

# 医学期刊编辑应注意的几个统计学处理问题\*

冉明会<sup>1)</sup> 罗萍<sup>1)</sup> 邓丹<sup>2)</sup>

(1)《重庆医科大学学报》编辑部;2)重庆医科大学公共卫生学院卫生统计教研室,400016,重庆)

**摘要** 为提高医学论文中统计学处理的编校质量,通过案例分析,指出医学编辑应注意论文中的统计分析软件的介绍、数据表达的方式、统计学方法的交代和检验水准的设置等问题。

**关键词** 医学期刊;编辑;统计学处理

**Editors of medical journals should notice some statistical treatment questions**//RAN Minghui, LUO Ping, DENG Dan

**Abstract** To enhance medical papers' editorial quality about statistical treatment, editors of medical journals should notice the introduction of statistical analysis software, data expression ways, statistical methods, setting levels of significance, and so on.

**Key words** medical journal; editor; statistical treatment

**First-author's address** Editorial Department of Journal of Chongqing Medical University, 400016, Chongqing, China

医学统计学在医学科研设计整个过程中发挥着非常重要的作用,因此,可将其作为评价医学论文质量优劣的重要依据;然而,医学期刊论文中普遍存在统计学误用和滥用问题<sup>[1]</sup>,几乎在统计学应用的各个环节上都有不妥甚至错误之处<sup>[2]</sup>。由于医学统计学具有内容涉及面宽、应用灵活、不易掌握等特点,作为一名非统计学专业的医学期刊编辑,是很难发现稿件中存在的统计学错误的。为提高医学论文中统计学处理的编校质量,本文通过案例分析,提出医学期刊编辑需要注意的几个统计学处理问题,与同人共同探讨。

医学论文所涉及的统计学问题通常会在全文第1部分“材料与方法”中以“统计学方法”作为一个比较独立的部分予以专门介绍。这一部分应该包含哪些内容,应该如何表达,目前没有统一标准,同时没有引起编辑的重视。笔者认为,它应该包含统计分析软件的介绍、数据表达方式的介绍、统计学方法的交代和检验水准的设置等至少4方面的内容。

## 1 统计分析软件介绍

科研论文中统计学处理部分通常位于第1部分“材料与方法”中的最后部分,以“统计学处理”或“统计学分析”形式将全文所涉及统计学方法作一介绍。首先应对文章使用的统计分析软件作比较详细的介绍,即注明软件名称、版本(严格来讲还应注明序列

号)。如使用 SAS 9.0 统计软件包进行统计分析,常见错误有:1)不介绍所用统计软件;2)只介绍统计软件,不介绍所用版本。

**例1 题名** 新型无支架猪主动脉瓣的体外加速疲劳试验<sup>[3]</sup>

文中“统计学处理:数据以 $\bar{x} \pm s$ 表示,采用 SNK- $q$  检验比较各组之间的差异”,未作统计分析软件及版本介绍。

**例2 题名** 细胞外调节蛋白激酶在雌激素促乳腺癌细胞 MCF-7 增殖中的作用<sup>[4]</sup>

文中“统计学处理:数据用 $\bar{x} \pm s$ 表示,采用多样本均数比较的方差分析、多重比较的  $q$  检验对数据进行分析”,对统计软件使用根本未提及。

## 2 数据表达方式的介绍

正确判别统计资料的性质是合理选择统计分析方法的重要前提。根据统计指标的性质,统计资料一般分为定量资料和定性资料2大类。所谓定量资料,指从每个观察单位(针对此资料,其观察单位是病例标本)上测得的指标是用具体的数值表示的。一般带度量衡单位,通常称为计量资料,如身高、体重、血压等。所谓定性资料,指从每个观察单位上测得的指标仅反映某一方面的性质,并不能用具体的数值表示。一般没有度量衡单位,通常称为计数资料,如血型、性别等。还有一种资料叫做等级资料,也叫有序资料,指将观察单位按某种属性的不同程度或次序分成等级后分组计数的观察结果。由于它是一种介于定量资料与定性资料之间的资料,有的教材将其归入定性资料一类,有的将其单列一类;但不管如何分类,它可根据研究者目的的不同处理成不同的资料形式<sup>[5]</sup>。

