

# 运用学术不端文献检测系统 检测医学论文存在的问题及对策

杨 晨 晨

《新疆医科大学学报》编辑部,830011,乌鲁木齐

**摘 要** 通过对医学论文文稿的检测,发现科技期刊学术不端文献检测系统(AMLC)在缩短查抄时间的同时,也存在一些问题:系统不能有效地检测图表和公式等的重复性;对论文发表的时滞性,出现漏检;不能有效地识别稍加变动的多源抄袭现象;存在语种转换的检测盲区;对于合理引用,系统仍会认为抄袭。应从以下方面加以注意:发挥编辑的主观能动性,客观分析检测结果;加强宣传教育,从思想上杜绝抄袭;加强对抄袭的警示;利用其他审稿方式进行辅助审查;增强审稿专家的作用,完善同行评议制度。

**关键词** 学术不端文献检测系统;存在问题;对策

**Problems and solutions of AMLC in detecting medical articles**//YANG Chenchen

**Abstract** Although the academic misconduct literature check system (referred to AMLC) can help editors shorten the time of examining manuscripts, there still exists some problems. The system cannot check charts and formulas effectively. The publication delay of papers causes the leakage of inspection; the system does not effectively identify slight changes in multi-source plagiarism; there is short of detection of similar or even same contents with different languages. In order to use this system sufficiently, attentions should be paid to the following aspects: full

playing editors' subjective initiative, analyzing the check results objectively; strengthening publicity and education; reinforcing warnings of plagiarism; using other auxiliary review methods; strengthening the role of peer reviewers, and improving peer review system.

**Keywords** AMLC; medical paper; problem; solution

**Author's address** Editorial Department of Journal of Xinjiang Medical University, 830011, Urumqi, China

近年来,学术界一稿多投、重复发表、抄袭、编造等学术不端行为,已严重影响了科技期刊的健康发展。为净化学术环境,震慑学术不端行为<sup>[1-2]</sup>,中国学术期刊(光盘版)电子杂志社与同方知网技术有限公司于2008年共同研发出科技期刊学术不端文献检测系统(简称AMLC),可检测抄袭、剽窃、伪造、篡改、不当署名等多种学术不端行为<sup>[3-4]</sup>。该系统以中国学术文献网络出版总库为全文比对数据库,使用者将任意一篇待测稿件的电子版上传,即可完成检测。笔者在使用中发现,该系统在缩短查抄时间的同时,也带来一些问题。作为编辑,需要参考检测结果,但也不能简单、机械地以重合

## 7 结束语

关键词是论文的检索标志,是表达文献主题概念的自然语言词汇,其标引是一项知识性和技巧性都极强的工作。

通过上述关键词标引的讨论,我们可以认为:1)选词既不必拘泥于文章自身,也不必在意是叙词还是自由词,只要选用的词能表达文章主题而且使用频率比较高即可;2)应当关注或迎合多数读者检索文献时的思路 and 习惯;3)要充分利用权威专业词典来规范关键词的选取,因为专业或行业的叙词表或主题词表更新极其缓慢而且在检索文献的读者中罕有。

总之,要让广大读者检索文献时有尽可能高的检准率和检全率。

## 8 参考文献

[1] GB 7713—1987 科学技术报告、学位论文和学术论文的

编写格式[S]//作者编辑常用标准及规范.3版.北京:中国标准出版社,2011:518

[2] 中国农业出版社辞书编辑室.英汉农业大词典[M].北京:中国农业出版社,1998

[3] GB 13190—1991 汉语叙词表编制规则[S].北京:中国标准出版社,1991

[4] GB/T 3860—2009 文献主题标引规则[S].北京:中国标准出版社,2009

[5] CAJ-CD 1—2006 中国学术期刊(光盘版)检索与评价数据规范[S].北京:中国学术期刊(光盘版)电子杂志社,2006

[6] 农业部情报研究所.农业科学叙词表[M].北京:中国农业出版社,1994

[7] 刘大乾. SCI 关于关键词的一般选取准则及词义库建设[J].中国科技期刊研究,2007,18(6):1073-1074

[8] 张建蓉,陈燕.学术论文中关键词标引的常见问题剖析[J].编辑学报,2003,15(2):104-105

(2013-08-27 收稿;2013-09-08 修回)

