

重视学术伦理是期刊编辑义不容辞的责任

——《浙江大学学报》(英文版)初探 CrossCheck 的工作模式和规范标准*

林汉枫** 贾晓燕** 张月红*** 张纯洁 金美青 张欣欣 翟自洋

收稿日期:2011-02-21

修回日期:2011-04-08

《浙江大学学报》(英文版)编辑部,310027 杭州市浙大路38号,E-mail:jzus_lhf@zju.edu.cn

摘要 从重视学术伦理的角度,阐述了本刊编辑是如何结合国际同行评审意见,以 CrossCheck(反剽窃文献检测系统)为辅助工具,对投稿论文从内容组织,到语言描述,用初步建立的检测流程,来查对投稿论文撰写的原创性比,把好学术诚信关。同时,结合《浙江大学学报 B 辑:生物医学与生物技术》(英文版)所遇到的一些典型案例,探讨了生物医学论文所出现的一些常见问题以及解决方法。最后,对本刊在 2010 年某时间段已经通过国际同行评审的 138 篇论文所做的 CrossCheck 查比结果,进行了稿源统计分析,强调了重视学术伦理、提高研究论文原创性的必要性。

关键词 学术伦理 工作流程 CrossCheck 案例分析 稿源分布

1 引言

随着全球科技论文数量的迅猛上升,国际学术界和出版界关于学术不端的报道已是屡见不鲜,因此,重视学术伦理是学术界和出版界义不容辞的责任^[1-5]。1997 年由英国几家医学期刊编辑成立的国际出版伦理委员会(Committee on Publication Ethics, COPE),已经发展到今天的六千多家会员。一些国际大型科学出版集团,如爱思维尔(Elsevier)、威立-布莱克威尔(Wiley-Blackwell)、施普林格(Springer)、泰勒弗朗西斯(Taylor & Francis)和威科集团(Wolters Kluwer),纷纷为旗下的期刊申请成为 COPE 会员。COPE 旨在对学术出版中遇到的各种学术伦理问题提供建议和解决方法,从 2001 年起每年召开学术诚信专题研讨会^[6]。2007 年和 2010 年分别在墨西哥里斯本和新加坡召开的第一届和第二届国际学术诚信会议,吸引了世界各地的科学家、大学校长、研究机构负责人以及编辑出版者积极参会^[7,8]。

此外,2008 年,由国际出版链接协会(PILA)研发,与几家全球出版集团共同参与试验的一个世界范围内防止学术剽窃功能性的国际项目 CrossCheck(反剽窃文献检测系统)一举获得全球非赢利出版商同业协会(Association of Learned and Professional Society Publishers, ALPSP)所颁发的“全球最佳出版创新奖”^[9]。CrossCheck 的工作原理简单,

易于操作,只要将检测文件上传至网络,与其储备数据资源库或者互联网开放共享资源进行查比检测,最后会生成一个“相似度报告”,显示该文件与其他内容匹配的相似百分比,并以不同的颜色标出各相似的文本内容。编辑人员可根据这些信息,结合审稿意见和相似文本的具体内容,对文章是否涉及抄袭,做出相对公正客观的判断^[10]。CrossCheck 目前会员高达五万四千多名,其中不仅包含期刊,许多学术会议和书籍的出版也在使用 CrossCheck 对其内容进行检查。这些事实充分说明学术伦理是 21 世纪学术和出版界焦点问题之一^[11]。

《浙江大学学报》(英文版)在 2008 年获得国家自然科学基金资助后,申请成为国内第一家 CrossCheck 会员。在检测论文原创性方面做了一些探索性的工作,如对论文的多种学术不端现象进行解剖和分类等^[5,10,12]。工作中,我们以国际同行评审为主,并利用 CrossCheck 检测工具,对来稿的内容和语言表述进行查比。期间,我们在体会到使用 CrossCheck 的直观便捷的同时,也产生了一些疑惑:即如何对 CrossCheck 查对的相似度报告进行客观的分析?同样的相似度数据是否能得出同样的分析结果?不同的学科,甚至相同学科的不同体裁或内容,其相似度分析结果是否有不同的衡量标准?对已确定的不合理的相似度结果,如何与作者沟通并处置?

