

如何正确把握实验设计的随机与对照原则

胡良平 关雪

我们知道,实验设计的要点除了“三要素”之外,还应正确把握“四原则”,即随机、对照、重复与均衡。实验设计坚持这四个原则是实验成功的关键,其原理在于它能有效地排除非处理因素的干扰和影响,使实验误差降到最低限度^[1]。无论违背了哪一个原则,实验结果都是不能令人信服的。在上期讲座中我们介绍了实验设计的三要素,本期将结合脑血管病科研方面的实例重点阐述如何正确把握随机与对照原则,包括如何运用 SAS 程序实现随机化。

随机原则

一、什么是随机原则

所谓随机原则,就是在抽样或分组时,必须做到使总体中任何一个个体都有同等的机会被抽取进入样本,以及样本中任何一个个体都有同等机会被分配到任何一个组中去^[1]。其作用就是使样本具有极好的代表性,使各组受试对象在重要的非实验因素方面具有极好的均衡性,提高组间实验资料的可比性。在受试对象的选取和分组时,必须严格按这一原则实施。

随机化应贯穿于实验研究全过程,在受试对象的抽样、分组以及实验实施过程中均应遵循随机化原则。随机化应该体现在如下三个方面^[2]:①抽样随机:每一个符合条件的实验对象参加实验的机会相同,即总体中每个个体有相同的会被抽到进入样本之中。它可保证得到的样本具有代表性,使实验结论具有普遍意义;②分组随机:每个实验对象分配到不同组(通常为对照组、不同处理组)的机会相同。它可保证大量难以控制的非处理因素在对比组间尽可能均衡一致,以提高各组间的可比性;③实验顺序随机:每个受试对象先后接受处理的机会相等,它使实验顺序的影响也达到均衡。

二、几种常见的随机化方法

1. 随机抽样

(1)简单随机抽样:就是从总体中不加任何分组、划类、排队等,完全随机地抽取受试对象。特点是每个个体(样品)被抽中的概率相等,各样品之间完全独立,彼此间无一定的关联性和排斥性。简单随机抽样是其他各种抽样形式的基础。通常只是在总体样品之间差异程度较小和数目较少时,采用这种方法。

(2)分层抽样:先依据一种或几种特征将总体分为若干个子总体,每个子总体称作一个层,然后从每层中随机抽取一个子样本,这些子样本合起来就是总体的样本。分层抽样所得的样本具有较好的代表性,抽样误差较小。

(3)系统抽样:先按照某种顺序将总体中的个体(样品)从1~N相继编号,并计算抽样距离 $K = N/n$ 。式中N为总体样品总数,n为样本容量。然后在1~K中抽一随机数*i*,作为样本的第一个被抽样品,接着取*i*+K,*i*+2K,⋯,*i*+(n-1)K,直至抽够n个样品为止。值得注意的是,当总体的被观察样品按顺序有周期趋势或增加(减少)趋势时,容易产生偏倚。

(4)整群抽样:先将总体分成若干群,从中随机抽取几个群;抽中群内的全部个体(样品)组成样本。其优点是便于组织调查,节省经费,容易控制调查质量。缺点是样本量一定时,抽样误差一般大于单纯随机抽样。

2. 随机分组

(1)完全随机化:直接对受试对象进行随机化分组,常通过掷硬币或随机数字表,或用计算机产生随机数来进行随机化,在事先或者实施过程中不作任何限制和干预或调整。分组后各组受试对象的例数不一定相等。其具体步骤如下:①编号:将N个实验单位从1到N编号。动物可按体重大小,患者可按预计的样本量编号;②获取随机数字:从随机数字表中任意一个数开始,沿同一方向顺序获取随机数,每个实验单位一个随机数字;③求余数:随机数除以组数求余数;④分组:按余数分组;如分两组时,可规定遇到随机数字为偶数时将对应的受试对象分入实验组、遇到随机数字为奇数时将对应的受试对象分入对照组(反过来规定也可以);再如分三组时,可事先规定,凡随机数字除以3得余数为1者分入第一组,余数为2者分入第二组,余数为3者分入第三组。当然,也可以规定其他的分组规则,但规则必须事先确定下来,一旦确定不应随意改动。⑤调整:假如共有n例待调整,需要从中抽取1例,续抄一个随机数,除以n后将得到的余数作为所抽实验单位的序号。