度高低作为稿件取舍的标准,应发挥自己的专业优势和职业敏感性,对稿件进行甄别,否则就不能达到真正抵制不端行为的目的,反而有可能失去一些优秀稿件。

本文对如何合理使用 AMLC 提出建议,以与同人交流。

## 1 AMLC 存在的问题

### 1.1 系统不能有效地检测图表和公式等的重复性

该系统对于含有数学公式和图的 PDF、Word 格式的文件,检测效果不理想,有很大的局限性<sup>[5]</sup>。

例如医学论文中的计量资料常采用均数  $\pm$  标准差 ( $\bar{x} \pm s$ ) 的形式,如体质量为  $(55.4 \pm 6.4)$  kg、空腹血糖为  $(7.0 \pm 1.2)$  mmol/L 等,类似于这种含有公式的数据,系统不能识别。

再如分子生物学科论文中常出现 PCR 凝胶电泳图、聚丙烯酰胺凝胶 (SDS-Page) 电泳图,以及病理学论文中的免疫组织化学染色图,系统在检测时会将这些图自动跳过,不作检测。

一些作者利用检测系统的这一缺点,在文章中加入大量图表及公式,以减少文章的检测重合度。更有甚者,将别人文章中的图复制过来,作为自己的实验结果。如本刊一篇论文不端行为检测符合要求;但是审稿专家指出论文中的流式细胞仪检测图不是作者的实验结果,原图标的是检测淋巴细胞的,而文中描述检测的则是另一种细胞,属于抄袭行为。流式细胞仪的使用属于基础医学免疫学专业的范围,专业性很强,编辑对此不熟悉;所以对这类图不能识别,而系统也不能识别,导致这种不端行为不能被发现。

**1.2 对于论文发表的时滞性出现漏检** 论文印制成纸质版发表后或在论文发表的同时,刊社将论文的电子版上传至数据库,数据库要对论文进行一定的汇编处理,导致中国学术文献网络出版总库中的论文有一定的滞后性。对于已公开发表但尚未被总库收录的论文,该系统无法检测。同时,系统只能将投稿论文与已发表的论文进行比对检测,而对各编辑部大量正在处理的稿件也无法检测。有些作者为了评职称赶时间,选择一稿多投,在论文正式发表之前,系统更无法识别。

### 1.3 不能有效地识别稍加变动的多源性抄袭

AMLC检测一段文字原封不动的抄袭效果极为理想,而对于文字顺序上稍加修改或小幅度调整,系统检测能力则明显降低。例如,在一段文字前加上一些过渡词,如“本研究显示”“研究报道”“近年来”等,或在一个句子中间加上一些过渡词语,如“同时”“另一方面”“但是”等,将句子断开,系统在检测时就不会标红,不会判定为抄袭。

一些作者了解到编辑部采用 AMLC,为顺利通过检测,想尽办法对文字加以处理,如对文字顺序作调整、变更词语或变动表达方式,玩一些文字功夫,逃避 AMLC 的检测。这实际上属于一种升级版的造假。

**1.4 存在语种转换的检测盲区** 中国学术文献网络出版总库收入的均为中文文献。一些作者将外文文献整篇翻译成中文后发表,或者将自己的稿件作中英文互译,生成 2 篇论文。这种手段的高明之处就是利用了系统存在语种转换检测的盲区。

**1.5 对于合理引用系统仍会认为是抄袭** 一篇文章中使用的文字有可能在多篇文章中找到原文。在医学论文中,多为普遍公认的试验原理、方法和概念,如类似于“健康(清洁级) Wistar (SD) 大鼠  $\times$  只,体质量  $\times$  g,雌雄不拘,购于  $\times$  动物实验中心,明暗周期 12 h/12 h,自由摄食、饮水。随机分为  $\times$  组与  $\times$  组……”这样的表述经常出现在文中。AMLC 在进行比对时就会将上述表达雷同的段落标红,判断为段落抄袭<sup>[5]</sup>。这些表述实际为合理引用,应不属于抄袭。