* 国家自然科学基金重点学术期刊资助项目(No. 30824802)和国际出版伦理委员会(COPE)研究基金资助项目(Dec-2010)

** 共同第一作者:林汉枫、贾晓燕

*** 通讯作者:张月红,编审

针对以上的种种疑问,本文以《浙江大学学报 B 辑:生物医学与生物技术》(英文版)(JZUS-B)为例,介绍了我们如何借助 CrossCheck 为辅助手段,考量学术论文撰写的原创性比。同时,以本刊 2010 年某时间段 138 篇通过国际同行评审意见的生物医学论文为样本,结合 CrossCheck 查比报告,进行稿源等统计分析,解析了一些投稿作者对学术伦理的认知差异。

2 CrossCheck 辅助检查的工作流程

工作中,我们执行“文前审查”与“文后跟踪”双向工作流程,重视论文撰写的原创性比,即使综述论文也应在参考大量文献的基础上,作者用自己独特的观点归纳提炼,撰写成行业引领性的论文,而不是原搬照抄,堆积文献成文。“文前审查”就是以国际审稿和 CrossCheck 辅助工具对来稿在投稿/录用前和发表前分别进行同行评议和撰写文本的检查,力求刊发具有科学性和原创性的论文,保护期刊和作者的著作权。同时,我们还在网站上开辟了论文发表后同行评议栏目,以“文后跟踪”的形式让所刊发的论文接受读者的评议。由于 CrossCheck 查比要按上传文章数目付费,为了节约成本,我们前两年多采用在国际审稿后进行 CrossCheck 的查比。同时摸索出一套 CrossCheck 辅助检查论文原创性的工作流程(见图 1)。同时,我们意识到 CrossCheck 只能作为一种辅助性的检测手段。编辑人员对查比结果,必须结合国际审稿意见和相似文本的具体内容做具体分析,以求合理地发挥它的作用。

2.1 论文投稿/录用前的初查

工作中,我们将国际同行评审和 CrossCheck 两者有力地结合,对投稿论文从学术质量和原创性两方面进行把关。对论文进行严格的国际同行评审的同时,也要通过 CrossCheck 对其写作内容和语言方面的原创性进行查比。只有当两者都符合要求,稿件才能进入录用发表的环节。其中,用 CrossCheck 查比时,编辑需要对 CrossCheck 的相似度报告,结合各个学科的不同特点进行分析判断,以求公正客观。

CrossCheck 检测报告结果有总相似度和单篇论文相似度两个不同的概念。总相似度是查比所出现的单篇论文相似度的总和;单篇论文相似度是指单篇文章中内容重合百分数,这是一个重要的参考值,它涉及文章内容和语言相似度的合理性。总相似度是一个宏观判断指标,需要进一步结合单篇相似度结果进行分析。经过对多篇论文的 CrossCheck 检查结果的统计研究,我们发现论文正文内容(不包括参考文献)总相似度在 30% 和 50% 以上是两个重要的参考值(非绝对)。总相似度超过 50% 的稿件,多数存在单篇相似度超过 15% 的现象,通过仔细对比,发现大部分内容或语言与其查对的已发表的论文重复,有一稿多投、重复发表、自我抄袭、文字抄袭、思想抄袭、图像抄袭等之嫌^[2,10,12]。此类稿件

超过了可接受的范围,因此多数予以退稿处理。总相似度低于 30% 的大部分稿件,其相似程度相对可以接受,一般可以要求作者修改后发表。但不排除也有少数稿件存在单篇论文相似度颇高的情况,多见于课题组系列论文的范畴,需要与作者交流后再做处理。而总相似度在 30% 和 50% 之间的稿件,情况比较复杂。有个别文稿单篇相似度高达 25%,其表现也各有不同,包括自我抄袭、文字抄袭、拼凑抄袭、引用抄袭等^[2,10,12]。将这些不合理的相似度报告反馈给作者,如果作者能给出合理的解释并做出适当的修改,修改稿能通过 CrossCheck 复查分析,编辑可以考虑录用,否则予以退稿处理。具体工作流程见图 1a。

