

• 讲座 •

如何正确进行生物医学科研设计

IV. 如何合理选择三因素设计

李长平, 胡良平, 周诗国, 高辉

编者按

生物医学科研的研究对象是生物体, 而生物体具有极大的变异性。科研的目的就是要从表面上看似杂乱无章的事物或现象中找出规律性的东西来, 以便正确地解释和揭示事物或现象的变化、发展乃至消亡的规律, 从而达到认识人类社会并促进其发展、认识自然、改造自然以至于征服自然之目的。人类要完成此重任, 仅靠热情和苦干是不够的, 必需加上巧干。所谓巧干, 就是要充分利用人类已经积累起来的宝贵经验, 加上人类特有的创造性思维和勇于实践的精神, 在确立了可行的奋斗目标之后, 对未知事物或现象进行调查或试验。要想使调查或试验得出的结论真实可信, 同时花费的人力、物力、财力和时间又比较少, 就需要有科学完善严谨的科研设计方案的指导。本刊特邀军事医学科学院生物医学统计学咨询中心主任胡良平教授, 以“如何正确进行生物医学科研设计”为题, 撰写系列统计学讲座。希望该系列讲座能对生物医学科研工作者制定科研设计方案有所帮助。

在实际工作中, 当科研工作者进行的某项实验研究涉及到 3 个因素时, 该如何选择恰当的实验设计类型, 才能使实验达到准确、可靠、高效的研究目的呢? 最关键的问题在于: 是否选择了正确的三因素实验设计类型, 同时, 还应注意受试对象的分组要随机化、重要非实验因素对各组的影响要均衡。三因素实验设计类型有很多, 适合于不同的场合; 对于同一个问题, 解决的途径往往也不只一条, 这就需要根据现有的实验条件, 以统计学实验设计的知识指导实际工作, 解决如何从这些可供选择的实验设计类型中找出合适的实验设计类型。在本讲中, 主要从三因素实验设计的类型、特点、区别及联系等方面阐述如何合理选择三因素实验设计类型问题。

1 与三因素实验设计相关的概念

1.1 三因素实验设计的概念及类型

三因素实验设计就是实验中涉及到 3 个因素的一类实验设计, 这 3 个因素可以同时是实验因素, 也可以包含 1 个或 2 个区组因素。常见的三因素设计类型及其区别与联系见表 1。

表 1 仅仅列出了三因素实验设计大致的分类框架, 仅仅通过表 1 还不能完全准确、快速地判断该选择何种实验设计类型, 这是因为每一大类中还有很多可供选择的具体的实验设计类型, 每一种具体的类型之间还有一些不同。例如交叉设计和拉丁方设计, 虽然配对二阶段交叉设计的结构可看成是若干个 2×2 拉丁方设计的纵向排列, 而且二者都适合于涉及 1 个实验因素和 2 个区组因素且要求 3 个因素间的交互作用可以忽略不计的场合以及实验因素的各水平间在同一个体上分不同阶段实施时前后无“携带效应”(指同一个受试对象在接受前后 2 次处理时前一处理对后一处理所产生的影响)的场合, 然而二者又有不同之处。拉丁方

设计要求 3 个因素均要为 k 个水平, 而交叉设计不要求实验因素和区组因素水平数必须相等。以下将结合具体实例, 着重介绍较常见的具有 1 个重复测量的两因素实验设计、具有 2 个重复测量的两因素实验设计、交叉设计、拉丁方设计、三因素系统分组(或嵌套)设计^[1]。至于三因素析因设计、具有 1 个区组因素的两因素析因设计和三因素分割(或裂区)设计原则上同三因素系统分组设计一样^[2], 只是在各自两因素的基础上根据研究的目的增加了 1 个区组因素或 1 个实验因素, 在本讲中将不再赘述。

1.2 适合选用具有 1 个重复测量的两因素设计的场合

当某项实验研究中, 从专业上需要了解随时间推移或部位改变时, 观测指标的动态变化时, 需要选用重复测量设计。所谓重复测量设计就是同一个受试对象(如个体、标本等)在接受某种处理后, 在几个不同的时间点上重复获得指标的观测值或是从一个个体的不同部位(或组织)上重复获得指标的观测值。当实验中涉及 2 个实验因素, 而其中 1 个因素是重复测量因素时, 就可选用具有 1 个重复测量的两因素设计。值得注意的是: 重复测量设计中隐含了 1 个“区组因素”, 故具有 1 个重复测量的两因素实验设计其实包含了 3 个因素: 2 个实验因素(欲考察的重要实验因素、重复测量因素)和 1 个区组因素(个体编号)。