符合正态分布的计量资料表达应该采用均数 $\pm$ 标准差( $\bar{x} \pm s$ ),偏态分布的计量资料表达通常采用中位数( $M$ ) $\pm$ 四分位间距( $Q$ )。计数资料的表达常用的有比、构成比、率。这部分常见的错误是:1)数据表达根本未提及;2)偏态计量资料误用均数 $\pm$ 标准差( $\bar{x} \pm s$ )表示。

**例3 题名** 致病性大肠杆菌引发宿主细胞 A/E 损伤分析方法的建立<sup>[6]</sup>

文中“统计学处理:采用 SPSS 13.0 统计软件,对组间进行单因素方差分析,以 LSD 法对不同组别之间

\* 中国高校自然科学学报研究会基金项目(GBJXC0828);重庆市高校学报研究会科研项目(XB200902)

的均数进行比较”未提及数据表达,但从文中数据来看,采用的是均数 $\pm$ 标准差( $\bar{x}\pm s$ )的形式来表达计量资料。

**例4** 题名 脉冲Nd:YAG激光对口腔溃疡组织中NO及NOS含量的影响<sup>[7]</sup>

文中“统计学分析:数据以 $\bar{x}\pm s$ 表示,采用SPSS 10.0统计软件进行单因素方差分析”指明是用 $\bar{x}\pm s$ 来表示数据,据此我们可以判断文中数据是符合正态分布的计量资料;但文中表1的数据显示NOS数据并不符合正态分布,因此不应用它来表达。

“平均数 $\pm$ 标准差”是用来表达呈正态分布的资料的,说明其数据分布的集中趋势和离散趋势。表1中带“\*”数据的标准差大于平均数或几乎与平均数相等,基本上可以认为此资料服从偏态分布,不适合用正态分布法说明此资料的集中趋势和离散趋势。

描述偏态分布资料的集中趋势应该选用中位数,描述其离散程度应该选用第1( $Q_1$ )和第3( $Q_3$ )四分位数间距,数据表达成“ $M(Q_1\sim Q_3)$ ”的形式。由于原作者用 $\bar{x}\pm s$ 表示出来,掩盖了数据呈偏态分布的真相,进而选用方差分析处理资料,使之一错再错。可以看出,作者数据表达方式的错误,进而导致统计分析方法的错误。

### 3 统计分析方法的交代

对文中所用统计学方法的交代应非常清楚,不应含糊其词。不仅应该交代清楚采用的是什么样的统计分析方法,而且要注意实验设计类型,还要指明文中的哪些数据采用该统计分析方法。这样才会让读者一目了然,同时能够对作者采用的方法是否正确作出准确判断。

**例5** 题名 热休克小鼠肝癌细胞( $H_{22}$ )源Exosomes的制备及其蛋白组成的初步研究<sup>[8]</sup>

文中“统计学处理:SAS 8.0软件进行统计学分析,采用 $t$ 检验”。从作者的介绍中我们只能得知文章处理定量资料用了 $t$ 检验,至于这些统计分析方法选用得是否正确则不得而知。通常情况下,比较各平均值之间的差别是否具有统计学意义,可能会用到的 $t$ 检验有3种。它们之间的区别体现在定量资料所对应的“实验设计类型”上。讲 $t$ 检验时,应注明是“单组设计定量资料的 $t$ 检验”“配对设计定量资料的 $t$ 检验”,还是“成组设计定量资料的 $t$ 检验”。显然,如果只是说明统计分析用 $t$ 检验或定量资料统计分析用 $t$ 检验是很不够的,一定要注明实验设计类型。

