

“非可被引文献”的引证特征及其对科技期刊影响因子的贡献*

刘雪立^{1,2)} 盖双双^{2,3)} 张诗乐^{2,3)} 周晶^{2,3)}

1) 新乡医学院期刊社; 2) 河南省科技期刊研究中心; 3) 新乡医学院管理学院; 453003, 河南新乡

摘要 介绍可被引文献和“非可被引文献 (non-citable document, NCD) 的概念。对 NCD 的引证特征进行统计分析, 证实了 NCD 是可被引用的, 甚至还可能有极高的被引频次。通过文献计量学分析, 探讨 NCD 对期刊影响因子的贡献度。

关键词 可被引文献; 非可被引文献; 科技期刊; 影响因子

Citation characteristics of non-citable documents and their contribution to journal impact factor // LIU Xueli, GAI Shuangshuang, ZHANG Shile, ZHOU Jing

Abstract The concepts of citable document and non-citable documents (NCD) are introduced in this paper. Statistical analysis of citation characteristics of NCD are then carried out in detail. We hold that NCD can also be cited, and even may have high citation frequency. Through bibliometric analysis, we discuss the contribution of NCD to journal impact factor.

Keywords citable document; non-citable document; sci-tech journal; impact factor

First-author's address Henan Research Center for Science Journals, 453003, Xinxiang, Henan, China

众所周知, 影响因子是指某期刊前 2 年发表的全部文献在统计当年被引用的总次数除以该刊前 2 年发表的可被引文献 (citable document) 数^[1-2], 而非可被引文献数仅包括论文 (article) 和综述 (review)^[3-4], 其他许

多类型的文献是不被统计在内的, 我们把这部分文献称为“非可被引文献” (non-citable document, NCD)。NCD 是相对于可被引文献而言的, 这些文献依然有可能被频繁引用, 如《Hepatology》2011 年发表的一篇述评^[5] 目前已被引用 1 025 次, 《Nature Methods》2011 年发表的一篇信稿^[6] 被引用达 1 093 次。

最近, 武夷山^[7] 根据 Heneberg^[8] 的一项研究提出了一条重要的建议: 影响因子的定义最好调整一下。他认为, 既然“不可引文献”其实是可引的, 影响因子定义中的分母就应该改为全部文献而不仅仅是可被引文献。Heneberg 文中数据表明, 某些期刊不可引文献已经异化为人为提升影响因子的手段。我们前期对《Nature》及 10 种国际权威期刊影响因子构成分析的研究表明, 不同期刊 NCD 对其影响因子贡献有较大差异, 介于 0 ~ 15% 之间^[9-10]。

早在 1995 年, Moed 等^[11] 就建议改善影响因子计算的准确性。他通过对大量 SCI 期刊的实证研究得出结论, 认为 ISI 的 JCR 中多数期刊影响因子的计算是不准确的, 尤其是个别高影响因子期刊。不准确的主要原因是“对可被引文献”不合理的界定。他以《The Lancet》和《Nature》为例说明影响因子计算的不合理

[5] 科技期刊怎么活? [EB/OL]. [2014-11-05]. http://news.xinhuanet.com/newmedia/2012-10/30/c_123887382.htm

[6] 王登峰, 崔红. 心理卫生学 [M]. 北京: 高等教育出版社, 2003

[7] 柯进. 2013 中国科技论文含金量几何 [EB/OL]. [2014-11-05]. http://www.jyb.cn/world/gjgc/201410/t20141014_601002.html

