

• 继续教育园地 •

《中国骨伤》杂志论文中存在的统计学错误辨析(6)

郭晋, 胡良平, 高辉

(军事医学科学院生物医学统计学咨询中心, 北京 100850)

本文对《中国骨伤》杂志 2006 年与 2007 年第 6 期所载论著、临床研究、基础研究文章进行仔细阅读, 对文章中所选用的统计分析方法进行研究, 针对实验设计和统计分析两个方面存在的统计学失误进行辨析并提出修改建议, 旨在提高骨科科研中实验设计与统计分析水平, 促进骨科科研的良好发展。

1 在实验设计中违反实验设计基本原则

一篇学术论文, 衡量其科学性优劣的重要标志之一就是科研设计的质量。科研设计包括专业设计和统计研究设计两个方面。在统计研究设计(包括实验设计、临床试验设计和调查设计)中, 实验设计的理论和方法几乎是通用的。而实验设计的核心内容可以概括为: 三要素(受试对象、试验因素和试验效应), 四原则(随机、对照、重复和均衡), 统计研究设计类型(观察性研究设计和实验性研究设计)^[1]。

1.1 样本含量过小 例 1:《强骨宝方对体外培养成骨细胞影响的实验研究》^[2]表述: 实验动物, 1~2 日龄 SD 乳鼠 5 只, 购于福建医科大学实验动物中心。颅骨成骨细胞的分离与培养, 取 2 只 1~2 日龄 SD 乳鼠颅盖骨, 先以 0.25% 胰酶溶液振荡消化, …… , 以后每 2~3 d 换液 1 次, 细胞汇合后进行消化传代。分组, 取 5~8 代成骨细胞, 2×10^4 /ml 的细胞密度接种于 96 孔培养板, 经同步化后, 分对照组与 4 个实验组, 共 5 组, 见表 1。

表 1 不同浓度的强骨宝方提取液对成骨细胞增殖功能影响的比较($\bar{x} \pm s, n=6$)

组别	成骨细胞分泌 ALP 值	
	时间(h):	
	48	72
100 μ g/ml 组	0.367 7 \pm 0.012 9	0.576 3 \pm 0.032 7
50 μ g/ml 组	0.399 3 \pm 0.031 5	0.625 0 \pm 0.020 1
10 μ g/ml 组	0.371 3 \pm 0.018 3	0.562 0 \pm 0.043 6
5 μ g/ml 组	0.349 5 \pm 0.020 4	0.528 4 \pm 0.029 3
对照组	0.355 0 \pm 0.019 8	0.513 7 \pm 0.017 6

辨析: 原作者购进乳鼠 5 只, 只用了 2 只, 而后所培养传代的细胞均来自这 2 只乳鼠, 培养的细胞分到 96 孔板, 试验分为 5 组, 每组 $n=6$ 。原作者这种做法违反了实验设计的重复原则, 96 孔并非样本含量, 而乳鼠数量 2 才是真正的样本含量, 5 个实验组样本均来自于 2 只乳鼠的细胞, 这样的实验代表性不够好, 不能全面真实反映出这种动物间的个体差异对实验结果的影响, 应该增加样本数量, 分几批重复此实验, 以便充分揭示药物剂量和作用时间对定量观测结果的影响。

1.2 犯了“对照不全”错误 在下面的例 2 和例 3 中, 原作者

实际上犯了“对照不全”错误, 即有些实验组有与其对应的对照组, 而有些实验组则没有与其对应的对照组。

2 在分析定量资料时误判了实验设计类型, 导致统计分析方法的误用

2.1 将多因素非平衡组合实验误判为多个单因素两水平设计 例 2:《活血化瘀中药联合血管内皮生长因子基因转移促进股骨头坏死处新生血管形成的实验研究》^[3]表述: 动物模型制作除对照组, 模型组、中药组、基因组和综合组均按文献方法造模。成功后, 对照组、模型组生理盐水 25 g/kg 灌胃, 每日 2 次。中药组用上述中药液 25 g/kg 灌胃每日 2 次。基因组将重组 pcDNA/VEGF121 质粒髌关节腔内注射, 每只 200 μ g, 每 2 周 1 次; 股动脉注射, 每只 200 μ g, 每 4 周 1 次。综合组为中药灌胃+基因治疗。共治疗 8 周, 资料见表 2。实验数据用均数 \pm 标准差表示, 采用 SPSS 12.0 统计软件包进行各组间方差分析及 t 检验。

表 2 各组分区血管计数情况

组别	动物数	血管计数($\bar{x} \pm s$)	
		A 区	B 区
对照组	8	3.96 \pm 1.12	10.56 \pm 2.87
模型组	8	3.18 \pm 1.07	6.43 \pm 2.25
中药组	8	3.25 \pm 0.93	8.24 \pm 1.88
基因组	8	3.84 \pm 1.22	10.03 \pm 3.54
综合组	8	4.37 \pm 1.85	10.82 \pm 3.69

