世昌,陈雷音. 青少年攻击行为影响因素研究进展[J]. 四川精神卫生,2015,28(2):112-115.
- [10] 郑银佳,麦家铭,周燕玲,等. 精神分裂症患者家庭负担影响因素的研究进展[J]. 四川精神卫生,2015,28(4):306-309.
- [11] 余金鸣,张岱威,张建龙,等. 生活事件、社会支持与双相情感障碍患者症状复发的相关研究[J]. 四川精神卫生,2015,28(4):354-357.
- [12] 张少丽,李玉焕,董继承. 青少年心理健康的家庭影响因素研究进展[J]. 四川精神卫生,2015,28(5):400-402.
- [13] 胡良平. 统计学三型理论在实验设计中的应用[M]. 北京:人民军医出版社,2006:18,94-106,121-132.
- [14] 胡良平. 科研设计与统计分析[M]. 北京:军事医学科学出版社,2012:129-205.
- [15] 胡良平. 面向问题的统计学——(1)科研设计与统计基础[M]. 北京:人民卫生出版社,2012:42-80.

(收稿日期:2016-06-20)

(本文编辑:吴俊林)

《四川精神卫生》杂志邮发代号:62-283