

## · 讲座 ·

# 如何合理选择统计分析方法 处理实验资料 (I)

胡良平

在阅读生物医学杂志论文时,不难发现如下的现象:只要结果是定量资料,人们普遍运用“ $t$  检验”、“单因素方差分析”或“秩和检验”来处理。事实上,在人们用前述方法处理的定量资料中,有相当多的定量资料同时受到多个因素(通常包括实验因素和重要的非实验因素)的影响,即定量资料来自某种特定的多因素实验设计类型。这种用单因素设计定量资料的统计分析方法处理原本属于多因素实验设计的定量资料,其结论的可信度大为降低,有时,甚至会不可避免地得出错误的结论。同理,结果为定性资料时,很多人不管结果变量的具体情况,也不管原因变量有多少个,甚至不管统计分析的目的是什么,一律盲目套用  $\chi^2$  检验。在采用其他统计分析方法时,也存在盲目套用的问题。正因如此,使得我国乃至全世界生物医学杂志论文的质量令人担忧(错误率平均约为 80%),大大降低了科研工作的科学性和严谨性。

值得注意的是:统计分析方法选用不当,导致论文的可信度下降或得出错误结论,只是科研质量差的一种表面现象。更为严重的是相当多的科研课题的设计质量低劣,这是不可弥补的错误。事实上,在 1990—2003 年期间,在 *New England Journal of Medicine*、*Journal of American Medical Association* 和 *Lancet* 上发表的曾被引用 1000 次以上的 45 篇学术论文中,其结论约有 1/3 未经受得起时间的考验<sup>[1]</sup>,其“症结”就是科研设计不严谨、不科学所致。当今世界,人们普遍追求论文的数量、特别是追求在《科学引文索引(SCI)》上发表的论文数量,却很少有人追求学术论文的质量,这是应当引起全世界广泛关注的一件大事!

本文和后续的几篇文章将结合生物医学科研中的实际问题,介绍如何合理选用统计分析方法处理实验资料。

较不同条件下定量指标的平均值或平均秩之间的差别是否具有统计学意义所采用的统计分析方法常被称为“定量资料的统计分析”。如何才能合理选用统计分析方法处理定量资料呢?关键在于以下两点:第一,检查定量资料是否满足参数检验的前提条件;第二,正确判定定量资料所对应的实验设计类型。

参数检验通常指以正态分布作为理论根据的  $Z$  检验、以  $t$  分布作为理论根据的  $t$  检验和以  $F$  分布作为理论根据的方差分析。使用参数检验法时,要求定量资料满足 3 个前提条件,即独立性、正态性和方差齐性。对于单组设计和配对设计的定量资料,要求满足独立性和正态性 2 个前提条件;而对于成组设计、单因素多水平设计和各种多因素设计的定量资料,则要求满足前述 3 个前提条件。非参数检验(通常为秩和检验)对定量资料没有明确的要求,但它能解决的问题通常比较简单,即定量资料对应的实验设计类型为单因素设计或随机区组设计。当定量资料不满足参数检验的前提条件又没有相应的非参数检验方法可供选用时,就需要设法寻找合适的变量变换方法(如的对数变换、开方变换等),当变换后的定量资料满足参数检验的前提条件时,仍可采用相应设计定量资料的统计分析方法处理。

标准的实验设计类型有很多种,如单因素设计(包括单组设计、配对设计、成组设计和单因素多水平设计)、随机区组设计、拉丁方设计、交叉设计、析因设计、具有重复测量设计等。其中,使用频率最高的实验设计类型为“单因素设计”、“析因设计”和“具有重复测量设计”3 类。还有一类,看上去很像“单因素多水平设计”,其实,它既不是真正的“单因素多水平设计”、也不是某种“标准的多因素设计”,而是“多因素非平衡的组合实验”,就是仅凭专业知识而非科学的统计学知识确

## 1 正确分析定量资料的关键点

当人们关心的结果变量为定量变量时,为了比

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  
收稿日期:2007-01-15

定的若干个组，它们涉及到多个因素，但因素的水平组合并不完全。对于这种“七拼八凑”出来的实验分组，应采取“拆分组别”的技术进行重新组合，根据专业知识和统计学知识确认各种新组合所对应的“实验设计类型”。

