

• 读者 • 作者 • 编者 •

【编者按】 统计学方法在科研工作中的正确应用是决定科研结果的科学性和严谨性的重要因素之一,应引起读者和作者的高度重视。我们在审阅稿件的过程中,发现统计学方法的错误使用时时有发生,导致结果的科学性和严谨性受到质疑;尤其是一些投入了大量的科研经费的基金资助的论文也不例外。为此,邀请了生物医学统计学者对我刊 2006 和 2007 年两年发表的相关论文,采用连载的形式针对统计学方法的应用情况进行辨析,希望对读者和作者在科研设计和工作中能正确运用统计学方法有所帮助。

《中国骨伤》杂志论文中存在的统计学错误辨析(1)

柳伟伟,胡良平

(军事医学科学院生物医学统计学咨询中心,北京 100850)

Zhongguo Gushang/China J Orthop & Trauma,2008,21(2):158-160 www.zggszz.com

本文对《中国骨伤》杂志 2006 年与 2007 年第 1 期所载论著、临床研究和基础研究文章进行了仔细阅读,针对论文中统计学方法的应用情况做了细致的考察,就实验设计、统计描述和统计分析 3 个方面可能存在的问题进行了辨析,提出了改进的建议和意见,希望在提高科研工作的科学性和严谨性方面能够发挥一定的作用。

1 实验设计方面的错误辨析

实验设计对于结果的可靠与否、误差的控制和准确估计、实验中耗费的成本大小都有着至关重要的作用与意义,在设计过程中必须遵循的 4 个原则是随机、对照、重复、均衡,其中任何一者如果在实施中不严格,就有可能导致研究结论存在偏倚甚至错误。对于未开始的实验研究,需要选择合适的实验设计类型安排实验因素及其水平、控制重要非实验因素对结果的干扰和影响;而对于已完成的实验研究,则需要对研究的实验设计类型做出准确的判断,才能正确合理地选择统计分析方法处理资料。本文通过下面的例子说明何为“多因素非平衡组合实验”以及如何对其进行辨析。

例 1 《杜仲腰痛丸对大鼠非压迫性髓核突出神经根损伤的组织形态学研究》^[1]一文,将 50 只大鼠随机分为 5 组,每组 10 只,本文将具体分组情况整理为表 1。

表 1 50 只大鼠实验分组情况

编号	组别	例数
A 组	假手术组	10
B 组	模型组	10
C 组	伸筋丹组	10
D 组	杜仲腰痛丸高剂量组	10
E 组	杜仲腰痛丸低剂量组	10

辨析:本例中假手术组(A 组)未建立模型,而其余 4 组都经手术建立了模型,其中 C、D、E 3 组在建模的基础上分别使用了伸筋丹胶囊、含生药 0.8 和 0.4 g/kg 的杜仲腰痛丸;A、B 组用生理盐水灌胃。

这种类型的资料常常被误认为“单因素 5 水平设计”,进

而选择单因素 5 水平设计定量资料的方差分析处理所收集的定量资料,但实际上本例包含 3 个实验因素,即“是否建立模型”、“在模型基础上,药物种类(伸筋丹、杜仲腰痛丸)”和“杜仲腰痛丸剂量(高、低)”。在已设计的 5 个组中,没有将上述诸因素各水平做全面组合,而只是实施了其中的一部分,所以,本例属于多因素非平衡组合实验,不是一个标准的多因素设计。

对于这种资料,需要结合专业知识,将“组别”拆分成几个不同的组合。拆分的依据是所形成的每一种组合在专业上能解释得通,在统计学上又有相应的标准设计与之相对应^[2]。拆分结果有以下 4 种情况。

- 组合 1:A 组与 B 组;
- 组合 2:B 组、C 组与 D 组;
- 组合 3:B 组、C 组与 E 组;
- 组合 4:B 组、D 组与 E 组。

组合 1 属于成组设计;组合 2 与组合 3 是单因素 3 水平设计,相应的实验因素为“药物种类”,在模型基础上未用药组可理解为用了“安慰剂”;组合 4 也是单因素 3 水平设计,因素为“药物剂量”,具体剂量分别为 0、0.8 和 0.4 g/kg。

对上述 4 种组合,若观测结果为定量资料,应先检查定量资料是否满足参数检验的前提条件,再根据具体情况,选用相应设计定量资料的参数检验法(如 t 检验、方差分析,定量资料满足参数检验的前提条件),或非参数检验法(如秩和检验,定量资料不满足参数检验的前提条件)。

