

## · 讲座 ·

## 如何正确进行生物医学科研设计

## V. 如何合理选择 4 种多因素实验设计 (1)

李长平, 胡良平, 高辉, 周诗国, 郭晋, 金松华

## 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。

在前几讲中介绍了如何实现单因素、两因素及三因素实验设计, 那么当实验中涉及到的因素(因子)较多或因素(因子)的水平数较多时, 若进行析因设计(全因子实验), 全部因素的每种水平组合下都至少要做 2 次独立实验, 总实验次数可能多到实际工作者不能承受的程度。这就产生了 1 个问题: 在进行实验设计时如何做到既能减少实验次数, 只做全部水平组合中的一部分, 又能达到实验目的(能比较各因素的重要性且能比较准确地反映各水平组合的优劣)呢? 这就是全因子实验的部分实施问题。

当各因素间的交互作用不存在或可以忽略时, 可采用拉丁方设计、希腊拉丁方设计等进行部分实施。当因子间存在交互作用时, 可采用分式析因设计、正交设计、均匀设计等进行部分实施。

需要指出的是, 进行全因子实验的部分实施, 不同的实验设计之间并不等同, 它们之间既有联系又有区别。因此本讲中通过对析因设计、分式析因设计、正交设计、均匀设计 4 种实验设计的介绍, 阐述如何正确选择这 4 种实验设计类型。

## 1 4 种实验设计的特点及适用范围

## 1.1 析因设计

析因设计又称全因子实验设计, 当实验因素个数大于等于 2, 各因素在实验中同时实施且所处的地位基本平等, 因素之间存在一级、二级甚至更复杂的交互作用, 且需要加以全面考察时, 所采用的多因素实验设计类型。

析因设计有如下特点: 实验中涉及的实验因素的个数不少于 2 个; 实验中的实验条件是全部实验因素水平的全面组合; 在每个实验条件下至少要做 2 次独立重复实验; 各实验因素同时施加且地位平等。

当实验因素的个数不少于 2 个且因素之间的各阶交互

作用不可忽视时, 在研究者的时间、人力、经费等允许的情况下, 应选用此设计。一般来说当实验因素的个数大于 6 时, 不宜选用此设计。

## 1.2 分式析因设计

如果实验者能合理地判断各实验因素的某些高阶交互作用可被忽略, 则关于主效应和低阶交互作用的信息就可只做完全析因实验的一部分来得到, 此类设计被称为分式析因设计。

它有以下特点: 分式析因设计是析因设计的一部分; 分式析因设计可以投影到更大的由重要因素的子集组成的设计中去; 可将 2 个或多个分式析因设计序贯地组合成 1 个设计, 用来估计所感兴趣的因素效应及其交互作用。

当研究的实验因素或因素的水平数较多, 采用析因设计所需要做的实验次数很多, 在实际工作中不便实现时, 可采用分式析因设计。分式析因设计主要用于筛选实验。此类实验是要在众多因素中识别出对观测结果有较大的影响的因素。在随后的实验中, 采用析因设计对被识别的重要的因素进行更为深入细致的研究。

## 1.3 正交设计

正交设计是一种高效、多因素实验的设计。它是利用一套规格化的正交表将全部实验因素、因素之间水平的部分组合均匀搭配, 合理安排, 大大减少实验次数, 并提供较多的信息。

正交设计有 3 个最突出的特点: 其一, 由正交表挑出来的实验点在空间具有“均匀分散性”; 所谓均匀分散性,

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(李长平、胡良平、高辉、周诗国、郭晋); 100850 北京, 北京华素制药股份有限公司(金松华)

通讯作者: 胡良平, Email: lphu812@sina.com

就是实验点在空间分布得很均匀,无论从哪个角度看,都是有代表性的点被挑出来了,这个性质主要是由正交表中水平的排列规律决定的,这种规律叫“正交性”。其二,由正交表挑出来的实验点在统计分析时具有“整齐可比性”;其三,某些好的未包括在正交表中的实验点,可以通过统计分析将其发现。

