

# 医学论文中 2 组等级疗效比较问题分析与建议

刘霞<sup>1)</sup> 周晴<sup>2)</sup> 周英智<sup>1)†</sup>

1) 山东大学自然科学学报编辑部, 250061; 2) 山东大学基础医学院, 250012; 济南

**摘要** 列举医学论文中 2 组等级疗效比较时常见的问题: 将数据合并成一个四格表分析, 按多个四格表分析, 采用  $t$  检验, 秩和检验结果错误, 未做统计学处理。分析出现这些问题的原因并给出了建议。

**关键词** 医学论文; 等级资料; 疗效比较; 秩和检验

**Problems and suggestions on the comparison of ranked curative effect between two groups in medical papers** // LIU Xia, ZHOU Qing, ZHOU Yingzhi

**Abstract** The following problems in comparison of ranked curative effect between two groups are analyzed according to samples from medical papers: analyzing the data as a  $2 \times 2$  table, analyzing the data as several  $2 \times 2$  tables, using  $t$  test, reporting an error result of rank sum test, and not performing statistical analysis. The causes

are analyzed and some suggestions given on how to avoid these problems.

**Keywords** medical paper; ranked data; curative effect comparison; rank sum test

**First-author's address** Editorial Department of Journal of Shandong University ( Science and Technology ), 250061, Jinan, China

**DOI**: 10.16811/j.cnki.1001-4314.2017.02.007

医学研究中经常比较 2 种治疗方法的效果, 当结果为等级有序资料时, 如将疗效分为治愈、显效、好转、无效等, 属于有序效应标志的指标。当有序水平数大于 2 时, 如果采用  $\chi^2$  检验, 只能说明它们的内部构成

## 5 参考文献

[1] 文双全. 提高高校学报编辑加工质量的思考[J]. 安徽工业大学学报(社会科学版), 2015, 32(2): 149

[2] 刘洪华, 陈晓锋, 李朝前, 等. 做好科技期刊的编辑加工[J]. 长江大学学报(社会科学版), 2014, 37(8): 96

[3] 游俊, 赵燕, 胡小洋. 科技论文中数据问题的编辑审读技巧[J]. 编辑学报, 2012, 24(6): 536

[4] 师福东. 科技论文编辑加工探微[J]. 编辑学报, 2010, 22(增刊 2): 47

[5] 李艳. 学术论文编辑加工中的引文分析[J]. 沈阳大学学报(社会科学版), 2015, 17(3): 425

[6] 蒋鸿标. 学术论文的学术价值和创新性: 兼谈学术期刊编辑的责任[J]. 中国科技期刊研究, 2005, 16(6): 915

[7] 富群华. 医学论文实验设计科学性的编辑审读[J]. 中国科技期刊研究, 2013, 24(5): 997

[8] 陈浩元. 科技书刊标准化 18 讲[M]. 北京: 北京师范大学出版社, 1998

[9] 降小宁. 学术期刊编辑审稿的思维定势[J]. 内蒙古电大学刊, 2011(5): 46

[10] 张小兵. 论编辑甄选文稿和作者撰写文稿在思维认识上的区别[J]. 长沙铁道学院学报(社会科学版), 2012, 13(4): 261

[11] 郑一奇. 编辑的悟性[M]. 北京: 首都师范大学出版社, 2009: 101

[12] 李振荣. 论编辑评价稿件的四个维度[J]. 出版科学, 2012, 20(3): 45

[13] 邵益文. 编辑学当前需要研究的问题[J]. 中国科技期刊研究, 1998, 9(1): 3

[14] 陈著, 沈剑虹. 加强科技期刊精神层面内涵的思考[J]. 中国科技期刊研究, 2003, 14(5): 470

[15] 刘岭. 试论编辑加工和校对中的思维差异及转换[J]. 编辑之友, 2013(10): 87

[16] 胡浩志, 耿卓, 邓菁. 专家审稿制度下学术编辑的职业定位与发展[J]. 出版科学, 2016, 24(1): 45

[17] 周作新. 论期刊文稿的编辑学术把关[J]. 编辑学报, 2003, 15(3): 178

[18] 赵立芹, 邢宝山, 王健, 等. 科技期刊编辑流程工作中的审读及其意义[G] // 学报编辑论丛: 2006. 上海: 上海大学出版社, 2006: 90