(2)分层随机化:首先对可能影响实验过程和结果的主要混杂因素(如年龄、性别、病情、疾病分期等)进行分层,然后在每一层内进行完全随机化分组,最后分别合并为实验组和对照组。配对随机化和区组随机化可看成是分层随机化的实际应用。下面以区组随机化为例介绍分层随机化的主要步骤。①编号排序:将受试对象按照主要混杂因素相近的原则分层或分组,为同一层或区组内个体编号;②取随机数:可从随机数字表、计算器或计算机获得。每个受试对象可取两位数;③确定组别:根据事先设定的规则,按照每层内受试对象获得的随机数决定受试对象被分在哪一组。

(3)动态随机化:在临床实验研究中,研究者常习惯将患者分为实验组与对

照组。患者来医院就诊,若将先来就诊的分入实验组,后来就诊的分入对照组,是不正确的。因为患者来医院就诊,在患病严重程度、患病时间等重要非实验因素方面不一定是均衡的,可能带有某种程度上的“聚集性”。亦即某一段时间上来的患者多数为重病患者,另一段时间上来的多数为轻病患者。若规定第一个来就诊者分入实验组,第二个来就诊者分入对照组,这样交叉进行分组,最后也可能两组患者在某些重要非实验因素方面相差悬殊,组间缺乏均衡性。有一种简便易行的随机化分组方法叫“按不平衡指数最小的分配原则”分组^[3]。此法也称为动态随机化法。这个分组方法的做法如下:根据专业知识选取几个拟加以控制的重要非实验因素,假定一个是患者的“性别(分为男、女)”,另一个是患者的“病情(分为轻、中、重)”。将先来的两例患者在实验组与对照组各放1例。记下他们的性别和病情。记分的方法是每个因素的每个水平出现1次记1分。计算两组各因素对应水平的得分之差的绝对值,最后求出绝对值之总和。此“和”即为两组患者在两个重要非实验因素上的不平衡指数。若再来第三位患者,分别依次将此患者放入实验组、对照组各1次。每次都根据他(或她)的性别、病情累加到原有患者的基础之上。可以得到两个不平衡指数,取不平衡指数最小的那种分组方法,这样第三位患者的分组就定下来了。用同样的方法去分配以后来的该病患者,直到两组有了事先规定的样本含量时停止。现举例说明如下:例1:假定在对肩周炎患者分两组时考虑病情(轻、中、重)、患病时间(短、长)和日常运动量(少、多)3个重要的非处理因素,再假定治疗组和对照组各有了1例患者,他们的基本情况如表1所示。如果又来了一例新患者,其病情重、患病时间短、日常运动量多,应将此患者分入哪个组为好?

表12例患者的基本情况及分组后的平衡程度

实验分组	例 数							合 计
	病 情			患病时间		运 动 量		
	(轻	中	重)	(短	长)	(少	多)	
治疗组	0	1	0	0	1	1	0	
对照组	1	0	0	1	0	1	0	
差的绝对值	1	1	0	1	1	0	0	4

注:表中合计值“4”就是“不平衡指数”,下同

试着将新患者分别分入治疗组和对照组,取“不平衡指数”较小者所对应的分配方案,分别参见表2和表3。

表2 将新患者分入治疗组后的平衡程度的探索

实验分组	例数							合计
	病情			患病时间		运动量		
	(轻	中	重)	(短	长)	(少	多)	
治疗组	0	1	1	1	1	1	1	
对照组	1	0	0	1	0	1	0	
差的绝对值	1	1	1	0	1	0	1	5

表3 将新患者分入对照组后的平衡程度的探索

实验分组	例数							合计
	病情			患病时间		运动量		
	(轻	中	重)	(短	长)	(少	多)	
治疗组	0	1	0	0	1	1	0	
对照组	1	0	1	2	0	1	1	
差的绝对值	1	1	1	2	1	0	1	7

因 $5 < 7$, 故本例应将新患者分入治疗组。若再有新患者, 分配方法相同。假如有 k ($k \geq 2$) 个治疗组, 可分别与对照组比较, 分别求 $(k - 1)$ 个表中的“不平衡指数”, 取其中最小者对应的分配方案。

3. 用 SAS 实现随机化: 实现随机化的方法有多种, 如抽签或利用随机数。在实验设计中常通过随机数来实现随机化。获得随机数的方式有两种, 随机数字表(包括随机排列表)和计算机产生的伪随机数。现在也可以通过 SAS 来实现随机化。下面结合实例介绍如何运用 SAS 实现随机化。因篇幅有限, SAS 程序略, 仅给出输出结果。详细程序请参阅相关专著^[4]。在 SAS/STAT 模块中, 有一个与调查设计有关的 SURVEYSELECT 过程。该过程可用来实现各种随机抽样, 包括: 单纯随机抽样、系统随机抽样、分层抽样、无限随机抽样(有替换)、序贯随机抽样以及按规模大小成比例概率抽样(PPS)等。SAS/STAT 模块的 PLAN 过程, 可用于构建各种常见的实验设计并对设计方案中的实施顺序进行随机化, 也可用于产生数字的排列组合表。