在可以应用析因设计的实验研究中,若高阶或部分低阶交互作用可以忽略不计,且因实验条件所限,希望减少实验次数时,可采用正交设计以达到减少实验次数的目的。

#### 1.4 均匀设计

均匀设计是基于数论方法(或伪蒙特卡罗方法)推导出来的实验设计方法。它是只考虑实验点在实验范围内均匀分布的实验方法。它借助于均匀设计表安排实验。

它有以下特点:均匀设计表每列中每个数字出现次数相同且任意 2 列同行数字构成的数对在全部行中出现的次数与其他数对出现的次数不一定完全相同;均匀设计保留并进一步增强了正交设计“实验点在空间具有‘均匀分散’”的优点,但这是以牺牲“实验点在数学上的正交性”为代价的;均匀设计因丧失了“部分正交性”而不具有“实验点在统计分析时的‘整齐可比性’”,所以均匀设计定量资料不能像正交设计那样可以采用相应的方差分析处理资料,而需要借助多重回归分析的方法来分析数据。

当实验中拟考察的实验因素很多且因素的水平也很多,即使使用正交设计仍感觉实验次数太多时,可选用均匀设计。它一般适用于全部因素为定量因素的实验研究场合,可进行多因素多水平的筛选实验,用最少的实验次数获得可能对实验结果有统计学意义的少数几个实验因素的信息。

## 2 4 种实验设计的异同点

### 2.1 相同点

分式析因设计、正交设计、均匀设计都属于析因设计的部分实施;均可考察多个实验因素对响应变量(即观测结果)的影响;各因素同时进入实验,且因素间地位平等;均可通过较少的实验次数来考察因素及因素间的全部或部分交互作用;无论是分式析因设计、正交设计还是均匀设计,其实验设计的实施均应以专业知识为依据,将感兴趣的实验因素罗列出来,根据研究的目的(如考虑哪些交互作用项等)、能承受的实验次数等,选择相应的实验设计类型。

### 2.2 不同点

无论是分式析因设计还是正交设计、均匀设计,其实验设计不同点的本质在于:从全因子实验点中选取其中部分实验点来构造 1 个设计方案所采用的方法不同。比如,从全因子实验点中,按正交表选取实验点,使实验点具有均匀分散性、整齐可比性,就称该实验设计为正交设计;若选取的实验点是根据分辨率的大小将实验点分成相应的几部分,称其中的任何 1 个部分的设计为分式析因设计。同理,均匀设计是按照均匀分散性从全因子实验点中挑选实验点组成的设计方案。

所需实验次数不同。不言而喻,析因设计是全因子实验,所需做的实验次数最多。比如有 A、B、C 3 个因素,每个因素有 2 个水平,全部水平组合即实验点是 8 个(称之为  $2^3$  设计)。它要估计的效应有 3 个单因素效应、3 个两因素间的交互效应和 1 个三因素间的交互效应。此时,每个实验点至少要做 2 次独立重复实验,故若进行全因子析因设计实验则至少需要做 16 次实验;若各实验点条件下实验结果非常稳定且仅需考虑部分两因素之间的交互作用,此时不需要进行重复实验,可用  $L_8(2^7)$  正交表来安排此实验;若 4 个交互作用项都无统计学意义可以忽略时(应有专业知识为依据),则可用较小的正交表  $L_4(2^3)$  或分式析因设计来进行  $2^3$  设计的 1/2 实施,即只做 4 个实验点,但各实验点上至少要做 2 次独立重复实验,否则,无法估计实验误差,即总共也需要做 8 次实验。这 2 种设计方案都是做 8 次实验,哪一个更好一些呢?若相同实验条件下观测结果很稳定,前一种方案优于后一种方案,因为它涉及的实验点全面;反之,后一种稍好一点,因为通过重复实验,可获得相当稳定一些的实验结果。相同的因素和水平下,析因设计所需实验次数最多,正交设计、分式析因设计次之,均匀设计所需实验次数最少。

实验中的作用不同。均匀设计最适合用于多因素的筛选实验,用很少的实验次数获得可能对实验结果有统计学意义的少数几个实验因素的信息,再利用分式析因设计或正交设计或析因设计对这些实验因素做进一步的验证性实验,从而达到预期的目的。