[19] 刘德敏, 姚萍, 徐永红, 等. 编辑审查学术论文内容错误的切入点及方法[J]. 科教导刊, 2015(1): 44

[20] 任辉. 论学术期刊的编辑审稿[J]. 十堰职业技术学院学报, 2010, 23(2): 97

[21] 徐超富. 论科技编辑审读的三原则[J]. 编辑学报, 2004, 16(3): 157

[22] 虞沪生, 张瑞清, 阎为民. 科技论文创新性的审读[J]. 编辑学报, 2006, 18(5): 333

[23] 王克黎. 浅议科技论文写作的逻辑构成[J]. 北京建筑工程学院学报, 2006, 22(1): 73

[24] 秦社彩. 运用逻辑学编辑高质量的学术论文[J]. 中国科技期刊研究, 2003, 14(5): 475

[25] 刘泉. 刘泉顾问谈出版物标准化[J]. 编辑学报, 2001, 13(1): 12

(2016-10-10 收稿; 2016-11-11 修回)

† 通信作者

比是否相同,而不能评价它们的效应是否相同。这种资料宜用2个独立样本的秩和检验评估反应指标的效应差异有无统计学意义,具体方法为Wilcoxon秩和检验或Mann-Whitney  $U$  检验<sup>[1-2]</sup>。这2种方法的 $Z$ 检验统计量和 $P$ 值完全一致<sup>[3]</sup>,撰写论文时,一般只报告 $Z$ 值、 $P$ 值即可<sup>[4]</sup>。医学论文中,对这类资料的统计学处理问题较多。编辑同人对医学稿件中的统计学问题进行了许多研究<sup>[5-6]</sup>;但就等级资料分析方法和结果表达问题的研究较少,因此,我们结合案例对2组等级疗效比较时出现的问题做一系统总结。所用案例均来自2016年CNKI收录的医学论文。

## 1 常见问题分析

### 1.1 将数据合并成一个四格表分析

例1<sup>[7]</sup> 将92例支原体肺炎患儿分为研究组和对照组,2组疗效比较见表1。按照检验水准 $\alpha = 0.05$ ,认为2组疗效差异有统计学意义。

表1 2组患儿治疗总有效率比较

组别	例数	痊愈	显效	好转	无效	总有效率/%
研究组	46	26	17	2	1	93.48
对照组	46	21	15	8	2	78.26

注:  $\chi^2 = 4.3894, P = 0.0361$ 。

上面 $\chi^2$ 和 $P$ 值是将痊愈与显效、好转与无效合计,作为一个四格表进行统计分析得出来的。将等级资料合并成四格表,没有充分利用资料信息,如本例痊愈与显效、好转与无效权重相同。此外,把好转未计入有效是否合适?正确的做法是采用秩和检验。

我们采用SPSS 22.0软件(后面案例测算亦采用该软件)计算的统计学处理结果为 $Z = -1.510, P = 0.131$ ,2组疗效差异无统计学意义,与文中报告的结果相反。

为了进一步验证秩和检验结果的可信性,将2组的痊愈、显效、好转与无效数据行 $\chi^2$ 检验,发现2组数据不符合 $R \times C$ 联列表 $\chi^2$ 检验的条件;因此使用Fisher确切概率法计算,得到 $P = 0.193$ ,显示它们的内部构成比差异无统计学意义,表明秩和检验的结果是可信的。

### 1.2 按多个四格表分析

例2<sup>[8]</sup> 将100例结核性脑膜炎患者分为观察组和对照组,采用 $\chi^2$ 检验比较2组疗效,结果见表2。

表2 2组疗效比较

组别	例数	痊愈	好转	无效	总有效
观察组	50	22	23	5	45
对照组	50	12	20	18	32
$P$	-	<0.05	<0.05	<0.05	<0.05

