

医学论文统计学的“度”问题分析及建议*

吴俊玲¹⁾ 周晴²⁾ 周英智^{3)†}

1)《山东中医药大学学报》编辑部,250355;2)山东大学基础医学院,250012;3)山东大学自然科学学报编辑部,250061;济南

摘要 结合典型案例,总结医学论文中统计学方法的选择、“统计学处理”与统计学结果描述中存在的与统计学的“度”有关的问题。分析出现这些问题的原因,并提出了改进措施。

关键词 医学论文;统计学;度;问题分析

Analysis and suggestions on the “degree” problem of statistics in medical papers//WU Junling,ZHOU Qing,ZHOU Yingzhi

Abstract The paper summarizes the problems in selection of statistical methods, and description of statistical processing and statistical results in medical papers combined with typical cases. The reasons for these problems are analyzed and the improvement measures are put forward.

Keywords medical paper; statistics; degree; problem analysis

First-author's address Editorial Department of Journal of Shandong University of Traditional Chinese Medicine, 250355, Jinan, China

DOI:10.16811/j.cnki.1001-4314.2017.04.012

国际医学期刊编辑委员会^[1]和各期刊的稿约中对医学论文统计学有详细的要求;但如何在论文中将科研过程中的统计学内容客观、准确地表达出来,仍存在许多问题,其中涉及统计学的“度”。编辑同人对医学论文中的统计学问题进行了大量研究,但尚未见专门针对统计学“度”的报道。笔者结合已经发表的论文,对如何把握医学论文统计学的“度”做一分析。由于医学论文的各部分几乎都含有统计学内容^[2],难以对各部分均做详细的阐述;因此,本文仅针对与论文统计学的“度”有关的重要方面,即统计学方法的选择、统计学处理与统计学结果的描述做一分析。

1 统计学方法选择中的问题

医学科研中,要根据研究目的、设计类型、资料性质等选择合适的统计学方法。选择统计学方法时,如果把握不好“度”,所得结果就会有偏差,甚至得出与实际情况相反的结果。

1.1 所用统计学方法不够全面

例1^[3] 为研究辛伐他汀联合辅酶 Q10 对慢性心衰大鼠心功能及心室重构的影响,将大鼠慢性心力衰竭模型随机分为模型组、辛伐他汀组、辅酶 Q10 组、联合用药组,另设假手术组,每组 8 只。采用析因设计资

料的方差分析,其中左室舒张末压结果见表 1。

表 1 各组大鼠左室舒张末压($\bar{x} \pm s$)

分 组	例数	左室舒张末压/mmHg
假手术组	8	5.60 ± 1.25*
模型组	8	20.38 ± 2.35
辛伐他汀组	8	16.16 ± 1.87**
辅酶 Q10 组	8	17.17 ± 2.13**
辛伐他汀加辅酶 Q10 组	8	13.80 ± 2.19*

注:与模型组比 * $P < 0.01$;与辛伐他汀加辅酶 Q10 组比 ** $P < 0.05$ 。

该研究 8 只大鼠无慢性心力衰竭,为假手术组。32 只大鼠为慢性心力衰竭模型,按照是否给予辛伐他汀和辅酶 Q10 分为 4 组。虽然文中说明采用析因设计资料的方差分析,但表 1 并未体现该方法的结果,该方法也不可能得出表 1 中的统计结果。

建议对这 5 组进行拆分,首先比较假手术组和模型组,分析建模是否成功。若满足正态性和方差齐性,则采用成组 t 检验;若满足正态性但方差不齐,则采用近似 t 检验或 Wilcoxon 秩和检验;若不满足正态性,则采用 Wilcoxon 秩和检验。后面 4 组为 2 因素 2 水平析因设计定量资料,如果符合参数检验的前提条件,则可采用析因设计资料的方差分析,观察辛伐他汀、辅酶 Q10 的单独作用和联合作用。

