

Measuring Academic Productivity and Changing Definitions of Scientific Impact

by Cathy C. Sarli, MLS & Christopher R. Carpenter, MD

Advances in digital technology afford numerous avenues to disseminate research findings and to document the diffusion of innovations.



Cathy C. Sarli, MLS, AHIP, is the Scholarly Publishing and Evaluation Coordinator, Bernard Becker Medical Library and Christopher R. Carpenter, MD, MSc, FACEP, FAAEM, AGSF, is Associate Professor, Emergency Medicine and Director, Evidence Based Medicine, at Washington University School of Medicine in St. Louis.
Contact: sarlic@wustl.edu

Abstract

This manuscript provides a brief overview of the history of communication of scientific research and reporting of scientific research impact outcomes. Current day practices are outlined along with examples of how organizations and libraries are providing tools to evaluate and document the impact of scientific research to provide a meaningful narrative suitable for a variety of purposes and audiences.

Introduction

As a measure of scientific impact, publication counts alone are an insufficiently descriptive reporting metric of scientific research activities for the public, physicians, scientists, academic institutions and funding agencies. Shrinking biomedical research funding, along with a growing emphasis by key stakeholders to demonstrate tangible and meaningful outcomes, have motivated stakeholders to devise alternative methods that more concretely quantify the impact of scientific research on knowledge diffusion, healthcare professional uptake, and public health outcomes. In order to establish strategic directions, university administrators simultaneously face increased pressure to analyze research productivity

and impact, as well as a return on investment from research. Concurrently, funding agencies face public demand via lawmakers to ensure judicious use of taxpayer supported research while promoting the transparency and availability of research findings. The objective of this manuscript is to briefly review existing and emerging means of reporting on and quantifying scientific research impact, ongoing trends towards harmonization of evaluation for reporting of impact, and to present some examples of how academic libraries provide support.

The Historical Context of Scientific Research Communication

Formal reporting of scientific research and the peer review process dates to 1665 with Henry Oldenburg, (See Figure 1), the Secretary of the Royal Society of London, publisher of the *Philosophical Transactions*, the earliest known scientific journal in continuous publication. [See: *Philosophical Transactions*, Royal Society Publishing: (<http://rstl.royalsocietypublishing.org/>).]

Scientists during the Renaissance were reluctant to share their scientific discoveries out of concern that others would claim their work. To address this concern, Oldenburg implemented

a series of practices that established the process of modern day peer review. He appointed members of the Royal Society as independent experts to review manuscripts before approving for publication. By registering authors and manuscripts, i.e., “time stamping” of new scientific findings, Oldenburg’s methods obligated others to “cite” findings in subsequent manuscripts and ensured a regular schedule for publication of the accepted manuscripts in *Philosophical Transactions*.¹ The Royal Society of London’s practices were the precursors to modern-day principles of scientific communication and peer review, of which scholarly journals operate in their role to communicate scientific research findings.

The capacity for physicians and scientists to find applicable scientific literature from the 17th to early 20th century was rudimentary at best. To keep abreast of contemporary scientific discoveries or clinical findings physicians and scientists relied on journal subscriptions, catalogs containing a list of items held by a library, case histories, medical society memberships or correspondence with others in their field.² In the United States (U.S.), print bibliographic indexes or bibliographies on a specific subject to locate scientific literature were not available until the latter part of the 19th century. Bibliographic indices are collections of references to the published literature generally arranged by subject and/or by author names. In 1879 *Index Medicus* (the print precursor to MEDLINE®/PubMed®) was introduced and allowed for an additional means of discovery of scientific research findings.³

Bibliometrics: One Step Forward

During the 20th century peer review was a proxy for impact, with the quantity of peer-reviewed journal articles or monographs serving as leading indicators of one’s research penetration and professional community recognition for evaluation purposes as exemplified by the “publish or perish” philosophy coined by Logan Wilson in 1942.⁴ The large influx of U.S. governmental funding for health research following World War II led to an increase in the number of scientific journals to meet the demand of scientific research reporting. The proliferation of journals and articles spurred the development of bibliographic tools



Figure 1
Henry Oldenburg in a portrait by Johannes van Cleef.
Image courtesy of the National Library of Medicine.

to manage and index peer-reviewed scientific publications. Eugene Garfield in his seminal work from 1955 suggested the possibility of a citation index based on mechanical means to control and track the scientific literature.⁵ This led to the development of an index of scientific literature in the early 1960s based not only on indexing of the literature on a specific subject, but also indexing citations to the literature, the Science Citation Index, precursor to the Thomson Reuters Web of Science database.⁶ Journal Citation Reports (JCR) followed in 1976, which introduced the Journal Impact Factor score for peer-reviewed journals.