例 1 某外科医生欲研究 2 种治疗方式对半月板损伤患者下肢力线的影响, 拟观测患者下肢力线长度在治疗后的变化情况。现有 22 例半月板损伤但站立时膝关节能伸直的患者且各患者在病情等方面具有可比性, 观测指标为下肢力线长度(mm)。请根据提供的背景资料选择恰当的实验设计类型。

设计类型的判定: 该研究中涉及 2 个实验因素: 治疗

作者单位: 100850 北京, 军事医学科学院生物医学统计学咨询中心

通讯作者: 胡良平, Email: lphu812@sina.com

表 1 常见三因素实验设计区别与联系列表

是否考察交互作用	含区组因素的个数	是否需进行 2 次或 2 次以上独立重复实验	实验因素是否分主次	因素施加是否有先后	是否含重复测量因素	可供选择的主要实验设计类型
是	0 或 1	是	否	否	否	三因素析因设计、含 1 个区组因素的两因素析因设计、三因素正交设计、三因素均匀设计、三因素反应曲面设计
是	1	是	否	否	是	具有 1 个重复测量的两因素实验设计、具有 2 个重复测量的两因素实验设计
是	0 或 1	是	否	是	否	三因素分割(或裂区)设计、含一个区组因素的两因素分割(或裂区)设计
否	0	否	否	否	否	三因素无独立重复实验设计
否	2	否	否	否	否	交叉设计(包括两阶段交叉设计、三阶段交叉设计、3×3 交叉设计)、三因素拉丁方设计(通常称拉丁方设计)
否	0 或 1	是	是	否	否	三因素嵌套设计、含 1 个区组因素的两因素嵌套设计

方法和测量时间。治疗方法为 2 种；测量时间为“治疗前”、“治疗后”2 个水平。由于需在同一例患者治疗前、后测量下肢力线长度值，所以测量时间属于重复测量因素，因此该研究宜采用具有 1 个重复测量的两因素设计。将 22 例患者随机均分到 2 种治疗组中去，每组 11 例患者，观测 2 组患者治疗前、后的下肢力线长度。实验设计的标准型如表 2 所示。由于治疗前、后测得的下肢力线长度值之间存在一定的相关性，因此在选用统计分析方法时，若资料满足参数检验的前提条件，宜选用具有 1 个重复测量的两因素设计一元定量资料方差分析。

表 2 半月板损伤患者不同方法治疗前后的下肢力线长度

治疗方法	患者编号	下肢力线长度 (mm)	
		治疗前	治疗后
方法 1	1

	11
方法 2	1

	11

1.3 适合选用具有 2 个重复测量的两因素设计的场合

当实验中涉及到的 2 个实验因素均为重复测量因素时，就可选用具有 2 个重复测量的两因素设计，在该实验设计中也隐含 1 个区组因素。

例 2 某研究者欲了解前十字韧带损伤与内侧半月板后角的生物力学量化关系，依据专业知识，需分析前十字韧带切断与不切断的情况下，给屈曲度不同(30°、60°、90°)的膝关节一定载荷后内侧半月板后角所能承受的压力载荷。现选择 5 具尸体的 5 对膝关节进行实验，请选择恰当的实验设计类型。

设计类型的判定：在该研究中涉及到 2 个实验因素：前十字韧带损伤程度和膝关节屈曲度。前十字韧带损伤程度为“切断”、“不切断”2 个水平；膝关节屈曲度为“30°”、

“60°”、“90°”3 个水平。每具尸体均有 1 对膝关节，每对中的 2 个膝关节被随机地分配到 2 个处理组(韧带切断组、韧带不切断组)。由于是对来自同一具尸体的左、右膝关节实施不同程度的损伤处理，故“损伤程度”属于重复测量因素。而对每一个膝关节，在每一种曲度时都要接受压力载荷的测量，因此“屈曲度”也属于重复测量因素。故该实验宜采用具有 2 个重复测量的两因素设计，而对子号就是区组因素。该实验设计标准型如表 3 所示。进行统计分析时，若该资料满足参数检验的前提条件，可采用具有 2 个重复测量的两因素设计一元定量资料的方差分析。

