

即时开放获取论文在时间和地理空间上的使用优势分析 ——以《新英格兰医学期刊》为例

■张光耀 姜春林 王贤文

收稿日期: 2019-05-11

修回日期: 2019-06-26

大连理工大学科学与科技管理研究所暨 WISE 实验室, 辽宁省大连市甘井子区凌工路 2 号 116024

摘要 【目的】从时间和地理空间两个视角比较即时开放获取(Open Access, OA)论文和延时 OA 论文的浏览量差异,以探索 OA 是否在时间和地理空间维度上都存在优势。【方法】选取 2011—2018 年发表在《新英格兰医学期刊》(*The New England Journal of Medicine*, NEJM) 中类型为 Article 的 2338 篇论文为研究对象,包括 135 篇即时 OA 论文与 2203 篇延时 OA 论文,从 nejm.org 收集论文使用数据进行文献计量和统计检验分析。【结果】OA 会增加论文的浏览量,促进知识扩散,扩大读者的地域范围,特别是使发展中国家的学者获益。延时 OA 政策虽然能够在一定程度上增加论文浏览量,但仍然无可超越即时 OA 论文在刚发表后迅速累积的浏览优势。【结论】即时 OA 是促进科学知识传播、提高论文影响力,进而提升期刊影响力的最有利方式,实现即时 OA 对增强期刊影响力和扩大期刊影响范围有着积极意义。

关键词 使用数据; 即时 OA; 延时 OA; OA 优势

DOI: 10.11946/cjstp.201905110360

开放获取(Open Access, OA)是学术界和出版界为推动科研成果在公共互联网免费传播和获取而采取的行动,其目的在于促进科学交流,提高科研效率,提升科研成果的公共利用程度。相对于传统的学术出版模式,OA 论文具有更短的出版周期,减轻了科研机构的财政压力,为科研人员提供了更多获取学术资源的机会^[1]。从兴起到现在,OA 运动不断引起关注并持续向纵深发展,已经成为不可阻挡的大趋势,其意义不仅在于知识的开放和共享,更重要的是它或许将改变科学研究范式和知识交流生态。欧洲研究理事会和 12 个国家的主要研究机构和资助者于 2018 年 9 月启动了 OA 出版计划——Plan S,要求从 2020 年起,由公共资助的研究产生的科学出版物必须在 OA 期刊或平台上发布。虽然 OA 促进了科研成果传播,但是在 OA 期刊上发表论文是否存在引用优势仍存在争议,从而影响了学者向 OA 期刊投稿的积极性^[2]。而且 OA 直接冲击到专业学术期刊的商业利益,并会增加版面费从而引起稿源流失,成为出版商推动 OA 进程的直接阻碍^[3]。

已有大量证实 OA 优势的研究主要基于发文和引文视角,如 Eysenbach^[4]对发表在 *Proceedings of*

the National Academy of Sciences of the United States of America(PNAS)的 OA 和非 OA 论文进行文献计量分析;Hua 等^[5]对 PubMed 肿瘤学引文数据进行统计分析;Yan 等^[6]使用 CiteScore 作为期刊影响力指标,对 OA 和非 OA 期刊的影响力差异进行了检测。这些研究都在一定程度上证明了 OA 能加速研究成果的传播和吸收。国内学者如武庆圆^[7]采用科学计量学方法对 OA 期刊的知识交流行为进行分析,以验证 OA 期刊知识交流的有效性;陈娟^[8]采用引文分析法、多元统计法分析 Scopus 和 Web of Science(WoS)收录的 OA 期刊分布情况,发现 Scopus 和 WoS 收录的各学科 OA 期刊的影响因子均低于非 OA 期刊,而 OA 期刊在影响因子的提升速度上更具有优势。

在 OA 对知识在地理空间上的扩散和对发展中国家获取科学资源的影响的研究上,Evans 等^[9]使用根据论文引用建立的数据集,测算了 OA 对低收入国家科学传播的影响,发现 OA 论文对发展中国家的影响是发达国家的两倍。Tang 等^[10]通过比较 OA 论文与同一期刊上发表的非 OA 论文,测算了 OA 对 *Journal Citation Reports*(JCR)中 46 种生态学期刊载

基金项目: 国家自然科学基金(71673038);中央高校基本科研业务费项目(DUT18JC32)。

作者简介: 张光耀(ORCID: 0000-0002-0505-1160),博士研究生, E-mail: sdgyzhang@163.com;姜春林,副教授,所长。