1.2 使用了不必要的统计学方法

例2^[4] 制作脓毒症急性肺损伤模型大鼠 60 只,随机平均分为甘草黄酮组 and 对照组,术毕分别经腹腔注射甘草黄酮和等剂量生理盐水,观测 2 组不同时间氧分压等指标的差异。

该文“统计学方法”中交代:“经检测,本研究所用统计分析均符合正态分布,均采用 t 检验进行统计分析并以 $\bar{x} \pm s$ 表示”。其中氧分压结果见表 2。

表 2 2 组大鼠氧分压比较($\bar{x} \pm s$)

组别	鼠数	氧分压/mmHg		
		术前	术后 1 h	术后 24 h
对照组	30	85.4 ± 11.3 ^{ab}	60.7 ± 8.8	55.2 ± 9.4
甘草黄酮组	30	84.3 ± 12.3 ^{ab}	61.3 ± 9.3 ^b	71.3 ± 11.2
t 值		0.321	0.213	8.973
P 值		0.563	0.762	<0.001

注:2 因素重复测量方差分析;时间(P 值) < 0.001,时间 × 分组(P 值) = 0.103,分组(P 值) < 0.001。与术后 1 h 比较, * $P < 0.05$;与术后 24 h 比较, ^b $P < 0.05$ 。

表 2 数据涉及注射液及时间 2 个因素,其中时间

* 山东中医药大学“名科工程”项目(ZYDXW1310)

† 通信作者

为重复测量因素。由表中数据可知,作者进行了重复测量设计资料的方差分析;但根据“统计学方法”及表中内容可以推测,作者在3个时间点对2组做了 t 检验,每组3个时间点之间也进行了 t 检验,也就是进行了9次 t 检验。该研究应采用重复测量设计资料的方差分析,在SPSS软件中,时间变量如果总体差异有统计学意义,则可以同时显示不同时间的两两比较结果。该研究旨在观察甘草黄酮对脓毒症急性肺损伤的疗效,重复测量设计资料的方差分析结果已经达到研究目的,不必再进行 t 检验,否则反而割裂了设计的整体性,也增加了假阳性错误的概率^{[5]226}。

另外,“统计学方法”中的部分内容表述欠妥,还应补充说明采用重复测量设计资料的方差分析,以及“时间”因素总体比较后两两比较的方法;即使能够采用 t 检验,也应明确说明是成组 t 检验还是配对 t 检验;结果中应补充重复测量设计资料的方差分析的 F 值。

2 “统计学处理”中的问题

“统计学处理”在医学论文中往往作为一项独立的内容予以介绍,包括统计软件及版本、统计描述与统计推断、检验水准等。这部分与统计学的“度”有关的问题可归纳为下述2个方面。

2.1 内容不够全面具体

例3^[6] 为探讨2型糖尿病肾病患者白细胞介素-9、白细胞介素-17和转化生长因子-1水平的变化及其临床意义,设立了糖尿病非肾病组、糖尿病肾病组和健康对照组,分析3组上述指标的水平差异。

“统计学处理”仅下面一句话:采用方差分析与 q 检验及 χ^2 检验。

该文“统计学处理”存在下列问题:1)缺少统计软件及其版本、检验水准,根据文中内容,检验水准 $\alpha=0.05$;2)缺少定量资料统计描述方法,从结果部分可以看出,其实是以 $\bar{x}\pm s$ 表示;3)所列统计学方法用于哪些指标的分析没有交代。从原文可见,性别比较采用 χ^2 检验;其他指标总体比较采用单因素方差分析(原文交代的“方差分析”不够具体),两两比较采用 q 检验。

假设该研究采用的软件为SPSS 22.0,定量资料符合正态分布和方差齐性。“统计学处理”可修改如下:

应用SPSS 22.0软件,定量资料以 $\bar{x}\pm s$ 表示。3组性别构成比较采用 χ^2 检验;定量资料符合正态分布和方差齐性,总体比较采用单因素方差分析,两两比较采用 q 检验。检验水准 $\alpha=0.05$ 。