表 3 5 对膝关节前十字韧带不同损伤情况、不同屈曲度下的压力载荷

对子号	压力载荷 (N)							
	切断:	30°	60°	90°	不切断:	30°	60°	90°
1
2
3
4
5

1.4 适合选用拉丁方设计的场合

如果欲考察 1 个实验因素和 2 个区组因素(重要的非实验因素)，当这 3 个因素水平数(k)相等且因素间的交互作用可以忽略时，可选用 k×k 拉丁方设计。k×k 拉丁方阵是此设计的核心部分，k×k 拉丁方阵中的每一个字母代表实验因素的 1 个水平；让 2 个区组因素也各取 k 个水平，并把它们分别放在 k×k 拉丁方阵的横向和纵向上，由 2 个区组因素便形成了 k×k 种水平组合，每种水平组合下伴有实验因素的 1 个水平，这 3 个水平的每一种组合便构成了 1 个特定的实验条件，每个实验条件下做 1 次实验(若数据间变异很大，建议做 2 次以上独立重复实验)。

例 3 某研究者欲研究家兔注射某种药物后不同部位所生疱疹的大小，考虑到家兔个体差异、注射顺序对观测指标可能有一定的影响，实验设计时最好将这 2 个重要的非处理因素一并安排，并尽可能节省样本。根据专业知识得知，这

3 个因素之间的交互作用可忽略不计。现有 6 只动物，其编号分别为 I、II、III、IV、V、VI；注射部位有 6 处，其代号分别为 A、B、C、D、E、F，且部位间有一定间隔，所生疱疹相互影响可忽略不计；注射顺序用 1、2、3、4、5、6 来表示。请选用合适的实验设计方案。

设计类型的判定：在该研究中涉及到 1 个实验因素：注射部位，有 6 个水平；2 个区组因素：家兔编号和注射顺序，各有 6 个水平。根据专业知识得知，注射部位与家兔编号、注射顺序间的交互作用可忽略不计，定量的观测指标为“疱疹大小 (cm²)”。根据前面介绍的拉丁方设计的适用条件，可知该实验设计可选用 6×6 拉丁方设计。该实验设计标准型如表 4 所示。

表 4 家兔注射某种药物后不同部位所生疱疹大小

家兔 编号	注射部位代号及疱疹大小 (cm ²)					
	注射顺序： 1	2	3	4	5	6
I	B (X)	C (X)	A (X)	D (X)	F (X)	E (X)
II	E (X)	D (X)	B (X)	C (X)	A (X)	F (X)
III	C (X)	F (X)	E (X)	A (X)	B (X)	D (X)
IV	A (X)	E (X)	C (X)	F (X)	D (X)	B (X)
V	F (X)	B (X)	D (X)	E (X)	C (X)	A (X)
VI	D (X)	A (X)	F (X)	B (X)	E (X)	C (X)

注：A、B、C、D、E、F 表示注射部位，X 表示疱疹大小观测值

这里需要指出的是，对于每一只家兔，前一个部位的注射对下一个部位注射的影响是微弱的或无影响的（即不存在“携带效应”），此时采用拉丁方设计不仅可以达到减少受试对象的目的，还可以分析 2 个重要非处理因素对观测结果的影响。但若实际条件不便在同一只动物身上做重复实验，同一行上要用不同的动物做实验时，即每个区组上的 k 个数据来自 k 个个体，则不存在“携带效应”，但并未达到减少样本含量的目的。若资料满足参数检验的前提条件，可采用 6×6 拉丁方设计定量资料的方差分析来进行处理；若不满足，则可采用变量变换的方法使其满足参数检验的前提条件，然后再对变换后的资料采用 6×6 拉丁方设计定量资料的方差分析来进行处理，也可直接采用 6×6 拉丁方设计定量资料的秩和检验来进行处理。

1.5 适合选用交叉设计的场合

当欲考察 1 个实验因素、2 个区组因素（即个体差异、测定顺序）对实验效应的影响时，若要求实验因素的 2 个水平先后施加于同一个受试对象，则每一个受试对象在该实验因素 2 个水平施加先后的间隙中要有一个“洗脱期”，以便前一种处理的效应不影响下一种处理的效应，尽可能消除或降低“携带效应”的影响。3 个因素间交互作用可以忽略时，可选用交叉设计。根据涉及到的实验因素的水平数和实验阶段数，可将交叉设计分为 3 类：二阶段交叉设计、三阶段交叉设计及 3×3 交叉设计。前 2 种交叉设计通常适合于考察 1 个具有 2 个水平的实验因素，同时考察 2 个