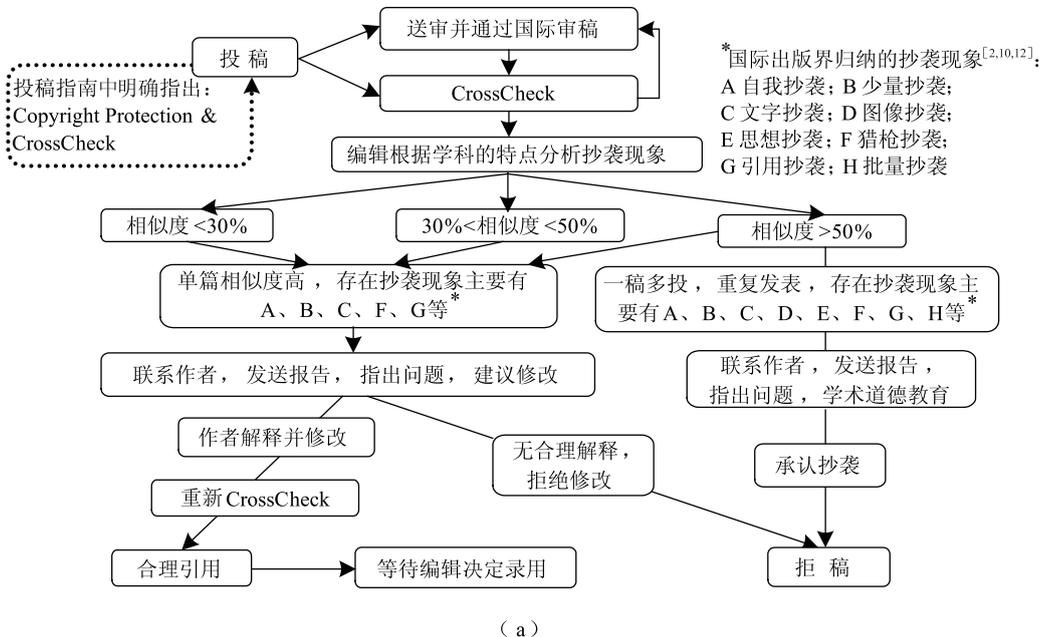
2.2 论文发表前的复查

稿件发表前需要再次进行 CrossCheck 检查,以避免因数据库更新或者时滞带来的检测数据缺失^[10,12]。如图 1b 所示,将复查报告与初查报告相对比,如果变化不大,准予接受发表,并与作者签署版权转让协议,进一步从法律上保障著作权。如果复查与初查结果相差大,相似度增加,其原因多为增加了新的高相似度论文。这种情况多为作者一稿多投、重复发表,或者由于数据库更新扩充而发现新的抄袭等。此时,我们会将产生的新报告与对方刊物作进一步确认,同时联系作者,在确认作者侵犯版权后,严正拒稿。