通信作者: 王贤文(ORCID: 0000-0002-7236-9267),博士,教授,教育部青年长江学者, E-mail: xianwenwang@dlut.edu.cn。

《中国学术期刊(光盘版)》电子杂志社编者注: 本文中涉及台湾的“国家”均应为“国家(地区)”,“countries”均应为“countries (regions)”。

文被引频次的影响,发现 OA 论文相对于非 OA 论文获得了更多引用,这种引用优势在北美、欧洲、亚洲、非洲和大洋洲的国家均存在。Ezema 等^[11]以 African Journals Online(AJOL) 中的 65 种 OA 和 69 种非 OA 的医学健康类期刊为研究对象,对两类期刊的 h 指数、被引频次、篇均被引频次进行分析,发现 OA 期刊相对于非 OA 期刊均存在显著的优势,从而证实 OA 对非洲的医学学术交流具有重要的影响。

在互联网技术高速发展的当下,OA 呈现出更加旺盛的发展态势,同时,论文使用数据和 Altmetrics 数据的开放大大拓展了 OA 的研究空间。相比于传统引文数据的滞后性,使用数据有着即时、量大的特点,而且可以在动态时间模式下进行对比分析。如 Alperin^[12]使用 SciELO 和 Altmetric.com 中的数据构建数据集,通过测度发展中国家和地区的论文覆盖率来反映 Altmetrics 在世界范围的变化情况;Wang 等^[13]将 Altmetrics 数据进一步拓展到论文使用和社交媒体层面,OA 论文具有更高的下载量;王真等^[14]跟踪记录 PLoS Biology 和 PLoS Medicine 2016 年 11 月发表的 38 篇研究论文的浏览量数据,利用统计方法探索了 PLoS 平台 OA 学术论文在网络媒体中浏览量的累积规律。本研究根据论文使用数据,从时间和地理空间视角探索 OA 的优势,为 OA 优势提供新的研究视角。

在获取《新英格兰医学期刊》(The New England Journal of Medicine, NEJM) 独特的论文计量数据(包括每篇文章每天更新的浏览量和地理分布)后,可以从时间和地理空间的角度来审视 OA 的优势。本研究主要关注以下两个问题:(1) 从地理空间角度来看,OA 论文的地理分布是否比非 OA 论文更广? OA 是否会显著扩大读者的地域范围,尤其是能否让发展中国家的读者受益?(2) 从时间的角度来看,非 OA 论文和 OA 论文的下载量的差距是否随时间扩大?特别是当论文在发表 6 个月后由非 OA 转为 OA,差距是否会缩小?

1 数据与方法

1.1 期刊的选取

本研究选取 NEJM 在 2011—2018 年间发表的研究论文(Article)作为研究对象。NEJM 是 Massachusetts Medical Society 出版的医学周刊。自 1812 年创刊以来,它是最有声誉的经同行评审的医学期刊,也是最早连续出版的期刊之一。

<http://www.cjstp.cn>

本研究选取 NEJM 发表的论文作为研究对象有以下 3 个原因:

(1) NEJM 提供自 2010 年 7 月以来所有文章的 Metrics 数据(图 1),包括浏览量(Page Views)、被引频次(Citations)、媒体报道(Global Media Coverage)、社交媒体(Social Media)。浏览量包括数量和地理分布,因此,可以研究论文在地理维度上的扩散情况。更重要的是,Metrics 还通过交互式全球地图显示浏览量的地理分布,这使得 Metrics 数据非常独特,可以从空间角度研究 OA 优势。

(2) 在发表 6 个月后,NEJM 在 nejm.org 免费提供所有研究论文的全文,因此,可以从时间角度来评估 OA 前后论文浏览量的变化。

(3) NEJM 发表的论文既包括即时 OA 论文也包括延时 OA 论文(延时 OA 论文是指期刊完全按照传统出版模式出版后,出版方在一定时间后将期刊发表的过刊文献在互联网上公开,读者可以免费获取,如 NEJM 在出版 6 个月后将非 OA 论文转为 OA 论文),因此可以对二者进行比较分析。



图 1 NEJM 论文的 Metrics 数据页面

(数据来源于 <https://www.nejm.org/doi/metrics/10.1056/NEJMoa1714284>)

