

医学期刊编辑应重点审核的统计学问题分析及建议*

郭瑞 宋国营 张媛 周阳 朱瑶 霍永丰 刘雪松 陈新石[†]

中华医学会杂志社《中华医学杂志》编辑部, 100710, 北京

摘要 针对医学论文中经常出现的统计学问题, 审读近3年中华医学会系列期刊论文, 通过实例来探讨医学期刊编辑在稿件退修和编辑加工过程中应重点审核的统计学问题, 分析出现问题的原因, 并提出相应对策及建议。

关键词 统计学描述; 统计学方法; 医学研究论文

Analysis and advices on key statistical problems for editors of medical journals//GUO Rui, SONG Guoying, ZHANG Yuan, ZHOU Yang, ZHU Yao, HUO Yongfeng, LIU Xuesong, CHEN Xinshi

Abstract Taking into account the truth that the statistical problems do occur in medical papers, this study reviewed a series of papers published by the Chinese Medical Association in recent three years and explored the statistical problems, on which the editors of medical journals should focus in the process of manuscript editing. In addition, the paper further analyzed the causes of the problems, and thus put forward corresponding countermeasures and suggestions.

Keywords statistical description; statistical analysis; medical research articles

Authors' address Editorial Department of National Medical Journal of China, Chinese Medical Association Publishing House, 100710, Beijing, China

DOI:10.16811/j.cnki.1001-4314.2019.06.009

规范的统计学描述, 准确的分析方法, 清晰的统计数据呈现和结果的合理解释是高质量医学研究论文的重要组成部分, 也是研究证据链中的关键一环, 编辑在论文统计学问题的审核和把关中应起到积极的作用。尽管在编辑业务培训和杂志质量管理中, 统计学的重要性被一再重申, 但医学论文中的统计学问题仍频繁出现, 全面提升医学期刊编辑的统计学素养需要长期不懈的努力。本研究通过审读近3年来中华医学会系列期刊论文, 从稿件退修和编辑加工时应重点审核的几个方面系统解读医学论文中常见的统计学问题, 并提出相应对策及建议, 以供编辑同人借鉴。

1 统计学方法描述常见问题

1.1 描述不完整 统计学方法介绍部分应包括统计软件名称及版本、资料类型及相应的数据表达方式、不

同资料使用的统计分析方法、单/双侧检验及检验水准, 但很多论文中统计学方法未能完整描述。

例1^[1] 该文统计学方法部分: “统计方法为描述性统计、*t* 检验、相关和多元回归分析”描述较为笼统, 未描述不同资料类型的数据表达方法, 何种资料采用何种相应的分析方法; 相关和回归的方法有多种, 应描述清楚具体使用何种方法。

例2^[2] 该文统计学方法部分虽提及统计软件及版本, 以及不同数据类型所使用的统计分析方法, 但未说明采用单侧还是双侧检验, 也未描述检验水准。

解析 完整清晰的统计学方法介绍是一篇优秀医学研究论文的重要组成部分, 亦是论文撰写规范的要求, 能更好地还原研究者的统计分析过程和评估研究的科学性; 因此编辑应要求作者根据实际情况完整介绍其使用的统计学方法。

1.2 描述与实际不一致 部分论文中, 方法部分提及采用某种或某些统计分析方法, 但结果中并无相应的统计数据呈现; 或者结果中出现某种统计方法的统计量及 *P* 值, 但统计学方法介绍部分并未描述。

例3^[3] 该文统计学方法中提到“所有数值以均数 ± 标准差的方式表示, 正态分布的计量资料组间比较采用方差分析”, 但文章的结果部分仅为简单的病例资料描述, 未见以均数 ± 标准差表示的数据, 也未见方差分析的统计分析结果。

例4^[2] 该文统计学方法中提到采用方差分析, 但结果部分未出现方差分析结果; 而结果部分出现 logistic 回归分析和 Kaplan-Meier 分析结果, 前文统计学方法介绍中却未提及, 前后描述不一致。