辨析: 该实验涉及 3 个因素, 即“是否建模”、“是否用中药液灌胃”、“是否注入基因组质粒”, 每个因素有 2 个水平, 3 个因素全面组合共有 8 个组。原作者只列出 5 个组, 因此, 在实验设计时犯了对照不全错误, 导致该实验成为多因素非平衡组合实验。原作者误认为 5 个组(注: 综合组为中药组+基因组)是“单因素 5 水平设计”, 这种认识混淆了因素与水平之间的关系, 既不能达到实验因素的合适组合也无法考察因素之间的交互作用。对于多因素非平衡组合实验应采取“拆分”的方法重新组合, 以便在并非正确的实验分组条件下找到基本合理的统计分析方法处理实验资料。

根据原文实验, 对上述 5 组进行“拆分”, 在实验分组方面, 对“是否用中药灌胃”和“注入基因组质粒”两个因素可构造出一个两因素析因设计结构, 再结合被测定的两个区域(属于重复测量因素), 可构造出“具有一个重复测量的三因素设计”, 建议列成表 3 形式。

对“是否建模”这一因素而言, 从实验分组角度看, 可构造

表 3 是否用中药灌胃和注入基因组质粒各组分血管计数情况

是否用中药 液灌胃	是否注入基 因组质粒	动物数	血管计数($\bar{x}\pm s$)	
			A 区	B 区
是(综合组)	是(综合组)	8	4.37±1.85	10.82±3.69
是(中药组)	否(中药组)	8	3.25±0.93	8.24±1.88
否(基因组)	是(基因组)	8	3.84±1.22	10.03±3.54
否(模型组)	否(模型组)	8	3.18±1.07	6.43±2.25

出一个单因素两水平设计(常简称为成组设计),结合重复测量因素(即区域),可构造出“具有一个重复测量的两因素设计”,见表 4。

表 4 是否建模各组分血管计数情况

是否建模	动物数	血管计数($\bar{x}\pm s$)	
		A 区	B 区
是(对照组)	8	3.96±1.12	10.56±2.87
否(模型组)	8	3.18±1.07	6.43±2.25

对表 3 和表 4 资料,若定量资料满足参数检验的前提条件,应选用与实验设计类型对应的方差分析处理为宜。

例 3:《古方稀莪丸对膝骨性关节炎模型家兔关节液中 IL-1 α TNF- β 含量及关节软骨细胞形态学的影响》^[4]表述:取家兔 24 只,按上述造模方法造模,另设一正常对照组 8 只。造模 6 周后开始给药。将芬必得与酒蒸稀莪草的药粉分别以 1:200 和 1:5 的比例添加于饲料配方中自制成饲料,每日早上 1 次给足一天的剂量的含药饲料,待含药饲料吃饭后再添加普通饲料,正常饮水。每晚少给普通饲料。A 组(正常组)、B(模型组)同样方法只给普通饲料,资料见表 5。

表 5 古方稀莪丸对家兔膝骨性关节炎模型关节液中 IL-1 β 、TNF- α 含量的影响($\bar{x}\pm s$)

组别	动物数(n)	IL-1 β 含量(ng/ml)	TNF- α 含量(ng/ml)
正常组	8	3.00±2.79	0.11±0.03
模型组	8	31.17±4.50	38.39±13.46
芬必得组	8	18.51±5.43	19.06±8.38
稀莪丸组	8	24.02±5.81	24.26±7.38

统计学处理:用 SPSS 11.0 统计软件进行数据处理,实验结果均数±标准差表示,采用成组设计定量资料的 *t* 检验对关节液中 IL-1 β 、TNF- α 含量进行统计学分析。结果表明:与正常组比较,模型组关节液中 IL-1 β 和 TNF- α 的含量明显升高,差异有显著性意义($t=15.06$ 及 $t=8.04$, $P<0.01$)。芬必得组与模型组比较差异有显著性意义($t=5.078$ 及 $t=3.45$, $P<0.01$)。稀莪丸组较模型组关节液中 IL-1 β 和 TNF- α 含量降低,有统计学意义($t=2.75$ 及 $t=2.60$, $P<0.01$)。

辨析:显然,组别之下含有两个因素,其一是“是否建模(有两个水平:未建模、建模)”,其二是“在模型基础上用药种类(有三个水平:安慰剂、芬必得、稀莪丸)”。这两个因素的水平组合并不完全,故叫做两因素非平衡组合实验,而不是一个标准的两因素实验设计类型。需要对“组别”进行拆分,如:组合 1:正常组与模型组,为单因素两水平设计;组合 2:模型组、