Eugene Garfield and Irving Sher, founders of the Institute for Scientific Information (later absorbed by Thomson Reuters)

proposed the JCR Impact Factor score in 1963 as a means of comparing peer-reviewed journals regardless of size. This metric was also used as a journal selection tool for inclusion in the Science Citation Index and later as an acquisitions tool by libraries.⁷ The Science Citation Index and Journal Citation Reports were ground-breaking resources that provided new means of quantifying the scientific literature and paved the way for new proxies for impact: citation counts and the JCR Impact Factor score. Tenure and securing external funding gradually became associated with publishing in “high impact” peer-reviewed journals and how frequently an investigator’s publications are cited.⁸

The development of automated systems for management of publication data and methods of analysis fostered new areas of study, in particular bibliometrics, a term introduced by Alan Pritchard in 1969.⁹ Studies in bibliometrics that outlined the applications of publication and citation data for measuring scientific impact spurred the U.S. government and the National Science Foundation (NSF) to adopt metrics available from Science Citation Reports and Journal Citation Reports for reporting purposes.^{10,11} Academic institutions soon followed suit.

For varied reasons, the JCR Impact Factor score evolved into a proxy for individual author’s impact or influence of their published works, however unintentional.¹² The higher the JCR Impact Factor score of a journal, the more prestigious any manuscript in that journal was deemed

to be. Garfield stressed that the JCR Impact Factor score was designed as a metric for journal performance and warned against its use to evaluate scientific articles and authors.¹³ One reason for wide-spread use of the JCR Impact Factor score is that it was, and still is, an easy-to-find single numeric score, and does not require extensive knowledge of database searching. Another reason favoring use of the JCR Impact Factor score as a universal metric for author impact is that the interfaces of the early citation databases were crude. Many physicians and journal editors lack any bibliometric database training or familiarity. To construct a search query for citation analysis required third-party mediation from experts (usually medical librarians) who were familiar with formulating queries, reconciling author name variants, data idiosyncrasies, and capable of interpreting the results.¹⁴

Moving Beyond Citations and the JCR Impact Factor Score

Although citations and the JCR Impact Factor score have been used as an indicator of influence and impact for decades, the landscape is changing.^{15, 16} Advances in computer and digital technology along with the general availability of the Internet spurred the development of additional resources. In 2004, two new citation data resources were introduced: Google Scholar (<http://scholar.google.com/>), and Elsevier Scopus (<http://scopus.com>). In 2005, Hirsch introduced the h index which is derived from a formula using publications and citations to provide “an estimate of the importance, significance, and broad impact of a scientist’s cumulative research contributions.”¹⁷ The h index is based on a formula that includes the “X” number of an author’s publications and the number of citations that have been cited at least “X” times. For example, an author with an h index of ten has ten publications that have been cited at least ten times. Although the h index is increasingly

recognized as a viable and even preferable alternative to the JCR Impact Factor score and raw citation counts to quantify academic productivity, it is not a perfect measure of one’s academic portfolio. First, the h index ignores bedside clinical instruction, journal editing, mentoring, and textbook authorship without which academic medicine would cease to exist. Second, the h index is simply a construct based upon citations, which does not necessarily measure clinical relevance.¹⁸

The attempt to develop one-size-fits all metrics to measure productivity and impact for all disciplines and authors remains elusive despite the numerous attempts to do so. Among the many derivatives are: the v index¹⁹ which includes the proportion of time devoted to research to normalize for clinical academicians who may devote only 40 to 50% of their time to research; the Absolute index (Ab index)²⁰ which takes into account the impact of research findings while weighting the physical and intellectual contributions of the researcher; and the hi-5 index!²¹ which is the h index over a five year period, to name a few.