(1) 完全随机化的 SAS 实现 例2: 用 SAS 程序将 18 只小鼠随机分为 A、B 两组。①对 18 只小鼠进行编号, 范围为 1 ~ 18。②编写 SAS 程序(略)。③输出结果如下:

group1	mouse1	group2	mouse2
A	4	B	1
A	5	B	2
A	7	B	3
A	9	B	6
A	11	B	8
A	12	B	10
A	13	B	14
A	15	B	17
A	16	B	18

结果解释：“mouse1”和“mouse2”列代表小鼠的编号，“group1”和“group2”列代表小鼠的分组。18 只小鼠被随机等分为 A、B 两组。

(2)区组随机化的 SAS 实现 例 3:为研究现有 4 窝不同种系的大鼠,每窝 4 只,共 16 只。请以大鼠种系作为区组因素,将这 16 只大鼠随机分配到 4 个组中去。①对大鼠进行编号,第一窝编号为 1~4,第二窝编号为 5~8,第三窝编号为 9~12,第四窝编号为 13~16。②编写 SAS 程序(略)。③输出结果如下:

brood	A	B	C	D
wobie1	2	4	3	1
wobie2	5	7	8	6
wobie3	9	10	11	12
wobie4	15	14	13	16

结果解释：“brood”列代表大鼠的窝别,其 4 个水平(wobie1 , wobie2 , wobie3 , wobie4)分别代表 4 窝不同种系的大白鼠。A、B、C、D 代表 4 个组,其下的数字为大鼠的编号。

4. 随机分配方案的隐匿与盲法:在临床试验中,当审核受试者入选条件的研究人员知道下一个(随机数字所对应的)受试者治疗方案时,研究者可能会根据下一个患者的特征和自己对不同治疗方案的好恶,人为地决定入选或排除该患者;患者也会因此人为地决定是否参与研究。这样的分组不能实现随机分组的根本宗旨,无法起到控制偏倚的作用^[5]。为避免这种选择性偏倚,应防止随机分组方案被分配者或患者事先知道,即随机分配方案的隐匿。进行随机分配方案的隐

匿,首先要求产生随机分配序列和确定受试对象合格性的研究人员不应该是同一个人。其次,如果可能,产生和保存随机分配序列的人员最好是不参与实验的人员^[6]。其方法有按顺序编码不透光密封的信封、中心随机系统、编号或编码的瓶子或容器、中心药房准备的药物等^[7]。后两种方法目前在临床中并不常用。由于治疗和组别的关系,患者很可能中途失访或退出治疗,从而破坏组间的可比性。盲法会在一定程度上帮助降低这些事件在组间发生的不均衡性,从而维持组间可比,也可以避免研究者的主观偏倚。盲法是一种蒙蔽治疗分组的措施,就是在治疗和追踪随访期间,保密每一个研究对象的治疗分组,使参与研究的人员不知道分组情况。常用的方法有单盲、双盲和三盲。随机方案的隐匿与盲法不同^[8]。

①实施的时间不同。前者在临床试验最后一名患者完成分组后即告结束;盲法需要在整个治疗和随访过程中保持盲的状态,直到实验干预和结局测量完成后才结束。②适用条件不同。随机分配方案的隐匿在任何临床试验中都能进行,无论是分配前或者在分配时;而并非所有的临床试验中都能进行盲法。例如,比较针灸和药物两种疗法治疗某种疾病的疗效,盲法难以实施,而随机分配方案的隐匿却是可行的。

对照原则

一、什么是对照原则

所谓对照,是指在实验研究中,要想得出什么处理或治疗更佳,必须交代“与谁比较”,比较至少要在两个事物之间进行。在统计学上通常是两个组之间进行,即实验组和对照组。设立对照实际上就是寻找一个参照物或对比的基础。对照原则的作用是提高实验的鉴别能力,增强结论的说服力。需要强调一下,对照组的概念还必须予以拓展,在多因素的实验设计中,为了考察因素间的交互作用,往往要求实验各组都有对应的对照组,否则,如果对照设置不全就不能把它们一起进行分析。因此,并非在任何实验研究中设立1或2个对照组,就一定能达到研究目的^[9]。