每种实验设计本身的特点决定了其结果的可靠性。全因子实验结果最可靠。均匀设计以牺牲部分正交性、最大化地追求均匀性来达到减少实验次数的目的,当实验次数少于因素个数时,采用多重回归分析方法来研究响应变量随自变量变化的回归方程不是唯一的,并且其在实验设计时并没有考虑交互作用,只有通过回归分析时引入因素之间的交叉乘积项等来进行探索。因此,其结果不够稳定。而正交设计比均匀设计更严格些,通常实验次数要多一些,统计分析方法较简单(通常可用相应实验设计定量资料的方差分析),结果较稳定。

## 3 析因设计的部分实施

在 2 中对 4 种实验设计异同点做了简要介绍,下面以  $2^3$  设计的 1/2 部分实施为例,并结合实例介绍如何实现析因设计的部分实施。

如果 1 个实验设计中有  $k$  个因素,每个因素有  $p$  个水平,全部水平组合有  $p^k$  个。这种析因设计称为  $p^k$  设计。当  $p$  的数值较大时,每增加 1 个因素,  $p^k$  设计的实验次数就要大量增加。因此在实际问题中,使用  $p^k$  设计时,一般高阶交互作用都可忽略不计,而且,有时还有部分两因素交互作用也可忽略不计。此时,可采用  $p^k$  设计的部分实施。如  $2^{3-1}$ 、 $2^{4-1}$ 、 $2^{4-2}$ 、 $2^{5-1}$ 、 $2^{5-2}$ 、 $2^{5-3}$ 、 $2^{6-1}$ 、 $2^{6-2}$ 、 $2^{6-3}$ 、 $2^{6-4}$ , 同理还有  $3^{4-1}$ 、 $3^{4-2}$ 、 $3^{5-1}$ 、 $3^{5-2}$ 、 $3^{5-3}$ 、 $3^{6-1}$ 、 $3^{6-2}$ 、 $3^{6-3}$ 、 $3^{6-4}$  设

计等都是分式析因设计。

现设 A、B、C 为 3 个 2 水平因素，其全部水平组合有 8 个实验点，可称为  $2^3$  设计，若仅作一半实验点，称为原设计的 1/2 实施，记作  $2^{3-1}$  实施。请看例 1。

【例 1】某化工厂为提高某化工产品的转化率，经分析研究决定选取的因素与水平如下：

水平	反应温度 (°C) A	反应时间 (min) B	用碱量 (%) C
1 (-)	80	90	5
2 (+)	85	120	6

欲通过实验找出影响转化率（越高越好）的各因素最佳水平组合及各因素的主次顺序。通过预实验发现因素之间的交互作用可以忽略不计，希望实验次数尽可能少一些，那么应该如何设计实验？

分析：该实验中涉及到了 3 个实验因素，每个实验因素有 2 个水平，若选择  $2^3$  析因设计，它有 8 种组合，每种组合至少要做 2 次独立重复实验，总实验次数至少为 16 次。现因条件限制，要尽量地减少实验次数，且通过预实验发现因素之间的交互作用可以忽略不计，选择实验设计类型时可有如下选择。

(1) 首先可构造分式析因设计 首先 3 个因素为 A、B、C，用“-”表示它们的第 1 水平，用“+”表示它们的第 2 水平。为方便起见，把实验点按各种组合 a、b、c、abc、ab、ac、bc、(1) 写出（见表 1 第 1 列），然后再把表 1 分解，考虑从中选择何种组合。

表 1 加减符号表达的各因素的水平

处理 组合	用加减符号表达的各因素的水平						
	因素效应: I	A	B	C	AB	AC	BC
a	+	+	-	-	-	-	+
b	+	-	+	-	-	+	-
c	+	-	-	+	+	-	-
abc	+	+	+	+	+	+	+
ab	+	+	+	-	+	-	-
ac	+	+	-	+	-	+	-
bc	+	-	+	+	-	-	+
(1)	+	-	-	-	+	+	-

说明：第 1 种处理组合记为“a”，A、B、C 的水平分别为“+”、“-”、“-”。表明第 1 次实验仅涉及 A 因素的第 2 水平（反应温度为 85 °C）、B、C 因素均为第 1 水平，同理可以理解其他处理组合的含义。最后 1 种处理组合(1)表明 A、B、C 均为 1 水平。表中第 2 列“I”为单位列，用它来标示何时最高阶交互作用（该例为 ABC）皆取“+”号、何时 ABC 皆取“-”号。显然，前 4 行 ABC 皆取“+”号，简记为“ $I=ABC$ ”、后 4 行 ABC 皆取“-”号，简记为“ $I=-ABC$ ”。故称  $I=ABC$  为该设计的定义关系。按 ABC 是否取相同的符号将其分成 2 个部分。可以看出，取前 4 行与取后 4 行是完全等价的。一般来说，分式析因

