

科技论文中数据问题的编辑审读技巧

游俊 赵燕 胡小洋

《湖北大学学报》编辑部, 430062, 武汉

摘要 编辑初审及刊后审查中常发现科技论文中存在一些数据方面的问题, 有些数据错误甚至很严重。针对科技论文中常见的3大类8小类数据问题, 建议编辑按一定的审读程序, 综合采用比对全文各处、利用科学常识、审查方法设计、进行简单计算、借助权威软件、依托各种资源等方法来加强编辑审读, 以最大化地避免和有效解决科技论文中的数据错误。

关键词 科技论文; 数据问题; 编辑实践; 审读技巧

Reviewing skills for data problems in sci-tech papers // YOU Jun, ZHAO Yan, HU Xiaoyang

Abstract During the editor's first review and re-review after publication, many data problems can often be found in some sci-tech papers, and some data errors are even serious. Aiming at eight kinds of common data problems, a combination of skills should be adopted to strengthen the editor's reviewing, so as to maximally avoid the data errors. The skills include: contrasting throughout the full text, using scientific knowledge, checking the design scheme or methods, carrying out a simple calculation, having the aid of authority software, and relying on all kinds of resources.

Key words sci-tech paper; data problem; editing practice; reviewing skill

Authors' address Editorial Department of the Journal of Hubei University, 430062, Wuhan, China

在编辑实践中, 我们发现许多论文稿件存在着数据方面的问题, 有些数据错误相当严重, 不仅影响了对论文结果的分析 and 讨论, 有的甚至得不出作者想要表达的结果和结论。编辑同人的相关研究^[1-3]也表明, 即便是正式出版的论文中仍存在着很多的数据问题, 特别是生物学、农业科学、医学等学科论文中的数据错误更为严重。

科技论文中存在的数据问题, 已引起了编辑同人的关注^[4-5]。论文中出现的 data 问题尽管多种多样, 但编辑仍然是可以有针对性地予以解决的, 如王有登等^[6]提出的可保证数据科学性和有效性的数据检验。在编辑实践中我们遇到过多类 data 问题, 借鉴他人经

围, 提高社会效益; 对于刊登广告的医药企业采取提供多元化服务的方法吸引订购, 借助该企业临床用药推广的渠道开拓潜在的作者市场, 在扩大期刊社会效益的同时带来了可观的经济利益。同时, 为满足不同读者的需求, 本刊针对读者群体进行市场细分, 为持有不同需求的读者发行单行本。合理、有效的发行方式满足了医学科研管理人员对医学技术发展方向的宏观把握, 满足了科研人员学术交流的需求, 满足了医药厂商产品介绍与推广的需求。

4 结束语

“推动科技发展, 维护学术尊严”, 是医学期刊编辑社会价值与精神的重要体现^[11]。“不崇洋、不唯书、只求实”的箴言丰富了医学期刊编辑工作者素质培养内涵, 值得推崇与弘扬。在国内医学期刊行业蓬勃发展的时代, 期刊人应深刻体会这一精神, 勇于质疑、积极探索、求真务实, 并怀有强烈的民族荣誉感与责任感, 真正践行能推动国内医药卫生行业发展的办刊宗旨, 体现出期刊的存在意义与价值。

5 参考文献

[1] 百度百科网. 黎磊石 [N/OL]. <http://baike.baidu.com/>

view/98418.htm

- [2] 游苏宁, 石朝云. 中国医药卫生期刊六十年 [J]. 中国科技期刊研究, 2009, 20(5): 777-782
- [3] 关卫屏, 游苏宁, 石朝云. 我国科技学术期刊分级评价及管理机制探讨 [J]. 编辑学报, 2012, 24(2): 103-107
- [4] 黄政, 郝希春, 汪峰. 编辑应重视对科技论文参考文献的审核 [J]. 编辑学报, 2009, 21(4): 310-311
- [5] 王荣兵. 论医学期刊编辑的专业主义精神 [J]. 编辑学报, 2009, 21(1): 78-80
- [6] 吴涛, 袁天峰, 王睿, 等. 医学期刊, 严谨但不能拘谨: 谈研究 SCI 名刊所感 [J]. 中国科技期刊研究, 2011, 22(1): 85-87
- [7] 马芳莲. 医学术语规范化的必要性: 兼谈几个常用词的辨证 [J]. 科技术语研究, 2000, 2(4): 6-9
- [8] GB/T 6447—1986 文摘编写规则 [S]. 北京: 中国标准出版社, 1986
- [9] 乔闻钟, 陈德刚. 学术论文“摘要”人称之分析 [J]. 编辑之友, 2008(2): 88-89
- [10] 科技期刊成“软肋” 院士力推科技期刊突破发展僵局 [N/OL]. (2010-06-09) [2012-06-01]. <http://scitech.people.com.cn/GB/11822484.html>
- [11] 刘建超. 科技期刊编辑的社会价值及其实现 [J]. 编辑学报, 2011, 23(2): 101-102

(2012-06-09 收稿; 2012-07-23 修回)