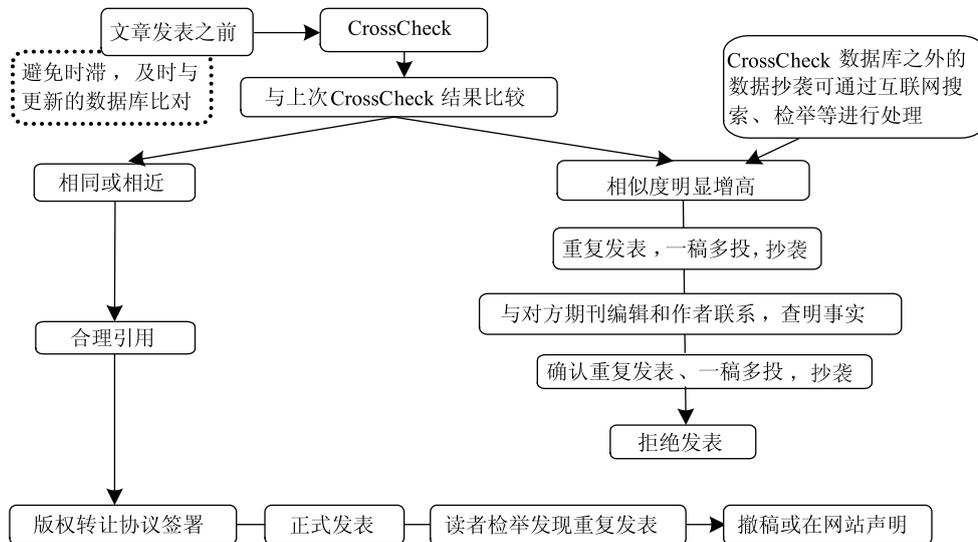
客观地讲,CrossCheck 也有缺憾。它只能对语言文字进行比对,而公式、图像以及表格等内容目前还不能进行直观有效地比对。有时会出现对所能检测到的数据资源库外内容遗漏或因“高级”抄袭技巧等原因造成的检查失真等现象。因此,我们常常需要借助其他辅助手段,以求刊发论文的真实性。如编辑会对其中的关键词句等进行 Google 搜索,时而有一些意外发现。比如,在 CrossCheck 无法访问的一个网站 <http://www.jatit.org/>,我们发现了一篇通过同行评审和 CrossCheck 查比而录用的论文,实属一稿两投。此外,也发现一篇综述论文,其后半部分 70% 的内容与别人已发表的内容一样,只是将抄袭的内容顺序重新整理,并添加了小部分新内容,而 CrossCheck 查比时却未发现这部分内容的相似。在以后的工作中,我们将进一步完善 CrossCheck 的工作程序力求将工作做细,维护科学研究的真实性。

2.3 论文发表后的追踪反馈

论文发表之前的重重把关,使大多数论文的原创新性得到了客观体现。但是,我们认为论文发表后其学术价值和原创性的体现还需要进一步的文后跟踪。于是我们在网站上开辟了论文发表后最新引用、下载和同行评议等栏目(<http://www.zju.edu.cn/jzus/rankall.php> 5),促进了读者、作者和编辑之间的互动交流。如果读者对论文的原创新性有任何疑问,或认为有思想抄袭之嫌,可以在网页上直接评论或举报。我们将对这些评论进行及时查证,一经证实学术不端,会立即将该文屏蔽撤稿,并给出郑重声明。这就是本刊 2010 年启



(a)



(b)

图1 CrossCheck 辅助检查论文撰写原创性比的工作流程

(a) 投稿/录用前 CrossCheck 初查; (b) 发表前 CrossCheck 复查

动的“文后跟踪”又一新举措。

3 生物医学论文查比中常见的问题与讨论

CrossCheck 查比应该以单篇论文的相似度为判断依据, 那么相似度在什么范围才合理呢? 这是我们工作中一直摸索的难题。事实上, 目前国际上也没有一些绝对的规范标准可参考, 要视具体情况而定。如不同学科的结构方式, 相同学科的共性表达等等。多数情况下, 我们目前以单篇论文相似度在 5% 为相对合理范畴, 当然这并不是绝对的标准, 需要与同行, 以及相关专业人士进行探讨。

如何界定生物医学论文中相关部分为合理的相似范围? 我们发现这类论文在试验材料与方法部分相似较多, 其次是引言和讨论部分。而试验结果部分, 多为作者的原创内容, 相似度较低。即使同属于相同的试验方法, 在不同的情况下也有不同的标准。正如 *The Journal of Bone and Joint Surgery* 的编辑 Cathy Griffin 谈到“以 10% 的总相似度为例, 如果由 10 处不同的 1% 相似内容组成, 每处内容都是他人独特的观点表达且没有引用, 那么这种重复不可接受; 相反, 即使仅仅是 1 处内容有 10% 的相似, 如疾病特征、传统放射检查程序, 或者行业常用的经

