

医学论文中等级资料统计方法误用的案例辨析*

蒲素清 罗云梅 李纓来†

四川大学华西医院《中国普外基础与临床杂志》编辑部,610041,成都

摘要 医学论文内涉及大量的等级资料,等级资料统计方法的误用情况较多。作者列举公开发表的文章以及在编辑文章过程中所涉及等级资料的统计方法误用的多个案例,辨析了等级资料作为组别和作为效应量时的统计方法误用类型,并提出了正确的使用方法,以供医学编辑同人参考。

关键词 医学论文;等级资料;案例辨析

Misuse of statistical method for ordinal data in medical papers and case discrimination//PU Suqing, LUO Yunmei, LI Yinglai

Abstract Ordinal data is common in medical papers, and there is some misuse of statistical method for ordinal data in medical papers. This paper elaborates on some cases about the misuse of statistical method for ordinal data and put forward the correct usage method in order to make reference to the editors of medical journal.

Keywords medical paper; ordinal data; case discrimination

Authors' address Editorial Department of Chinese Journal of Bases and Clinics in General Surgery, West China Hospital, Sichuan University, 610041, Chengdu, China

DOI:10.16811/j.cnki.1001-4314.2019.05.010

医学论文中涉及大量的等级资料。等级资料即有序分类资料,具有半定量的特征。医学论文中的 TNM 分期、分化程度、心功能分级、美国麻醉医师协会分级(ASA 分级)、疗效(如治愈、有效、无效)等即是常见的等级资料;然而,等级资料作为“效应量”或组别时,采用的统计方法是不同的,因此,对于等级资料的统计方法,绝不仅仅是简单的秩和检验就可完成的。目前,对于等级资料,有学者^[1-4]进行了相关的误用辨析,但针对的基本都是单一成组设计类型的等级资料,并不全面。笔者列举公开发表或笔者编辑的涉及等级资料统计方法误用的多个案例,辨析等级资料作为组别和作为效应量时的统计方法差异,以供医学编辑同人参考。

1 单向有序资料

单向有序资料,即行和列项目中仅有 1 项为等级资料,此类型数据包括 2 种情况:仅效应量为等级资料和仅组别为等级资料。

1.1 仅效应量为等级资料 这是医学期刊中最常涉及的等级资料类型,如比较 2 组患者接受某种治疗后的疗效、2 种组织标本中某种蛋白表达程度的差异、2 组患

者 ASA 分级程度的差异等。对于此类资料,正如其他学者^[1-4]所言,最常见的方法误用即是采用成组 χ^2 检验进行分析。采用成组 χ^2 检验处理此种等级资料的缺陷在于:它只能得出组间的“分布”不同,而不能得出组间“程度”不同,而医学研究中,研究者往往想知道的是诸如何种方法更有效等关于“程度”的问题。但笔者在 CNKI 数据库中查阅了 2019 年发表的多篇文献,发现采用成组 χ^2 检验分析等级资料的情况其实已比较少见,此类资料的分析方法总体已经呈现出新的特点:文章中展示各等级数据,但在分析时进行了合并,使等级资料退化为二分类资料,再采用成组 χ^2 检验进行统计分析。

如《米曲菌酶联合伊托必利辅助治疗慢性胃炎伴反流性食管炎的疗效观察》^[5]一文中,研究者探讨了中西医结合治疗重症肺炎合并呼吸功能衰竭的疗效,得到结果见表 1。根据表 1,该文作者的结论是“治疗组总有效率显著优于对照组($P < 0.05$)”。仔细观察表 1 数据,可发现存在 2 个问题:1)作者将显效和好转定义为有效,然而数据的合并无疑会损失信息;2)比较 2 组的总有效率,统计量也应是 χ^2 值而非 Z 值(即使采用的是秩和检验,结果也不正确)。笔者采用秩和检验对该数据进行分析,得出 $Z = -1.234, P = 0.217$,与原文结论有所不同。由此可见,对于数据的合并,可能得出完全相反的结论。

表 1 治疗组和对照组患者的临床疗效对比($n = 30$)

