

利用文中数据识别统计学错误

周英智 靳光华

山东大学自然科学学报编辑部, 250061, 济南

摘要 利用医学论文中的数据, 举例说明下列统计学问题: 绝对数或相对数计算错误, P 值计算错误, 误用 t 检验处理定性资料, 用 $\bar{x} \pm s$ 描述偏态分布资料。分析出现这些问题的原因, 并就如何避免给出建议。

关键词 医学论文; 统计学错误; 数据

Identification of statistical errors using data in the papers //

ZHOU Yingzhi, JIN Guanghua

Abstract This paper lists and analyzes some statistical errors in medical papers using the data in the text: errors in calculating absolute and relative measures, error in calculating P -value, analysis of qualitative data using t test, and description of skewness distribution using $\bar{x} \pm s$. The reasons causing these errors are analyzed, and some suggestions are given on how to avoid these errors.

Keywords medical paper; statistical error; data

Authors' address Editorial Department of Journal of Shandong University (Science and Technology), 250061, Jinan, China

DOI:10.16811/j.cnki.1001-4314.2016.01.010

统计学处理是医学论文的一项主要内容, 而其中的错误非常普遍, 严重影响了论文的学术及编辑质量^[1-3]。由于不少论文只提供统计学方法及结果, 不交代原始数据及具体计算过程, 许多问题难以发现; 但如果认真分析论文内容, 仍然能根据文中提供的数据识别一些错误。我们阅读了近年发表的多篇医学论文, 对发现的相关问题做一总结, 以供编辑同人借鉴。

1 根据文中数据发现的统计学错误

1.1 绝对数或相对数计算错误 例 1^[4] 为探讨高频超声对小器官疾病的临床检出效果, 对经病理检查确诊为相关小器官疾病的患者进行高频超声检查。论文“结果”部分如下:

86 例小器官疾病患者中: 甲状腺疾病 34 例, 超声检出率 94.1%; 乳腺疾病 21 例, 超声检出率 85.7%; 眼科疾病 33 例, 超声检出率 96.9%; 男性外生殖器疾病 19 例, 超声检出率 100.0%。具体情况见表 1。

上述内容存在下列问题:

1) 文字介绍患者总计 86 例, 而 4 类疾病合计 107 例, 表中“例数”一列合计 108 例, 3 个数字不一致。

2) 文字部分甲状腺疾病 34 例, 而表中 4 种甲状腺疾病合计为 35 例。

3) 根据表中数据, 超声检出甲状腺疾病 33 例。

表 1 86 例小器官疾病患者的超声检出率

疾病	例数	检出例数	检出率/%	
甲状腺疾病	甲亢	6	5	83.3
	结节性甲状腺肿瘤	11	11	100.0
	亚急性甲状腺炎	15	15	100.0
	甲状腺肿瘤	3	2	66.7
乳腺疾病	乳腺增生	9	9	100.0
	乳腺炎	6	4	66.7
	乳腺纤维腺瘤	4	4	100.0
	乳腺癌	2	1	50.0
眼科疾病	白内障	15	15	100.0
	玻璃体混浊	6	6	100.0
	视网膜脱落	11	10	90.9
	眼内肿瘤	1	1	100.0
男性外生殖器疾病	睾丸扭转	8	8	100.0
	结核	2	2	100.0
	附睾炎	4	4	100.0
	囊肿	4	4	100.0
	睾丸肿瘤	1	1	100.0

根据表中受检总数 35 例计算, 检出率为 94.3%; 根据文字部分受检总数 34 例计算, 检出率为 97.1%。而文字部分甲状腺疾病检出率为 94.1%, 与上面的 2 个检出率无一相符。