Article-Level Metrics

Sophisticated publisher platforms and social media applications have resulted in a new set of metrics beyond citation counts that provide for tracking of a work (journal articles, books, slides, software, conference papers, data sets, figures, etc.) based on usage at the document level unit of analysis. Article-level metrics represent “tallies” based on usage and the social or public engagement of a work that can be captured in order to determine how a work is shared among others, commented upon, recommended, viewed, downloaded, cited in bibliographic databases, or saved in online reference managers.^{22, 23} Some article-level metrics are more scholarly in nature, and perhaps more meaningful within the context of end-user uptake since they are documented in the literature (i.e., citations) or either tied with specified technology parameters (i.e., downloads and views). Other article-level metrics remain to a point, anonymous and transient, but nonetheless, can serve as an early harbinger of the potential influence of a work (i.e., comments, mentions, favorites, bookmarks, recommendations). See Table 1 for examples of Article-Level Metrics.

Article-level metrics are available from various publisher sources and platforms, software applications, and databases. These metrics can serve as complementary measures of impact to citations, empowering authors to highlight multiple examples of scholarly output and reach beyond the traditional peer-reviewed journal article.

Table 1
Examples of Article-Level Metrics

- Citations from works in bibliographic databases and repositories
- Online downloads of a work
- Online views of a work
- Bookmarks to a work from online reference managers such as Mendeley
- Mentions of a work in social network sites such as Twitter or Facebook
- Discussions of a work in blogs or by mass media technologies
- Favorites/Recommendations of a work in platforms for sharing of works such as in Slideshare, Figshare or YouTube
- Comments/annotations for a work noted in online commenting platforms such as PubMed Commons

Recent Trends for Reporting of Scientific Impact by the Government, Funding Organizations and Publishers

The trend from bibliometric-based measures to quantify the overall value of research outputs is slowly shifting towards more meaningful outcomes of measurable impact. The U.S. Government and funding bodies are taking notice of performance and impact measures with an emphasis on outcomes that transcend bibliometrics. The National Institutes of Health (NIH) (<http://grants.nih.gov/grants/glossary.htm>) currently defines ‘impact’ as “the likelihood for the project to exert a sustained, powerful influence on the research field(s) involved.” In 2012, the NIH began implementation of a new standardized research performance progress report, (RPPR), (<http://grants.nih.gov/grants/rppr/>), as a means of harmonizing the reporting of federally-funded research across all governmental agencies that disburse extramural funding. One section of the RPPR is “Products” and includes not only publications, but also other products such as websites that disseminate the results of research activities, inventions, technologies, patents, software, databases, etc. Another section is “Impact” and grantees are instructed to report on ways that their research has had an impact. The National Science Foundation’s Biographical Sketch includes a section titled “Synergistic Activities” to allow for listing of examples that demonstrate the broader impact of an individual’s professional and scholarly activities (http://www.nsf.gov/pubs/gpg/nsf04_23/2.jsp).

Agencies such as the National Institute of Environmental Health Sciences (NIEHS) have implemented strong evaluation programs that emphasize reporting of qualitative-based outcomes and produced a manual: Partnerships for Environmental Public Health: Evaluation Metrics Manual, (http://www.niehs.nih.gov/research/supported/assets/docs/a_c/complete_peph_evaluation_metrics_manual.pdf). The Centers for Disease Control (CDC) has developed the Science Impact Framework, (<http://www.cdc.gov/od/science/impact/>) which utilizes a combination of quantitative and qualitative indicators to measure impact towards health outcomes, through five levels of influence: disseminating science, creating awareness, catalyzing action, effecting change, and shaping the future. Of particular interest is the inclusion of metrics that reflect indicators of “internal” impact such as new collaborations or partnerships that reflect on the organization and investigators themselves as opposed to external impact indicators such as public health outcomes.

Research organizations and universities also face increased pressure to report on research outcomes and to demonstrate a return on investments. Research

organizations and universities have joined with funding agencies to develop methods that enhance the transparency of research findings and to document tangible outcomes for the public. One effort is the Science and Technology for America’s Reinvestment: Measuring the Effect of Research on Innovation, Competitiveness and Science, or STAR METRICS project, (<https://www.starmetrics.nih.gov/>), launched in 2010. STAR METRICS is an effort led by the NIH and the NSF under the auspices of Office of Science and Technology Policy (OSTP), in collaboration with research organizations and universities. The objectives for STAR METRICS are to establish uniform and auditable measures of the impact of science spending and to develop measures of impact on scientific knowledge, social outcomes, workforce outcomes and economic growth. Specific metrics and testing of the metrics are still in development as of this writing.