对于定量资料是否满足参数检验的前提条件，可以借助专业知识（如判断是否具备独立性要求）和统计分析软件予以实施；而如何判定定量资料所对应的实验设计类型，则需要使用者具备较专门的统计学知识。因此，在上述两个关键点中，第二个关键点显得尤为重要，正确掌握其技能具有相当大的难度，故本文在解决每个实际问题时，将把重点放在实验设计类型判定上，不再赘述如何选择统计分析方法。因为定量资料前提条件的检查和实验设计类型的判定一旦得到正确落实，只需要选用相应设计定量资料的统计分析方法处理即可。

实际工作者在撰写论文时务必要注意：在交代所使用的定量资料统计分析方法之前，一定要加上实验设计的名称，不应仅说用了“秩和检验”、“*t* 检验”或“方差分析”。应该采用如下的方法来表述：例如，①采用成组设计定量资料的 Wilcoxon 秩和检验；②采用单因素多水平设计定量资料的 Kruskal-Wallis 秩和检验；③采用随机区组设计定量资料的 Friedman 秩和检验；④采用成组设计定量资料的 *t* 检验；⑤采用单因素多水平设计定量资料的方差分析；⑥采用三因素析因设计定量资料的方差分析；⑦采用具有一个重复测量的两因素设计定量资料的方差分析。当需要对多因素设计中的某个因素各水平下的平均值进行两两比较时，若因素之间的交互作用项具有统计学意义，不应简单地采用单因素多水平设计定量资料的 *q* 检验法或 LSD 检验法或 Dunnett 检验法，而应建立在相应设计定量资料的方差分析基础之上，具体算法比较复杂，请参阅有关专著。

2 实例及解析<sup>[2]</sup>

例 1 请根据表 1 中的定量资料辨析实验设计类型  
分析与解答：表格设计不够规范，不便看出实验是怎样做的，因此，也就不便于辨析实验设计类型。若 8 行数据测自 8 个样品，应在第 1 列前加上“样品编号”列，表格应改成表 2 的形式，它是含有 2 个定量观测指标的定量资料，其实验设计名称为“自身配对设计”。这是一个“自身配对设计的二元定量资料”。

表 1 不同色谱进样方式测定 PFIB 染毒浓度

注射器进样分析		自动采样分析	
峰面积	PFIB (mg/ml)	峰面积	PFIB (mg/ml)
14337	0.120	14648	0.123
15874	0.131	16384	0.135
17325	0.141	17078	0.140
23007	0.182	22178	0.176
24980	0.196	24475	0.192
26160	0.204	27380	0.213
29598	0.228	30304	0.233
30439	0.234	30590	0.235

表 2 不同色谱进样方式测定峰面积和 PFIB 染毒浓度

样品 编号	峰面积 (单位)		PFIB (mg/ml)	
	注射器进样	自动采样	注射器进样	自动采样
1	14337	14648	0.120	0.123
2	15874	16384	0.131	0.135
3	17325	17078	0.141	0.140
4	23007	22178	0.182	0.176
5	24980	24475	0.196	0.192
6	26160	27380	0.204	0.213
7	29598	30304	0.228	0.233
8	30439	30590	0.234	0.235

若在表 1 中，不是用两种进样方法平行测定 8 份样品中 2 个定量观测指标的数值，而是将 16 份样品完全随机地均分成 2 组，分别用两种进样方法独立地各测定 8 份样品中 2 个定量观测指标的数值，则应将表 2 中第 1 列“样品编号”改为“编号”，其列表格式见表 3。

表 3 不同色谱进样方式测定峰面积和 PFIB 染毒浓度

编号	峰面积 (单位)		PFIB (mg/ml)	
	注射器进样	自动采样	注射器进样	自动采样
1	14337	14648	0.120	0.123
2	15874	16384	0.131	0.135
3	17325	17078	0.141	0.140
4	23007	22178	0.182	0.176
5	24980	24475	0.196	0.192
6	26160	27380	0.204	0.213
7	29598	30304	0.228	0.233
8	30439	30590	0.234	0.235

表 2 与表 3 看上去几乎完全相同，其实，在实验设计类型上是不同的。表 2 对应的实验设计叫“自身配对设计”，即每行的同一个样品被一分为二，用两种进样方法检测 2 个定量指标的数值；而表 3 对应的实验设计叫“成组设计”，每行代表两个独立样品，分别用两种进样方法检测 2 个定量指标的数值。这是一个“成组设计的二元定量资料”。

例2 请根据表4中的定量资料辨析实验设计类型

表4 BMS与对照组的年龄、性别构成比较				
组别	例数		平均年龄(岁)	
	性别: 男	女	男	女
BMS组	21	66	53.57±11.55	52.27±10.67
对照组	22	60	48.27±14.36	50.82±10.20