**例6** 题名 细胞外调节蛋白激酶在雌激素促乳腺癌细胞MCF-7增殖中的作用<sup>[4]</sup>

文中“统计学处理:数据用 $\bar{x}\pm s$ 表示,采用多样本均数比较的方差分析、多重比较的 $q$ 检验对数据进行分析”。同样,我们只知道作者使用了方差分析;但方差分析有“单因素多水平设计定量资料的方差分析”“随机区组设计定量资料的方差分析”“拉丁方设计定量资料的方差分析”“交叉设计定量资料的方差分析”“ $x$ 因素析因设计定量资料的方差分析”“具有 $x$ 个重复测量的 $x$ 因素设计定量资料的方差分析”等10种之多,作者使用的何种设计类型我们不得而知,当然更无从判断其统计方法使用是否正确。

### 4 检验水准 $\alpha$ 的设定

在统计学处理介绍中还应包含检验水准 $\alpha$ 的大小,它是预先规定的判断小概率事件的概率尺度,表示在研究设计时根据不同研究目的预先规定的允许犯I类错误概率的最大值<sup>[5]105</sup>。根据不同的研究目的, $\alpha$ 设定的大小不同,通常设定是 $\alpha=0.05$ 或 $0.01$ 。这部分常见的错误是:1)对检验水准不作介绍;2)将 $P$ 值与检验水准等同;3)以 $P$ 值的大小与统计差别大小相联系。

**例7** 题名 乙型肝炎病毒X基因对c-met基因启动子区的调控作用<sup>[9]</sup>

文中“统计学分析:数据以均数 $\pm$ 标准差( $\bar{x}\pm s$ )表示,用SPSS 10.0统计软件对数据进行统计学处理,荧光素酶活性值采用ANOVA单因素方差分析,侵袭细胞数统计采用 $t$ 检验”对检验水准没有提及,从文中 $P$ 值均小于 $0.01$ ,猜测其检验水准是 $0.01$ 。

**例8** 题名 姜黄素对大鼠血管平滑肌细胞增殖和凋亡的影响<sup>[10]</sup>

文中“统计学处理:计量资料采用 $\bar{x}\pm s$ ,统计软件采用SPSS 13.0分析,多个样本均数比较采用方差分析,两两比较用Dunnet- $t$ 检验,显著性标准采用 $P<0.05$ ”。

$P$ 值在统计学上指的是在假设检验中由 $H_0$ 所规定的总体作同样的重复试验,获得等于及大于(或等于及小于)当前检验统计量的概率<sup>[5]105</sup>。虽然 $\alpha$ 与 $P$ 均表示概率,但两者是不能等同的。前者是事先设定的标准,后者是根据该标准作的假设检验后得到的概率值,并与之相比较得出其差别是否具有统计学意义。上述例子中将检验水准等同于 $P$ 值显然是错误的。

**例9** 题名 NF- $\kappa$ B和HMGB1在大鼠肠缺血再灌注肺损伤中的表达及白藜芦醇的保护作用<sup>[11]</sup>

文中“统计学处理:数据均以均数 $\pm$ 标准差( $\bar{x}\pm s$ )表示,采用SPSS 12.0统计软件进行两组间均数比较, $t$ 检验及方差分析, $P<0.05$ 为差异有显著意义, $P<0.01$ 为差异有非常显著意义”。如前所述, $P$

# 期刊封一上的条码上方无需附注 ISSN

郭 莉

(西北化工研究院《工业催化》杂志社,710600,西安)

GB/T 16827—1997《中国标准刊号(ISSN 部分)条码》规定了中国标准刊号(ISSN 部分)条码的代码结构、条码符号技术要求和印刷位置。笔者对具有中国标准刊号(ISSN 部分)的685种科技期刊进行统计分析,发现有653种期刊条码上方附注国际标准连续出版物号(ISSN),占95.3%,仅有32种期刊条码上方没有ISSN,占4.7%。

GB/T 16827—1997规定,期刊条码代码由主代码和附加码共15位数字组成。主代码由前缀码、数据码、年份码和校验码4部分13位数字组成:最前面的3位前缀码“977”是期刊专用前缀代码;随后7位是取自ISSN中除去校验位以外的流水号(品种号);再往后2位是条码的年份码,标志期刊的出版年份,以公历

