应慎重使用期刊的影响因子评价科研成果

任胜利① 王宝庆② 郭志明③ 金碧辉④

(①国家自然科学基金委员会科学基金杂志部，北京 100083；②中国科学杂志社，北京 100717；③中国科学院出版图书情报委员会，北京 100864；④中国科学院文献情报中心，北京 100080. Email: rensl@rose. nsfc.gov.cn)

摘要 通过对ISI数据库中科技期刊和论文的被引证情况分析，认为期刊的影响因子受时间、研究领域、SCI源期刊库及参与统计的引文条目等因素的制约，进而提出科研成果评价中应慎重使用期刊的影响因子。

关键词 影响因子 SCI 科研成果评价

1955年美国科技信息研究所(Institute for Scientific Information, ISI)所长Garfield首次阐明可通过论文的被引证频次来测度期刊的影响[1]，并在1963年ISI出版的“1961年度科学引文索引”中正式提出和使用影响因子(Impact Factor，IF)这一术语. 自1975年开始，ISI在“科学引文索引”(Science Citation Index, SCI)的基础上每年发行上一年度世界范围的“期刊引证报告”(Journal Citation Reports, JCR)[2]，从而使得期刊的定量评价成为可能.

随着科学引文数据的积累，英国、荷兰、意大利等国于80年代中期开始尝试使用SCI来评价科学研究工作，并因此引起广泛的关注和讨论[3～11].

我国自1989年开始有组织地分析和整理SCI的相关资料[12]，并于90年代中期开始倡导利用SCI来评价科研工作[13]. 由于缺乏对SCI足够的研究和认识，目前，在我国少部分高校和研究机构已出现片面强调SCI及由其所派生的影响因子在科研评价中的作用，如有些单位在对科研人员进行考评时，硬性规定必须有若干篇论文发表于国外的、影响因子为某数值以上的期刊上. 这种对SCI统计数据不加区分地使用在很大程度上引导我国科技工作者把自己的科研成果投向为数不多的高影响因子刊物，一方面可能因为论文和刊物的主题相关性较差而降低了投稿的命中率；另一方面，即便稿件被录用刊出，也可能由于刊物与作者研究领域的差异而影响到论文在同行中的传播范围，进而可能影响到论文的被引证频次.

1 影响因子的影响因素

影响因子是SCI中最重要的参数，它决定了各期刊在JCR中的排序和级次. 期刊在某年的影响因子是指该年引证该刊前两年论文的总次数与前两年该刊所发表的论文总数之比. 影响因子的计算方法导致其因学科的不同而产生较大差异. 下面试从影响因子的计算方法及不同研究领域的特点等方面来讨论制约影响因子的几个主要因素.

1.1 时间因素

在进行影响因子计算时，期刊被引频次中两年的时间限制可导致不同刊物中论文的被引证次数有较大的差异. 主要表现为：

(1) 出版时滞较短的刊物更容易获得较高的影响因子. 这是因为对于研究成果相同或相似的论文来说，首先被公开发表的更有可能引起较大的影响或被引证，并且，由于引证行为

中的“因循守旧”性(Conformism)，这些论文会被更多地引证或转引[8]. 尤其是对于研究成果时效性较强、持续时间较短的热点研究领域的论文，出版时滞显得更为重要. 如统计表明，美国临床医学方面的论文有 83%在发表的当年即被美国的科学家引证(包括自引)[14]，这与美国该学科论文的出版时滞大多只有几周，且临床医学的研究热点持续时间大多较短，无疑有密切的联系.

此外，由于影响因子的计算中只统计相应刊物的新近两年中的论文被引频次，若相关刊物中论文的出版时滞较长，则相当一部分的引文因为文献老化(超过两年)的原因而不能被统计参与影响因子的计算，从而降低了影响因子值.

(2) 不同学科论文的引证行为有所不同. 这是因为不同研究领域或研究主题的成果在完善或验证过程中经历的时间段可能很不相同，如地球科学的研究成果多具区域性特点，且印证时常需要一段时间的野外考察或实验室工作才能得出结论，因而该学科的论文需经过较长的时间方可获得影响；但对于分子生物学或天文学方面具创新研究成果的论文来说，则有可能很快地开展进一步的实验观察研究，因而很快引起较大影响并被大量引证.