1.2 数据获取

选择 2011 年 1 月 6 日至 2018 年 12 月 13 日发表的论文作为研究对象,共获得 2338 篇论文,包括 135 篇即时 OA 论文和 2203 篇延时 OA 论文。本研究收集了每个研究样本的 Metrics 数据,包括浏览量及其地理分布数据。

2 分析结果

对数据进行统计分析,从国家分布、浏览量变化、即时 OA 与延时 OA 对比、即时 OA 与延时 OA 论文的日均浏览量变化等方面呈现分析结果。

2.1 国家分布

本研究统计了每个国家对所有论文的浏览量,并在表 1 中列出了浏览量最多的前 30 个国家和地区。美国以绝对优势位居第一。在 30 个国家和地区中,发达国家与发展中国家或地区(在表 1 中以

表1 NEJM 论文浏览量最高的30个国家和地区

| 位次 | 国家和地区 | 浏览量 | 位次 | 国家和地区 | 浏览量 | 位次 | 国家和地区 | 浏览量 |
|----|-------|----------|----|-------|---------|----|-------|---------|
| 1 | 美国 | 68559456 | 11 | 西班牙 | 3372110 | 21 | 印度尼西亚 | 1306706 |
| 2 | 中国 | 9396228 | 12 | 印度 | 2899847 | 22 | 以色列 | 1205300 |
| 3 | 日本 | 7511684 | 13 | 墨西哥 | 2418425 | 23 | 秘鲁 | 889027 |
| 4 | 英国 | 5557526 | 14 | 哥伦比亚 | 2281177 | 24 | 爱尔兰 | 862015 |
| 5 | 加拿大 | 5167137 | 15 | 泰国 | 2248705 | 25 | 瑞典 | 858016 |
| 6 | 法国 | 4502411 | 16 | 中国台湾 | 2081379 | 26 | 比利时 | 855589 |
| 7 | 德国 | 4457772 | 17 | 荷兰 | 2060717 | 27 | 奥地利 | 774686 |
| 8 | 巴西 | 3758636 | 18 | 瑞士 | 1924389 | 28 | 伊朗 | 758502 |
| 9 | 澳大利亚 | 3576717 | 19 | 韩国 | 1861162 | 29 | 新加坡 | 757301 |
| 10 | 意大利 | 3441060 | 20 | 阿根廷 | 1651843 | 30 | 土耳其 | 714488 |

下划线标示)的比例为3...2,前10名中只有中国和巴西为发展中国家。

从所有国家和地区的论文浏览量分布来看,北美、西欧、东亚和大洋洲等经济较为发达地区的论文浏览量较高,而除巴西以外的南美洲国家和非洲、中亚等发展中国家和地区的浏览量较低。

2.2 每年论文的统计

2.2.1 浏览量统计(绝对值)

根据发表年份对2203篇延时OA论文进行分类,并分析论文浏览量的国家分布,结果如图2所示。美国的数据绘制在主(左)轴上,而其他国家的

数据绘制在次(右)轴上。由图2可知:美国在所有年份的论文浏览量上一直保持绝对领先;中国对发表在2018年的论文的浏览量相对较低,对2015年前发表的论文的浏览量较高,从2011—2015年发表论文的浏览量来看,中国的浏览量大幅超过大多数国家。

发达国家和发展中国家的论文浏览量的年度分布情况如图3所示。对于2016—2018年发表的论文,发达国家与发展中国家之间的浏览量存在明显差距,而2011—2014年间的浏览量差距明显小于2015—2017年间的浏览量差距。