[8] 刘雪立. 中文核心期刊评价指标体系: 演进·问题·建议 [J]. 编辑学报, 2014, 26(1): 92

[9] 顾秀瑛, 丁玉松, 冯刚玲, 等. 新疆沙湾县哈萨克族居民营养健康状况调查 [J]. 中国公共卫生, 2014, 30(9): 1117

[10] 潘曙东, 庄勋, 成鹏, 等. 常州市五年制高职在校生网络信息成瘾现状及相关因素 [J]. 中国学校卫生, 2014, 35

(7): 1011

[11] 吕迎春. 科技期刊编辑心理健康状况调查分析 [J]. 山东精神医学, 2005, 18(4): 238

[12] 张秀红, 吴琼. 辽宁省学术期编辑心理健康状况调查分析 [J]. 辽宁师范大学学报 (自然科学版), 2011, 34(2): 243

[13] 张勋, 栾奕, 库雪飞, 等. 医学期刊编辑心理健康状况调查研究 [J]. 中国当代医药, 2012, 19(10): 175

[14] 刘腾飞, 栾奕. 编辑工作者心理健康状况研究现状 [J]. 中国煤炭工业医学杂志, 2011, 14(6): 912

[15] 吴红艳, 李春华, 白发秀. 科技期刊编辑的心理健康状况与个性特征分析 [J]. 编辑学报, 2003, 15(4): 306

[16] 丁仲元. 忻州市男性出租车司机体质状况现状调查及对策研究 [J]. 宜春学院学报, 2010, 32(4): 90

[17] 王金道. SCL-90 量表使用的现状及检测心理健康的异议 [J]. 中国心理卫生杂志, 2004, 18(1): 41

* 国家自然科学基金 (15BTQ061); 河南省高等学校哲学社会科学基础研究重大项目 (2015-JCZD-013); 新乡医学院科学研究培育基金 (人文社科类) 重点项目 (2013ZD201)

性:假设仅考虑可被引文献的被引次数计算《The Lancet》的影响因子,那么其影响因子将下降40%;如果将 Letter 计入影响因子的分母,《Nature》杂志的影响因子将下降30%。

为了进一步认识 NCD 对期刊影响因子的影响,这里,我们探讨 NCD 的引证特征,并计算不同期刊 NCD 对影响因子的贡献。

1 研究方法

1.1 NCD 的界定 NCD 是指 ISI 定义的可被引文献以外的所有文献。早期的文献^[12-13]中把 Article、Review 和 Note 定义为可被引文献,而近期的文献^[3,14]中仅把 Article 和 Review 定义为可被引文献。事实上,早在 1997 年后,Web of Science 数据库中已经检索不到 Note 类型的文献,可能是已合并到 Article。本文中所指的 NCD 包括述评(editorial)、信稿(letter)、重印文献(reprint)、科技新闻(news item)、书评(book review)、人物传记(biographical item)和更正(correction)等。

1.2 文献数据的获取

1.2.1 不同类型文献量 登录 Web of Science 数据库,检索 2011—2012 年的全部文献,对文献类型进行“精炼”,获取述评、信稿、重印文献、科技新闻、书评、人物传记和更正等各类文献数量。检索日期为 2014 年 12 月 10 日。

1.2.2 文献被引量和被引文献量 通过 Web of Science 数据库的“创建引文报告”功能,获取不同类型文献在 2013 年的被引量及截至检索日期的被引文献量,记录不同类型文献单篇最高被引量。篇均被引量 = 文献被引量/文献量,文献被引率 = 被引文献量/文献量。

1.2.3 h 指数的确定 通过观察由数据库创建的引文报告,确定各类型文献 2013 年的 h 指数(即根据 2011—2012 年论文在 2013 年的被引频次确定的)。

1.2.4 NCD 对期刊影响因子的贡献 通过对述评和信稿 2 类文献的分析,确定 2011 和 2012 年发表述评和信稿 ≥ 800 篇和被引频次 ≥ 20 次述评和信稿至少 20 篇以上的期刊,计算 NCD 对各期刊影响因子的贡献值和贡献率。计算方法如下:

$$F_{1,NCD2013} =$$

$$\frac{\text{2011—2012 年某期刊 NCD 在 2013 年的被引频次}}{\text{该刊 2011—2012 年可被引文献数}};$$

$$R_{NCD2013} = \frac{F_{1,NCD2013}}{F_{I2013}}。$$

式中: $F_{1,NCD2013}$ 为某期刊 NCD 对该刊 2013 年影响因子的贡献值; $R_{NCD2013}$ 为某期刊 NCD 对该刊 2013 年影响因子的贡献率; F_{I2013} 为该刊 2013 年的影响因子。