Publishers are also stressing the need for improving the methods of evaluating and reporting on impact from scientific research. The San Francisco Declaration on Research Assessment (DORA), (<http://am.ascb.org/dora/>), recently issued a set of recommendations urging funding bodies, publishers and institutions to avoid use of the JCR Impact Factor score as a means of assessing research impact or scientific quality. DORA also stressed the use of other metrics to shift the focus towards the scientific content of an article rather than the publication metrics of a journal. Among other metrics suggested by DORA are article-level metrics, the scientific content of a publication, the influence of a work on policy and practice, and the h index. DORA also emphasizes the recognition of research outputs beyond the peer-reviewed journal article.

The Role of Libraries

Evaluation of scientific research findings and activities is an increasingly important effort by academic medical libraries. New resources and evolving recognition by funding agencies allow medical libraries to demonstrate transformative service models as essential consultants by leveraging expertise of literature searching (published and unpublished) to retrieve information that quantifies scientific impact based on bibliometrics and other measures. Evaluation and consultation to assess productivity and impact can occur at the individual author level; the department level; the research group level, including physical or virtual research groups; the institutional/university level; or for a transient population such as scholars/trainees in which longitudinal tracking is required for reporting purposes. Some libraries are going beyond use of traditional bibliometric evaluation methods by using social network

Note to Readers

If you are not affiliated with an institution that has a subscription to a citation database such as Elsevier Scopus or Thomson Reuters Web of Science, consider using Google Scholar (<http://scholar.google.com/>), a freely available resource. Google Scholar allows for searching of scholarly literature and citations, and includes a feature for authors to create a personalized profile that contains a listing of publications and citations, affiliation information, and contact information to allow for discovery of your works. By creating a profile, your profile appears when a search of your name is executed in Google Scholar. After a profile is created, publications and citations will automatically be added to your profile and metrics such as the h index and the i10-index are available. Privacy settings for the Google Scholar profile are controlled by the individual.

tools to illustrate impact in the translational environment of the millennial generation.²⁴

Medical librarians possess skill sets that are well-suited for conducting evaluation of research findings: familiarity with various database and resources, knowledge of the scholarly processes for dissemination of scientific research, formulating search queries, reconciling author variant names, capturing data from databases, and providing reports based on publication data. Librarians can also provide consultation and reports for specific purposes such as benchmarking, tenure/promotion, recruiting, performance, and funding applications and renewals as well as recommending bibliographic resources or other databases.²⁵ Some libraries are creating frameworks for scientists to identify qualitative outcomes beyond publication data for documenting and quantifying meaningful health outcomes. One example of a framework developed by a library is the Becker Model, a framework for assessment of research impact that includes a list of over 300 examples of biomedical-based outcomes including bibliometric measures.²⁶ The outcomes are grouped under five pathways represented by the research cycle with multiple examples noted for some outcomes. Evaluation services provided by libraries have parlayed into invaluable partnerships with campus units with some librarians serving as official members on tracking and evaluation teams affiliated with Clinical and Translational Sciences Awards (CTSA).²⁷ The Becker Model is currently being used by the Washington University Institute of Clinical and Translational Sciences (<http://icts.wustl.edu/>) for evaluation purposes. See the Assessing the Impact of Research website: (<https://becker.wustl.edu/impact-assessment/>) for more information.

Conclusion

Crafting a narrative of scientific research impact is a daunting task. Strides have been made in recognizing that impact transcends publication counts. Impact includes both improvement in public health outcomes and other outcomes correlated with the diffusion of knowledge such as new research collaborations focused on a specific area of study, synthesis into clinical applications, or influence on public policy. These advances in the quantification of “impact” are

occurring in tandem with efforts to harmonize reporting of research activities and outputs. The future holds great promise for a more complete and illuminating narrative of the multilevel impact of scientific research. Advances in digital technology afford numerous avenues to disseminate research findings and to document the diffusion of innovations. The capacity to measure and report tangible outcomes can be used for a variety of purposes and tailored for various audiences ranging from the layperson, physicians, investigators, organizations, and funding agencies.

