

• 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(7)

李长平, 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)

统计学思想贯穿于科研工作的各个环节, 具有科学完善的统计学思想可使科研工作达到事半功倍的效果。本文以《中国骨伤》杂志 2006 年与 2007 年第 7 期所刊载的临床研究(共 14 篇)、基础研究(共 6 篇)、经验交流(共 18 篇)等相关文章为基础, 就如何在生物医学科研工作中正确运用统计学, 从实验设计、统计分析和统计描述 3 个方面加以论述, 旨在帮助广大科研人员提高统计学应用的水平。

1 实验设计时需要把握的要点

前 6 期中已经对实验设计的相关内容进行了详细的介绍, 总体来说, 实验设计关系到科研的成败。若对“三要素”、“四原则”、“设计类型”掌握不好, 其科学研究可能前功尽弃。

1.1 随机分组不等于“随意”分组 例 1: 某临床研究^[1]中选取 2003 年 3 月至 2004 年 11 月, 均为本院骨伤科腰腿疼专科门诊确诊的急性 L_{4,5} 节段腰椎间盘突出症患者 98 例。全部患者按入组先后分成治疗组与对照组。

释疑: 在选取受试对象或对其分组时必须采用随机化方法, 如“随机数字表”、查“随机排列表”、用计算机产生“伪随机数字”, 这样做可使样本具有代表性, 并尽可能使组间均衡, 即尽量使已知或未知的重要非试验因素对试验因素各水平组的影响保持一致, 否则得出的结论没有说服力。在该研究设计中, 原作者将先入组的患者分入治疗组, 后入组的患者分入对照组, 是不符合随机原则的。也就是说, 是假随机。因为这样分组很可能掩盖了“病情轻重不等”这一重要非实验因素对结果产生的影响。在具体实施时, 可采用“不平衡指数最小的原则”进行分层随机化^[2]。

1.2 回顾性总结临床经验不等于临床试验研究 例 2: 某临床研究^[3]中选取 1996 年 6 月至 2003 年 10 月四肢长骨良性肿瘤并发骨折患者 25 例, 男 16 例, 女 9 例, 年龄 10~79 岁。根据骨折移位情况分: 轻度 14 例, 中度 8 例, 重度 3 例。进行 I 期外固定, II 期肿瘤刮除植骨术方法治疗。结果发现: 25 例中, 行 I 期外固定、II 期肿瘤刮除植骨术 13 例, 骨折均 I 期愈合, 复发 1 例, 再骨折 1 例。

释疑: 该研究未遵循“对照”原则。欲考察一种疗法的疗效如何, 必须设立对照组, 因为没有比较就没有鉴别。而在该研究中只选用了一种治疗方法(I 期外固定, II 期肿瘤刮除植骨术), 描述 25 例患者采用该方法进行治疗后骨折的愈合情况, 这并不能说明采用上述治疗方法效果的优劣, 因为对于四肢长骨良性肿瘤并发骨折的治疗方法有多种, 而该研究却没有以标准或公认的疗法作为对照。另外, 在该研究中, 选取的受试对象年龄跨度很大, 其病情、病程及预后情况应该相差比较大, 因此, 不能认为他们来自同一总体, 盲目地将他们归入一个实验组不妥。可见该项研究在设计上违反了对照和均衡的原则, 因而结论的可信度较低。故该研究只能属于回顾性临床经验总结, 不能算作临床研究。

1.3 一定要遵循重复原则 例 3: 在某基础研究^[4]中, 研究者欲应用计算机三维仿真模拟技术, 探讨骨盆受到冲击载荷作用的力学行为特征, 为临床分析及判断骨盆骨折类型、力学分布、冲击载荷的影响提供力学基础。为此, 研究者选取 1 名女性志愿者(39 岁)进行了相关实验。

释疑: 生物医学研究的现象常带有变异性, 只有在相同实验条件下进行多次独立重复实验, 随机现象的变化规律性才能正确地显露出来。在该研究中, 只选取了 1 名 39 岁的女性志愿者, 分析出来的结果只是在该名志愿者本身机体上体现的个性, 并不能体现共性, 其相关结论不能随意外推。

2 选择统计分析方法时应把握的两个关键点

在分析定量资料时, 要把握好定量资料统计分析的两个关键点: 第一, 正确判定定量资料所对应的实验设计类型; 第二, 检查定量资料是否满足参数检验的前提条件。在本文中主要针对重复测量设计定量资料的统计分析方法进行探讨。