解析 统计学方法介绍部分所描述的分析方法应与结果部分的统计分析数据相一致, 如只报告阳性结果或者研究者期望的结果, 则会产生选择性报告偏倚。描述与实际不一致的现象在论文中时有发生, 不排除部分作者为了描述规范而套用其他文章的统计学方法作为模板, 而并未根据实际所用的方法进行描述; 因此, 编辑在退修稿件时应核查统计学方法和结果的一致性, 以保证论文的学术质量。

1.3 统计学方法的表达与描述不准确、不规范 准确、规范地描述统计学方法是一项研究科学、严谨的重要体现, 模糊、不准确的表述可能使读者产生疑问或引

* 中文科技期刊精品建设设计项目(2018KJQK005); 中国科协精品科技期刊工程——学术质量提升项目(IV-XSZL-111)

[†] 通信作者

起歧义。最为普遍的例子是绝大多数论文中对于检验水准的描述通常为“ $P < 0.05$ 为差异有统计学意义”，但在研究中会遇到一部分 P 值等于 0.05 的情况。实际上，当 $P = 0.05$ 时，仍然意味着在无效假设 (H_0) 成立的前提下发生了“小概率事件”^[4]，因此，更推荐表述为检验水准 $\alpha = 0.05$ 。论文中其他方法学描述不准确、不规范的现象也较为常见，举例如下。

例 5^[5] 该文统计学方法表述为“偏态分布的计量资料以 M (范围) 表示，多组间比较采用秩和检验，两两比较采用非参数检验，计数资料多组间比较采用精确 χ^2 检验”。

解析 秩和检验即为一种非参数检验，因此“多组间比较采用秩和检验，两两比较采用非参数检验”的描述较笼统且不规范，即使多组间比较采用秩和检验，也应具体说明是哪一种检验方法，例如 Kruskal-Wallis H 检验。另外，“精确 χ^2 检验”的表述也不规范，不存在精确 χ^2 检验这种方法，正确表述应为 χ^2 检验或 Fisher 精确概率法。

例 6^[6] 该文统计学方法中提到“计数资料组间比较采用 χ^2 检验，两两比较时，其检验水准进行 Bonferroni 校正后为 $P = 0.017$ ”。

解析 两两比较时，检验水准采用 Bonferroni 校正，该文中“Bonferroni”拼法有误；另外，“校正后 $P = 0.017$ ”的写法有误，应表述为校正后的检验水准 $\alpha = 0.017$ ，或者两两比较时 $P \leq 0.017$ 为差异有统计学意义。

1.4 统计方法使用不恰当 研究应基于设计类型、数据性质和分析目的，选用合适的统计分析方法，而不应盲目套用常见的统计方法，否则有可能通过错误的统计方法得出有偏差的结论。

例 7^[7] 该文中描述 2 组患者之间性别构成比的比较使用秩和检验，组间计量资料比较均采用配对 t 检验。

解析 性别构成比比较使用秩和检验的描述有误，应为 χ^2 检验；该研究并非配对设计，2 组之间比较无法采用配对 t 检验，应采用独立样本 t 检验；且从结果数据中可以发现，该研究涉及不同时间点重复测量资料的比较，每个时间点均采用 t 检验比较会增大假阳性错误概率，因此宜采用重复测量资料的方差分析。

例 8^[8] 该文中正态分布资料组间比较采用 Kruskal-Wallis H 检验。

解析 正态分布资料组间比较应采用参数检验，而 Kruskal-Wallis H 检验为非参数检验。在符合参数检验的条件时采用非参数检验会损失很多数据信息，降低检验效能，虽然不认为该方法绝对错误，但明显不