4) 根据表中数据, 眼科疾病超声检出率为 97.0%, 而文字部分为 96.9%。

5) 检查人数 10 例以下者, 检出率不主张用百分数, 最好用绝对数直接表示。

6) 应补充各种检出率的 95% 可信区间。

7) 文字部分只交代了受检例数及检出率, 建议补充检出例数。

此外, 统计表编排及具体疾病的分类、排序等方面也值得商榷, 此处不再赘述。

例 2^[5] 为观察老年综合评估 (CGA) 在老年冠心病患者住院期间的应用效果, 将 143 例老年冠心病患者分为 CGA 组 (69 例) 与对照组 (74 例), 分别给予 CGA 护理和常规护理, 对 2 组的心脏不良事件、疾病应对方式、治疗、护理满意度、住院时间及住院费用等进行比较。其中枢绞痛或心前区不适发作情况见表 2。

2 组日间发作均为 2 例, 与括号内的百分比 40.6、32.4 不符合。全天发作合计与日间、夜间发作例数之和吻合, 但百分比吻合。根据样本例数与日间发作的百分比, 或者根据夜间发作与全天发作合计的数值,

推测对照组和CGA组发作例数分别为28、24。对2组日间发作情况行 χ^2 检验,与 $\chi^2 = 1.024, P = 0.312$ 一致,进一步验证了该推测的正确性。

表2 2组患者心绞痛或胸前区不适发作比较

组别	例数	日间发作	夜间发作	全天发作合计
		例(%)		
对照组	69	2(40.6)	32(46.4)	60(87.0)
CGA组	74	2(32.4)	15(20.3)	39(52.7)
χ^2	-	1.024	11.029	19.667
P -	-	0.312	0.001	0.000

此外,统计软件计算结果有时会出现 $P = 0.0000$,这是由于 P 值太小,统计软件自动四舍五入,在论文中可写作 $P < 0.0001$;因此,表2“全天发作合计”一列的 P 值应改为 < 0.0001 ^[6]。

1.2 P 值计算错误 例3^[7]为探讨新辅助化疗在进展期食管癌治疗中的意义,某研究分析了272例进展期食管癌患者的资料,其中112例行新辅助化疗+手术治疗,160例行单纯手术治疗,比较围手术期并发症发生率、手术切除率及术后生存期。其中新辅助化疗组术后并发症发生率为34.8%(39/112),单纯手术组为29.4%(47/160)。采用 χ^2 检验行组间比较,得出 $P = 0.50$ 。

为了便于分析,将上述资料整理成表3。

表3 2组患者术后并发症比较

组别	有并发症	无并发症	合计
新辅助化疗组	39	73	112
单纯手术组	47	113	160
合计	86	186	272

对2组资料行 χ^2 检验,结果为 $\chi^2 = 0.904, P = 0.342$,并不是0.50。那么,0.50的 P 值是如何得到的呢?请看表4。

表4 2组患者术后并发症比较

组别	有并发症	无并发症	合计
新辅助化疗组	39	112	151
单纯手术组	47	160	207
合计	86	272	358

表5 各组血清中MASP、HBeAg、HBsAg、AST、HBV-DNA水平比较($\bar{x} \pm s$)

组别	n	MASP/(ng/mL)	HBeAg/(S/CO)	HBsAg/(S/CO)	AST/(U/L)	HBV-DNA/(拷贝/mL)
A组	20	74.45 ± 14.91	10.60 ± 16.75	200.54 ± 101.47	35.21 ± 20.26	1.12 × 10 ⁶ ± 2.36 × 10 ⁶
B组	20	111.87 ± 21.79	0.22 ± 0.21	121.26 ± 106.03	78.95 ± 47.99	1.25 × 10 ⁶ ± 3.84 × 10 ⁶
⋮	⋮	⋮	⋮	⋮	⋮	⋮
F组	20	80.54 ± 12.02	0.39 ± 0.91	117.76 ± 102.88	212.97 ± 23.87	4.34 × 10 ⁶ ± 9.46 × 10 ⁶
G组	20	134.02 ± 26.56	0.39 ± 0.13	0.88 ± 0.29	21.84 ± 7.36	-