设计的定义关系总是与单位列 I 相等的各行构成的集合。故取  $I=ABC$  所对应的那些处理组合即表 1 前 4 行为  $2^3$  析因设计的部分 1/2 实施的分式析因设计。

并且，在该实验中，交互作用可以不考虑，因此，可选用表 1 中前 4 行进行实验，仅需做 4 个实验点。若所有的一级、二级交互作用不考虑的话，仅需做 4 个实验点，每个实验点做 2 次独立重复实验更好，即共做 8 次实验就可达到本例的研究目的。

将表 1 各符号转换为因素的真实水平，使设计看上去简单明了，易于实施，可将表 1 转换为表 2 形式。

表 2 分式析因设计安排 3 个实验因素

实验号	1	2	3	4	5	6	7
	反应温度 (°C)	反应时间 (min)	用碱量 (%)				
1	85	90	5	-	-	+	+
2	80	120	5	-	+	-	+
3	80	90	6	+	-	-	+
4	85	120	6	+	+	+	+

当只考虑 3 个实验因素不考虑因素间交互作用时，在实验的实施过程中可将 4~7 列去掉（仅在采用手工计算时有用）。可见只需做 4 个实验点就可满足研究者的要求。

这里需要说明的是，该分式析因设计中分辨度为 3。分辨度就是人们根据定义关系中所包含的最短词的字母个数 R 来命名的，可用来度量混杂程度。也就是说，根据分辨度将析因设计的表来进行分解的。通常 R 取 3、4、5。分辨度太大太小均不符合实际。在该例中，定义关系  $I=ABC$ ，故  $R=3$ ，也就是说，该分式析因设计是分辨度为 III 的设计。可由该研究看出，分辨度为 III 的设计中，只有当两因素间的交互作用可以忽略不计时，才能估计 3 个主效应。这是因为在估计因素 A 的效应的同时，包含了 BC 的效应（因为这两列对应行上的“+”、“-”号相等）；在估计 B 因素效应的同时，包含了 AC 的效应；估计 C 因素效应的同时，包含了 AB 的效应。这就说明，当我们估计 A、B、C 效应的同时，实际上估计的是“ $A+BC$ ”、“ $B+AC$ ”、“ $C+AB$ ”的效应。所以说，只有当两因素间的交互作用可以忽略不计时，才能估计 3 个主效应。若在分析 3 个因素主效应的同时还要准确分析因素间的交互作用，这就意味着要增加实验次数、提高分辨度。

常用的分式析因设计由 2 水平或 3 水平的 k 因素析因设计演变而来，简记为  $2^{k-P}$  设计或  $3^{k-P}$  设计。它们分别代表“k 个 2 水平因素析因设计的  $2^P$  分之一的分式析因设计”或“k 个 3 水平析因设计  $3^P$  分之一的分式析因设计”。若加上每个设计所对应的分辨度可简记为  $2^R$  或  $3^R$ 。这里 R 代表 III、IV、V 3 种分辨度。

对于该分式析因设计，根据研究目的，若结果是定量指标且资料满足参数检验的前提条件，可采用与其实验设计相对应的方差分析来处理，若在表 2 的 4~7 列安排了其他

的实验因素或欲探索性地分析因素间的交互作用,则不能采用方差分析的方法,需采用多重线性回归分析来处理。

(2)也可采用正交设计来实现:当 3 个两因素间交互作用、1 个三因素间交互作用都无统计学意义时(例 1 中已做预实验辨别此判断可信),欲减少实验点数,可用较小的正交表  $L_4(2^3)$  实现研究目的。(从表 1 中前 4 行的第 3~5 列可以发现,该 3 列与正交表  $L_4(2^3)$  是等价的。)如表 3 所示。