是最恰当的分析方法。因此，编辑在发现此类问题时，应提醒作者使用参数检验。

2 统计分析结果的表述和解释中常见问题

2.1 统计结果表述不规范、不准确 统计结果的清晰呈现和合理解释是论文证据链的重要环节，亦是搭起讨论架构及得出可靠结论的基础。

例 9^[9] 该文统计分析结果术中透视时间两组比较 P 值为 0.000。

解析 作者将 P 值记为 0.000 往往是由于统计软件分析得出的结果中 P 值太小，输出结果的窗口中只显示至小数点后 3 位，而实际 P 值并不为 0，如果记为 0.000，不但写法本身不规范，且影响了数据的科学性，此时应表述为 $P < 0.001$ ^[10]。虽然例文只列出了 1 篇，但该问题非常普遍，多数期刊长期存在此现象，在今后工作中值得注意并严格把关。

例 10^[11] 该文 Pearson 相关性分析 $P < 0.05$ 时即描述为“呈显著相关性”，而未考虑 r 值。

解析 相关性分析的关键评价指标是相关系数 r 值，该文中部分 r 值在 0.1 ~ 0.4 之间，虽然 $P < 0.05$ ，提示存在相关性，但此种相关程度很弱，如果将结果解释为“呈显著相关性”就偏离了真实情况。

2.2 统计结果明显错误

例 11^[12] 该文采用多因素 logistic 回归分析房颤消融术后复发的危险因素，分析结果显示房颤病程是术后复发的危险因素，比值比 (OR) = 0.538，95% 置信区间 (CI) : 1.003 ~ 1.021， $P = 0.013$ ，其中 OR 值为 0.538，而 95% CI 却为 1.003 ~ 1.021，二者相互矛盾；另外，表 1 中 2 组左心室射血分数比较， P 值为 33，明显错误，因为 P 值的含义为原假设为真时得到样本结果或更极端结果的概率，而概率只能在 0 ~ 1 范围内。

例 12^[12] 该文表 3 为幽门螺杆菌感染的多因素 logistic 回归分析结果，回归系数 $\beta = -3.995$ 时，OR 值为 3.205，此处存在较明显的错误，因为 β 值为负数时，OR 值应小于 1。

例 13^[13] 该文表 2 中 OR 值 (95% CI) 为 0.227 (0.059 ~ 0.882) 时， P 值为 0.051，明显有误，OR 值的 95% CI 在 1 的一侧时， P 值应小于 0.05。

解析 明显的统计结果错误会严重影响论文的学术质量，使读者质疑研究的真实性，甚至误导临床决策。论文出现明显错误的统计结果可能原因包括：1) 作者笔误；2) 因疏漏在誊抄统计结果时看错数据；3) 捏造数据等学术不端行为。因此，为保证统计结果真实呈现，编辑在退修稿件时宜核对文中的统计数据，及时发现明显的统计错误。

3 论文中频现统计学描述和分析问题的原因

3.1 统计学专业知识欠缺 绝大多数医学期刊编辑并非统计学专业出身,可能未接受过系统的统计学专业知识培训,对于统计学描述和分析中出现的问题缺少足够的鉴别能力,面临拿不准、不确定的问题时容易产生回避倾向,从而将问题遗留在论文中。

3.2 审稿人无法全盘审核 多数医学论文所采用的统计分析方法为常见方法,且考虑到统计学审稿专家资源紧张,一般只有涉及复杂的统计分析方法时编辑才会考虑将稿件呈送统计学专家评阅。即使统计学专家审阅稿件,也主要专注于审核统计分析方法是否使用得当,而难以对文中具体的统计学描述和分析结果进行一一把关。

3.3 重视程度不足 医学期刊审稿专家和编辑往往更关注于研究的创新性是否足够,是否为当前热点,是否为学科前沿,是否有实际临床价值,而忽视了高质量的研究应基于严谨的科研设计和恰当的统计分析。如果研究设计不合理,或是分析方法和分析结果有偏时,研究结论常难有实际推广价值,甚至会误导临床诊治方案。

4 对策及建议

4.1 加强制度建设 杂志社可制定相应编辑加工规范,对论文中的统计描述、统计结果的解释和表述提出具体的规范性要求,使编辑在工作中有据可依,有制可循,从思想上重视论文中的统计学问题。中华医学会系列杂志目前执行的为2008年中华医学会第5次杂志工作会议制定的编排规范,为满足现阶段工作的需要,中华医学会杂志社于2019年开始进行编排规范的修订工作,通过召开一系列修订小组会议,对现有的编排规范条目进行逐条讨论、修改,其中对于统计学描述和分析部分的要求也进一步细化和明确。