二、对照的种类

对照有很多种类型,可根据研究目的和内容加以选择。

1. 安慰剂对照:安慰剂是指外观与受试药物相同,且无药理活性的物质。在临床研究中安慰剂对经常用来代替受试药物,以排除精神心理等非药物因素的影响。安慰剂还可消除疾病自然进程的影响,分离出实验药物所引起的真正效应,从而直接度量实验药物和安慰剂之间的差异。安慰剂的使用需慎重,应以不损害患者健康为前提,适用于所研究的疾病尚无有效药物治疗,或使用安慰剂后对该病病情、临床过程、预后无影响或不利影响小的情况,一般与盲法结合使用。

2. 空白对照:即对照组不接受任何处理因素,在动物实验和实验室方法研究

中常见,常用于评价测量方法的准确度,评价是否处于正常状态等。空白对照简单易行,但在以人为受试对象的研究中涉及伦理问题,且实施过程中容易引起对照组和实验组在心理上的差异,从而影响实验结果的可靠性^[10]。空白对照可用于以下两种情况:一是由于处理手段非常特殊,安慰剂盲法实验无法执行,或执行起来比较困难。如实验组为手术治疗或放射治疗等;二是实验药的不良反应非常特殊,以至于无法使研究者处于盲态。

3. 实验对照:对照组不施加处理因素的真实水平。如研究饲料中某种营养成分的价值,在这个实验中,该营养成分是处理因素(是否含有此营养成分)的真实水平,饲料则是实验组与对照组的“本底”或称“载体”,两组除是否含有这种营养成分外,其余条件都一致,这样才能分析这种营养成分存在与不存在之间的差异。因此,当处理因素的施加需伴随其他“载体”(本例为“饲料”)时,而这种“载体”可能影响实验结果时,应设立实验对照(即仅有“载体”),以保证组间的均衡性。

4. 自身对照:对照与实验在同一受试对象身上进行。但应注意实验措施在不同时间实施,对实验结果的影响。此时,可以考虑在实验中另外设立一个平行的对照组,用实验组和对照组处理前后效应的差值来进行比较。

5. 标准对照:用现有标准方法、常规方法、标准值或参考值作为对照。实验研究中常用于某种新方法是否可代替传统方法。但是不设立对照组,仅用现有标准值或参考值做对照是不提倡的,因为实验组和对照组的时间、地点和环境不同。

6. 历史对照:历史对照是将研究者以往的研究结果或文献上他人的研究结果与本次研究结果作对照。这种对照一般用于考核时间因素带来的变化。

值得一提的是,有时在同一项实验研究中需要同时运用多种对照形式,如安慰剂对照和实验对照同时并存,有时甚至还会出现阴性对照、阳性对照等对照形式。此时,所对应的实验安排常常不是一种标准的实验设计类型,而叫做“多因素非平衡组合实验”更为贴切。

实际工作者在执行随机与对照原则方面易犯的错误

一、用简单随机化代替动态随机化

在临床试验中,实际工作者往往用简单的随机化对受试者进行分组,在很多场合下,这样的做法是错误的。下面将以实例加以说明。例4:为评价三七通舒胶囊预防脑梗死复发的有效性及安全性以及对血液流变学的影响,以阿司匹林肠溶片为对照。研究病例为我院2007年8月至2008年7月收治的100例脑梗死患者。将100例脑梗死患者随机分为两组,治疗组给予三七通舒胶囊,对照组口服阿司匹林肠溶片。治疗组和对照组患者年龄、性别、既往史及入组时的NIH评分差异均无统计学意义。随访一年,观察脑梗死复发率及血液流变学变化。对差错的辨析与释疑:研究者在文中仅简单描述了将100例脑梗死患者随机分为两

组,未描述随机分组的方法。最后对两组患者的性别、年龄、病情等进行了均衡性检验,发现差异无统计学意义。但是,若发现这些重要的非实验因素在两组间有差异,已无法弥补。因为在患者入院时,就已经被决定分入哪一组。正确的做法是根据患者的年龄、性别、既往史及入组时的 NIH 评分,按不平衡指数最小的分配原则分组。才能确保两组患者在几个重要非实验因素方面均衡可比。

二、未进行随机分配的隐匿

例 5:为探讨尤瑞克林治疗急性脑梗死的临床疗效及安全性。选择起病 48 h 内的急性脑梗死患者 78 例,随机分为实验组(53 例)和对照组(25 例)。实验组给予尤瑞克林 0.15 PANU,1 次/d,对照组给予丹红 20 ml,1 次/d,两组疗程均为 14 d。对差错的辨析与释疑:有研究表明^[7],一般情况下,与充分使用随机隐匿的临床实验结果相比,不充分随机隐匿的或者随机隐匿方案不清楚的临床实验结果往往会夸大 37% 的效应估计值。原文中未提及是否使用随机分配方案的隐匿。研究者很可能在分组时将病情容易好转的患者分入治疗组,从而得出不可信的结论。

