

期刊影响因子的历史和意义

Eugene Garfield¹⁾著 冷怀明²⁾编译

(1) Thomson Scientific 3501 Market St Philadelphia PA 19104 USA; 2) 第三军医大学学报编辑部, 400038, 重庆)

1955年,我第一次在《科学》杂志上提出了影响因子的概念,接着在美国国家健康研究所的支持下,尝试性地出版了遗传学引文索引,继而在1961年出版了科学引文索引(SCI)。H. Irving和我创立期刊影响因子来帮助选择增加源期刊。我们认识到那些为数不少且被大量引用的期刊应该收录到科学引文索引中。如果仅依靠期刊的发文量来选择期刊,一些发文量小的期刊将不可能被收录;所以,我们创立了期刊影响因子(JIF)。

1 影响因子的概念及相关因素

期刊的影响因子基于2个因素:1)分子,该刊前2年发表的论文在统计当年被引用的次数;2)分母,该刊前2年发表的实质性论文和述评的数量。影响因子可以仅基于前1年的论文发表数量,这样能更好地估计迅速变化的领域的权重。影响因子应当考虑更长的统计周期和更广的统计范围,但是这样做的时效性差。

“影响因子”逐步被用来描述期刊和作者的影响。期刊影响因子一般涉及大量的发表论文数和引用次数。在《期刊引证报告》(JCR)出版以前,我们尝试对1969年的SCI期刊按影响因子排序。如今,《JCR》包括5000多种期刊中的每种期刊引用情况(每年大约有来自100万篇文献源的1500万条引文)。影响因子的精度是个疑问,但报道影响因子的前3位数能减少同样影响排位的期刊数量。

影响期刊影响因子的因素很多,其中参考文献著录错误较为普遍,但大多数错误并不影响期刊的影响因子。不过,参考文献中期刊名缩写错误对计算影响因子的影响较大。

期刊引证报告报道影响因子实际上显示《BMJ》《JAMA》《柳叶刀》《新英格兰医学》杂志等期刊的栏目设置分类清晰。这些期刊大量出版一些在引文分析中认为没有实质性内容的文章,如通讯、信函、注释、观察、新闻故事、讣告、社论、采访、颂词等。这些文章可能被引用,特别是在发表当年。正因为如此,它们对影响因子的计算影响不大。

2 科学计量学和期刊学

引文分析已经盛行了40多年。现在已有国际科

学计量学和情报计量学协会。BMJ的前编辑 Stephen Lock恰当地将文献计量学在期刊发展中的应用命名为“期刊学”。

所有的引文研究都应当考虑如专业性、引文密度、半衰期等可变因素(变量)。引文密度是每篇源文献引用参考的平均数。数学类期刊的引文密度明显低于分子生物学类期刊。生理学期刊的半衰期比物理学期刊的半衰期长。对于某些领域,期刊引证报告采用2年期限计算影响因子不如采用5年或10年期限反映全面,然而,当按类研究期刊时,基于1年、7年或15年的影响因子没有显著差别。当研究跨领域期刊时,随着统计年的累加,生理学期刊的排名会提高,但同一种类期刊中,排名变化不大。与此相似,Hansen和Henriksen报告:临床生理学和核医学论文的累计引用频次和期刊的影响因子之间有较好的一致性。

3 学科范围与引用密度

人们普遍认为,期刊面向的科学社会范围的大小对期刊的影响因子影响较大。这种假设忽略了这样一个事实:大量的作者产生了大量的引用,但要由大量的被引用的论文来分享。大多数文章较少被引用,但某些文章可能有不寻常的跨学科影响。

众所周知,大多数领域的引用分布是不平衡的,也就是“二八”现象,即20%的论文占了80%的引用。影响因子的关键因素不是某个领域论文或作者的数量,而是引文密度和被引文献的年龄;然而,一个领域的大小将决定“超级被引”论文的数量。一些经典的方法学论文会超越引用的高阈值,而其他成千上万的方法学和述评论文则不然。发表普普通通的综述文章很难扩大期刊的影响。某些超级引用的经典如Lowry方法,被引用30万次,E. M. Southern发明的Southern印迹技术,被引3万次,粗略估计,有60余篇被引1万次以上的论文都是几十年前发表的。它们不影响当前影响因子的计算。事实上,从1900—2005年被引用的3800万篇论文中仅有0.5%的论文被引200次以上。一半的论文根本没有被引用过。大约1/4的论文是不真实的。