典方法描述等,那么这部分内容的重复是相对合理的或为允许的”^[13]。下面详细探讨我们工作中常见的典型案例,以及它们的相似度问题。

3.1 关于生物医学论文中常用的方法描述相似度问题

一般来说,如果是作者研制的一个新方法,其原创性一定显而易见。但若只是借鉴已建立的方法去做不同物种的试验,这一部分多见其相似度较高。此时,我们则建议作者用自

己的语言完整地表述,并适度地借鉴引用,而不应该照抄他人的语言为己用。但是,如果描述一些常用的标准操作程序,如聚合酶链式反应、高效液相色谱和电泳分析、统计分析方法,由于其操作是固定或者相似的,因此对其语言的描述多见大同小异,出现相对高的相似现象尚为合理。如图2所示的实时聚合酶链式反应和结果统计分析方法的描述,虽然没有出现参考引用,但是这种适度重复的语言描述是可接受的。

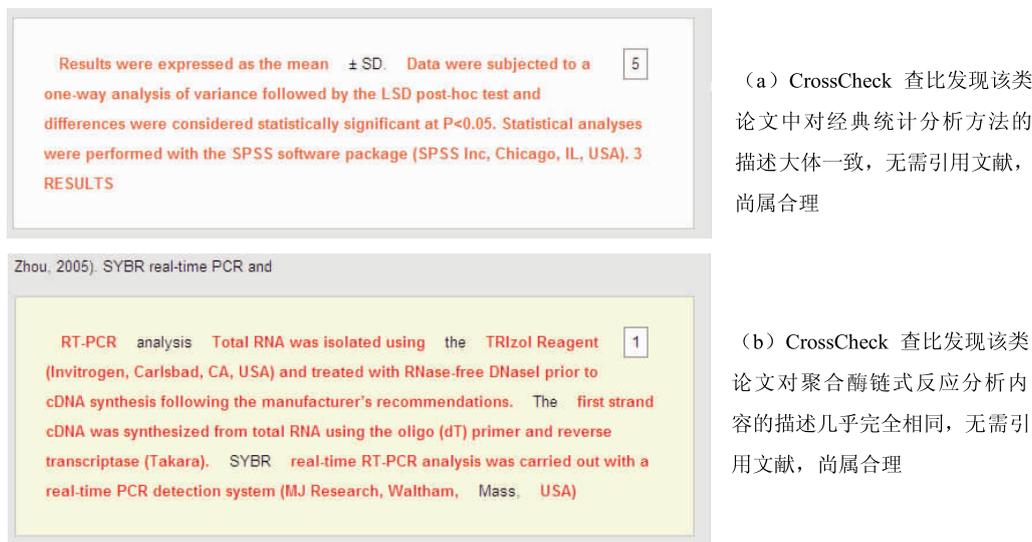


图2 常用的经典方法(以统计分析和聚合酶链式反应为例)的描述的相似尚属合理

3.2 关于系列试验方法和材料描述的相似度问题

有些作者之前已经发表过同系列试验的研究论文,所使用的材料和方法是相同或相似的,因此在描述这些内容的时候常出现较高的相似度。对这种情况,我们应当根据相似内容的性质,采取不同的措施,或保持原样,但要求作者补充参考引用,或在补充参考引用的同时建议采用简洁的语言重新描述。常见的表达方式有“... as described previously (references). Briefly, ...”, “... following the method reported previously (references) with a minor/slight modification”, “... with the method described by (references)”等。切忌大段落内容复制,以免涉及自我抄袭之嫌疑,而被同行批评。如图3的CrossCheck查比结果显示某论文的实验方法的描述和作者已经发表过的系列论文相关部分完全一致,这样的行为是不可取的。