表中数据以 $\bar{x} \pm s$ 表示,但 $\bar{x} \pm s$ 用于描述正态分布资料。许多指标的 \bar{x} 与 s 相近,甚至 s 远大于 \bar{x} ,初步可以判断为偏态分布。应首先进行定量测定指标数

对2组资料行 χ^2 检验,结果为 $\chi^2 = 0.466, P = 0.495 \approx 0.50$ 。可见,作者在行 χ^2 检验时,误将各组的总例数当作了无并发症的例数。此外,作者仅给出了 P 值,未给出 χ^2 值,应予补充。

1.3 误用 t 检验处理定性资料 例4^[8]为探讨子宫动脉化疗栓塞术联合刮宫术治疗剖宫产瘢痕妊娠(CSP)的可行性和安全性,将82例CSP患者分为2组:A组39例,行氨甲蝶呤+刮宫术;B组43例,行子宫动脉化疗栓塞术+刮宫术。比较2组刮宫术成功率、术中出血量、 β -HCG下降至正常的时间、住院时间及术后6个月性激素水平和月经周期恢复时间等。其中刮宫术成功率A组为89.74%(35/39),B组为100%(43/43),差异有统计学意义($t = 2.093, P < 0.05$)。

刮宫术成功与否的资料属于定性资料,而 t 检验适用于分析定量资料,上述资料应该用 χ^2 检验分析。由于总频数 > 40 ,且有2个期望频数 $1 \leq E < 5$,故选用连续性校正 χ^2 检验^[9]。其结果为, $\chi^2 = 2.689, P = 0.101$ 。2组结果差异无统计学意义,与该文作者所得结论相反。

1.4 用 $\bar{x} \pm s$ 描述偏态分布资料 例5^[10]为研究甘露糖结合凝集素相关的丝氨酸蛋白酶(MASP)在病毒不同感染程度的乙肝患者中的表达及其临床意义,根据血清中丙氨酸转氨酶(ALT)和HBeAg的水平,将140例乙肝患者分为6组:A组(ALT正常,HBeAg阳性)、B组(ALT正常,HBeAg阴性)、C组(ALT轻度/中度升高,HBeAg阳性)、D组(ALT高度/重度升高,HBeAg阳性)、E组(ALT轻度/中度升高,HBeAg阴性)和F组(ALT高度/重度升高,HBeAg阴性)。另设20例体检人员作为健康对照组(即G组)。比较各组MASP、HBeAg、HBsAg、AST、HBV-DNA水平。“统计学处理”中的部分内容摘录如下:

本研究的统计学处理采用SPSS 16.0软件,以 $\bar{x} \pm s$ 来表示所得的数据,采用单因素方差分析来比较计量数据,用LSD法对差异有统计学意义的组间进行进一步两两比较,检验水准依据 $\alpha = 0.05$ 。

部分结果见表5。

据的正态分布检验,然后确定合适的描述分析指标。若为偏态分布,应改为中位数及四分位间距描述^[11]。

从统计学角度分析,还存在以下问题:

1)“以 $\bar{x} \pm s$ 来表示所得的数据”不够确切,因为文中既涉及定量资料,也涉及定性资料。即使资料符合正态分布,也只能说以 $\bar{x} \pm s$ 表示定量资料。

2)方差分析的应用条件之一是资料符合正态分布,如果资料为偏态分布,则不宜用方差分析。即使资料符合方差分析的条件;但由于本研究对象包括乙肝患者与非乙肝患者,其中乙肝患者包括6个组,属于2因素(ALT和HBeAg)3×2水平的析因设计,所以应首先采用双向析因设计的方差分析,分析2个因素与乙肝的关联Y度,同时可以分析二者有无交互效应,而不能采用单因素方差分析。

3)比较乙肝患者与正常对照组有关指标的差异,如果资料符合单因素方差分析的条件,则可采用单因素方差分析,比较7个独立组间总的差异有无统计学意义。如果有统计学意义,最佳的两两比较方法是Bonferroni方法,该方法对两两比较的检验水准进行了校正,以避免增加犯I类假设检验的累积错误;若总体差异有统计学意义,采用LSD-*t*检验进行组间两两比较则需要对检验水准进行校正。