验和结合稿件实际,我们采用如下的程序和方法纠正了不少稿件中的数据错误。

1 比对全文各处,避免相异数据错误

针对论文中各部分数据内容不一致的问题,编辑审读的第一步可考虑全文比对,中英文摘要,图表,结果,结论与讨论是要重点比对的部分。通过这种简单的方法就可以避免许多最常见的低级错误。如近期处理的一篇关于高校食堂含油废水的稿件,在作者修改后的稿件中进行全文比对时,笔者发现其中文摘要中的“化学需氧量(COD)平均值为4 170 mg/L(2 000 ~ 5 000 mg/L),油脂质量浓度平均值为229.07 mg/L(100 ~ 300 mg/L),分别超出GB 8978—1996《污水综合排放标准》三级标准的4 ~ 10倍和1 ~ 3倍”,与英文摘要“were nearly 40 and 3 times of those in the GB 8978-1996' third standard allowances”的倍数关系数据明显相异,与第一原稿也相差很大。作者的投稿稿件中为“COD(平均值为4 170 mg/L)和油脂质量浓度(平均值为299.07 mg/L)分别是GB 8978—1996规定的排放标准的4 ~ 40倍和3 ~ 20倍,远远超过了国家标准”。相互沟通后作者的解释是“原稿中采用的是国家最大允许浓度和标准浓度,修改稿中采用了三级排放标准,餐饮废水在排入有污水处理厂的城市管网前,需治理达到三级标准。中英文数据相异是因为修改中文后英文摘要部分忘记同步修改”;因此,修改稿中的中文摘要及正文数据是正确的,英文部分应该修改为“4-10 and 1-3 times”。

2 利用科学常识,核查简单数据错误

对于浓度单位及指数倍数错误,部分大于整体数据,或同组分项中百分比总计大于1等错误,可以借助普通科学常识来发现和处理。还有一些通过简单计算产生的衍生数据,若与本学科的公认、公知相去甚远,又无特别说明的,必须利用专业学科常识来判断。

关于科学常识的积累和储备,需要编辑平时加强学习,善于总结;有了这些科学积累,在工作实践中只要用心观察,是很容易发现明显不合常理的数据错误的。在如表1(部分截取)所示的作者提供的数据表中,我们根据有关动物生长的常识来判断——在正常情况下,随着时间的延长,动物的体质量在成长期应该是不断增加的,而表1中用黑体排出的5个数据应该是不正确的。经查作者的投稿,我们发现,原表格中平均体质量用g作单位,24个数值中有19个小于0.1,数据太小。作者在修改稿中将平均体质量单位更改为mg,使数据的取值更合理;但是表1中用黑体排出的5

个数据在调整时小数点向右少移了1位,应该分别更改为125.3、111.9、147.5、135.1、116.9才正确。

表1 “查虫清”对4龄菜青虫幼虫生长发育的影响

处理质量浓度/ (mg/mL)	平均体质量/mg			
	处理前	24 h	48 h	96 h
0(对照)	56.5	90.5	12.53	14.75
12.5	53.7	80.9	11.19	13.51
25.0	60.9	84.6	89.9	11.69
50.0	56.0	68.1	86.8	92.4
100.0	47.2	49.7	53.5	78.0
200.0	58.0	59.8	60.4	64.5

3 透过方法设计,审查数据是否可信

为了实现学术交流与传播的目的,科技论文在报道研究结果及结论的同时,应注重报道该研究是“如何完成”的,反映在论文的方法学部分,要用恰当的语言,详细、清晰地描述整个研究的过程,特别是样本数、抽样方法、对照设置、试验重复、数据含义、取值范围、测定方法、分析软件等一定要交代清楚,让读者能够判断作者的研究方法是否准确、结论是否可靠,从而可否借鉴,也可让小同行能够以这些方法描述为依据重复此项研究,进而对论文研究质量“评头论足”。

张卫华^[7]指出,我国医学期刊上刊载的论文,方法学部分普遍存在着内容过于简单、关键细节描述不清甚至残缺的问题;不仅如此,其他专业的论文也存在同样的问题。

我们曾收到一份关于用95%乙醇提取大蒜中的黄酮类化合物的稿件,作者通过研究提取过程中不同料液比、不同提取时间、不同提取温度对提取率的影响,确定最佳的提取工艺条件为:固液比1:16,80℃浸提3h。其结论是:与传统的水提法相比,此方法不仅提取率高(样品液中总黄酮含量为3.6%),而且速度快、节能。虽然此论文的选题有一定实用价值,但因作者在修改稿中仍没有提供重复试验数据以及没有对各种条件下的数据进行统计分析,我们作了退稿处理;因为我们不能判断作者简单的表格(如表2)中提供的数据是否真实地影响着提取过程,是试验误差还是本质差异。

在此类工艺过程研究类论文中,对工艺条件本身的要求应该是很精准的,表2中提取时间也许还可多设计些取样点,如5h、6h,也许黄酮含量还有更多变化。既然提取4h时的黄酮含量比较高,为何选择3h,若现实中选择4h对生产周期和成本增加也许没有多大影响。不仅如此,作者文中的其他几个表格也存在类似问题。由于数据不能充分说明结论,因此,我们建议作者重新设计试验并补充相关数据后再投稿。在稿件初审中,由于缺少重复试验、缺乏对照或对照设置