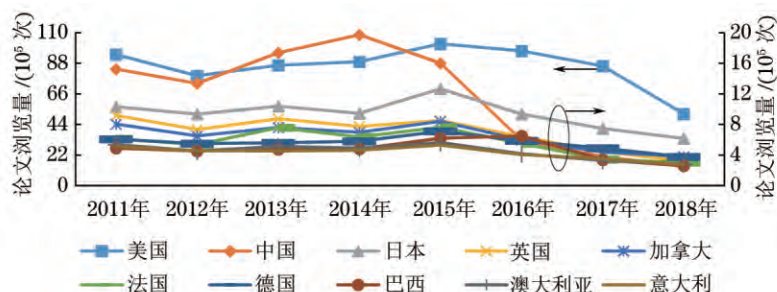


图2 前10位国家在不同年份的论文浏览量分布

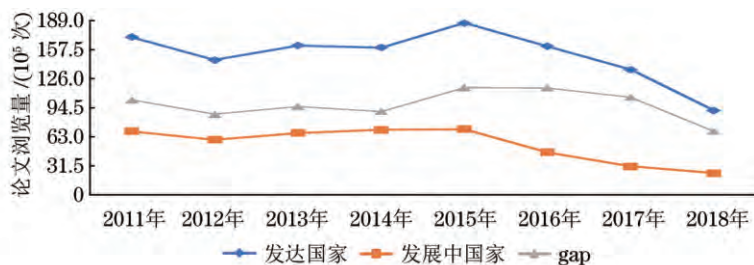


图3 发达国家和发展中国家不同年份的论文浏览量分布

2.2.2 浏览量统计(相对值)

本研究计算了前10位国家的论文年浏览量与所有国家的年浏览量之和的比值。在图4中,美国的数据绘制在主(左)轴上,而其他国家的数据绘制在次(右)轴上。美国仍在所有年份一直保持着绝对领先的优势,但随着论文发表时间的向前推移,美

国、日本和加拿大等发达国家的比例呈现持续下降趋势,而中国的比例呈明显上升趋势。

图5所示为发达国家和发展中国家浏览量相对量的变化。为了更好地显示和比较浏览量绝对值和相对值(比例)的变化趋势,以浅色标记绝对值(论文浏览量),并以深色标记相对值(比例)。相对值

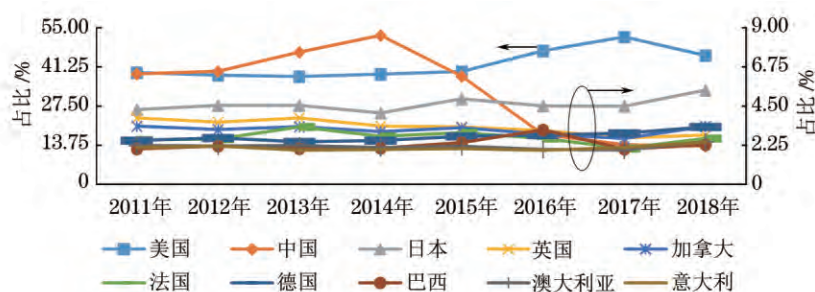


图4 前10位国家不同年份的论文浏览量分布

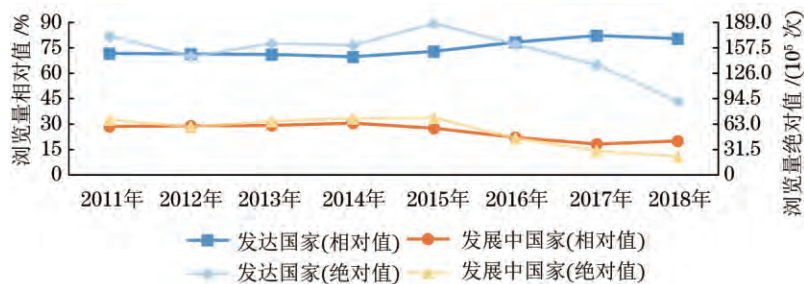


图5 发达国家和发展中国家在不同年份的浏览量分布

的数据绘制在主轴上,而绝对值的数据绘制在次轴上。对于发达国家,绝对值的变化趋势与相对值的变化趋势有很大的不同。随着论文发表时间的向前推移(2018年到2011年),发达国家的相对值呈明显下降趋势,而其他三条曲线的变化趋势几乎相同。值得注意的是,发达国家的浏览量占比在所有年份中始终保持在70%以上。

2.2.3 国家浏览量排名变化

图6和图7所示分别为发达国家与发展中国家和地区的浏览量排名变化。发展中国家对2011—2014年发表论文的浏览量排名高于2015—2018年的排名;美国始终排名第1,日本在2011—2014年排名第3,在2015—2018年排名第2,中国在2011—2014年排名第2,但在2015—2018年的排名明显下降,哥

伦比亚2011年排在第12位,但2018年排在第20位。德国、瑞士、荷兰、以色列和其他发达国家在2015—2018年的排名均高于其在2011—2014年的排名。

2.3 即时OA与延时OA

本研究选择论文浏览量、每篇论文的浏览国家数2个指标来比较即时OA和延时OA论文的影响差异。根据最初的出版模式为即时OA或延时OA,将2338篇论文进行分组。由表2可知,即时OA论文的平均值和中位数明显高于延时OA论文,每篇论文的浏览国家数也高于延时OA论文。数据不符合正态分布,因此,使用非参数检验(Wilcoxon检验)对即时OA和延时OA的论文浏览量和浏览国家数进行检验, P 值分别为0.003和0.018,可以证明二者的差异具有统计学意义。