组别	治疗效果(例)			总有效率/例(%)	Z 值	P 值
	显效	好转	无效			
治疗组	13	15	2	28(93.33)	-2.007	0.045
对照组	10	14	6	24(80.00)		

1.2 仅组别为等级资料 有时研究者的研究标的不是等级资料,而是不同等级的某种效应量的差异,如《Mdm 2 在 ER α 阳性乳腺癌组织中的表达及其 siRNA 对 MCF-7 细胞生物学行为的影响》^[6]中,研究者比较了不同 TNM 分期 ER α (雌激素受体 α)阳性乳腺癌组织中 Mdm 2 表达阳性率的差异,此时 TNM 分期包括 I、II、III、IV 期,是等级资料,类似于组别,而 Mdm 2 的表达状态是“效应量”(表 2)。作者最开始来稿时采用 $R \times C$ 列联表资料的 χ^2 检验进行组间总体比较,再进行组间两两比较;但多次的组间两两比较无疑会增

* 四川省软科学研究计划项目(2018ZR0257)

† 通信作者

大犯 I 型错误的概率,且也增加了工作量,而趋势 χ^2 检验(线性趋势检验)可完美解决这个问题,因此笔者建议作者斟酌统计分析方法,最终作者在修回稿件中采纳了笔者的建议,得出 $P = 0.027$,可以认为 Mdm 2 的表达阳性率随 TNM 分期的增高而增加。

表 2 ER α 阳性乳腺癌组织中 Mdm 2 表达与患者临床病理特征的关系(节选)

TNM 分期	n	Mdm 2 表达阳性	Mdm 2 表达阴性
		例(%)	
I	22	7(31.8)	15(68.2)
II	25	12(48.0)	13(52.0)
III	26	15(57.7)	11(42.3)
IV	5	4(80.0)	1(20.0)

2 双向有序资料

双向有序资料,即行和列项目均为等级资料,此类数据包括 2 种情况:等级资料属性不同和属性相同。

2.1 属性不同 在《PARP-1 和 caspase-3 在胰腺癌及癌旁组织中的表达及意义》^[7]一文中,研究者检测了 66 例胰腺癌组织中 PARP-1 [聚(腺苷酸二磷酸核糖)聚合酶 1] 和 caspase-3(半胱氨酸天冬氨酸蛋白酶 3) 的表达,得到表 3。该研究者在整理表格时未纳入 PARP-1 表达阴性的情况,且采用的统计分析方法也简单描述为 χ^2 检验。该处理欠妥,一是损失了阴性表达信息(表 3 中笔者已进行了补充),二是统计方法不明:是配对 χ^2 检验还是成组 χ^2 检验? 该例中,分化程度和 PARP-1 表达都是等级资料,分析方法应依据研究目的而定:若研究目的为分析不同分化程度组患者 PARP-1 表达程度的差异,可采用 Kruskal-Wallis H 检验,因 Kruskal-Wallis H 检验可以明确表达强度高低的问題;若研究目的为分析 2 个有序分类变量间是否存在相关关系,宜采用等级相关分析或 Pearson 积差相关分析;若研究目的为分析 2 个有序变量间是否存在线性变化趋势,宜用有序分组资料的线性趋势检验(test for linear trend)^[8]。无论研究者的目的如何, χ^2 检验都是错误的。笔者认为,该研究以有序分组资料的线性趋势检验为最佳选择,能得到随分化程度的增高,PARP-1 表达是否呈现线性变化趋势的结论(增强/减弱),这无疑更符合研究实际,且得到的信息量更大(相关,且关系是线性的)。

表 3 胰腺癌分化程度和 PARP-1 表达的关系(节选)