3.3 关于他人原创方法或试验结果的引用问题

当论文中引述他人原创方法或者发现时,无论其相似度高,都应当标注引用,如图4所示,该文作者比较规范地引用了他人的方法,否则会涉及侵犯他人版权。另外,在引用注释后,是否就可以大段复制他人的方法和研究发现呢?答案是“不”。作者借用他人的语言描述时,应当把握“度”,既要引证他人的观点,又要力求有自己原创性的独到见解。

3.4 关于论文刊发前数字在线的问题

网络发达的今天,一些作者在论文未发表前,将其放到个人主页或者机构网站开放浏览,使自己的研究成果尽早地与同行交流传播。此时,使用CrossCheck检查会发现摘要甚至全文大部分的内容相似。在这种情况下,论文本身非正式发表,版权仍属于作者,与一稿多投性质不同,可以发表。但是正式发表后,我们便会要求作者将之前公开浏览的网页进行更新,以免造成作者一稿多投之嫌。

4 138篇通过国际同行审稿的论文CrossCheck查比统计分析

我们以2010年某时间段138篇已经通过国际同行评审的生物医学论文为样本,对论文的CrossCheck查比结果与稿件来源做了详细的统计分析。以相似度30%为界限,按照不同的稿件来源[浙江大学、中国大陆(除浙大)、其他国家和中国港澳台地区]进行分类。结果发现,总相似度高于30%的论文量高达50%(69/138),且稿源地区分布各异。本刊编辑曾在Nature发表通讯谈到此类问题^[14]。图5具体列出了相似度高于30%论文的稿源分布情况。其中浙江大学和中国大陆(除浙大)的稿源占据相当的比例,均为40.6%。但由于浙江大学的本身投稿量大,总相似度高于30%的论文量占自身的比例(43.8%)相对低于中国大陆(除浙大)稿源(50%)。

(Cutlip et al., 2007) 2 Materials and methods 2.1

Criteria for study selection Randomized clinical trials that compared PES with BMS for management of STEMI were included if results of more than six months follow-up were reported. 2.2 Data sources and search strategy We carried out this meta-analysis in accordance with the standard protocol recommended by the Quality of Reporting of Meta-analyses group (QUOROM) and the guidelines of Cochrane Handbook of Systematic Reviews of Interventions 4.2.6. We searched PubMed, EMBASE, ISI Web of Science, and the Cochrane Central Register of Controlled Trials for randomized trials comparing PES with BMS in STEMI patients. 3

In addition, we searched conference proceedings from the American College of Cardiology, American Heart Association, and European Society of Cardiology. 1

Relevant reviews and editorials from major medical journals published within the last three years were identified and assessed for possible information on trials of interest. Internet-based sources of information on the results of clinical trials in cardiology (<http://www.cardiosource.com>, <http://www.clinicaltrials.gov>, <http://www.clinicaltrialsresults.org>, <http://www.clinicaltrialsresults.org>). 3

outcomes and definitions The efficacy and safety outcomes were evaluated. 1

Clinical endpoints were mortality, recurrent myocardial infarction (rMI), repeat revascularization

(RR), and stent thrombosis (ST). RR included target vessel revascularization (TVR) or target lesion revascularization (TLR). TVR was defined as repeated PCI or bypass grafting of the target vessel, driven by clinical symptoms of myocardial ischemia, a positive stress test, electrocardiographic evidence of ischemic changes at rest attributable to the target vessel, or an in-lesion stenosis by visual estimate. TLR was defined as percutaneous or surgical revascularisation of the stented lesion and 5-mm segments immediately proximal and distal to the stent. Angiographic confirmation of the presence of 1

Funnel plots as well as the adjusted rank correlation test, according to the method of Begg and Mazumdar, were used to assess publication bias. 1

All P values were two-sided, and P <0.05 was considered significant. Statistical analyses were performed with Revman software version 4.2.8 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen). 3 Results 1