2 结果

2.1 NCD 的引证特征 2011—2012 年各类型 NCD 的引证特征见表 1。2 年中,SCI 数据库收录的 NCD 数量最多的是述评,达 15 万 8 031 篇,其次是信稿(8 万 415 篇),而后是科技新闻和更正(都超过 2 万篇),其他文献如重印文献、传记和书评相对较少。从引证效率来讲,述评和信稿的篇均被引量和文献被引率也是最高的,重印文献、科技新闻和更正篇均被引量和文献被引率明显低于述评和信稿,传记和书评引证效率极低。各类文献 2 年内 h 指数及单篇最高被引量与篇均被引量和文献被引率相比,呈现出较为一致的文献计量学特征。

表 1 2011—2012 年各类型 NCD 的引证特征

文献类型	文献量 ¹⁾	被引量 ²⁾	最高被 引量 ²⁾	篇均 被引量	h 指数	被引文 献量 ³⁾	文献 被引率
Editorial	158 031	94 591	1 025	0.599	76	61 172	0.387
Letter	80 415	36 997	1 093	0.460	52	28 560	0.355
Reprint	459	129	12	0.281	7	119	0.259
News item	36 879	6 436	372	0.175	33	4 436	0.120
Correction	23 530	2 885	241	0.123	20	3 271	0.139
Biographical item	6 453	226	11	0.035	6	381	0.059
Book review	6 212	79	11	0.013	4	155	0.025
合计	311 979	141 343	—	0.453	—	98 094	0.314

注:1) 文献量是指 2011 和 2012 年 2 年文献量之和;2) 被引量和最高被引量均指相应类型文献在 2013 年的被引量;3) 被引文献量是指相应类型文献发表后截至检索日期被引量 ≥ 1 的文献数量。

2.2 被引频次最高的述评和信稿 从表 1 可以看出,述评和信稿是 NCD 中引证效率最高的 2 类。那么,这 2 类文献被引频次究竟能高到什么程度?单从篇均被引量是无法表征的;因此,我们列出了这 2 类文献被引频次最高的前 10 篇及其被引频次(被引频次均指 2011—2012 年的文献在 2013 年的被引频次,都参与了相应期刊影响因子的计算),见表 2 和表 3。2011—2012 年部分述评和信稿在 2013 年的被引频次相当高,对相应期刊影响因子的贡献是不可忽视的。

2.3 述评和信稿对部分期刊影响因子的贡献 为了准确认识述评和信稿对期刊影响因子的贡献,我们从 SCI 数据库筛选出 2011—2012 年发表述评和信稿 ≥ 800 篇或者发表的被引频次 ≥ 20 次的述评和信稿 20 篇以上的期刊,详细统计各期刊论文、综述、述评、信稿和其他类型文献的文献数及被引频次,计算各期刊述评和信稿对其影响因子的贡献值和贡献率(见表 4)。

符合规定条件的期刊共 24 种,其中医学期刊 17 种(70.8%),还有一定数量的生物学期刊。不少期刊在国际上具有较高的知名度,如《Nature》《Science》、《New England Journal of Medicine》(N Engl J Med)、《Lancet》、《Journal of the American Medical Associa-

tion》(JAMA)、《British Medical Journal》(BMJ)、《Archives of Internal Medicine》《Circulation》等。

表2 2011—2012年SCI数据库中被引频次最高的10篇述评文献

第一作者	缩略题名	来源期刊	机构	国家	出版年	被引频次
Bruix	Management of Hepatocellular Carcinoma; An Update	Hepatology	Univ Barcelona	Spain	2011	1 025
Ghany	An Update on Treatment of Genotype 1 Chronic Hepatitis C	Hepatology	NIDDK	USA	2011	485
Harriss	Update - Ethical Standards in Sport and Exercise Science Research	International Journal of Sports Medicine	Liverpool John Moores Univ	England	2011	483
Zhou	to Metal-Organic Frameworks	Chemical Reviews	Texas A&M Univ	USA	2012	426
Roger	Executive Summary: Heart Disease and Stroke Statistics……	Circulation	Anonymous	Anonymous	2012	317
Llovet	EASL-EORTC Clinical Practice Guidelines……	Journal of Hepatology	EASL	Switzerland	2012	312
Levey	The definition, classification, and prognosis of chronic kidney disease	Kidney International	Tufts Med Ctr	USA	2011	308
Yizhar	Optogenetics in Neural Systems	Neuron	Stanford Univ	USA	2011	258
Ross	The 2011 Report on Dietary Reference Intakes for Calcium……	Journal of Clinical Endocrinology & Metabolism	Penn State Univ	USA	2011	257
Cerami	The cBio Cancer Genomics Portal	Cancer Discovery	Mem Sloan Kettering Canc Ctr	USA	2012	251