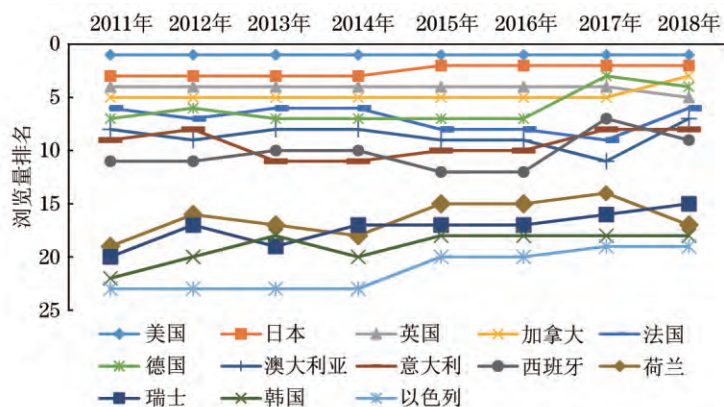


图6 发达国家的年度浏览量排名变动

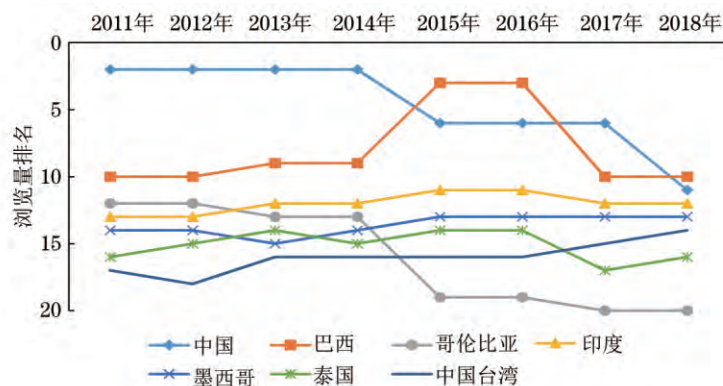


图7 发展中国家和地区的年度浏览量排名变动

表2 即时OA与延时OA的评价指标对比

| 评价指标 | 指标类型 | 即时OA | 延时OA | 即时OA/ 延时OA |
|----------------|------|----------|----------|---------------|
| 浏览量 | 平均值 | 89635.09 | 69340.11 | 1.29 |
| | 中位数 | 57768 | 45609 | 1.27 |
| 每篇论文的 浏览国家数 | 平均值 | 166.0519 | 161.9483 | 1.03 |
| | 中位数 | 168 | 162 | 1.04 |

2.4 论文日浏览量的变化趋势

发表6个月后论文由非OA转为OA,论文的浏览量是否会发生变化?图8所示为即时OA、延时OA论文发表360天内的浏览量变化情况,其中gap为即时OA与延时OA论文的浏览量差值。论文发表2天后达到浏览量的峰值,之后不断降低,但前10天的浏览量仍大幅超过其他时间段的浏览量,说明论文阅读集中于发表最初几天,即时

OA论文的浏览峰值和在360天内的日浏览量均高于延时OA论文,同样证明了OA的优势。在第1~180天内,论文浏览量呈明显的下降趋势,即时OA论文和延时OA论文间的差值也有缩小趋势,但从第181天开始,延时OA论文出现了小幅波动上升的趋势,即时OA与延时OA论文间的差距却并未缩小,说明论文从非OA转为OA确实会增加浏览量,但是与此同时,论文浏览量的上升幅度较小且浏览高峰往往集中于论文发表的最初几天,因此,即时OA相比于延时OA更有益于知识传播。约从第271天起,即时OA与延时OA论文的浏览量均呈下降趋势,且二者差值也在不断缩小,但此时距离论文发表将近1年,论文时效性已大大降低。