1.2 源期刊库的缺陷

由于 SCI 仅对其所收录的期刊进行数据采集和评价，而人力、财力等诸多因素对参与统计的期刊数目有很大的限制，这就不可避免地造成 SCI 的源期刊库存在以下几方面的主要缺陷：

(1) 收录的期刊数过少. 如 1996 年 SCI 和 SSCI(Social Science Citation Index)收录的期刊数约 4 500 种[2]，而世界上总期刊数则约有 126 000 种[6]，即世界上只有约 3.6%的期刊被 ISI 的源期刊库收录，显然，该期刊库很难全面、公允地反映不同国家、不同学科、不同语种期刊的情况.

(2) 各国被 SCI 收录的期刊数极不平衡. 如 1997 年 SCI CD ROM 选用的 3 449 种期刊中，美国、英国、荷兰和德国被收录的期刊数目分别为 1 392、732、339 和 266 种，计约占总刊数的 80%，其中在该年度被收录的中国刊物有 9 种，只占总刊数的 0.2%[12]. 而同国科学家之间因研究成果传播的快捷性、研究主题的相关性等种种因素而倾向于相互引证. 如美国科学家之间的相互引证可提高美国论文被引证频次的 30%[6]. 显然，这在一定程度上导致了不同国家期刊影响因子的差异.

(3) 被收录期刊的语种分布极不平衡. 研究表明，同语种刊物间相互引证的概率较大[6,15]，而 SCI 倾向于收录英文刊，这就使得 SCI 源期刊库中为数不多的其他语种期刊的影响因子相对较低. 如 Bauin 等人[15]研究发现，1986~1987 年 SCI 收录法国植物基因研究方面的论文中约有 15%是以法文形式发表的，这些论文的平均被引证次数为 0.58，而另外约 85%的以英文形式发表的论文的平均被引证次数为 4.33. 可见被 SCI 收录论文的语种比率和不同语种论文被引证频次的比率具有较相似的倍数关系. 这种语种分布不平衡的情况在 SCI 收录的中国论文中表现得更为突出，如 1997 年 SCI 收录的 10 033 篇中国论文中只有 10 篇是以中文形式发表的[12]，这对于提高中国论文的被引证频次显然是一个很不利的因素. 可以认为，

目前 SCI 数据库中期刊的组成十分有利于美、英等以英语为母语国家的刊物获得高影响因子，相反，对于其他国家的非英文刊物来说，获取高影响因子的难度则相对较大.

1.3 引文条目统计的差异

在影响因子的计算中，引证次数(分子)统计了相应刊物中所有论文被引证的总次数，而刊载论文数(分母)则只统计论文(Articles)、简讯(Notes)和评述(Reviews)类栏目的文章数，对评论(Editorials)、来信(Letters)、通讯(Correspondence)和其他一些常被引证的栏目的文章数目则不进行统计. 实际上，这些未被统计部分的被引证频次对影响因子的贡献很大[2, 5, 7]. 例如，1992 年 The Lancet 的“正确”影响因子(即论文、报告和评述类论文的实际平均被引证率)比 JCR 公布的要低 43%[5]；又如，Moed 等人[5]对 1994 年 SCI 源期刊中 130 种化学类期刊影响因子的“正确”计算表明，在该年度的 JCR 中，按影响因子排序的化学类期刊中有 89.2%的刊物排列次序有变化，其中变动次序为 10 以上的刊物达 4.6%.

1.4 学科性质的差异

不同学科间由于发展速度和成熟度的不同，期刊基数的差异很大，从而致使 SCI 收录的不同学科的期刊数目差别很大(即 SCI 所检索的不同学科的论文数和引文数相差较大)，而相同或相近研究领域的论文倾向于相互引证，这又反过来在很大程度上导致不同学科间刊物的影响因子产生较大差异. 如 1996 年 SCI 和 SSCI 的约 4 500 种期刊中，大约有 3 000 种与生物医学有关[2]，显然，这是 SCI 中生物医学方面刊物的影响因子普遍较高的主要原因. 又如，1997 年生物化学和分子生物学类刊物中最高影响因子值为 40.782，而数学类刊物中最高影响因子值仅为 2.070；该年度影响因子为 1.0 的刊物在生物化学和分子生物学类刊物中位于第 194 位，在数学类刊物中则位于第 7 位[16].