表3 2011—2012年SCI数据库中被引频次最高的10篇信稿

第一作者	缩略题名	来源期刊	机构	国家	出版年	被引频次
Petersen	SignalP 4.0: discriminating signal peptides	Nature Methods	Tech Univ Denmark	Denmark	2011	1 093
Robinson	Integrative genomics viewer	Nature Biotechnology	Massachusetts Inst Technol & Harvard	USA	2011	656
Darriba	2: more models, new heuristics and parallel	Nature Methods	Univ Vigo	Spain	2012	511
Huang	Heritable gene targeting in zebrafish	Nature Biotechnology	Peking Univ	China	2011	191
Tesson	Knockout rats generated by embryo microinjection	Nature Biotechnology	INSERM	France	2011	182
Sander	Targeted gene disruption in somatic zebrafish cells	Nature Biotechnology	Massachusetts Gen Hosp	USA	2011	167
Prinz	Believe it or not: how much can we rely	Nature Reviews Drug Discovery	Bayer HealthCare	Germany	2011	156
Howarth	Methane and the greenhouse-gas footprint	Climatic Change	Cornell Univ	USA	2011	149
Mackenzie	A harmonized classification system for FTLT-TDP	Acta Neuropathologica	Univ British Columbia	Canada	2011	148
Araujo	Climate change threatens European conservation areas	Ecology Letters	CSIC	Spain	2011	122

2年间发表述评和信稿最多的期刊是《BMJ》,达到了3 169篇。述评和信稿文献数超过1 000篇的还有《Lancet》《New Scientist》《N Engl J Med》《Nature》《Journal of Urology》《Science》《JAMA》《International Journal of Cardiology》《Oil Gas Journal》《Neurology》《Annals of Thoracic Surgery》《Circulation》等。

《New Scientist》的NCD(包括述评、信稿和其他NCD)对其影响因子的贡献率最大,达到57.9%,共有4种期刊NCD对其影响因子的贡献率超过了30%,分别是《New Scientist》《International Journal of Cardiol-

gy》《BMJ》《Medical Journal of Australia》。另外有5种期刊,其NCD对影响因子的贡献率介于20%~30%。《Science》和《Nature》的述评和信稿数量较多,被引频次也很高;但对影响因子的贡献率并不高,分别为4.7%和7.2%(NCD对其影响因子总的贡献率分别为6.1%和9.1%),与这2刊2012年影响因子的构成相比,述评和信稿对其影响因子的贡献率(4.8%和7.3%)^[9-10]相当。