References

- Spier, R., The history of the peer-review process. *Trends Biotechnol*, 2002. 20(8): p. 357-8.
- Hook, O., Scientific communications. History, electronic journals and impact factors. *Scand J Rehabil Med*, 1999. 31(1): p. 3-7.
- Brodman, E., *The Development of Medical Bibliography*. 1954, Washington, DC: Medical Library Association.
- Wilson, L., *The Academic Man: A Study in the Sociology of a Profession*. 1942, New York: Oxford University Press.
- Garfield, E., Citation indexes for science; a new dimension in documentation through association of ideas. *Science*, 1955. 122(3159): p. 108-11.
- Garfield, E., The evolution of the Science Citation Index. *Int Microbiol*, 2007. 10(1): p. 65-9.
- Garfield, E. and I. Sher, New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, 1963. 14(3): p. 195-201.
- Holden, G., et al., Should decisions about your hiring, reappointment, tenure, or promotion use the impact factor score as a proxy indicator of the impact of your scholarship? *MedGenMed*, 2006. 8(3): p. 21.
- Pritchard, A., Statistical Bibliography or Bibliometrics?. *Journal of Documentation*, 1969. 17(1): p. 348-349.
- Pendlebury, D.A., The use and misuse of journal metrics and other citation indicators. *Arch Immunol Ther Exp (Warsz)*, 2009. 57(1): p. 1-11.
- Narin, F., *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*. 1976, Cherry Hill, New Jersey: Computer Horizons, Inc.
- Alberts, B., Impact factor distortions. *Science*, 2013. 340(6134): p. 787.
- Garfield, E., The history and meaning of the journal impact factor. *JAMA*, 2006. 295(1): p. 90-3.
- Borgman, C., *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. 2007, Cambridge, MA: MIT Press.
- Cone, D.C., Measuring the measurable: a commentary on impact factor. *Acad Emerg Med*, 2012. 19(11): p. 1297-9.
- Cone, D.C. and C.R. Carpenter, Promoting stewardship of academic productivity in emergency medicine: using the h-index to advance beyond the impact factor. *Acad Emerg Med*, 2013. 20(10): p. 1067-1069.
- Hirsch, J.E., An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A*, 2005. 102(46): p. 16569-72.
- Carpenter, C.R., et al., Best Evidence in Emergency Medicine (BEEM) rater scores correlate with publications' future citations. *Acad Emerg Med*, 2013. 20(10): p. 1004-12.
- Sheridan, D.J., Reforming research in the NHS. *BMJ*, 2005. 331(7528): p. 1339-40.
- Biswal, A.K., An absolute index (Ab-index) to measure a researcher's useful contributions and productivity. *PLoS One*, 2013. 8(12): p. e84334.
- Hunt, G.E., I.S. McGregor, and G.S. Malhi, Give me a hi-5! An additional version of the h-index. *Aust N Z J Psychiatry*, 2013. 47(12): p. 1119-23.
- Lin, J. and M. Fenner, Altmetrics in evolution: Defining and redefining the ontology of article-level metrics. *Information Standards Quarterly*, 2013. 25(2): p. 20-26.
- Priem, J., H. Piwowar, and B.M. Hemminger, Altmetrics in the Wild: Using Social Media to Explore Scholarly Impact. eprint arXiv:1203.4745, 2012.
- Hunt, J.D., E.C. Whipple, and J.J. McGowan, Use of social network analysis tools to validate a resources infrastructure for interinstitutional translational research: a case study. *J Med Libr Assoc*, 2012. 100(1): p. 48-54.
- Hendrix, D., An analysis of bibliometric indicators, National Institutes of Health funding, and faculty size at Association of American Medical Colleges medical schools, 1997-2007. *J Med Libr Assoc*, 2008. 96(4): p. 324-34.
- Sarli, C.C., E.K. Dubinsky, and K.L. Holmes, Beyond citation analysis: a model for assessment of research impact. *J Med Libr Assoc*, 2010. 98(1): p. 17-23.
- Holmes, K., et al., Library-based clinical and translational research support. *Journal of Medical Library Association*, 2013. 101(4): p. 326-325.

Disclosure

None reported.

MM