芬必得组、稀莪丸组,为单因素三水平设计。

若定量资料在符合独立性、正态性、方差齐性检验后应对其分别采用单因素两水平设计定量资料 *t* 检验和单因素 3 水平设计定量资料方差分析。

方差分析可得多组之间的差别是否具有统计学意义,但是对于每两组间差别是否具有统计学意义尚不能作出明确判断,对多组均数进行两两比较方法颇多,常用 SNK(Student-Newman-Keuls)法,即 *q* 检验。

2.2 将具有重复测量的多因素设计误判为多个单因素两水平设计 例 4:《接骨续筋丸影响骨折愈合过程的临床研究》^[5]表述:骨痂 BMD 测定,采用 SD-1000 骨矿测定仪,241 镭射线测定,于骨折后第 30 天及第 60 天,对包含全部骨折的最小面积区域的 BMD 进行测定,同时测定骨折远端区和骨折近端区的 BMD 进行对照。为保证复查时同一患者测定分析区与上次完全一致,复查时使用骨折远端区和骨折近端区普通区域分析对照,资料见表 6,7。统计学分析采用 SPSS 11.5 软件包,数值变量以表示,同组治疗前后比较用配对设计定量资料的 *t* 检验,两组间比较用成组设计定量资料的 *t* 检验。

表 6 治疗前后两组 X 线评分和骨痂 BMD 变化比较($\bar{x}\pm s$)

组别	例数	X 线评分(分)		骨痂 BMD(g/cm ²)	
		骨折后 30 d	骨折后 60 d	骨折后 30 d	骨折后 60 d
治疗组	35	1.732±0.385	2.487±0.443	0.423±0.125	0.684±0.116
对照组	33	1.706±0.390	2.018±0.535	0.419±0.180	0.551±0.103

表 7 治疗前后两组血清 ALP 和 GH 含量变化($\bar{x}\pm s$)

组别	例数	血清 ALP 含量(U/L)		血清 GH 含量(ng/ml)	
		骨折后 30 d	骨折后 60 d	骨折后 30 d	骨折后 60 d
治疗组	35	185.62±18.83	80.43±12.76	3.58±2.89	2.66±2.52
对照组	33	180.69±17.31	115.66±11.28	3.47±2.51	2.34±2.47

辨析:原文资料有四个观测指标,设计类型属于具有一个重复测量的两因素设计,两因素分别为“组别”和“时间”,原作者对两个因素分别使用 *t* 检验进行统计分析,这是不合适的,多次使用 *t* 检验会增加犯假阳性错误的概率。针对该类型资料在符合独立性、正态性、方差齐性检验后应选择具有一个重复测量的两因素设计定量资料方差分析。方差分析归纳起来有 3 个优点:①不受对比的组数之限制;②可同时分析多个因素的作用;③可分析因素间的交互作用^[6]。

例 5:《益气活血法预防老年患者髋部术后下肢深静脉血栓形成》^[7]表述:预防组和对对照组均常规应用抗生素预防感染,必要的西医疗合并症(如高血压病、糖尿病等)。预防组(益气活血组):术前 1~3 d,术后当天至术后 7 d 给予静脉滴注参麦注射液 40 ml 加丹参注射液 20 ml。对照组(抗凝组):术前 12 h 及术后 12 h、1~3 d 每日腹壁皮下注射低分子肝素(LMWH)100 IU/kg,术后 4 d 起每日给药 150 IU/kg,预防用药至术后 7 d。术中、术后常规输血、补液,维持水电解质和酸碱平衡,术后早期开始下肢功能锻炼,资料见表 8,9。

表 8 两组血红蛋白含量($\bar{x}\pm s$, g/L)

分组	例数	血红蛋白含量			
		术前	术后 1 d	术后 4 d	术后 10 d
预防组	52	131±12.8	101±10.2	98±11.6	121±15.2
对照组	50	129±11.1	98±8.8	90±12.3	94±13.8

表 9 两组血小板数量变化($\bar{x}\pm s$, $\times 10^6/L$)

分组	例数	血小板数量			
		术前	术后 1 d	术后 4 d	术后 10 d
预防组	52	190±43	177±65	213±61	207±57
对照组	50	197±52	163±72	191±49	172±73

统计学处理:采用 SPSS 10.0 统计软件,两组间计数资料比较用检验,计量资料比较用 *t* 检验。

辨析:“时间”因素为重复测量因素,该实验设计类型为具有一个重复测量的两因素设计,在符合独立性、正态性、方差齐性前提条件下应选用具有一个重复测量的两因素设计定量资料方差分析,选用 *t* 检验分析该资料不妥。