《N Engl J Med》《Science》《Nature》这3刊NCD的被引频次超过4 000,分别为4 961、4 849、4 377次,还有

表4 述评和信稿类文献对部分期刊2013年影响因子的贡献值和贡献率

刊名	EL 论文数 ¹⁾	Article		Review		Editorial	
		文献数	被引次数	文献数	被引次数	文献数	被引次数
New Scientist	2 531	412	16	0	0	755	11
International Journal of Cardiology	1 412	626	1 649	73	251	48	85
British Medical Journal	3 169(27) ⁵⁾	441	4 743	106	495	1 946	1 611
Medical Journal of Australia	948	345	880	5	13	539	282
Nature Biotechnology	319(30)	163	4 858	13	309	215	279
Archives of Internal Medicine	884(10)	283	2 645	26	490	357	318
Plastic and Reconstructive Surgery	975	870	2 398	32	118	219	115
Journal of The American Medical Association	1 444(61)	417	9 124	35	1 358	705	2 048
Nature Methods	389(24)	269	5 694	10	308	332	183
Oil Gas Journal	1 225	593	27	0	0	1 200	3
New England Journal of Medicine	2 441(157)	630	30 460	79	2 685	907	3 459
Lancet	2 673(40)	548	18 737	41	857	1 575	1 781
Clinical Infectious Diseases	519(22)	875	6 911	46	574	256	627
Journal of Clinical Oncology	903(24)	1 232	15 956	89	1 240	464	943
Journal of the American College of Cardiology	940(18)	779	10 045	83	11 66	601	745
Neurology	1 154	985	7 119	39	473	952	719
Science	1 551(115)	1589	43 230	115	5 102	1 133	3 626
Circulation	1041	1 116	14 943	58	569	723	695
Nature	2 263(81)	1 639	61 533	71	5 546	1 703	2 891
Critical Care Medicine	841	624	3 444	49	389	569	207
Journal of Urology	2 077	1 218	4 209	41	197	1 917	106
Neurosurgery	970	843	2 072	67	204	817	42
Angewandte Chemie International Edition	187(36)	4 004	37 569	225	7 137	180	753

刊名	Letter		其他		影响因子 ²⁾	贡献值 ³⁾	贡献率 ⁴⁾
	文献数	被引次数	文献数	被引次数			
New Scientist	1 776	0	1 223	11	0.092	0.053	0.579
International Journal of Cardiology	1 364	1 074	494	3	4.381	1.662	0.379
British Medical Journal	1 223	177	3 366	595	13.932	4.356	0.313
Medical Journal of Australia	409	112	133	1	3.680	1.129	0.307
Annals of Thoracic Surgery	269	40	212	753	4.202	1.174	0.279
Nature Biotechnology	104	1 140	189	190	38.500	9.142	0.237
Archives of Internal Medicine	527	540	20	0	12.922	2.777	0.215
Plastic and Reconstructive Surgery	756	184	148	367	3.528	0.738	0.209
Journal of The American Medical Association	739	499	972	219	29.310	6.119	0.209
Nature Methods	57	1 070	24	5	26.022	4.509	0.173
Oil Gas Journal	25	2	15	0	0.054	0.008	0.156
New England Journal of Medicine	1 534	1 395	169	107	53.746	6.997	0.130
Lancet	1 098	712	305	60	37.601	4.334	0.115
Clinical Infectious Diseases	263	324	108	3	9.163	1.036	0.113
Journal of Clinical Oncology	439	261	10 981	905	14.614	1.597	0.109
Journal of the American College of Cardiology	339	257	5 982	363	14.589	1.584	0.109
Neurology	202	32	5 267	46	8.192	0.778	0.095
Science	418	222	2 255	1 001	31.210	2.846	0.091
Circulation	318	32	8 671	336	14.118	0.905	0.064
Nature	560	455	1 269	1 031	41.787	2.560	0.061
Critical Care Medicine	272	36	1 263	2	6.059	0.364	0.060
Journal of Urology	160	10	2 342	89	3.662	0.163	0.044
Neurosurgery	153	24	124	26	2.602	0.101	0.039
Angewandte Chemie International Edition	7	18	122	20	10.758	0.187	0.017

注:1)指述评(Editorial)和信稿(Letter)文献数;2)是根据WOS数据库引证分析计算的影响因子,与JCR报告的影响因子有一定误差;3)指NCD贡献的影响因子值;4)指NCD对期刊影响因子的贡献率;5)括号内数字是该刊发表的被引频次 ≥ 20 的述评和通信文献数。

另外4种期刊NCD的被引频次超过2000次,分别是《JAMA》《Lancet》《BMJ》《Journal of Clinical Oncology》。这7种期刊中,《Nature》和《Science》是综合性期刊,其他均为医学期刊。虽然《Nature》和《Science》的NCD被引频次都在4000次以上,但二者的NCD对各自期刊影响因子的贡献率均不超过10%,其影响因子主要还是论文和综述贡献的。而另外5种期刊NCD对期刊影响因子的贡献率都超过10%,《BMJ》甚至达到31.3%。