分化程度	n	+++	++	+	-
高分化	20	3	6	5	6
中分化	32	18	6	1	7
低分化	14	10	3	1	0

2.2 属性相同 当配对资料 2 种属性相同,且均超过 2 类时,就是配对资料的扩展了。医学论文中的常见形式为:2 种检测方法对同一样品的测定结果分级、同一组织中 2 种蛋白的表达(表达程度分类相同)等。在《X 线、CT 与 MRI 检查在早期强直性脊柱炎骶髂关节病变诊断中的价值探究》^[9]一文中,研究者的目的是比较 CT 检查和 X 线检查对于早期强直性脊柱炎骶髂关节病变的诊断价值,也就是说比较的是 2 种检查方式诊断结果的差异,叙述成统计语言就是比较 2 种检查诊断结果的概率分布是否相同。作者采用的方法是,在不同分级下进行成组 χ^2 检验,有 6 个分级就进行了 6 次成组 χ^2 检验(表 4)。此部分存在 2 个问题:一是并没有体现出来是配对设计,表格的呈现方式有误;二是统计方法选择有误,将本质是配对设计的资料,人为拆分成多个分类下的成组设计资料。该研究的正确统计方法应选用配对 $R \times C$ 列联表资料的 χ^2 检验。由于笔者无该文章的准确数据,无法列出该研究的正确表格,因此呈现了另一篇相似文章的正确表格(表 5^[10])。

表 4 X 线和 CT 检查对 110 例强直性脊柱炎的检查结果(节选)

检查方式	0 级	I 级	II 级	III 级	IV 级
	例(%)				
X 线	49(44.55)	11(10.00)	14(12.73)	25(22.73)	11(10.00)
CT	27(24.55)	22(20.00)	26(23.64)	23(20.91)	12(10.91)
χ^2	9.730	4.314	4.440	0.107	0.049
P	0.001	0.037	0.035	0.743	0.824

表 5 X 线与 CT 诊断骶髂关节病变级别的分级情况^[10]

X 线分级	CT 扫描分级					合计
	0	I	II	III	IV	
0	17	6	7	0	0	30
I	8	30	28	16	0	82
II	0	6	32	22	0	60
III	0	0	8	66	4	78
IV	0	0	0	2	20	22
合计	25	42	75	103	24	272

此外,针对等级资料属性相同的资料,笔者再延伸一点。有时候研究的目的并不是分析 2 种检查方法的概率分布差异,而是观察 2 种检查方法结果的一致性(特别是有一种方法是金标准时),此时应采用一致性检验即 Kappa 检验。Kappa 检验是一致性检验,配对 $R \times C$ 列联表资料的 χ^2 检验是判断概率分布,属于不同的方法,达到不同的目的;因此对于同一资料,有时 2 种方法的检验结论不一致,如一种方法的 P 值小于检验水准,而另一种方法得出的 P 值大于检验水准。

综上所述,单向有序资料中等级资料作为效应量、

作为组别时,采用的统计分析方法不同;双向有序资料中,等级资料属性相同和不同时的统计分析方法也有所不同。在编修涉及等级资料的稿件时,编辑应根据作者的研究目的,核查统计方法是否正确。

3 参考文献

- [1] 韩宏志,官鑫,李欣欣,等. 医学论文中 χ^2 检验的误用及案例分析[J]. 编辑学报, 2018, 30(3): 257
- [2] 姚应水. 医学科研论文中 χ^2 检验误用分析[J]. 皖南医学院学报, 2005, 24(1): 80
- [3] 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): 501
- [4] 鲁立,傅万明,李风华,等. 医学论文统计学问题的编辑审查[J]. 编辑学报, 2006, 18(5): 338
- [5] 张长青,张葵玲,王育斌,等. 米曲菌酶联合伊托必利辅助治疗慢性胃炎伴反流性食管炎的疗效观察[J]. 中

国现代应用药学, 2019, 36(12): 1558

- [6] 时云,王耕,周坤,等. Mdm 2 在 ER α 阳性乳腺癌组织中的表达及其 siRNA 对 MCF-7 细胞生物学行为的影响[J]. 中国普外基础与临床杂志, 2016, 23(9): 1070
- [7] 蒲竞,李汛,朱骏,等. PARP-1 和 caspase-3 在胰腺癌及癌旁组织中的表达及意义[J]. 临床肝胆病杂志, 2016, 32(2): 337
- [8] 孙振球,徐勇勇. 医学统计学[M]. 2版. 北京:人民卫生出版社, 2008: 153
- [9] 吴泽健. X线、CT与MRI检查在早期强直性脊柱炎骶髂关节病变诊断中的价值探究[J]. 医学理论与实践, 2019, 32(10): 1573
- [10] 马丽,艾脉兴,赵孟君,等. X线和CT对强直性脊柱炎骶髂关节病变的诊断价值比较[J]. 中日友好医院学报, 2007, 21(6): 337