CrossCheck 检测到的相似论文

3 157 words / 4% - Internet from Dec 13, 2009
www.cmj.org

show in web page next match

Association IF in 2008: 0.858 Chinese Medical Journal, 2009, Vol. 122 No. 1: 88-92 • Meta analysis • LogIn/LogOut • Fulltext PDF(181K) Free • Abstract download TXT | XML • Articles in CMJ by PAN Xiao-hong ZHONG Wen-zhao • Articles in PubMed by PAN XH ZHONG WZ • Put into my bookshelf • Email to Friend • Email to author • Visit:2408 • Download:929 • Advanced Search • Related Articles • Change font size: • Cannot read some characters Clinical outcomes of serolimus-eluting stents versus bare metal stents in ST-segment elevation myocardial infarction patients: a meta-

单篇相似论文 1 和 3 属于同一被抄袭文本

1 465 words / 11% - Internet from Sep 6, 2009
www.cmj.org

next match

88 Chin Med J 2009;122(1):88-92 Meta analysis

Clinical outcomes of serolimus-eluting stents versus bare metal stents in ST-segment elevation myocardial infarction patients: a meta-analysis PAN

结论：这篇 2010 年投稿论文，其整个实验方法 (Materials and methods) (左图) 的语言描述与作者 2009 年发表在《中华医学杂志》(英文版) (*Chinese Medical Journal*) 的论文 (右图) 相关内容完全一致，因此我们认为属于不合理现象。

图 3 同系列研究论文中实验方法的描述完全相似的案例

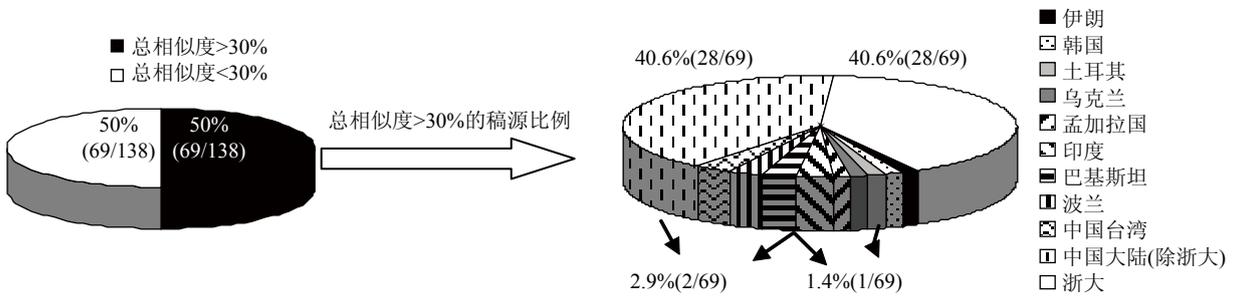
虽然其他国家和中国港澳台地区总相似度高于 30% 的论文仅占 18.8%，但本身通过国际审稿的论文总量小，因此总相似度高于 30% 的论文占自身的比例较高 (72.2%)，高于中国大陆的稿源，而且多分布在亚洲和非洲，如印度、伊朗、巴基斯坦、土耳其，尼日利亚等。由此可见我们的作者，包括我们自己，重视学术伦理，提高期刊论文的学术原创影响力刻不容缓。

control agents because they spread diseases worldwide Mendonsa et al. (1996) reported

that both first and fourth instar larvae of mosquito Aedes aegypti can be killed within 48 hours with the help of a crude preparation from M. verrucaria. Though 100% mortality had been observed 2

in 48 hours, but this was for the

图 4 引用他人独创方法或试验结果的案例



| 稿源分布 | 通过国际同行 评审的论文数 | 总相似度>30%的论文 | |
|--------------|------------------|-------------|------|
| | | 论文数 | 百分数 |
| 浙大 | 64 | 28 | 43.8 |
| 中国大陆(除浙大) | 56 | 28 | 50.0 |
| 其他国家及中国港澳台地区 | 18 | 13 | 72.2 |