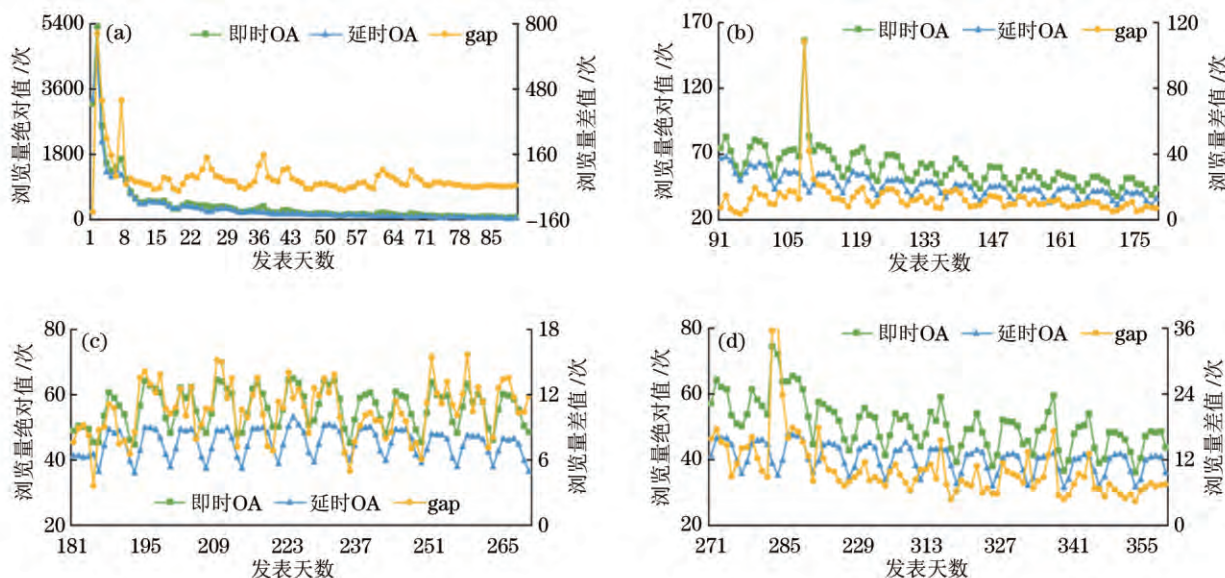


图8 论文发表后360天内的浏览量变化情况

(a) 第1~90天; (b) 第91~180天; (c) 第181~270天; (d) 第271~360天

3 结论

从地理空间角度来看,论文浏览量集中在经济发达地区,来自发达国家(如欧美地区)的读者在论文出版后的短时间内即可阅读到新发表的论文,而来自发展中国家(如中国)的读者在获取科学资源方面存在劣势,这类读者往往在论文发表后较长的一段时间才能获取相关论文。与非OA论文相比,OA论文有着更广泛的读者,其读者来自更多国家。从时间角度来看,论文在出版6个月后由非OA转为OA,论文浏览量有所增加,但增量较小,即时OA与延时OA的浏览量差值也并未缩小,说明虽然延时OA可以使更多读者受益,但是相比于即时OA仍存在不足。结合时间和地理空间角度来看,论文浏览可能存在从发达国家向发展中国家扩散的趋势,OA扩大了读者的范围,并使发展中国家的学者从中获益。延时OA虽然在开放运动中迈出了坚实的一步,但其目的仍在于保护期刊利益,如此一来,会在很大程度上影响期刊出版的时效性,从而影响文献的使用价值,因此,逐渐提高即时OA论文的比例,不断向即时OA方向努力,对促进学术资源的传播和利用有着重要的意义。即时OA是开放运动的最终实现模式,同时也是促进科学知识传播、提高论文影响力,进而提升期刊影响力的最有利方式。期刊出版机构应提高OA意识,充分认识到即时OA对促进科学传播和学术交流,以及提高期刊影响力的重要作用,从实践上促进OA体系完善和健全,更重要的是在实现OA的基础上进一步实现即时OA。期刊应积极探寻适合自身发展的OA出版模式,如可通过增值服务、开放仓储建设费等回收实行OA的成本;在渠道环节优化期刊网站建设,提高网站的易用性,强化文献检索和查询功能,丰富期刊资源获取方式,提供在线阅读、直接下载等论文获取方式,进一步降低即时OA的门槛。

本研究采用NEJM的论文浏览量数据,从时间和地理空间两个角度比较了即时OA和延时OA之间的差异,以探索OA的优势。相比于以往采用发文量和被引频次研究OA优势的思路,本研究提出了新的研究视角,即利用论文使用数据从时间和地理空间的角度分析OA的优势,因此,可以进行以往仅使用发文量和被引频次数据无法实现的研究。相比于发文量和被引频次数据,论文的使用数据具有反应及时、效果立竿见影的特点,如今大量的学术期