2.1 不要将某种特定的多因素设计误判为单因素多水平设计 例 4: 某基础研究^[5]表述如下: 实验分为材料试验组(G-PHB 组)、材料对照组(G-PHB 组及 G 组)和空白对照组。经过一系列处理后, 观测不同时间点细胞蛋白合成值的含量。其中 PHB 代表聚羟基丁酸酯, G 代表明胶, G-PHB 代表明胶-羟基

表 1 各组不同时间细胞蛋白合成值(μg/ml, n=8, $\bar{x} \pm s$)

分组	第 1 天	第 2 天	第 4 天	第 8 天
①G-PHB	26.56±7.65*	37.30±6.09*	93.33±14.01*	144.78±13.64*
②PHB	27.13±9.11	39.25±11.87	90.90±12.38	143.40±9.93
③G	27.07±10.32	39.54±3.16	92.75±7.34	142.30±10.76
④空白对照	26.77±5.47	38.90±5.51	93.12±7.80	142.57±11.27

注: 在细胞接种后第 1、2、4、8 天, G-PHB 组与其他对照组相比, *P>0.05

表 2 各组不同时间细胞蛋白合成值

PHB 用否	G 用否	细胞蛋白合成值(μg/ml, n=8, $\bar{x} \pm s$)				
		时间(d):	1	2	4	8
①用	用		26.56±7.65	37.30±6.09	93.33±14.01	144.78±13.64
②	否		27.13±9.11	39.25±11.87	90.90±12.38	143.40±9.93
③否	用		27.07±10.32	39.54±3.16	92.75±7.34	142.30±10.76
④	否		26.77±5.47	38.90±5.51	93.12±7.80	142.57±11.27

丁酸酯。实验结果如表 1 所示。采用单因素方差分析得出:G-PHB 组与各对照组比,差别无显著性意义。

释疑:在表 1 中“分组”实际上包含了 2 个因素:“G 用否”、“PHB 用否”,每个因素都有用与不用 2 个水平。另外表中还包含了“时间”这一因素。显然,“分组”是“G 用否”与“PHB 用否”复合而成的,它应被分解成 2 个因素,即有 2 个实验分组因素,全部受试对象或样品被完全分成 4 个独立的组。

“时间”因素很值得推敲,若每个试验分组中的受试对象(特别是活体动物或人)又被完全随机地分成 4 个小组,在每个时点仅观察 1 个小组中的受试对象,则“时间”因素也是一个实验分组因素。若 3 个实验分组因素对定量观测结果的影响地位平等,则此时可判定该定量资料所对应的实验设计类型为三因素析因设计;若 3 个实验分组因素对定量观测结果的影响地位不平等(能排出主要、次要、最次要的顺序),则此时可判定该定量资料所对应的实验设计类型为三因素嵌套(或叫系统分组)设计。若定量资料满足参数检验的前提条件(独立性、正态性和方差齐性),应选用与设计类型对应的定量资料方差分析。

若每个实验分组中的每个受试对象(特别是样品或细胞)可在 4 个时间点上被重复观测量指标的取值,此时,该实验设计应叫做“具有一个重复测量的三因素设计”。因为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”。因此该定量资料最好采用“具有一个重复测量的三因素设计定量资料的方差分析”处理。

人们习惯于对每个时点采用单因素方差分析,则无法考查“G 用否”、“PHB 用否”与“时间”之间的交互作用。实际工作在撰写论文时务必要注意的是,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,如“采用成组设计定量资料的 *t* 检验”、“采用单因素多水平设计定量资料方差分析”,不应仅说用了“*t* 检验”或“one-way ANOVA”。因为对不同的实验设计定量资料进行方差分析时对总离均差平方和的分解方法是不同的以及计算检验统计量时所采用的误差均方也是不同的,故正确判定实验设计类型是获得正确结论的重要前提^[6]。

另外,表 1 的编制很不规范,横纵标目安排均不合理,这样易给读者和数据处理器造成误解,不利于统计分析方法的正确选用。应依据“统计学三型理论”^[7-8]将资料由表现型转化为“原型”,进而转化为“标准型”,使资料的本质清晰地呈现出来。本例的“原型”与“标准型”是统一的,见表 2。

2.2 不要将某种特定的多因素设计误判为成组设计 例 5:某基础研究^[9]欲研究缺氧对成人成骨细胞增殖及分化的影

响。方法:将手术中取得的成人髂骨骨块,用酶消化法进行培养,当细胞传至第 2 代时,建立成骨细胞缺氧模型,用 MTT 法观察各组成骨细胞骨钙素(OCN)的变化。实验中,取 2 代细胞,分别设正常组和缺氧 24、48、72 h 组作为受试对象,经过一系列处理后接种于 96 孔板。将成骨细胞分别传至 1、2、3 d 后测定 OCN 含量。测定结果如表 3 所示。请问:选用成组设计定量资料的 *t* 检验处理此定量资料,错误的实质是什么?