## 2 建议

**2.1 发挥编辑主观能动性,客观分析检测结果** AMLC 毕竟只是一套机械系统,操作过程比较死板,鉴别剽窃抄袭还需人力辅助求证。论文质量的好坏决定于其是否有创新性、先进性等多种因素。在实际工作中,应将技术与人工手段有机结合起来,编辑的判别能力在遏制不端行为中起到很大作用。编辑要有严谨的工作态度、诚实的道德品质及相关的专业水平。作为医学编辑,在阅读有关于编辑专业的期刊如《编辑学报》《中国科技期刊研究》的同时,平时工作中也应该阅读相关医学类杂志,如《柳叶刀》《科学》以及中华系列杂志,了解医学领域内的国内外最新发展趋势、基本的研究方法和理论,提升编辑的鉴别能力。

例如,一篇稿件投至本刊时,通过了不端行为检测,审稿专家也同意全文刊用。笔者学的是临床专业,在辑校过程中发现文中描述的为其医院收治的脑出血患者的 CT 表现,而在 CT 图片上的英文信息显示,该 CT 图片是在内地另一家医院拍摄的。与作者沟通后,作者承认其复制了另一篇文章的图片。如果没有一定的临床经验,这一问题是很难被发现的。

对于编辑审稿中发现的可疑之处,可要求作者提供原始资料。特别是在遇到文字复制比为 0 或接近 0 的文章,极有可能是作者直接翻译英文文献而完成的。

本刊有一篇稿件,编辑在初审时查阅相关文献,发现有关报道较少,认为稿件有一定创新性,并且系统检测后复制比仅为 2%,远远低于本刊要求的 20%,即送

专家审稿。审稿专家正好阅读过这篇文献,从而指出文章出自英文译文。因指出及时而避免了严重后果。

若发现低复制比的文章,可怀疑其为外源论文,则采用相应的外文数据库如 Medline、PubMed 等进行搜索,或请作者提供原始资料<sup>[6]</sup>,如病历资料、原始图或原始数据。同时,在编辑部加强职业道德教育,使编辑认识到编辑出版工作是学术不端行为产生的领域,自身承担着抵制不端行为的社会责任。

**2.2 加强宣传教育,从思想上杜绝抄袭** 学术不端行为的泛滥与当前社会中的浮躁风气有关。学术期刊具有科研成果的刊载、传播、评价功能,在成果评奖、职称评定、论文答辩等活动中扮演着重要角色,发挥着重要作用;因此,学术期刊应当加强内部各项制度的建设,采取切实措施,在遏制学术不端行为中有所作为。编辑应及时与作者沟通,告之一稿多投、重复发表的处理办法。

作为研究生导师,也应加强对学生的思想教育,正确引导学生的价值取向。李淑娟等<sup>[7]</sup>建议,可在高校设立科研诚信教育课程,加大关于学术不端行为对个人和社会危害的宣传,增强作者及读者自觉抵制不端行为的意识,营造良好的学术氛围。

**2.3 加强对抄袭的警示** 本刊编辑部自2010年起,凡拟录用稿件,都要与作者签署《著作权转让授权书》,由作者声明来稿资料真实、无剽窃他人作品及一稿多投的问题,要求作者签字,并加盖单位公章,并提醒作者学术不端行为有侵犯他人著作权的危险性。如发现来稿中有抄袭内容,则将作者列入“黑名单”,并通知作者,本刊将拒绝被列入“黑名单”作者的任何稿件。