(2019-07-02 收稿;2019-08-28 修回)

参考文献著录中应纠正的几个常见错误

GB/T 7714—2015《信息与文献 参考文献著录规则》发布已近4年,实施效果差强人意,但仍有不少书刊对一些要求性条款未严格执行。现择要分列如下,提请纠正。

1)用西文或汉语拼音书写的著者的姓的字母未采用全大写。标准规定:“个人著者,其姓全部著录,字母全大写”;“用汉语拼书写的姓名,姓全大写”。这一规定与ISO 690:2010《信息和文献 参考文献和信息资源引用指南》的要求一致。姓的字母仅首字母大写、其余小写是不规范的。

2)其他题名信息前的标识符未使用“:”。标准规定:“: 用于其他题名信息前”;“其他题名信息包括副题名,说明题名文字,多卷书的分卷书名、卷次、册次,专利号,报告号,标准号等”。常见错误是副题名前使用了“——”、对标准未采用“标准名称:标准号”格式著录等。

3)专著中析出文献的出处项前未使用标识符“//”。标准规定:“// 用于专著中析出文献的出处项前”;从专著中析出有独立著者、独立篇名的文献,“其析出文献与源文献的关系用‘//’表示”。这里的“专著”包括会议录(C)和汇编(G)。采用标识符“//”替代“见:”(汉语)、“In:”(英语)、“Julkaisussa:”(芬兰语)等,具有简明性、通用性的优点,也是我国标准的一个亮点。目前很多书刊不使用“//”的错误应予改正。

4)引文文献的“页码”未著录引用信息所在页。标准将参考文献分为引文文献和阅读型文献2大类,并规定“引文页码”“析出文献的页码”是必备的著录项目。对期刊还专列条款规定:“凡是从期刊中析出的文章,应在刊名之后注明其年、卷、期、页码。阅读型参考文献的

页码著录文章的起讫页或起始页,引文参考文献的页码著录引用信息所在页。”标准还规定:多次引用同一文献,采用顺序编码制的“在序号的‘[]’外著录引文页码”,采用著者-出版年制的“在‘()’外以角标的形式著录引文页码”。因此,凡是作为引文文献引用的专著和汇编、会议录、期刊中析出的文章,不著录引用信息所在页是错误的。目前这类著录错误普遍地存在,包括很多编辑同人撰写的文章,应该认真加以纠正。《编辑学报》此项的著录规范是:引文文献著录引用信息所在页,阅读型文献著录文章起始页。

5)不设卷的期刊其“年”与“(期)”间错加了“,”。标准规定:“() 用于期刊年卷期标识中的期号”上,其前不加其他标识符号。例如:2019年第15卷第6期,应著录为“2019, 15(6)”;2019年第6期,应著录为“2019(6)”。将后者著录为“2019,(6)”是违反规则的。

6)将刊名组成部分的“××××版”误作为其他题名信息著录。GB/T 7714—2015纠正了GB/T 7714—2005的失误,指出期刊的“题名按著录信息源所载的内容著录”,并列出示例“北京师范大学学报(自然科学版)”。“(自然科学版)”是法定刊名的组成部分,将其视为说明题名文字或副题名则是错误的。

我国的出版物,无论是科技还是社科论著,无论是中文还是英文书刊,参考文献著录都应执行与国际标准基本接轨的GB/T 7714—2015。有人以遵循“国际惯例”为遁词不执行国家标准,是不利于提高我国参考文献著录标准化水平的,其实国际上各国、各期刊的参考文献著录格式是很不统一的,不值得我们盲目迷信。(诸仁)