2 期刊的影响因子与论文被引证频次之间的关系

研究表明[2,3,6]，期刊的影响因子并不能反映其中每篇论文的实际被引证频次. 如 Seglen 通过分别统计 J Biol Chem，Biochem J 和 Biochem Biophys Acta 的被引证情况后发现[3]，约 15%

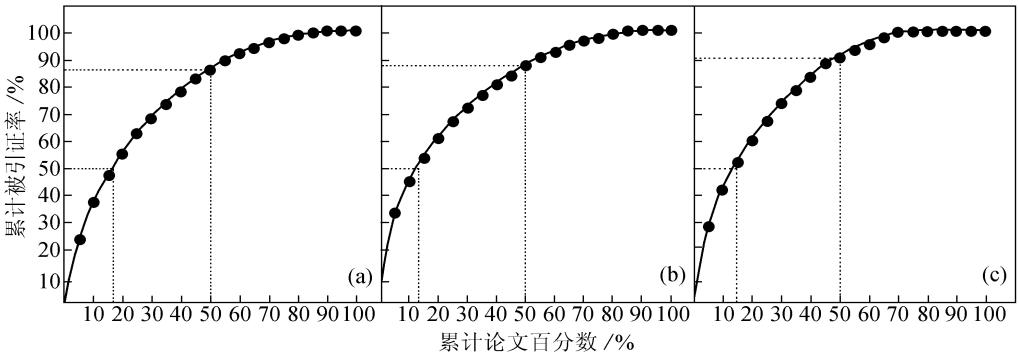


图 1 1984 年 3 种生物化学类刊物中论文在 1987 年的累积被引证频次分布图[3]

(a) J Biol Chem, IF=6.4, 统计的论文数为 403; (b) Biochem J, IF=3.8, 统计的论文数为 447; (c) Biochem Biophys Acta, IF=2.0, 统计的论文数为 323

的论文贡献了刊物总被引证频次的约 50%，约 50%的论文对总被引证频次的贡献高达 90%(图 1)，即刊物中 50%被引频次较高的论文的平均被引证次数大约是其余 50%论文的 10 倍. 这表明，好的刊物并不能提高一般水平论文的被引频次，即论文相对于刊物来说，不存在“免费搭车”(free ride).

研究表明[2,6,7]，同一刊物中不同论文间被引频次的差异大体由以下几种原因引起：(ⅰ)论文类型的差异：即评述类论文和某些介绍新方法论文的被引证频次通常高于研究成果类论文[2]；(ⅱ)论文性质的差异：即有争议的论文或具广泛兴趣的论文的被引证频次常高于普通论文[7]；(ⅲ)论文所涉及研究领域的差异：这种差异一方面表现为快速发展的或较新研究领域的论文比相对较成熟研究领域的论文的被引频次高；另一方面表现为不同研究领域间的相互引证并不等价,如基础医学方面的研究论文常引证临床医学方面的论文，反之则较少见[3].

不同研究领域间论文的被引证频次的差异甚至在同一学科内部也表现得十分明显. 如 Seglen[3]研究发现，尽管 1984~1987 年 J Biol Chem 的平均影响因子为 6.93，但该刊物中不同研究领域的论文平均被引证频次则分别为：钙相关研究 14.06，受体 13.57，蛋白激酶和磷酸酶 6.92，线粒体 5.21，血浆蛋白质 5.00，其他酶 4.74，细菌 4.30，植物 1.71.

可见，在对具体的科研工作者或研究群体进行评价时，不仅要考虑其论文所在刊物的影响因子值，同时也要综合考虑具体论文的被引证频次和学科特点.

3 讨论

综上分析，可将 SCI 中期刊和论文的被引证特点总结如下：(ⅰ)影响因子的计算与论文的学术水平并无直接联系，即论文被引证次数的多少实际上说明的是论文的被利用频次，其本身并不能反映论文的学术水平；(ⅱ)期刊影响因子的高低与其所属的学科或研究领域密切相关，不同学科间期刊的影响因子没有可比性；(ⅲ)同一刊物中不同论文的被引次数差别很大，论文的被引证频次决定了期刊的影响因子，但论文并不因为发表于高影响因子的刊物上就一定获得较高的被引频次.