2.2 含有未使用的统计学方法

例4^[7] 研究肝硬化食管静脉曲张破裂出血患者

门静脉血流参数的变化,“统计学处理”如下:应用SPSS 11.5软件,结果以 $\bar{x}\pm s$ 表示,计量资料采用成组设计的 t 检验,计数资料采用 χ^2 检验, $P<0.05$ 为差异有统计学意义。

其中提到了 χ^2 检验,而文中没有 χ^2 检验的内容,因此“计数资料采用 χ^2 检验”应删去。

3 统计学结果描述中的问题

医学论文“结果”部分以文字、统计表、统计图等形式给出主要的统计学结果,与“度”有关的常见问题概括为以下3种。

3.1 没有给出主要的统计学结果

例5^[8] 观察生长抑素在治疗恶性肠梗阻及术后早期炎性肠梗阻中的效果,结果部分“辅助检查情况对比”小标题下的内容如下:

治疗前2组患者的胃肠减压及C反应蛋白量对比不存在明显差异($P>0.05$),不具统计学意义。经临床治疗后,观察组患者胃肠减压及C反应蛋白的下降量明显多于对照组($P>0.05$),观察组患者的水电解质以及酸碱平衡水平、影像学X线片的恢复正常时间明显短于对照组($P>0.05$),上述差异均具有统计学意义。

上述内容涉及胃肠减压、C反应蛋白量、水电解质以及酸碱平衡、影像学检查等,但缺少具体的数据。2组比较的结果仅给出 $P>0.05$,没有检验统计量和具体 P 值。另外,后面的2个 $P>0.05$ 与“明显多于”和“明显短于”相矛盾,“明显”一词建议删掉。“不存在明显差异”改为“差异无统计学意义”。前面已经说明孰高孰低,并给出了 P 值,“上述差异均具有统计学意义”可删去。

3.2 给出了不必要的统计学结果

例6^[9] 观察外用康复新液治疗紫杉醇渗漏致局部皮肤溃烂的效果,“结果”部分先以“经治疗后两组患者转归情况”为小标题进行下列统计描述:根据同样标准判断2组患者转归。见表3。

表3 2组患者转归情况[例数(占比/%)]

转归	试验组	对照组
痊愈	18(60.0)	5(16.7)
好转	10(33.3)	21(70.0)
无效	2(6.7)	3(10.0)
进展	0(0)	1(3.0)

然后以“2组患者转归情况的非参数检验”为小标题介绍统计推断结果:采用2独立样本非参数检验分析2组患者转归情况,试验组平均秩次为24.00,对照组平均秩次为37.00。Mann-Whitney U 统计量为255.0, Wilcoxon W 统计量为720.0, Z 检验统计量为

$Z = -3.212, P = 0.001$ 。2组的转归情况有显著差异, 试验组治疗效果更好。

该文对统计学软件给出的结果均做了描述, 是不必要的, 只要给出 Z 和 P 值即可。另外, 统计描述和统计推断作为2部分分开描述也不合适, 如果保留统计表的话, 统计学结果可作为备注给出。“根据同样标准判断2组患者转归”“采用2独立样本非参数检验分析”属于方法部分的内容, 在结果中不必描述。“2独立样本非参数检验”也不够具体, 应给出具体的统计学方法。绝对数很小时, 如1、0用百分数描述也不合适。“显著差异”应改为“差异有统计学意义”。该文结果较为简单, 可以去掉统计表, 仅用文字描述即可。建议改为:

试验组痊愈18例, 好转10例, 无效2例; 对照组痊愈5例, 好转21例, 无效3例, 进展1例。试验组疗效好于对照组($Z = -3.212, P = 0.001$)。

3.3 统计表与文字重复

例7^[10] 研究多层螺旋CT诊断肺栓塞的应用价值, 在“患者CT检查后各血管显示数与栓塞数”小标题下列出了肺叶动脉、肺段动脉、主肺动脉、亚段动脉、其他分支动脉的显示数、栓塞数、栓塞率: 肺叶动脉显示数198支、栓塞数为69支, 栓塞率为34.85%, 肺段动脉显示数为……其他分支动脉显示数为4587支, 栓塞数为40支, 栓塞率为0.87%。