2.3 将具有重复测量的多因素设计误判为单因素多水平设计 例 6:《肌腱转移固定修复扁平足的生物力学研究》^[8]表述:力学试验,每个标本依上述顺序被重建后,分别将足固定于 CSS-44020 生物力学试验机(长春智能仪器设备研究所,河北骨科研究所提供)上,小腿上部 1/3 处用牙托粉固定与试验机杠杆连接,分别给予 100、300 及 600 N 的力进行加载,测量足底下移距离,资料见表 10。

表 10 修复的扁平足模型的载荷—足底下移距离

扁平足模型	足底下移距离($\bar{x}\pm s$)			
	加载受力(N):	100	200	300
腓骨长肌腱法		3.013±0.172	5.986±0.188	7.378±0.128
胫前肌腱法 1		3.190±0.103	6.787±0.148	9.550±0.243
胫前肌腱法 2		3.220±0.144	6.745±0.130	9.467±0.163
异体跟腱法		3.160±0.152	6.979±0.173	9.609±0.222

本试验所得数据以 Stata7.0 软件在电脑上处理。用单因素 4 水平设计的方差分析进行统计处理,先进行正态性检验,实验数据非正态分布,用秩变换检验法处理。统计检验的显著水准均定在 $\alpha=0.05$ 。

辨析:若每个模型只测定一个部位,则“修复方法”为一个实验分组因素,“加载受力”为一个重复测量因素,因此该实验设计类型应为:具有一个重复测量的两因素设计;若每个模型可分别测定四个部位,则“修复方法”和“加载受力”都是重复测量因素,与表 5 对应的列表格式要发生改变,只能写成 1 行 12 列,在每个部位之后套着 3 个“加载受力”,如:腓骨长肌腱法(100 N、200 N、300 N)、...、异体跟腱法(100 N、200 N、300 N)。这个表格太宽了,此处从略。此时的实验设计类型叫做“具有两个重复测量的两因素设计”,而原文作者误将此实验设计类型判为单因素 3 水平设计,因而误用了统计分析方法。

参考文献

- 胡良平.口腔医学科研设计与统计分析.北京:人民军医出版社,2007.78.
- 苏友新,郑良朴,陈智能,等.强骨宝方对体外培养成骨细胞影响的实验研究.中国骨伤,2007,20(6):394-396.
- 李峻辉,吴亚玲,叶建红,等.活血化瘀中药联合血管内皮生长因子基因转移促进股骨头坏死处新生血管形成的实验研究.中国骨伤,2007,20(6):391-393.
- 郭礼跃,胡慧华,米健国.古方稀莖丸对膝骨性关节炎模型家兔关节液中 IL-1 β 、TNF- α 含量及关节软骨细胞形态学的影响.中国骨伤,2006,19(6):377-378.
- 柳景红.接骨续筋丸影响骨折愈合过程的临床研究.中国骨伤,2007,20(6):373-375.
- 郭祖超.医用数理统计方法.北京:人民卫生出版社,1987.259.
- 刘军,万豫尧,许树柴,等.益气活血法预防老年患者髋部术后下肢深静脉血栓形成.中国骨伤,2006,19(6):344-347.
- 韩长伶,姚双权,张奉琪,等.肌腱转移固定修复扁平足的生物力学研究.中国骨伤,2006,19(6):351-352.

(收稿日期:2008-05-13 本文编辑:李为农)

本刊关于参考文献著录的要求

本刊参考文献按 GB7714-7《文后参考文献著录规则》采用顺序编码著录。参考文献必须以作者亲自阅读过的近年主要文献为限,并由作者对照原文核定。参考文献中的作者,1~3 名全部列出,3 名以上只列前 3 名,后加“等”。外文期刊名称用缩写,中文期刊用全名。每条参考文献均须著录起止页,只占 1 页的文献,给出所在页即可。将参考文献按引用先后顺序,用阿拉伯数字排列于文末。

期刊:作者.文题.刊名,年,卷(期):起页-止页.例:周辉,彭亮,韩勇,等.体外充气复位结合椎体成形术治疗胸腰椎压缩性骨折.中国骨伤,2007,20(3):155-157.

专著:作者.书名.版次(第 1 版不标注).出版地:出版者,出版年.起页-止页.例:刘云鹏,刘沂.骨与关节损伤和疾病的诊断分类及功能评定标准.北京:清华大学出版社,2002.30-31.

专著中析出文献:析出责任者.析出题名.见:原文献责任者.原文献题名.版次.出版地:出版者,出版年.起页-止页.例:孙树椿,张清.手法治疗.见:孙树椿,孙之镐.临床骨伤科学.北京:人民卫生出版社,2006.72-83.

《中国骨伤》杂志社