分析与解答: 表4中性别是定性资料, 需要用 $\chi^2$ 检验进行统计分析。在对年龄进行统计分析时, 组别是一个因素, 含2个水平: BMS组和对照组; 性别是另一个因素, 具有2个水平: 男和女。关于组别和性别两个因素对“年龄”的影响哪一个重要, 本文尚无专业知识为依据, 所以对年龄进行统计分析时, 实验设计类型是两因素析因设计。

例3 请根据表5中的定量资料辨析实验设计类型

表5 甲乙两种方法处理的样品在3种不同稀释度条件下做实验测得的结果				
处理方法	OD <sub>405</sub>			
	稀释度	1/4	1/8	1/16
甲		0.283	0.288	0.307
		0.291	0.311	0.291
		0.312	0.288	0.299
		0.311	0.311	0.320
乙		0.344	0.322	0.320
		0.353	0.344	0.331
		0.352	0.340	0.315
		0.384	0.363	0.332

分析与解答: 表5中包含两个实验因素。一个因素是处理方法, 具有2个水平: 甲和乙; 另一个因素是稀释度, 具有3个水平: 1/4、1/8和1/16; OD<sub>405</sub>是定量观测指标。关于处理方法与稀释度两个因素对“OD<sub>405</sub>”的影响哪一个重要, 本文尚无专业知识为依据, 所以该实验设计类型是两因素析因设计。

例4 请根据表6中的定量资料辨析实验设计类型

表6 肝纤维化SSS计分值治疗前后的情况及比较			
组别	治疗前	治疗后	治疗前后差值
A组	6.76±6.67 (25)	4.72±5.63 (25)	2.04±2.59 (25)
B组	4.13±2.83 (24)	6.33±4.04 (24)	-2.21±3.72 (24)

分析与解答: 人们处理表6中定量资料的习惯的做法是: (1) 分别在A组、B组中采用配对设计定量资料的 $t$ 检验; (2) 再分别在治疗前、治疗后两个时间点上, 采用成组设计(也有人称为两独立样本)定量资料的 $t$ 检验。这样分析资料, 割裂了整体设计, 每次分析时资料的利用率低, 犯假阳性

错误的概率增大, 无法考察“处理”与“时间”之间的交互作用对结果的影响情况。

其实, 表6中包含两个因素, 一个因素是组别, 有2个水平: A组和B组; 另一个因素是时间, 具有2个水平: 治疗前和治疗后。此定量资料所对应的实验设计类型是“具有一个重复测量的两因素设计”, 最好选用“具有一个重复测量的两因素设计定量资料的方差分析”处理资料为宜。

值得注意的是, 具有重复测量设计的定量资料不满足“独立性”要求, 所选用的统计分析方法应有助于削弱同一个体多次重复观测数据之间的“相关性”, 因此, 其统计分析方法相当复杂, 必要时, 可请统计学专业人员协助处理为宜。

例5 请根据表7中的定量资料辨析实验设计类型

表7 各组实验侧与对照侧神经元数目比较( $\bar{x} \pm s$ )				
组内比较	例数	实验侧	对照侧	%/对照侧
F组	7	9.8020±5.85	12.5198±6.11	78.3
N组	6	6.8318±2.85	11.7290±3.68	58.1
C组	7	5.8000±2.58	10.9160±3.69	53.2

分析与解答: 本例的关键是要弄清F、N、C这3组的真正含义, 若这3组确实是一个实验因素的3个水平, 则整个实验研究所对应的实验设计类型就叫做“具有一个重复测量的两因素设计”, 因为要观测每个受试对象实验侧与对照侧上神经元数目, 这里的重复测量因素就是“部位”, 实验分组因素为“F、N、C”的总称, 如药物种类或治疗方法; 若这3组中涉及两个因素但组合不完全, 则实验分组时就犯了“对照不全”的错误, 需要结合实际情况将3组拆分成“F与N组”、“F与C组”或“N与C组”, 这3种组合中的任何一种与“实验侧、对照侧”同时考察时, 其对应的实验设计类型都叫做“具有一个重复测量的两因素设计”。

例6 请根据表8中的定量资料辨析实验设计类型

表8 化合物9714对大鼠动静脉旁路血栓的影响( $\bar{x} \pm s$ )				
组别	鼠数	剂量(mg/kg)	血栓湿重(mg)	血栓干重(mg)
模型对照组	14	-	124.21±144.93	27.79±11.84
阿司匹林组	14	40	89.36±23.86	18.07±4.95
化合物9714组	14	10	80.64±42.70	16.43±10.03
化合物9714组	14	20	74.93±37.04	15.43±9.88
化合物9714组	14	40	66.50±29.78	12.14±7.43