年份的最后2位数字表示(00~99);最末1位(第13位)是校验码,其值是通过前12位数字经一定的公式计算得出的,在用微机软件制作条码时计算机会根据公式自动生成。附加部分(附加码)由2位数字组成,表示该刊当期期次号或月次号。

如果在条码上方附注ISSN,那么,在一幅条码中就会出现2次ISSN流水号,加之封面右上角已经排印了ISSN,造成ISSN重复标注,而且增加了条码的高度,字体字号混乱,影响期刊的美观度,不符合GB/T 16827—1997的要求。若条码排印在封一,则建议将条码上方的ISSN剪切掉。

(2009-06-12 收稿;2009-08-22 修回)

值在统计学上指的是在假设检验中由 $H_0$ 所规定的总体作同样的重复试验,获得等于及大于(或等于及小于)当前检验统计量的概率。也就是说, $P$ 值的大小只表示总体之间的差别是由抽样误差引起的可能性的大小, $P$ 值越小,这种可能性也就越小,那么总体之间有差别的可能性就越大,但并不表示总体之间的差别就越大。“ $P \leq 0.05$ 为差异有显著意义, $P < 0.01$ 为差异有非常显著意义”就是认为 $P$ 值越小,总体之间的差别就越大,很明显这是错误的。

那么这部分究竟应该如何正确表示呢?通常是直接说明检验水准 $\alpha = 0.05$ (或者 $\alpha = 0.01$ )。还有一种表示方法是以 $P < 0.05(0.01)$ 表示差异有统计学意义。

## 5 结束语

医学论文中统计学处理应该如何表达,应该包含哪些内容,是医学期刊编辑不可回避的问题。尽管没有统一的国家标准将其规范,但笔者认为应该包含软件介绍、数据处理、统计方法详细介绍、检验水准这4个方面的内容,这样才能让编辑、审稿专家和读者对文章的统计学处理一目了然,详细了解,从而便于判断。

## 6 参考文献

[1] 胡良平.再谈《中华口腔医学杂志》已发表论文的统计学质量问题[J].中华口腔医学杂志,2003,38(4):309-312

[2] 胡良平,李子建.医学统计学基础与典型错误辨析[M].

北京:军事医学科学出版社,2003:4

[3] 贺强,杨剑,魏旭峰,等.新型无支架猪主动脉瓣的体外加速疲劳试验[J].第四军医大学学报,2009,30(8):694-696

[4] 李苏华,邓华瑜,陈黎.细胞外调节蛋白激酶在雌激素促乳腺癌细胞MCF-7增殖中的作用[J].第二军医大学学报,2009,30(4):395-399

[5] 颜虹,徐勇勇,赵耐青.医学统计学[M].北京:人民卫生出版社,2008:3

[6] 王海光,顾江,余抒,等.致病性大肠杆菌引发宿主细胞A/E损伤分析方法的建立[J].第三军医大学学报,2009,31(8):659-662

[7] 杨苏平,刘鲁川,刘娜,等.脉冲Nd:YAG激光对口腔溃疡组织中NO及NOS含量的影响[J].第三军医大学学报,2009,31(8):703-705

[8] 杨麟,沈宜,李静,等.热休克小鼠肝癌细胞(H22)源Exosomes的制备及其蛋白组成的初步研究[J].重庆医科大学学报,2008,33(6):669-672

[9] 谢斌,唐春,陈平,等.乙型肝炎病毒X基因对c-met基因启动子区的调控作用[J].中华肝脏病杂志,2009,17(4):292-296

[10] 舒砚文,王天红,王祥.姜黄素对大鼠血管平滑肌细胞增殖和凋亡的影响[J].第四军医大学学报,2009,30(8):697-700

[11] 杨廷芳,王兴勇,胡语航,等.NF- $\kappa$ B和HMGB1在大鼠肠缺血再灌注肺损伤中的表达及白藜芦醇苷的保护作用[J].重庆医科大学学报,2009,34(1):24-28

(2009-06-01 收稿;2009-08-06 修回)