3 结论

1)虽然ISI把述评、信稿、重印文献、科技新闻、更正、书评、传记等界定为NCD,在计算期刊影响因子时这些文献也不计入分母,但事实上这些文献也是可以引用的,甚至有些述评和信稿被引用量还较大。

2)在NCD中,述评和信稿对期刊被引频次及影响因子的贡献明显优于其他文献。在计算2013年影响因子有效的2年引证时间窗口内,述评和信稿的最高被引量、篇均被引量、h指数、被引文献量及文献被引率等都有较好的表现,是提升期刊影响因子和影响力不可忽视的文献类型。

3)不少期刊发表了大量的NCD,例如:2011—2012年《Nature》发表的论文和综述共1710篇,同期发表的述评和信稿就达到2263篇;《Science》同期发表的述评和信稿为1551篇,略少于其发表的论文和综述文献。发表述评、信稿等NCD较多者主要集中于医学期刊,这些期刊多是国际著名的、有较高影响力的期刊,如《N Engl J Med》《Lancet》《JAMA》《BMJ》等。因此,我们是否可以这样认为,文献类型的多样化是学术期刊走向成熟的重要标志,是提升学术期刊影响力的必然选择。

4)多数期刊的NCD对其影响因子的贡献是正常的,只有个别期刊NCD对期刊影响因子贡献率较高。我们选择的研究对象是发表述评和信稿较多的或发表的述评和信稿被引频次较高的期刊,因此,这些期刊的NCD对影响因子的贡献率相对较高。不排除还有其他期刊,虽然发表述评和信稿数量不多,被引频次也不是很高;但由于其论文和综述被引频次也很少,其NCD对期刊影响因子的贡献率也很高。

4 参考文献

- [1] Campanario J M. The effect of citations on the significance of decimal places in the computation of journal impact factors [J]. *Scientometrics*, 2014, 99(2): 289
- [2] 刘雪立. 基于Web of Science数据库预测SCI期刊影响因子的方法[J]. *科技与出版*, 2014(2): 87
- [3] Citrome L. How we rate: is impact factor the most important measure? [J]. *International Journal of Clinical Practice*, 2013, 67(9): 819
- [4] Zupanc G K H. Impact beyond the impact factor[J]. *Journal of Comparative Physiology*, 2014, 200(2): 113
- [5] Bruix J, Sherman M. Management of hepatocellular carcinoma: an update[J]. *Hepatology*, 2011, 53(3): 1020
- [6] Petersen T N, Brunak S, von Heijne G, et al. SignalP 4.0: discriminating signal peptides from transmembrane regions [J]. *Nature Methods*, 2011, 8(10): 785
- [7] 武夷山. 影响因子的定义最好调整一下[EB/OL]. [2015-01-02]. <http://blog.sciencenet.cn/blog-1557-806325.html>
- [8] Heneberg P. Parallel worlds of citable documents and others: inflated commissioned opinion articles enhance scientometric indicators[J]. *Journal of the Association for Information Science and Technology*, 2014, 65(3): 635
- [9] 盖双双, 刘雪立, 张诗乐. SCI来源期刊影响因子预测和结构分析方法: 以Nature杂志为例[J]. *中国科技期刊研究*, 2014, 25(8): 980
- [10] 刘雪立. 10种国际权威科技期刊影响因子构成特征及启示[J]. *编辑学报*, 2014, 26(3): 296
- [11] Moed H F, Vanleeuwen T N. Improving the accuracy of institute for scientific informations journal impact factors [J]. *Journal of the American Society for Information Science*, 1995, 46(6): 461
- [12] Jones A W. Mode of classification of source material as citable items skews journal impact factor calculations[J]. *Scandinavian Journal of Clinical & Laboratory Investigation*, 2005, 65(7): 623
- [13] Campanario J M, Gonzalez L. Journal self-citations that contribute to the impact factor: documents labeled "editorial material" in journals covered by the Science Citation Index [J]. *Scientometrics*, 2006, 69(2): 365
- [14] Borrmann L, Marx W. The effect of several versions of one and the same manuscript published by a journal on its journal impact factor[J]. *Scientometrics*, 2012, 92(2): 277

(2015-01-07 收稿; 2015-02-10 修回)