三、缺乏必要的对照组

例 6:为探讨脑梗死大鼠外周血促炎/抗炎细胞因子的动态变化规律。建立成年雄性 SD 大鼠大脑中动脉阻塞(MCAO)模型,分为 MCAO 后 1 h、12 h、24 h、48 h、72 h、1 周和 2 周组。测定大鼠血清促炎细胞因子 IL-1 β 、INF- γ 和抗炎细胞因子 IL-10 的水平,以评估机体的免疫状态。对差错的辨析与释疑:该研究的目的是探讨脑梗死大鼠外周血促炎/抗炎细胞因子的动态变化规律,只是比较了建模后不同时间大鼠血清促炎细胞因子和抗炎细胞因子的水平,但不能确定这些变化是由于手术过程造成还是由于大鼠中动脉阻塞造成,因此应该再设立一个假手术组,即实验对照组,对大鼠进行手术,但不阻塞中动脉。这样得到的结果才更可信。

四、不合理的历史对照

例 7:观察某种药物对合并糖尿病的急性脑梗死患者治疗的临床效果。收集 2008 年 1 月至 2009 年 12 月的 63 例合并 2 型糖尿病的急性脑梗死患者,采用该药物治疗,与过去一年收治的 35 例合并 2 型糖尿病的急性脑梗死患者采用常规治疗的疗效进行比较。对差错的辨析与释疑:该研究采用的是历史对照,与过去采用常规治疗的患者进行比较。历史对照的缺点是,治疗组和对照组间除了研究药物外,有许多因素是不相似的。包括人口统计学特征、诊断标准、疾病的阶段及持续时间、伴随治疗以及观察的情况(如评价结果的方法、研究者的期望),很可能影响结果。因此正确的做法是,将同期合并糖尿病的急性脑梗死患者按照最小不平衡指数的方法分为实验组和对照组,一组给予这种新药物,一组给予常规治

疗,再比较两组的疗效。

五、对照方面的其他错误

实际工作者在执行对照原则时,还会犯其他错误,如未设置对照组、假对照、对照过剩、对照不全。所谓“对照不全”,就是有些组有相应的对照组,而另一些组却没有相应的对照组。例如,在某项实验研究中,每此实验都要涉及 A、B 两种药物,而研究者希望同时考察 A 药三种剂量的效果、B 药三种剂量的效果,原本有 9 种剂量组合,但研究者实际仅做了 6 种组合实验,缺少了其中的 3 种组合。此种实验结果很难得出令人信服的结论。

值得一提的是,在生物医学和临床试验研究中,人们经常会在随机与对照两个原则方面出现问题,从而,很容易导致组间缺乏“均衡性”。换一句话说,在科研设计中,人们最易在实验设计四个基本原则中的“随机、对照和均衡”原则方面犯错误,它们似乎有连带关系,这正是为什么很多科研课题的结论经不起时间和实践检验的根本原因。

参 考 文 献

- 1 胡良平. 统计学三型理论在实验设计中的应用. 北京:人民军医出版社,2006:20-22.
- 2 方积乾. 卫生统计学(第六版). 北京:人民卫生出版社,2008:258.
- 3 胡良平,刘惠刚,李子建,等. 检验医学科科研设计与统计分析. 北京:人民军医出版社,2004:4-8.
- 4 胡良平. SAS 统计分析教程. 北京:电子工业出版社,2010:674.
- 5 李立明. 流行病学(第6版). 北京:人民卫生出版社,2007:138.
- 6 李静. 随机分配方案的隐匿. 中国循证医学杂志,2004,4(10):714-715.
- 7 Schulz KF. Assessing allocation concealment and blinding in randomized controlled trials: why bother? Evidence Based Nursing,2001,4(1):4-6.
- 8 万霞,刘建平. 临床实验中的随机分组方法. 中医杂志,2007,48(3):216-219.
- 9 胡良平. 口腔医学科科研设计与统计分析. 北京:人民军医出版社,2007:79-82.
- 10 刘桂芬. 医学统计学(第二版). 北京:中国协和医科大学出版社,2007:167-169.

(收稿日期:2010-08-08)

(本文编辑:齐彤)

胡良平,关雪. 如何正确把握实验设计的随机与对照原则[J/CD]. 中华脑血管病杂志:电子版,2010,4(5):399-407.