作者做了4次 $\chi^2$ 检验,分别比较2组痊愈、好转、无效及总有效(痊愈、好转合计)情况。作者没有给出 $\chi^2$ 值,需补充。 $P$ 值只是笼统地给出<0.05,需要改为具体数值。需要说明的是,2组“好转”比较结果为 $\chi^2 = 0.367, P = 0.545$ ,并不是<0.05。

该研究的目的是比较2组的疗效,就4次比较而言,“痊愈”“无效”及“总有效”差异有统计学意义,而“好转”差异无统计学意义,到底2组有没有差异?结果很难解释,而且多次比较割裂了设计的整体性。秩和检验结果为 $Z = -3.010, P = 0.003$ ,可以认为2组疗效差异有统计学意义。

### 1.3 采用 $t$ 检验

例3<sup>[9]</sup> 将80例恶性肠梗阻及术后早期炎性肠梗阻患者分为观察组与对照组,2组疗效结果见表3。

表3 2组患者临床疗效比较

组别	例数	治愈	好转	无效	总有效
观察组	40	16	21	3	37
对照组	40	9	16	15	25
$t$	-	-7.89	8.97	-5.34	11.31
$P$	-	<0.000	<0.000	<0.000	<0.000

$t$ 检验用于定量资料的比较,而上表中的资料为等级资料,不符合 $t$ 检验的基本条件。此外,统计软件计算结果有时会出现 $P = 0.000$ ,这是由于 $P$ 值太小,统计软件自动四舍五入,在论文中可写为 $P < 0.001$ <sup>[10-11]</sup>。实际上,该论文“统计学处理”中明确说明“选择秩和检验对等级资料进行分析”,不仅误用 $t$ 检验,而且没有注意统计学处理与结果的一致性。秩和检验结果为 $Z = -2.839, P = 0.005$ ,可以认为2组疗效差异有统计学意义。

### 1.4 秩和检验结果错误

例4<sup>[12]</sup> 为观察美宝创疡贴联合高压氧治疗Ⅲ期压疮的疗效,将60例患者分为观察组和对照组:观察组痊愈16例,显效10例,好转4例;对照组痊愈10例,显效9例,好转7例,无效3例(1例因不能耐受高压氧治疗而退出研究)。采用秩和检验比较2组疗效,认为观察组疗效明显优于对照组,差异有统计学意义( $Z = -1.976, P < 0.05$ )。

我们测算的统计学结果为 $Z = -1.945, P = 0.052$ 。由此判断2组疗效差异无统计学意义,与作者报告的结果相反。需要强调的是,论文中一定要给出具体的统计量和 $P$ 值,该文给出了具体的 $Z$ 值,但未给出具体的 $P$ 值,需补充。

### 1.5 未做统计学处理

例5<sup>[13]</sup> 将40例糖尿病足且下肢血管重度狭窄或闭塞患者分为联合治疗组和单独移植组,2组疗效

比较结果如下:

治疗后60 d,临床治愈、显著有效、有效和无效患者例数,联合治疗组分别为9、8、3和0例,单独移植组分别为2、8、10和0例。联合治疗组临床治愈率(45%)高于单独移植组(10%),显著有效率(40%)2组相同,联合治疗组有效率(15%)低于单独移植组(50%)。

根据文中内容,与单独移植组相比,联合治疗组临床治愈率较高,有效率较低,显著有效率相同,这些数据很难说明2组疗效有无差异。秩和检验结果为 $Z = -2.824, P = 0.005$ ,可以认为2组疗效差异有统计学意义。

## 2 原因分析

从上面的例子可以看出,在比较2组等级疗效时,医学论文中出现多种错误。其原因如下:

1)对统计学重要性认识不足,认为统计学是可有可无的内容,有的作者甚至不做统计学处理,仅凭个人主观判断下结论。如例2中作者4次采用 $\chi^2$ 检验比较,所用方法并不合适。结果中没有报告 $\chi^2$ 值,4个 $P$ 值均 $<0.05$ ,有的 $P$ 值与实际测算结果不符,作者是否做过统计学处理值得怀疑。