4)结论应描述HBeAg阴性、阳性与ALT的关系,以便指导临床工作。

此外,“统计学处理”部分的描述还可精简,此处不再赘述。

2 原因分析

出现这些错误,有作者、审稿专家、编辑等各方面的原因,下面仅从编辑的角度予以分析。

1)没有系统地掌握统计学的基础知识,例如不清楚正态分布、偏态分布、计量资料、计数资料等的基本概念,不了解研究设计的类型,不注意单因素方差分析、*t*检验等常用方法的应用条件。

2)不能运用常用统计学软件,无法依据文中数据验证结果。

3)编辑工作不够仔细,如不注意测算各部分数字与合计值、相对数是否相符。

4)不注意论文相关内容之间的内在联系。如编辑稿件过程中,有一处数据发生了变化,与其相关的和、构成比、率等都要发生变化。如果没有整体观念,就会出现数据不一致的差错。

3 建议

几乎每一篇医学论文都离不开统计学,如果统计学方面存在问题,轻则结果不准确,重则得出与实际情况相反的结论。作为编辑,一定要重视统计学的正确运用,努力减少、避免发生统计学错误。根据本研究发

现的问题及我们的体会,提出以下建议:

1)系统学习医学研究设计的基本类型及统计学基础知识。统计学内容复杂,而且在不断发展,新理论、新方法不断涌现。只有努力学习,才能发现论文中的问题,给作者提供有针对性的修改意见。

2)掌握统计学软件的使用方法,养成利用文中数据验证结果的习惯。例如,大多数论文给出了用于 χ^2 检验的原始数据,只要能运用SPSS等软件,就能发现其中的错误。

3)充分发挥统计学专家的作用。对于一些涉及复杂统计学的论文,除了请相关专业的专家评审外,还应请统计学专家评审,对其科研设计、统计学分析及结果描述等做出评价。

4)树立“宏观编辑”的概念^[12]。论文题名、中英文摘要及正文的各个部分,都含有统计学内容^[13],编辑过程中一定要注意其中的联系,修改论文时避免顾此失彼,应保持相关内容的一致性。

4 参考文献

- [1] 李子建,沈东舜.医学期刊论文中常见统计学错误辨析[J].实用医药杂志,2013,30(12):1129
- [2] 章新生,刘璐,姚仁斌,等.医学稿件中统计学分析使用调查[J].中国科技期刊研究,2011,22(3):393
- [3] 潘明志,唐毓金,梁明佩,等.医学科技期刊统计学方法应用概况[J].传播与版权,2014(2):38
- [4] 晏诚.高频超声检查小器官疾病的临床检出率分析[J].大家健康,2014,8(1):85
- [5] 韩静,屈海宏.老年综合评估护理在老年冠心病患者住院康复中的应用效果[J].护士进修杂志,2014,29(21):1960
- [6] 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
- [7] 杨冉,韩金利,张卫民,等.新辅助化疗在进展期食管癌中的疗效分析[J].中国肿瘤临床,2014,41(2):119
- [8] 王泽鑫,李建明,贾广志.子宫动脉化疗栓塞术联合刮宫术治疗剖宫产瘢痕妊娠的临床疗效[J].中国介入影像与治疗学,2014,11(1):15
- [9] 颜虹.医学统计学[M].2版.北京:人民卫生出版社,2014:149
- [10] 蒋静,王拥军,薛黎,等.慢性HBV感染者血清中人MASP水平检测的临床意义[J].国际检验医学杂志,2014,35(1):9
- [11] 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,2003:260
- [12] 周英智,孙瑶,李伟.科技论文统计表的宏观编辑[J].科技与出版,2008(9):38
- [13] 胡良平,鲍晓蕾.呼吸系统科研设计与统计分析[M].北京:军事医学科学出版社,2013:249

(2015-07-05 收稿;2015-09-01 修回)