2 统计描述方面的错误辨析

对资料进行统计描述,可以选用统计图、统计表或者统计学指标。对它们的选用要根据资料的性质和分析的目的来进行,要注意应用它们的一些基本原则和要求,要辨别不同指标应用的场合和条件。统计描述中的常见错误主要包括编制统计表时纵、横标目颠倒,表中数据含义表达不清;绘制统计图时图形的选择与资料性质不符,坐标轴设置违背数学原则;应用相对数时将率与比混用,分母很小时也计算相对数;用均数和标准差说明呈严重偏态分布的资料。本文由以下几个实例

表 2 每日恢复额定压力外固定器轴向加压术后 15 d 在体骨断端压力变化($\bar{x}\pm s$)

Times (d)	Group A		Group B		Group C	
	Pressure(kg)	Surplus rate(%)	Pressure(kg)	Surplus rate(%)	Pressure(kg)	Surplus rate(%)
1	1.12±0.18	34	1.11±0.15	53	1.00±0.12	85
5	1.92±0.32	58	1.49±0.20	71	1.09±0.13	93
10	2.80±0.35	85	1.76±0.24	84	1.12±0.14	95
15	3.20±0.43	96	2.02±0.27	96	1.16±0.14	98

来进行说明。

例 2 《外固定器在体轴向加压骨断端压力长期变化的实验研究》^[3]一文,研究目的为定量分析外固定器在体轴向加压时不同加压量、不同时期骨断端压力变化规律,动物被随机分为 A、B、C 3 组,分别施加 1.5、1、0.5 倍动物体重的加压量,给出了术后 15 d 在体骨断端压力变化情况,见表 2。

辨析:一般来讲,表格的横标目应代表原因变量,而纵标目代表结果变量,而本例中轴向加压量为原因变量,压力为结果变量,表 2 的编制颠倒了纵、横标目,修改后统计表见表 3。

表 3 每日恢复额定压力外固定器轴向加压术后 15 d 在体骨断端压力变化($\bar{x}\pm s$)

轴向加 压量	Pressure(kg)			
	Times(d):1	5	10	15
Group A	1.12±0.18	1.92±0.32	2.80±0.35	3.20±0.43
Group B	1.11±0.15	1.49±0.20	1.76±0.24	2.02±0.27
Group C	1.00±0.12	1.09±0.13	1.12±0.14	1.16±0.14

例 3 《经皮穿针治疗髌前上棘骨折 17 例》^[4]一文中,采用经皮钢针挑拨复位固定的方法治疗髌前上棘骨折 17 例,随访观察疗效,其中优 15 例,良 2 例,优良率 100%。

辨析:计算百分率、构成比等相对数时分母不宜过小,如果观察例数小于 20,不宜计算相对数,直接给出各组的总例数和发生某现象的例数即可。本例共观察 17 例,故不宜计算优良率,这样不但不能正确反映事实真相,往往还会造成错觉。

例 4 《经皮鱼嘴钳固定治疗胫骨平台骨折》^[5]一文中,应用撬拨复位、经皮鱼嘴钳固定的方法治疗胫骨平台骨折 83 例,随访其中 81 例,得到膝关节功能评价结果见表 4。

表 4 治疗前后膝关节功能结果($\bar{x}\pm s$,分)

时间	疼痛	行走	伸膝	活动度	稳定性	功能总分
治疗前	0.2±0.4	0.2±0.4	0.4±0.8	1.0±0.8	1.3±1.3	2.7±2.0
治疗后	5.2±1.1	5.0±1.2	5.5±1.1	5.1±1.1	5.1±1.1	26.0±4.7

辨析:本例中治疗前部分指标的标准差是均数的 2 倍,明显呈偏态分布,而“均数±标准差”用来说明呈正态分布的资料,用在这里是不合适的。描述偏态资料的集中趋势应该选用中位数,描述其离散程度应选择四分位数间距。

3 统计分析方面的错误辨析

在对资料进行统计分析时,要根据资料的类型和资料所对应的实验设计类型,同时结合分析的目的选用对应的统计分析方法。对于定量资料而言,如果资料满足参数检验法的前提条件,可以进行参数检验,比如 *t* 检验和方差分析等;若不符合参数统计分析的要求(如不满足正态性或方差齐性),则