区组因素对实验效应影响的场合；第 3 种通常适合于考察 1 个具有 3 个水平的实验因素并同时考察 2 个区组因素对实验效应影响的场合。

使用频率最高的是二阶段交叉设计，它又可分为配对二阶段交叉设计和成组二阶段交叉设计。配对二阶段交叉设计是在配对设计的基础上，能多考虑实验顺序对观测结果的影响。它是先将 $2n$ 个受试对象按某些条件（通常为复合型的重要非实验因素）配成 n 对，然后随机地决定每对中 1 个受试对象接受 2 种处理的先后顺序；另 1 个受试对象接受处理的顺序正好相反。成组二阶段交叉设计是先将 $2n$ 个受试对象完全随机地均分成 2 组，然后随机地决定其中 1 组受试对象接受 2 种处理的先后顺序；另 1 组接受处理的顺序正好相反。因此，二阶段交叉设计又称为 1 次交叉设计或 2×2 交叉设计。

例 4 某研究者欲比较 A、B 2 种闪烁液测定吗啡依赖受试者血浆中 ³H-cGMP 含量 (pmol/ml) 的结果是否相同，考虑到个体差异、闪烁液测定顺序（记为 AB 顺序、BA 顺序）可能有影响，实验设计时最好将这 2 个重要的非处理因素一并安排，并尽可能节省样本。根据专业知识可知，闪烁液种类、个体差异、测定顺序 3 个因素间的交互作用可忽略不计，现有 10 名受试者，请选用合适的实验设计类型来安排此项实验。

设计类型的判定：在该研究中涉及到 1 个实验因素：闪烁液种类，有“A、B”2 个水平；2 个区组因素：个体差异和测定顺序，个体差异有 10 个水平（1~10 号个体），测定顺序有 2 个水平（AB 顺序、BA 顺序）；定量的观测指标为“³H-cGMP 含量”。根据前面介绍的交叉设计的适用条件可知该实验设计可选用二阶段交叉设计。可将 10 名受试者随机均分为 2 组（AB 顺序组、BA 顺序组），每组 5 人，例如 AB 顺序组受试者编号为 1、2、4、6、8，则 BA 顺序组对应的受试者编号为 3、5、7、9、10。该实验设计的标准型如表 5 所示。

表 5 2 种闪烁液测定 10 名吗啡依赖受试者血浆中 ³H-cGMP 含量的结果

测定 顺序	受试者 编号	³ H-cGMP 含量 (pmol/ml)		测定 顺序	受试者 编号	³ H-cGMP 含量 (pmol/ml)	
		闪烁液:				闪烁液:	
		A	B			A	B
AB	1	BA	3
	2		5
	4		7
	6		9
	8		10

若先按照某些条件（如年龄、性别）将 10 名受试者配成 5 对，然后采用随机的方法决定每对中的 1 个受试对象接受 2 种闪烁液测定的顺序，另 1 个按相反顺序测定血浆中 ³H-cGMP 含量。那么此时就为配对二阶段交叉设计了。

当资料满足参数检验的前提条件时,可采用二阶段交叉设计定量资料的方差分析来处理;若不满足此条件,则可采用变量变换的方法使其满足参数检验的前提条件,然后再对变换后的资料采用二阶段交叉设计定量资料的方差分析来处理,也可直接采用二阶段交叉设计定量资料的秩和检验来处理。

1.6 适合选用嵌套(或系统分组)设计的场合

当实验涉及到的 3 个因素间无交互作用且各个因素对观测指标的影响有主次之分(以专业知识为依据),即主要因素(A 因素)各水平之下,嵌套着次要因素(B 因素)的全部水平,次要因素(B 因素)各水平之下,又嵌套着更次要的因素(C 因素)的全部水平时,可采用三因素嵌套设计。或者 3 个因素中,有 2 个有主次之分的实验因素,而且还包含 1 个区组因素,此时,也可采用含有 1 个区组因素的两因素嵌套(或系统分组)设计。在此类设计中,B 因素或 C 因素可以取不同的水平,各水平组合之下(即每种实验条件下)可以重复作 2 次或 2 次以上不同数目的独立重复实验。

例 5 研究者欲比较不同种系的大鼠和小鼠经内毒素脂多糖(LPS)和 LPS 与角叉菜胶(Ca)联合(LPS/Ca)制备血栓形成模型的差异。其中,要考虑的因素有老鼠的类别(大鼠、小鼠),种系(大鼠欲选取 SD、Wistar 2 种;小鼠欲选取 KM、ICR、Balb/c 3 种)及用于造模的物质(LPS、LPS/Ca),实验组给予一定剂量的 LPS/Ca,对照组给予生理盐水,观察形成血栓的长度。现有不同种系的大鼠和小鼠各 10 只。请选用合适的实验设计类型。