之后又给出了题为“患者CT检查后各血管显示数与栓塞数”的统计表, 表中数据与文字完全重复。不仅如此, 在“讨论”部分又一字不差地将上述“结果”中的文字部分复制了一遍。

4 原因分析与建议

从上面的例子可以看出, 如何选择合适的统计学方法, 并恰当地描述统计学结果, 是医学期刊编辑面临的一项重要任务, 也是一项难题。之所以出现诸多问题, 其原因如下。

1) 对统计学的重要性认识不足。尽管很多专家反复强调统计学的重要性, 人们的意识越来越强, 但仍然有对数据不做统计学分析, 仅凭感觉下结论的情况。

2) 对于复杂的统计学设计, 如析因设计、重复测量设计, 习惯按照简单的统计学方法, 如按成组 t 检验、单因素方差分析的思路去考虑, 导致所选方法不正确, 或者做了一些不必要的统计学分析。更有甚者, 有的作者不管资料是否符合统计学方法的应用条件, 尝试多种方法, 直到取得自己“满意”的统计学结果为止。

3) 不能有重点地选择统计软件提供的结果。统计软件提供大量的信息, 有的作者只交代 $P > 0.05$ 或 $P < 0.05$, 而没有给出检验统计量和具体 P 值; 有的作

者又走向另一个极端, 把许多不必要的结果罗列出来。为避免出现上述问题, 兹提出以下建议。

1) 把统计描述与统计推断结合起来分析。如例2根据析因设计资料方差分析的结果, 结合统计表中的数据, 必要时可给出交互效应轮廓图^[11], 就可以对主因素及其交互作用做出合理的解释。

2) 选择合适的统计方法, 特别需要提醒的是不要用简单的方法替代复杂的方法, 如利用 t 检验处理析因设计、重复测量设计资料的方差分析。由于目前期刊中存在的问题很多, 不要盲目地照搬现有期刊的做法。

3) 制订统计学编辑规范。这样, 既避免漏掉重要的信息, 也不会把不重要的内容罗列进来。如《郑州大学学报(医学版)》设计了拟刊发稿件统计学审查表^[12], 不仅可以作为稿件统计学内容编辑加工的标准, 而且可供作者撰写论文及专家审稿时参考, 值得推荐。

5 参考文献

- [1] International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals[S/OL]. [2017-02-10]. <http://www.icmje.org/recommendations/>
- [2] 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
- [3] 于海荣, 张秀芝, 王济兴, 等. 辛伐他汀联合辅酶 Q10 对慢性心衰大鼠心功能及心室重构的影响[J]. 河北医学, 2016, 22(10): 1615
- [4] 温换芳, 苗杰. 甘草黄酮对脓毒症急性肺损伤大鼠的保护作用[J]. 安徽医药, 2016, 20(8): 1440
- [5] 孙振球. 医学统计学[M]. 3版. 北京: 人民卫生出版社, 2013
- [6] 吴强鹏, 刘红. 2型糖尿病肾病患者血清白细胞介素-9、白细胞介素-17和转化生长因子-1的水平及临床意义[J]. 川北医学院学报, 2016, 31(5): 663
- [7] 贺庆红, 黄蔚. 肝硬化食管静脉曲张破裂出血患者门静脉血流参数的变化[J]. 临床荟萃, 2014, 29(3): 295
- [8] 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
- [9] 李静, 张高高, 康瑞, 等. 紫杉醇渗漏致局部皮肤溃烂患者外用康复新液的效果观察[J]. 实用医药杂志, 2016, 33(3): 222
- [10] 莫云海, 冉隆富, 杜涛明, 等. 肺栓塞患者行多层螺旋CT诊断的临床应用价值[J]. 中国CT和MRI杂志, 2016, 14(1): 46
- [11] 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: 158
- [12] 姜春霞. 在办刊实践中培养编辑统计学素养[J]. 编辑学报, 2014, 26(3): 287

(2017-02-14 收稿; 2017-05-31 修回)