释疑:处理表 3 中定量资料时,很多人常采用以下做法:

表 3 两组成骨细胞 OCN 含量的比较(ng/ml, $\bar{x} \pm s$)

天	正常组	缺氧组
1	1.51±0.34	3.02±0.13
2	3.31±0.30	5.69±0.45
3	5.17±0.13	9.36±1.26

首先分别在正常组和缺氧组中采用成组设计定量资料的 *t* 检验;然后再分别在治疗第 1、2、3 天 3 个时点上,采用单因素三水平设计定量资料的方差分析。这样做是不妥的!这样分析资料割裂了整体设计,使犯假阳性错误的概率增大,并且无法考察“处理”与“时间”之间的交互作用对结果的影响情况。实际上表 3 中包含 2 个因素,一个因素是缺氧与否(有缺氧与不缺氧 2 个水平);另一个因素是时间(有 1、2、3 d 三个水平),而“缺氧与否”是分别在同一种成骨细胞(将 1 份细胞均分成 2 份)上进行试验的,而每组中的每份细胞又可分成 3 小份,分别传不同天数,故该实验中涉及到两个重复测量因素,所对应

的实验设计类型叫做“具有两个重复测量的两因素设计”,最好选用“具有两个重复测量的两因素设计定量资料的方差分析”处理该定量资料为宜。

另外,表 3 是表达该定量资料的“表现型”,其“原型”见表 4,其“标准型”见表 5。

表 4 两组成骨细胞 OCN 含量的比较

缺氧与否	成骨细胞 OCN 含量(ng/ml, $\bar{x} \pm s$)			
	时间(d):	1	2	3
缺氧		3.02±0.13	5.69±0.45	9.36±1.26
不缺氧		1.51±0.34	3.31±0.30	5.17±0.13

表 5 两组成骨细胞 OCN 含量的比较

成骨细胞 OCN 含量(ng/ml, $\bar{x} \pm s$)					
缺氧(1 d)	2 d	3 d)	不缺氧(1 d)	2 d)	3 d)
3.02±0.13	5.69±0.45	9.36±1.26	1.51±0.34	3.31±0.30	5.17±0.13

3 表达与描述资料时需注意的要点

合理的资料表达与描述可以清晰明了地呈现事物的本

质,便于人们对其有一个全面、准确而又生动的认识。不合理的表达与描述却可能掩盖事物的真相,误导人们去得出不准确甚至错误的结论。因此,统计学思想不仅体现在实验设计和统计分析上,还体现在资料的表达与描述上。

3.1 计算相对数时分母不宜太小 例 6:某种治疗方法疗效评价^[10]描述如下:本组病例优 13 例,良 5 例,中 1 例,以优良率为标准,本组病例总体满意率 94.7%。

释疑:分母过小不能反映事实真相,还易造成错觉。当观察例数<20 时,不宜计算相对数,而给出各组的总数和发生现象的例数即可。本例中,总例数 19 例,不宜计算相对数,仅给出 19 例中优、良、中分别为 13、5、1 例即可,没有必要计算总体满意率。

3.2 不要用复式条图表达适合用线图表达的资料 例 7:在前述例 4 研究中,在材料试验组(G-PHB 组)、材料对照组(G-PHB 组及 G 组)和空白对照组 4 种处理条件下测定了接种后第 1、2、4、8 天所测得的 ALP 水平,绘制成复式条图,见图 1。请问:采用复式条图表达该资料是否恰当^[3]?