需要使用非参数检验。定性资料常以列联表的形式给出,合理选用统计分析方法的关键在于正确地列联表资料分类,列联表资料通常可以分为 2×2 表、2×K 表或 K×2 表、R×C 表和高维列联表,不同的列联表对应的统计分析方法也有所不同,这点在实际应用中要特别加以注意^[6]。常见的某些统计学分析方法的误用由以下实例来说明。

例 5 在前述例 2 所提到的文章中,原文作者使用 *t* 检验对资料进行了统计分析,资料见表 2。

辨析:本例中实验因素为加压量,它有 3 个不同的水平,即 1.5、1、0.5 倍动物体重,同时还存在一个重复测量因素,也就是时间,故该资料属于具有一个重复测量的两因素设计定量资料。这种资料的分析有其特殊性,对同一个受试对象来说,它在不同时间点的测量结果往往存在一定的相关性,这时不能简单地用 *t* 检验来对资料进行统计分析,而应该根据特定问题中数据之间的相关关系选用相应设计定量资料的方差分析模型进行处理。

例 6 《酒精性骨质疏松症造模鼠的骨与生化代谢特征》^[7]一文中,研究目的为观察酒精中毒引发骨质疏松造模鼠的骨与生化代谢变化特征,将 120 只大鼠随机分为 3 组,实验组 A、B 分别给予不同剂量的酒精,对照组以生理盐水灌胃。分别于 8、16、24 周后处死取材。统计分析“行 *F* 检验及 *q* 检验”。全身 BMD 含量变化情况见表 5。

表 5 不同时期全身骨骼 BMD 含量变化($\bar{x}\pm s$,g/cm²)

Groups	Time(weeks)		
	8	16	24
Experiment group A	0.258±0.012	0.198±0.002*▲	0.163±0.014*▲
Experiment group B	0.260±0.022	0.223±0.004*	0.194±0.015*
Control group	0.278±0.025	0.269±0.001	0.257±0.017

注: *与对照组相比 *P*<0.05, *与对照组相比 *P*<0.01, ▲与实验组 B 相比 *P*<0.05

辨析:本例中的一个实验因素为酒精剂量,对照组的酒精剂量可以被认为是 0,与上例不同的是,时间在这里并非重复测量因素,因为原文中“8、16、24 周后处死取材”一句说明不同时间点的测量结果取自不同的动物,而每只动物只在某一个时间点进行测量,这点要特别引起注意。如果酒精剂量和时间这 2 个因素对观测指标的影响是同等重要的,即因素之间在专业上是地位平等的,那么本例为两因素析因设计;如果 2 个实验因素对观测指标的影响有主次之分,则该实验为系统分组设计(也称嵌套设计)。

原文中只提到“行 *F* 检验及 *q* 检验”,却并未对实验设计的类型做出说明,不同的设计类型在统计分析时所选用的方

差分析模型是不同的。此外,在对某一个因素的不同水平进行两两比较时,首先要考虑两个因素之间的交互作用,如果存在明显的交互作用,那么接着对单个因素的各水平进行两两比较是没有意义的。本文没有阐述对交互作用的分析,直接使用 q 检验进行两两比较,这样做显然是不合适的。应该先分析交互作用,如果不存在交互作用的影响,继而就可以进行某一因素不同水平平均值之间的两两比较。

例 7 上述例 4 一文中,统计学处理部分叙述如下:用单因素方差分析法计算治疗前后评分的均数和标准差,统计结果用 $\bar{x} \pm s$ 表示。用 F 检验分析治疗前后各指标差异的显著性意义,显著性水平。资料仍见表 4。

辨析:均数和标准差根据实验数据直接计算就可以了,不需要使用单因素分析法。本例中如果同一观察对象治疗前后的数据能够对应起来,那么应该属于自身配对设计。这种资料如果其差量满足参数检验的条件,应该使用配对设计定量资料的 t 检验;若不能满足此条件,则应该对差量进行变量变换或选用配对设计定量资料的符号秩和检验进行分析。

通过对例 4 的分析可知,表 4 中部分数据明显呈偏态分布,且一般来讲,这种评分形式的资料往往并不服从正态分布,这就很难保证治疗前后的差量满足正态性的要求,在这种情况下,要么对差量进行变量变换使其满足正态性要求,要么直接选用符号秩和检验,而非文中所讲的进行 F 检验。

综上所述,在医学科研工作中,对于统计学的应用确实还有待于进一步的加强和优化,实验设计的不尽完善,统计描述的不够准确,统计分析方法选择的偏颇,都有可能导致实验结论产生偏差。特别需要指出的是,统计学的使用应该贯穿于整个科研工作的始终,而不是仅仅局限于最后的实验数据处理分析阶段。