可以认为,期刊影响因子的高低较客观地反映了期刊和编辑吸引最好论文的能力的高低[2]. 一般说来，高影响因子的刊物具较高的“显示度”(High Visibility，即刊物更可能具较广泛的读者群)和较高的“有用性”(High Utility，即刊物被使用的概率较大). 并可由此推断，高影响因子的刊物吸引相关学科中优秀论文的能力较强，并因此导致其稿源相对较多，进而使其对所录用论文的筛选更为严格(即稿件可能要经过更为严格的同行评议). 因此，高影响因子的刊物相对同类的低影响因子的刊物来说，所刊载的相同研究领域的论文应当具更高的学术质量，尽管这种质量上的差别常难以绝对地加以定量化.

需要特别强调指出的是，目前 SCI 数据库对美、英、荷兰等以英语为母语或常用语国家的期刊和论文反映得较全面，而对于非英语国家的刊物和论文的反映是极为不够的. 如中国 1997 年的 4 300 余种科技期刊中只有 9 种入选 SCI CD ROM，并且，1997 年 SCI 收录的 10 033 篇中国论文中有 10 023 篇是以英文形式发表的[10](占中国被收录的论文总数的 99.9%). 也就是说，中国科学家的绝大多数优秀科研成果很可能不能为中国的同行首先获悉和引证. 因此，

我国在使用 SCI 评价科研成果时，更要充分认识到这种差别.

由于影响因子只是对期刊的大致评估，因此它在科研成果的评价中应结合研究者所担负的研究课题、科研成果产出、国际合作、研究论文的被引频次、研究者所从事专业学科的特点和同行评议或评审小组的评议等使用. 在此过程中最重要的应当是同行评议或评审小组的综合评议结果，正如 SCI 的创始人 Garfield[2]所说: “SCI 的发明者告诫人们对这些数据不要不加区分地使用，…(SCI 数据的使用)尽管很方便，但很危险. 虽然期刊的评估很重要，但对科研人员的评价更重要，因为其影响到个人的职业生涯. 影响因子不应作为人的评价的替代物.”

致谢 本工作为国家自然科学基金(批准号: 79870014)和中国科学院文献情报研究基金资助项目.

参 考 文 献

1 Garfield E. Science indexes for science: a new dimension in documentation through association of ideas. Science, 1955, 122: 108~111

2 Garfield E. How can impact factors be improved? British Medical Journal, 1996, 313: 411~413

3 Seglen P O. Evaluation of scientists by journal impact. In: Weingart P, Sehringer R, Winterhager M, eds. Representations of Science and Technology. Leiden: DSWO Press, 1992. 240~252

4 van der Meulen B. Indicators in a framework of judgment and control——the use of science indicators by Dutch Evaluation Committees. In: Weingart P, Sehringer R, Winterhager M, eds. Representations of Science and Technology. Leiden: DSWO Press, 1992. 57 ~ 74

5 Moed H F, van Leeuwen Th N. Impact factors can mislead. Nature, 1996, 381: 186

6 Seglen P O. Why the impact factor of journals should not be used for evaluating research. British Medical Journal, 1997, 314: 498 ~ 502

7 Mestecky P. Impact Factors——handle with care. Materials Today, 1998, (1): 8 ~ 12

8 West R. Impact factors need to be improved. British Medical Journal, 1997, 313: 1401

9 Lowy C. Impact factor limits funding. The Lancet, 1997, 350(4): 1035

10 Schwarz A W, Schwarz S, Tijssen R J W. Research and research impact of technical university——a bibliometric study. Scientometrics, 1998, 41: 371 ~ 388

11 Phelan T J. A compendium of issues for citation analysis. Scientometrics, 1999, 45: 117~136

12 中国科技信息研究所. 1989~1997 年中国科技论文统计与分析(年度研究报告)

13 师昌绪，田中卓，黄孝瑛，等. “科学引文索引”——国际上评定科研成果的一种方法. 科学通报，1997，42(8)：888~893

14 Braun T, Glanzel W, Grupp H. The scientometric weight of 50 nations in 27 science areas, 1989~1993. Part II：Life sciences. Scientometrics, 1996, 34: 207 ~ 237

15 Bauin S, Rothman H. “Impact” of journals as proxies for citation counts. In: Weingart P, Sehringer R, Winterhager M, eds. Representations of Science and Technology. Leiden: DSWO Press, 1992. 225 ~ 239

16 Institute for Scientific Imformation. Science Citation Index Reports——A Bibliometric Analysis of Science Journals in the ISI Database. 1997

(1999-08-13 收稿，1999-11-22 收修改稿)