表 3  $L_4(2^3)$  正交表

实验号	列号:	1 (A)	2 (B)	3 (C)
1		1	1	1
2		1	2	2
3		2	1	2
4		2	2	1

它的 4 个实验点对应着 4 个水平组合。实际实施时仅需将 1、2 列号的 2 个水平转换成对应的 3 个实验因素的各个水平进行实验。这个设计只做  $2^3$  设计中的一半实验,故称它为  $2^3$  的 1/2 实施。在正交表  $L_4(2^3)$  中,第 1、2 列分别安排因素 A、B,第 3 列原来是用来估计交互作用 AB 的,现在安排因素 C 是因为交互作用 AB 无统计学意义,可忽略不计,此列可用来估计因子 C 的效应。

该实验设计若同时分析 3 个因素的主效应,在正交表  $L_4(2^3)$  中安排后,已无空白列,则各行对应的实验条件下需要做重复实验。根据正交设计的特点,其正交设计的定量资料可采用其相应的方差分析来处理,也可采用多重回归分析的方法进行统计分析。

需要指出的是,析因设计( $p^k$ )及其部分实施,它的实验设计与实验结果的统计分析一般可采用正交表来实现。可以这样说,由 1 个正交表所挑选出来的实验点,相当于与之对应的析因设计所决定的全部实验点中最有代表性的那些点,若点数一样多,则二者完全等价(如表 1 中 3~9 列与  $L_8(2^7)$  正交设计等价)。例如,具有  $k$  次独立重复实验的  $2^2$  析因设计,与具有  $k$  次独立重复实验的  $L_4(2^3)$  正交设计完全等价,若不考虑独立重复实验,  $L_4(2^3)$  正交设计抽取的 4 个实验点是  $2^3$  析因设计所决定的 8 个实验点中最有代表性的 4 个实验点;  $L_8(2^7)$  正交设计等价于  $2^3$  析因设计(是否进行重复实验和重复次数必须相同)、是  $2^4$  析因设计的“1/2 实施”、是  $2^5$  析因设计的“1/4 实施”、是  $2^6$  析因设计的“1/8 实施”等等,详细介绍请参见有关著作<sup>[1-2]</sup>。

(3)也可采用均匀设计:由于 2 水平的均匀表就是 2 水平的正交表,故  $2^3$  的 1/2 实施采用  $L_4(2^3)$  正交设计和采用  $U_4(2^3)$  是相同的,故不再赘述。

参考文献

[1] Wang WZ. The design and analysis of experiments. Beijing: Higher Education Press, 2004:114-150, 327-330. (in Chinese).  
王万中. 试验的设计与分析. 北京: 高等教育出版社, 2004:114-150, 327-330.

[2] Hu LP. Application of triple-type theory of statistics in experimental design. Beijing: People's Military Medical Press, 2006:121-132. (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, 2006:121-132.

· 环球动态 ·

膝关节擦伤与肿瘤生长存在共同点

肿瘤细胞能够激发与切口或伤口相同的免疫细胞应答这种观点曾存在争议。但是目前越来越多的证据显示,参与表皮擦伤的膝关节愈合的细胞机制在预防肿瘤生长中也起作用。而最近的一项研究揭示了果蝇的肿瘤生长与组织损伤之间存在共同点,这项研究成果刊登在 2008 年 9 月 18 日出版的 *Disease Models & Mechanisms* 上。

耶鲁大学医学院徐天和他的同事通过检测果蝇血细胞和 1 种免疫细胞在遗传学因素诱发的肿瘤应答过程中的活性发现,肿瘤会激发外周循环中的血细胞复制并黏附在肿瘤表面从而抑制肿瘤生长。通过与受伤及伴有组织损伤的正常果蝇的血细胞活性对比发现,这些受伤果蝇也出现血细胞增殖现象,这与罹患肿瘤的果蝇的情况是一致的。此外,通过检测免疫应答反应激发的分子信号也发现,肿瘤细胞的自我分裂和对周围组织的损伤是诱发血细胞应答的部分原因。

这项研究为之前关于免疫应答与肿瘤之间有联系的报道提供了进一步的支持依据,且阐明了果蝇对肿瘤生长和组织损伤产生免疫应答的重要通路;然而人类也存在相同的通路,这一点表明果蝇可以作为研究具有增强人体自身免疫反应抵抗癌症作用的潜在药物的靶点。

(胡璧)