刊和学术资源数据库不断向公众开放,这为科研工作者提供了宝贵参考,也为改善期刊发展、扩大期刊影响力提供了新的思路。本研究同时存在局限,如对NEJM选择以OA形式进行出版的原因,如公共利益、作者的基金资助等未加以考虑,这使得在混杂因素的排除方面存在一定局限性,这些问题有待在今后的研究中进一步探索。

参考文献

- [1] 王贤文. 科学计量大数据及其应用[M]. 北京: 科学出版社, 2016: 52-23.
- [2] 张帆. 我国管理学期刊开放存取现状及认知态度调研[D]. 大连: 大连理工大学, 2011.
- [3] 周阳. 开放获取的国内用户认同与需求度调查分析——以安徽省高校为例[J]. 情报资料工作, 2009(6): 96-100.
- [4] Eysenbach G. Citation advantage of open access articles[J]. *PLoS Biology*, 2006, 4(5): e157.
- [5] Hua F, Sun H, Walsh T, et al. Open access to journal articles in oncology: Current situation and citation impact[J]. *Annals of Oncology*, 2017, 28(10): 2612-2617.
- [6] Yan E J, Li K. Which domains do open-access journals do best in? A 5-year longitudinal study[J]. *Journal of the Association for Information Science and Technology*, 2018, 69(6): 844-856.
- [7] 武庆圆. 开放获取期刊的知识交流研究[D]. 武汉: 华中师范大学, 2013.
- [8] 陈娟. Scopus和WoS收录开放获取期刊的定量研究[J]. 中国科技期刊研究, 2015, 26(4): 405-413.
- [9] Evans J A, Reimer J. Openaccess and global participation in science[J]. *Science*, 2009, 323(5917): 1025.
- [10] Tang M, Bever J D, Yu F H. Open access increases citations of papers in ecology[J]. *Ecosphere*, 2017, 8(7): e01887.
- [11] Ezema I J, Onyancha O B. Citation impact of health and medical journals in Africa: Does open accessibility matter? [J]. *Electronic Library*, 2017, 35(5): 934-952.
- [12] Alperin J P. Geographic variation in social media metrics: An analysis of Latin American journal articles[J]. *Aslib Journal of Information Management*, 2015, 67(3): 289-304.
- [13] Wang X W, Liu C, Mao W L, et al. The open access advantage considering citation, article usage and social media attention[J]. *Scientometrics*, 2015, 103(2): 555-564.
- [14] 王真, 马建华. 基于PLoS开放获取数据的单篇论文网络浏览量累积规律的数理统计及分析[J]. 图书情报工作, 2018, 62(12): 72-83.

作者贡献声明:

张光耀: 处理数据, 撰写论文;

姜春林: 指导论文修订;

王贤文: 设计研究框架, 指导论文修订。

Use advantages of immediate open access articles in time and geographic space: A case study of *The New England Journal of Medicine*

ZHANG Guangyao , JIANG Chunlin , WANG Xianwen

Institute of Science of Science and S&T Management and WISE Lab , Dalian University of Technology , 2 Linggong Road , Ganjingzi District , Dalian 116024 , China

Abstract [Purposes] In this study , we compare the difference in page views between immediate open access (OA) and delayed OA articles from two perspectives of time and geographic space to explore whether OA has advantages in both time and geographical space. [Methods] A total of 2338 articles published in *The New England Journal of Medicine* (NEJM) from 2011 to 2018 were selected as the research objects , including 135 immediate OA articles and 2203 delayed OA articles. Article metrics data were obtained directly from nejm.org for bibliometric and statistical analysis. [Findings] The OA will increase the page views of articles , promote the spread of knowledge , and expand the geographical scope of readers , especially for scholars in developing countries. Although the delayed OA policy can increase the number of page views of articles to a certain extent , compared with immediate OA , the rapidly accumulated advantages on page views of OA articles immediately after publication are still unsurpassable. [Conclusions] Immediate OA is the most advantageous way to promote the dissemination of scientific knowledge , improve the influence of articles , and further enhance the influence of journals. The realization of immediate OA plays a promoting role in enhancing the influence of journals and expanding the influence scope of journals.

Keywords: Metrics data; Immediate OA; Delayed OA; OA advantage

(本文责编: 梁永霞)