引用偏差是显而易见的,被引用是会延迟的。引用论文不在期刊引证报告的2年内将不被计算在影响

因子内。在同一期中的同一主题的论文可能会有正面的影响。为了更加精确,有必要逐条审核期刊的编辑条目,以至不同类型期刊栏目的影响因子的差别都能够考虑到。

有些编辑对那些高引用的论文情有独钟,以抵消其他的低影响因子论文的影响;另一些编辑则希望按地理或语言因素排名次,因为 SCI 有语言偏爱,尽管 SCI 包括了欧洲的医学期刊(绝大部分是德语、法语和西班牙语期刊)。

对影响因子的其他异议牵涉到期刊引证报告采用的分类编排的体系。汤姆生科学(原汤姆生 ISI)使用的方法并不完善。Pudovkin 和我最近试图客观地聚合期刊。我们依靠期刊之间的相互引用关系来聚合期刊,以减少期刊名称的主观影响,如《实验医学》杂志——前 5 名的免疫学期刊之一。

最近《JCR》增加了一项新的功能——基于期刊之间引证关系的密切程度编排期刊目次。例如《Controlled Clinical Trials》杂志和《JAMA》的关系远比我们第一印象中的更为密切。同样,我们发现(2004 年的数据)《New England Journal of Medicine》杂志是发表心血管研究成果的主要期刊之一。

4 期刊绩效显示器

在被称为期刊绩效器(JPI)的另一个汤姆生科学数据库中,期刊影响因子的许多固有的缺陷被消除了。不像《JCR》《JPI》数据库链接每一条自己引用的源文献条目;因此,影响计算更加精确,仅有那些实质性引用条目才能包括在分母内。这样可以获得较长时间段的累计影响测量。例如,1999 年《JAMA》发表文章的累计影响是 84.5,这得益于 1999 年出版的 370 篇论文在 1999—2004 年被引用了 3 万 1 257 次,而 1999 年《JAMA》发表了 1 905 篇文章,其中 680 篇是通讯,253 篇是社论。引用这些文章不计算在《JPI》内。

期刊的影响因子除了帮助图书馆员购买期刊以外,还有助于作者确定期刊投稿的方向。一般来说,高影响因子的期刊具有较高的声誉。大家将较高的声誉与高影响等同。

用期刊的影响因子代替论文的实际引用次数来评价作者或论文的质量存在较大争议。基金组织和其他政策制定机构希望有别的办法获取单篇论文和作者的引用数;但最近发表的论文可能没有足够的时间被引用,所以尝试用期刊的影响因子作为替代评价工具。

一篇论文被一高影响因子的期刊接受发表,暗示了论文的质量较高。

今天,网络计量学的作用正在增强,尽管没有什么事实证明这种方法比传统的影响因子更好。网上“点击”可能发生得早一点,但它并不等同于引用;因此,我们必须对新发表的论文被读者阅读或下载与被实际引用加以区分。但是,有研究显示:网上点击是将来引用的先兆。

最近发表论文的影响不能在 SCI 中被评价的假设是不完全正确的。尽管某些主题论文可能有几年的延迟,但那些影响较强的论文通常在发表后几个月至 1 年内才被引用。这种即时性使得汤姆生科学能识别“热点论文”并发表在其双月刊《Science Watch》中;然而,其高影响的完全确认一般在论文发表 2 年后。《科学家》杂志用 2 年时间来收集作者的评价之后选择热点论文,其中大多数最终成为“引用经典”。

最近《JAMA》发表的 2 例热点论文是关于绝经后妇女补充雌激素的益处和危险。第 1 篇论文在发表后 6 个月被引用了 132 次,在 2003 年被引用了 776 次,2004 年被引用了 862 次。第 2 篇是最近发表的,已经被引用了 300 次。

5 结论

关于影响因子的相冲突的观点,Hoeffel 简洁表述为:影响因子并不是评价论文质量的完美工具,但目前也没有比这个方法更好的方法,它的存在有其合理性;因此,它仍然是科学评价的好方法。经验表明,每一专业最好的杂志,其影响因子也是较高的,论文被其接受也是最难的。这些杂志的大多数在影响因子出现以前就已出版很久了。影响因子作为论文质量的测量器之所以被普遍接受,是因为它符合了我们的观点,即每个专业都有各自最好的期刊。

用期刊的影响来评价单篇的论文或某个作者存在固有的缺陷和危险。理想的做法是评价者阅读每篇文章来做出自己的判断;但是,最近的生物医学出版同行专家审稿国际会议研讨表明,要使评审者意见趋于统一相当困难。大多数的评价者都没有时间来阅读所有的相关论文。即使大家愿意那么做,他们的判断也会被他们看到的引文注释所左右。

(原文来自: Eugene Garfield The History and Meaning of the Journal Impact Factor, JAMA, 2006 295: 90-93)

(2006-0-收稿;2006-06-10 修回)