设计类型的判定:在该研究中涉及了 3 个实验因素:造模与否、老鼠的类别和种系。由研究目的可知,大鼠、小鼠均有不同的种系,即“类别”下嵌套着“种系”这一因素,每种种系的鼠又被随机地决定接受生理盐水还是 LPS/Ca。由专业知识可知,对观测指标的影响由大到小依次为:造模物质、老鼠类别、老鼠种系,故该实验应采用三因素系统分组设计。在实验中可将同一种系的 10 只大鼠或小鼠完全随机地分成 2 组,每组 5 只,随机地决定其中 1 组接受 LPS,另 1 组接受 LPS/Ca,经过一段时间后观察形成血栓及血栓长度。该实验设计的标准型如表 6 所示。

表 6 使用不同造模物质后不同类别和种系老鼠的血栓形成情况

造模物质	老鼠类别	种系	鼠数	血栓长度 (mm, $\bar{x} \pm s$)
LPS/Ca	大鼠	SD	5	...
LPS			5	...
LPS/Ca	Wistar		5	...
LPS			5	...
LPS/Ca	小鼠	KM	5	...
LPS			5	...
LPS/Ca	ICR		5	...
LPS			5	...
LPS/Ca	Balb/c		5	...
LPS			5	...

2 携带效应的考察是应用拉丁方设计、交叉设计等需重点关注的问题

洗脱期是用来消除前一种处理对后一种处理所带来的影响的一种重要手段和方法。以下 3 种情形均采用拉丁方设计,请比较 3 种情况下得出的实验结果的说服力大小。

情形 1:用 4 台秤分别测量 4 个人的体重,要求每个人都用 4 台秤测量,比较 4 台秤的测量结果是否一致。

情形 2:用 4 台血压计分别测量 4 个人的血压,要求每个人都用 4 台血压计来测量 1 次,比较 4 台血压计的测量结果是否一致。

情形 3:4 个人,要求每个人按不同的服药顺序先后服用 4 种降压药,比较 4 种降压药的疗效是否相同。

在这 3 种情形中,每一个人在使用下一台秤或下一台血压计或下一种降压药前,均要间隔一段时间(称为洗脱期),以减少或排除前一次处理对观测指标的影响。

以上 3 种情形表面上看似似乎均可采用 4×4 拉丁方设计,但并非 3 种情形采用拉丁方设计来安排试验都妥当。在情形 1 中,每个人均要用 4 台秤称体重,因为人的体重在短时间内变化很小,受心理、外界环境的影响很小,不同的秤测量人的体重也不会对体重造成什么影响,洗脱期可以很短。此时采用拉丁方设计效果很好。在情形 2 中,由于人的血压在短时间内有可能会有较大变化,也会受到心理、外界环境等的影响,若在很短时间用不同的血压计测量人的血压可能会存在携带效应,因此,采用拉丁方设计的效果会较情形 1 稍差。在情形 3 中,让每人先各服用 1 种降压药,洗脱期应根据临床上的预试验结果来决定,应尽可能使前一次服用药物的疗效接近于零,使患者血压恢复到服药之前的水平,否则,携带效应必然会影响到后续药物疗效的评价。因此,在情形 3 中,若洗脱期不够长,或降压药已造成血压发生不可逆转的变化,即使对携带效应进行了评价,也很难真正反映药物的独立疗效,因此,对情形 3 使用拉丁方设计效果最差,严格地说,是不妥的。

由此可见,专业知识及具体问题的实际情况在实验设计中起着关键性的作用。成功的实验设计必须以专业知识作为基础和依据,以统计学思想为指导。在科研工作中,应将专业知识和统计学知识有机结合起来,对具体问题做出系统、全面、科学、合理的安排,只有这样才能使研究结论具有较高的可信度,经得起实践和时间的检验。

参考文献

[1] Hu LP. Application of triple-type theory of statistics in experimental design. Beijing: People's Military Medical Press, 2006:64-65, 85-120. (in Chinese)
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, 2006:64-65, 85-120.

[2] Hu LP. Modern statistics and the application of SAS. Beijing: Press of Military Medical Sciences, 2000:74-77. (in Chinese)
胡良平. 现代统计学与 SAS 应用. 北京:军事医学科学出版社, 2000:74-77.