4.2 重视统计学培训和学习 从杂志社角度来说,应重视对编辑统计学专业知识的培训,在常规编辑培训班中增加统计学授课内容,或定期邀请统计学专家做讲座或答疑。从编辑自身角度来说,除了参加培训班外还可通过向统计学专家请教或研读统计学教材及相关论文来提升自身统计学知识水平。

4.3 认真审核,严格把关 退修稿件时应全面审核论文中的统计学问题,并向作者逐一提出修改建议,如遇到确实难以把握的问题,建议作者咨询统计学专家。

5 结束语

医学期刊编辑在论文统计学问题把关中具有重要

的作用,因此,医学编辑应熟悉常见统计学方法的适用范围,掌握相关统计学名词的规范表达。在退修稿件时对统计描述和统计分析中可能存在的问题进行全面审核,重点包括统计学方法描述是否完整,描述是否和实际一致,统计方法是否恰当,统计结果的解释和表达是否准确、规范,及时发现明显的统计学错误,以保证研究论文的真实性和科学性。

6 参考文献

- [1] 郑亚楠, 严厉, 龚茜. 医学生创业意愿及其影响因素的研究[J]. 中华医学教育杂志, 2018, 38(2): 173
- [2] 李腾, 詹贤章, 薛玉梅, 等. 持续性心房颤动个体化射频导管消融的临床研究[J]. 中华心律失常学杂志, 2017, 21(1): 48
- [3] 张省, 秦金保, 殷敏毅, 等. 静脉激光原位开窗技术治疗弓形病变[J]. 中华血管外科杂志, 2017, 2(1): 17
- [4] 颜虹, 徐勇勇, 赵耐青, 等. 医学统计学[M]. 第2版. 北京: 人民卫生出版社, 2010: 97
- [5] 赵志强, 李双喜, 李子禹, 等. 新辅助治疗对局部进展期胃上部癌 D2 根治术淋巴结获取的影响[J]. 中华消化外科杂志, 2018, 17(4): 374
- [6] 林莉, 魏玉梅, 王晨, 等. 孕晚期血红蛋白水平与妊娠结局的关系[J]. 中华围产医学杂志, 2018, 21(4): 262
- [7] 王蕾, 张晓敏, 李筱荣. 27G 与 25G 玻璃体切割手术治疗特发性黄斑前膜疗效对比观察[J]. 中华眼底病杂志, 2017, 33(1): 46
- [8] 蔡成, 裘刚, 龚小慧, 等. 连续性肾脏替代治疗在新生儿急性肾损伤救治中的应用[J]. 中华实用儿科临床杂志, 2019, 34(1): 31
- [9] 刘长青, 邵钦, 杨晨松, 等. 锁骨中段斜形骨折固定中使用拉力螺钉的疗效分析[J]. 中华手外科杂志, 2019, 35(2): 85
- [10] 肖骏. “0”的科学意义辨析与用法建议[J]. 编辑学报, 2018, 30(3): 254
- [11] 史本龙, 徐亮, 毛赛虎, 等. Lenke I 型青少年特发性脊柱侧凸后路矫形术后身高增加的影响因素分析[J]. 中华医学杂志, 2017, 97(33): 2565
- [12] 蔡玲, 伊比然根, 张玫, 等. 拉萨地区自然人群幽门螺杆菌感染的流行病学调查[J]. 中华消化杂志, 2018, 38(1): 5
- [13] 章达冠, 柯乐斌, 曹曙光, 等. 叉头样/翼状螺旋转录因子 3 基因多态性与克罗恩病易感性和疾病表型的关系[J]. 中华消化杂志, 2018, 38(1): 11

(2019-08-23 收稿; 2019-10-06 修回)