分析与解答: 表8中“血栓湿重(mg)”与“血栓干重(mg)”是两项定量观测指标; 很多人都认

为表 8 定量资料对应的实验设计类型为“单因素 5 水平设计”。的确,“组别”看上去很像一个实验因素的名称,其实这是一种假象,也就是说,“组别”一词常具有“欺骗性”。在本例中,“组别”实际上包含了两个因素,即在模型基础上“用何种药”以及用化合物 9714 的“剂量”,它们之间的水平搭配并非完全,故这是一个“两因素非平衡的组合实验”,需要进行合理拆分后,才可以判定各种组合所对应的实验设计类型。注意:根据专业知识,通常应在模型基础上用药。拆分的结果如下:

组合 1:模型对照组、阿司匹林组(剂量 40 mg/kg)和化合物 9714 组(剂量 40 mg/kg);

组合 2:模型对照组、化合物 9714 组(剂量 10 mg/kg)、化合物 9714 组(剂量 20 mg/kg)和化合物 9714 组(剂量 40 mg/kg)。

组合 1 叫做单因素 3 水平设计,3 组动物都需

要先建模,然后用不同的药物处理,第 1 组虽然没有用药,但可视为用了 1 种没有任何疗效的药物(如安慰剂),故这个因素可叫做“药物种类”;组合 2 叫做单因素 4 水平设计,4 组动物都需要先建模,然后用不同剂量的化合物 9714 治疗,第 1 组可视为用了零剂量的化合物 9714 治疗,故这个因素可叫做“化合物 9714 的剂量”。

#### 参考文献

- [1] Hu LP, Zhang TM. An analysis of the factors affecting the validity of research findings and papers. *Sci Focus*, 2006,1(4): 9-19. (in Chinese)  
胡良平,张天明.影响我国科研成果和学术论文质量的要因分析.科学观察,2006,1(4):9-19.
- [2] Hu LP. Triple-type theory of statistics and its application in the experimental designs. Beijing: People's Military Medical Press, 2006: 1-10, 139-214. (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,2006:1-10, 139-214.

## · 环球动态 ·

### 苏格兰科学家利用数学“预报”癌症

苏格兰科学家在利用数学建立癌症治疗预测模型的研究中取得的成果,将可能免除某些患有乳腺癌妇女切除乳房的痛苦。邓迪大学数学系的高级讲师 Sandy Anderson 博士开发出了一种类似天气预报系统的数学模型,可用于预测肿瘤的生长。Anderson 博士的数学预测模型表明肿瘤生长的组织环境可以影响癌症的侵袭性。Anderson 博士称:“我的研究表明,如果肿瘤的生长环境有氧或者营养充沛,它的侵袭性可能不大;如果生长环境恶劣或者化疗与放疗的强度大,反而可能把一些侵袭性强的恶性肿瘤留存下来,并继续扩散。”

Anderson 博士的这个发现将使人们重新思考当前的一些治疗方法。Anderson 博士说:“目前大多数治疗对策是尽可能将肿瘤生长环境变得恶劣,实际上这会杀死侵袭性弱的肿瘤细胞而保留那些侵袭性强的细胞。我的研究表明在摧毁肿瘤前需要慎重考虑其生长的环境。例如,患有乳腺癌的妇女可能不必切除整个乳房。”

目前该项目注入的研究资金已达 770 万英镑。合作研究者美国的 Vito Quaranta 教授是一位知名的肿瘤生物学家,他认为该项研究是将数学和实验室研究相结合的创举,将开辟攻克癌症的又一途径。他说:“癌症研究的新纪元已经来临。数学给我们的研究领域带来了全新的视点。肿瘤不再是难缠的恶魔,而是可以理性描述并最终征服的复杂过程。我们的长远目标就是利用数学模型和计算机模拟,结束肿瘤治疗猜谜似的反复试验过程。”

苏格兰乳腺癌学术研究基金会(Breakthrough Breast Cancer)主任 Norman Freshney 指出:“了解肿瘤细胞的扩散对控制和治疗癌症至关重要。越来越多的证据表明肿瘤生长的环境是影响肿瘤扩散的一个重要因素。”乳腺癌治疗发展中心主任 Lorraine Dallas 认为该研究给预测肿瘤生长带来了希望,并有望促成专家治疗计划的发展。

(苏格兰国际发展局供稿,陈睿译)