此外,使用 AMLC 的科技期刊均在使用前签订保密协议,其只限于刊社内部检测用。本刊使用 AMLC,仅仅是对作者投稿的一种检测,发现后应及时给予警告,并给予引导,提醒作者改正。考虑到对作者人格的尊重和保护,刊社也要遵循一定的保密原则,不宜公之于众或者随意泄密。

**2.4 利用其他审稿方式进行辅助审查** 编辑可以利用百度、谷歌等搜索引擎对论文进行搜索,同时,中国生物医学文献服务系统、中国期刊全文数据库等数据库也是辅助编辑工作的有效工具。对于 AMLC 系统未查出抄袭或只有少量抄袭的论文,编辑可通过关键词、文章题名在数据库和搜索引擎中审查稿件的科学性、先进性和创新性。相关主题的文章如果很多,提示有抄袭的可能。同时,注意区分论文中雷同内容究竟是合理引用还是抄袭,对于一些介绍研究背景、实验方法、实验设计方面的内容,只要注明出处(引用参考文献),应当视为合理引用。

**2.5 增强审稿专家的作用,完善同行评议制度** 审稿

专家的评审意见是决定一篇稿件能否刊用的重要依据,也是决定论文质量的主要因素<sup>[8]</sup>。同行评议是科学评价过程中采用的最重要、最普遍的方式,在论文评审、基金申报、职称评定、学术荣誉给予等众多科学评价活动中都被广泛采用,无论国内还是国际莫不如此<sup>[9]</sup>。本刊在请专家审稿时,将重合度在 15% ~ 25% 范围内的稿件的查重结果,即标红稿件与稿件一同提供给审稿专家,同时,可将编辑初审时查阅的相关文献提供给他们,供其审阅,以充分发挥同行评议的作用。

### 3 结束语

《新疆医科大学学报》在未使用 AMLC 时,2009 年先后有 2 篇稿件被审稿专家指出已经公开发表过,作者将姓名换成自己的。虽然此稿件未予刊登,但是引起了编辑部的重视。2010 年 7 月正式启用 AMLC 对来稿进行检测,当年发现有 3 篇稿件的重复率 > 80%,给予退稿;2011 年至今,“问题”稿件的数量逐年减少:说明编辑部的不端行为防范措施收到了一定的成效。目前,本刊投稿数量增多,出版时滞难以避免,针对此种情况,本刊已于 2013 年 6 月采用清样稿再次查重的措施,以求能充分做到合理防范和治理不端行为。

### 4 参考文献

- [1] 崔洁,谭华.对“学术不端文献检测系统”的审稿功能探析[J].中国科技期刊研究,2010,21(3):336-338
- [2] 江霞,方玉桂,陈伶俐,等.医学科技期刊应用《科技期刊学术不端文献检测系统》判断抄袭时应慎重对待的几种情况[J].中国科技期刊研究,2010,21(4):471-472
- [3] 刘清海,王晓鹰,孙慧兰,等.AMLC 检测医学论文的特点及期刊的应用对策[J].编辑学报,2009,21(6):526-529
- [4] 中国学术期刊(光盘版)电子杂志社,同方知网技术有限公司.科技期刊学术不端文献检测系统用户使用手册[EB/OL].[2013-05-05].<http://check.cnki.net/>
- [5] 杨晨晨.“科技期刊学术不端文献检测系统”在《新疆医科大学学报》审稿中的应用及判断[J].中国科技期刊研究,2012,23(6):1088-1090
- [6] 石朝云,游苏宁.诚信是学术报道之本[J].编辑学报,2013,24(3):206-209
- [7] 李淑娟,李鸿斌,李邦.运用 AMLC 检测硕博学位论文存在的问题及对策[J].陕西教育:高教,2012(10):8
- [8] 聂兰英,王钢,金丹,等.论科技期刊审稿专家队伍建设[J].编辑学报,2008,20(3):241-242
- [9] 张秀峰,王蔚,段佳,等.数字出版时代科技期刊应对学术不端的防范措施[J].编辑学报,2012,24(5):471-473

(2013-05-14 收稿;2013-07-19 修回)