不合理、样本含量不足、采样方法有缺陷、统计分析方法不当等我们直接退稿的事例也很多。

表2 提取时间对黄酮含量的影响

提取时间/h	1.0	1.5	2.0	3.0	4.0
黄酮含量/%	2.61	2.90	3.24	3.51	3.52

4 进行简单计算,查验数据是否准确

对科技论文中的数据,编辑要进行简单验算,判断是否有误。通常可以对图表或公式选择几个特殊条件点进行分析计算,或者是对怀疑之处进行简单计算。

在一篇分析黄河水中的有机污染物成分的论文中,我们就利用这一方法找出了稿件中经作者几次修改仍有错误之处。如表3(部分截取)所示,峰面积栏中用黑体排出的255.87用常识直觉判断,它是不正确的:一是此数据的小数点结构与其他数值不同,可能不对;二是根据相对含量栏目中对应数值大小,判断其可能有误。根据峰面积与相对含量的对应关系计算:若按59 087→2.17推算,应为259 493→9.53;若按292 308→10.71推算,应为260 102→9.53。如此推算,数据255.87可能为259 870才对。经与作者联系,他承认录入时多点了小数点,将原数值缩小到0.1%,而且将9错为5,也没有发现。正确的应为259 870。我们的简单推算得到确认。

表3 全程序空白组表

分子式	峰面积	相对含量/%
C ₁₁ H ₂₄	59 087	2.17
C ₁₂ H ₂₆	202 814	7.43
未检出	1 297 222	47.54
C ₁₂ H ₂₆	84 848	3.11
C ₁₂ H ₂₆	255.87	9.53
C ₁₃ H ₂₈	28 861	1.06
C ₁₄ H ₂₈ O ₂	103 897	3.81
C ₁₆ H ₂₂ O ₄	292 308	10.71
C ₂₅ H ₄₂	404 472	14.82

5 借助权威软件,验算统计分析数据

目前作者使用较多的大型专业统计分析软件,如SAS、SPSS、Statistics、DPS等,可根据预先录入的数据自行计算并给出分析结果。相对于昂贵的专业统计软件,刘清海等^[8]推荐使用Office组件中的Excel软件及其拓展软件WinSTAT。他们认为Excel及WinSTAT不仅具备专业统计软件的基本功能,而且更方便易得、操作简单,非常适合于非统计学专业的作者和编者使用。

针对统计学处理中存在的常见问题,肖丽娟等^[3]提出了较好的应对措施,包括指明所用统计软件的名称和版本,满足样本数目及抽样要求,弄清资料类型分类,正确描述统计资料,正确选用统计分析方法,正确使用统计学名词及符号,正确表达统计结果的解释、检验水平及置信区间等。

6 依托各种资源,寻求专业数据支持

作者对论文数据把握的水平参差不齐,而编辑应该具备基本的数据处理知识,有些简单的数据问题应该稍作分析或处理就能发现,较复杂的数据问题可查找专业文献。为了更好、更快地对数据进行审读,建议每个科技编辑准备一些经典的统计学参考资料,遇到问题可随时随手查询。譬如在审读生物类论文数据时我们经常查阅《生物统计学》及其他最新、被大众认可的专业文献。

关于数据的真伪,若编辑采用以上各种方法仍不能自己解决,还可以聘请权威专家审稿:请统计学专家从统计方面把关,查验理论计算是否符合逻辑,数据是否具有统计学意义;请学科领域专家从专业方面把关,验证数据是否符合行业标准及行业实际,以及是否具有现实应用价值。审稿专家大都具有良好的学术素养和准确的判断能力,在某种程度上他们也同样是作者和读者,他们更知道应该在哪些细节方面严格把关,才可尽量减少论文中的数据错误。

7 参考文献

- [1] 苏振伦. 编辑的信息意识:拒绝错误信息,厘清虚假信息,追求正确信息[J]. 中国科技期刊研究, 2011, 22(1): 138-141
- [2] 王音. 农业和生物学期刊应重视对表格中数据统计分析结果的审校[J]. 编辑学报, 2010, 22(5): 413-415
- [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): 500-502
- [4] 朱久法, 张彩虹. 学术论文中常见的一些数据问题及对编辑工作的要求[J]. 中国科技期刊研究, 2010, 21(4): 546-548
- [5] 吴学军, 廖粤新. 科技期刊编辑应注重统计数据的审核[J]. 编辑学报, 2010, 22(5): 416-417
- [6] 王有登, 李文川, 刘继宁. 科技论文中数据的审读与检验[J]. 编辑学报, 2002, 14(1): 21-22
- [7] 张卫华. 医学论文方法学部分的形式审查[J]. 编辑学报, 2010, 22(6): 497-499
- [8] 刘清海, 张恩健. EXCEL的统计功能在医学期刊编辑中的应用[J]. 中国科技期刊研究, 2004, 15(6): 690-692

(2012-03-28 收稿; 2012-05-17 修回)