参考文献

- 1 赵继荣,张思胜. 杜仲腰痛丸对大鼠非压迫性髓核突出神经根损伤的组织形态学研究. 中国骨伤,2007,20(1):34-36.
- 2 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. 7.
- 3 范里,唐新宇,李家元,等. 外固定器在体轴向加压骨断端压力长期变化的实验研究. 中国骨伤,2006,19(1):28-30.
- 4 李卫国. 经皮穿针治疗髌前上棘骨折 17 例. 中国骨伤,2006,19(1):49.
- 5 王战朝,黄霄汉,杨明路. 经皮鱼嘴钳固定治疗胫骨平台骨折. 中国骨伤,2007,20(1):26-28.
- 6 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. 8.
- 7 齐振熙,王明千. 酒精性骨质疏松症造模鼠的骨与生化代谢特征. 中国骨伤,2006,19(1):31-33.

(收稿日期:2007-12-05 本文编辑:李为农)

本刊对来稿中统计学处理的有关要求

本刊针对来稿中经常出现的统计学方法处理不当和误用的情况,对来稿中涉及统计学处理的稿件提出以下有关要求。

1. 统计研究设计:在文稿的资料与方法中应围绕 4 个基本原则(随机、对照、重复、均衡)概要说明稿件的统计研究设计,尤其要交代如何控制重要非试验因素的干扰和影响并说明统计研究设计的名称和主要做法。如调查设计(分为前瞻性、回顾性或横断面调查研究);实验设计(应交代具体的设计类型,如自身配对设计、成组设计、交叉设计、正交设计等);临床试验设计(应交代属于第几期临床试验,采用了何种盲法措施等)。

2. 资料的表达与描述:用 $\bar{x} \pm s$ 表达近似服从正态分布的定量资料,用 $M(QR)$ 表达呈偏态分布的定量资料;用统计表时,要合理安排纵横标目,并将数据的含义表达清楚;用统计图时,所用统计图的类型应与资料性质相匹配,并使数轴上刻度值的标法符合数学原则;用相对数时,分母不宜小于 20,要注意区分百分率与百分比。

3. 统计分析方法的选择:在文稿中要求单独立项,说明统计学处理的方法。对于定量资料,应根据所采用的设计类型、资料所具备的条件和分析目的,选用合适的统计分析方法,不应盲目套用 t 检验和单因素方差分析;对于定性资料,应根据所采用的设计类型、定性变量的性质和频数所具备条件以分析目的,选用合适的统计分析方法,不应盲目套用 χ^2 检验。对于回归分析,应结合专业知识和散布图,选用合适的回归类型,不应盲目套用简单直线回归分析,对具有重复实验数据的回归分析资料,不应简单化处理;对于多因素、多指标资料,要在一元分析的基础上,尽可能运用多元统计分析方法,以便对因素之间的交互作用和多指标之间的内在联系进行全面、合理的解释和评价。

4. 统计结果的解释和表达:①当 $P < 0.05$ (或 $P < 0.01$) 时,应说明对比组之间的差异有统计学意义,而不应说对比组之间具有显著性(或非常显著性)的差别;②应写明所用统计分析方法的具体名称(如:成组设计资料的 t 检验、两因素析因设计资料的方差分析、多个均数之间两两比较的 q 检验等);③统计量的具体值(如 $t=3.45, \chi^2=4.68, F=6.79$ 等)应尽可能给出具体的 P 值(如 $P=0.0238$);④当涉及到总体参数(如总体均数、总体率等)时,在给出显著性检验结果的同时,再给出 95% 置信区间;⑤如得出有相关性的结论,必须提供相关系数 r 。

5. 统计学符号:按 GB3358-82《统计学名词及符号》的有关规定书写,常用如下:①样本的算术平均数用英文小写 \bar{x} (中位数仍用 M);②标准差用英文小写 s ;③标准误用英文小写 S ;④ t 检验用英文小写 t ;⑤ F 检验用英文大写 F ;⑥卡方检验用希腊文小写 χ^2 ;⑦相关系数用英文小写 r ;⑧自由度用希腊文小写 ν ;⑨概率用英文大写 P (P 值前应给出具体检验值,如 t 值 χ^2 值、 q 值等);以上符号均用斜体。

《中国骨伤》杂志社