图5 2010年JZUS-B的138篇论文通过国际同行评审后的CrossCheck查比统计分析

5 建立客观的规范标准,争取国际话语权

今天,遵守学术道德伦理、纯化学术环境、规范科技论文的发表,是我们科技期刊编辑的义不容辞的责任。目前,我们已采用CrossCheck等现代反抄袭工具来辅助查对期刊论文,力求弘扬科研的原创成果,杜绝学术不端行为。需要注意的是,应用CrossCheck比和分析时,相似度报告数据仅作为一个参考值,而不是一个标准。应该视具体情况而定,以求做到客观判断。2010年12月,我们首次申请COPE的研究基金课题“对不同学科典型抄袭案例的分析指南(CrossCheck Guidance: An Analysis of Typical Cases of Plagiarism in Different Disciplines)”,已经通过国际评委的严格评审立项,荣获了2011年度的首项基金资助项目(<http://publicationethics.org/research/projects-funded>)。这是COPE自2008年设立研究基金以来,全球资助的第6项课题,也是非欧美国家的期刊编辑首次荣获国际出版伦理委员会(COPE)的研究资助项目^[15-17]。我们希望通过大量的案例调查,研究建立一套国际上相对客观的CrossCheck查比规范,为科技期刊编辑和作者提供一些参考数据和标准,同时,这也是中国编辑在国际出版行业争取话语权的时机,我们一定努力做!

参考文献

- 1 OSA Board of Editors. From the Board of Editors: On Plagiarism. *Journal of Optical Networking*, 2005, 4(3):142-143
- 2 Errami M., Garner H. A tale of two citations. *Nature*, 2008, 451(24):397-399
- 3 Long TC., Errami M., George AC., Sun ZH., Garner HR. Responding to possible plagiarism. *Science*, 2009, 323(6):1293-1294
- 4 Jaschik S. Plagiarism prevention without fear. 2010. [2010-12-08].

- 5 张月红. 重视学术伦理是学术界和出版界共同的责任. 科学时报. [2011-01-04]. <http://news.sciencenet.cn/htmlnews/2011/1/242310.shtm?id=242310>
- 6 Committee on Publication Ethics (COPE). [2010-12-10]. <http://publicationethics.org/about>
- 7 1st World Conference on Research Integrity. [2010-12-20]. <http://www.esf.org/index.php?id=4479>
- 8 2nd World Conference on Research Integrity. [2010-12-20]. <http://www.weri2010.org/index.asp>
- 9 ALPSP Awards 2008-Winners. [2010-05-06]. http://www.alpssp.org/ngen_public/article.asp?aid=19852
- 10 林汉枫,张欣欣,翟自洋,伍秀芳,张月红. 遏止学术不端行为 保护科研原创成果——《浙江大学学报(英文版)》作为CrossCheck中国第一家会员的实践和体会. 中国科技期刊研究, 2009, 20(4):580-584
- 11 张月红. 透过几种现象看学术不端. 科学时报. [2009-05-07]. <http://news.sciencenet.cn/htmlnews/2009/5/219040.html?id=219040>
- 12 Zhang Helen. CrossCheck: an effective tool for detecting plagiarism. *Learned Publishing*, 2010, 23:9-14
- 13 Cathy GRIFFIN. The Journal of Bone & Joint Surgery's CrossCheck experience. *Learned Publishing*, 2010, 23:132-135
- 14 Zhang Helen. International, not "campus", please. *Nature*, 2010, 467:789
- 15 COPE Grant Awarded for First Time to Recipients from China. *Ethical Editing*, 2011, 3(1):3
- 16 COPE Projects Funded. [2010-12-16] <http://publicationethics.org/research/projects-funded>
- 17 张月红. 中国编辑首获国际出版伦理委员会研究基金资助. 科学时报. [2011-03-01]. <http://news.sciencenet.cn/sbhtmlnews/2011/3/241945.html>