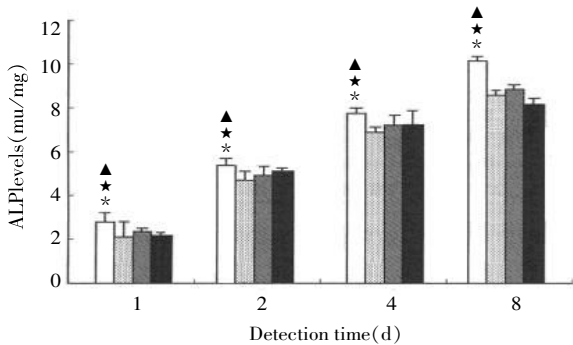


图 1 各组细胞 ALP 水平(n=8)

注:G-PHB 组与 PHB 组比较,*P<0.05;与 G 组比较,*P<0.05;与空白组比较,▲P<0.05

释疑:复式条图适合表达具有 1 个统计指标,2 个定性分组因素的资料。横轴上先安排其中 1 个分组因素,在其不同水平处再安排另一个因素。而本资料中的 1 个分组因素(即 4 种处理方法)是定性的,另一个分组因素(即接种后时间)却是定量的,是一连续型变量,因而不适合用复式条图来表达。

若希望表达一个或多个事物或现象随着时间的推移,若想反映数量的增减幅度,宜用普通线图;若反映数量的变化速度,宜选用半对数线图。横轴通常表示时间,纵轴表示观察指标或其对数值。用复式条图来表达此资料就不太合适了,因为它割裂了各时间点之间的“连续性”。可根据不同的处理在图中绘出 4 条折线。因缺原始数据,修改后的结果从略。

3.3 编制统计表中应注意的问题

3.3.1 “组别”作为横标目易造成混淆 例 8:在例 4 中,统

计表 1 的编制有何不妥之处?

释疑:首先,在表 1 中用“分组”作为横标目的总称,不能反映出问题的实质,容易使人误认为该实验设计类型为“单因素多水平设计”。标目的设置应使用能揭示各组共性特征的一个或多个名词作为标目的总称。在表 1 中,“分组”包含了“G 用否”、“PHB 用否”两个实验因素,不应将其放置在一起,应分别作为一个横标目;其次,在表 1 中缺少总的纵标目用来说明表中数据所代表的含义,总的纵标目应当是“细胞蛋白合成值(ng/ml)”,而不应当是“接种后时间”。“接种后时间”应该是总标目下细分的纵标目。为了使统计表达到“简单明了”的效果,建议将表 1 修改成表 2 的形式为宜。

3.3.2 统计表绘制时,不能“主谓倒置” 例 9:例 5 中,统计表 3 的编制有何不妥之处?

释疑:首先,在表 3 中“主谓倒置”。表格的横标目应该代表某种“原因”,即统计表的主语,在本例应是试验分组(即缺氧组、不缺氧组),正确的做法是将其放在统计表的左边,用以从左向右说明问题;其次应在表的上部加上总的纵标目即成骨细胞 OCN 含量(ng/ml),用来说明表中数据所代表的含义,即统计表的谓语,用以从上向下说明表中数据的真实含义,而与各组对应的重复测量的不同时间应为总纵标目下细分的纵标目,对表 3 进行修改结果见表 4。

参考文献

- 1 黄仕荣,詹红生,石印玉.单穴电针对腰椎间盘突出症患者腰腿痛及外周血血栓素 B2、前列环素的影响.中国骨伤,2006,19(7):398-400.
- 2 胡良平.医学统计应用错误辨析.北京:军事医学科学出版社,2000.12-13.
- 3 李东升,张志勇,黄满玉,等.四肢长骨良性肿瘤并发骨折的手术时机.中国骨伤,2006,19(7):401-402.
- 4 苏佳灿,管华鹏,张春才,等.冲击载荷作用下骨盆三维有限元分析及其生物力学意义.中国骨伤,2007,20(7):455-457.
- 5 易诚青,刘建湘,刘日光,等.骨髓基质细胞与可降解明胶-聚羟基丁酸酯膜复合培养的生物合成功能分析.中国骨伤,2006,19(7):413-416.
- 6 胡良平.如何合理选择统计分析方法处理试验资料(I).中国医药生物技术,2007,2(1):67-70.
- 7 胡良平,刘惠刚.统计学三型理论及其在生物医学科研中的应用.中华医学杂志,2005,85(27):1936-1940.
- 8 胡良平,刘惠刚.统计学思想与三型理论在生物科学研究中的应用.中西医结合学报,2007,5(2):216-219.
- 9 姚琦,张元和,姜华,等.缺氧对成人成骨细胞增殖及分化的影响.中国骨伤,2006,19(7):417-419.
- 10 许超,张万忠,何滨,等.可膨胀髓内钉治疗长骨干骨折 19 例.中国骨伤,2006,19(7):430-431.

(收稿日期:2008-06-10 本文编辑:李为农)