2)对资料性质分辨不清,如有的作者将等级资料作为定量资料采用 $t$ 检验,有的作者把等级资料当作无序资料按 $\chi^2$ 检验处理。

3)尽管医学期刊对统计学越来越重视,但目前存在的问题仍然很多,许多作者模仿已发表论文的统计学处理模式,有可能出现新的统计学错误。

## 3 建议

2组等级疗效比较时,若统计学处理不当,会导致论文报道的结果与实际情况不符,甚至得出相反的结论,很可能对临床治疗产生误导,其后果的严重性必须引起重视。为避免这类资料统计学处理的错误,提出以下建议:

1)期刊社或编辑部制订统计学基本要求,严格要求作者按照正确的统计学方法处理资料,并在论文中正确描述统计结果<sup>[14]</sup>。有的期刊要求投稿时注明统计学处理人员的单位、姓名,可以强化作者对统计学的重视,避免或减少统计学错误<sup>[15]</sup>。

2)聘请统计学专家审稿,特别是对涉及复杂统计学内容的稿件务必请统计学专家把关。越来越多的期刊聘请统计学专家担任编委或审稿专家,有的期刊还在网站公开统计学专家的信息,这种做法值得效仿。

3)医学论文的统计学处理大多使用SPSS软件,编辑应掌握该软件的常用方法。论文中一般会给出等级资料的频数分布,可根据文中提供的数据验证统计学结果,给作者提出意见或建议<sup>[16]</sup>。

## 4 参考文献

- [1] 颜虹. 医学统计学[M]. 2版. 北京:人民卫生出版社, 2014:173
- [2] 肖丽娟,孙茂民. 医学论文中统计学处理常见问题及应对措施[J]. 编辑学报,2010,22(6):500
- [3] 陈平雁,黄浙明. IBM SPSS 19 统计软件应用教程[M]. 2版. 北京:人民卫生出版社,2012:297
- [4] 孟泳,刘晓光. 中药离子导入配合药物治疗成人支气管肺炎疗效观察[J]. 郑州大学学报(医学版),2014,49(3):417
- [5] 陈文娟,汤雷,马莉. 医学期刊常见 $t$ 检验应用错误及案例分析[J]. 编辑学报,2016,28(3):237
- [6] 尚永刚. 编辑应重视医学期刊中数据的审查与检验[J]. 编辑学报,2016,28(1):26
- [7] 廖洪彬,文启芹,明怀志,等. 阿奇霉素联合红霉素对支原体肺炎患儿临床疗效研究[J]. 辽宁医学院学报,2016,37(1):29
- [8] 徐齐峰,宋晓东,孟婵. 糖皮质激素辅助治疗成人结核性脑膜炎的疗效[J]. 包头医学院学报,2016,32(8):29
- [9] 谢锋伟,陈娟,招婷. 生长抑素在治疗恶性肠梗阻及术后早期炎性肠梗阻中的应用效果观察[J]. 辽宁医学院学报,2016,37(1):44
- [10] 吴艳妮,周春兰,江霞,等. 国内护理学统计源期刊论文中报告精确 $P$ 值常见错误: $P = 0.000$ [J]. 编辑学报,2016,28(2):133
- [11] 冯国双,罗凤基. 医学案例统计分析与SAS应用[M]. 北京:北京大学医学出版社,2011:55
- [12] 丘明志,贲吕金,陈坚. 美宝创疡贴联合高压氧治疗Ⅲ期压疮的疗效观察[J]. 中国烧伤创疡杂志,2016,28(3):205
- [13] 韩雪立,刘东,张秀娟,等. 自体骨髓干细胞移植联合经腔内血管成形术治疗糖尿病足的效果评价[J]. 吉林大学学报(医学版),2016,42(1):125
- [14] 姜春霞. 论医学期刊编辑的统计学审核[J]. 中国科技期刊研究,2014,25(6):782
- [15] Baishideng Publishing Group Inc. Guidelines for manuscript preparation, submission, and manuscript format: Basic study[EB/OL]. [2016-10-09]. <http://www.wjnet.com/bpg/GerInfo/218>
- [16] 周英智,靳光华. 利用文中数据识别统计学错误[J]. 编辑学报,2016,28(1):29

(2016-10-27 收稿;2017-01-19 修回)