

Author self-citation in the diabetes literature

Apoor S. Gami, Victor M. Montori, Nancy L. Wilczynski, R. Brian Haynes

§ See related article page 1929

Abstract

Background: Author self-citation is the practice of citing one's previous publications in a new publication. Its extent is unknown. We studied author self-citation, choosing the major clinical field of diabetes mellitus to represent the general medical literature.

Methods: We identified every article about diabetes mellitus in 170 hand-searched clinical journals published in 2000. For every article, we recorded the bibliographic citation and publication type (original or review article) and assessed the methodologic rigour. Citation information was obtained from the ISI Web of Knowledge in April 2003.

Results: Of 49 028 articles, 289 were about diabetes mellitus and had citation information. Citation counts ranged from 0 to 347 (median 6, interquartile range [IQR] 2–12). Author self-citation counts ranged from 0 to 16 (median 1, IQR 0–2). Author self-citations accounted for an average of 18% (95% confidence interval [CI] 15%–21%) and a median of 7% (95% CI 5%–11%) of all citations of each publication that was cited at least once ($n = 266$). Original articles had double the mean proportion of author self-citations compared with review articles (19% v. 9%; median 7% v. 0%, difference 7%, 95% CI 0–10%). Methodologic rigour and review type were not significantly associated with subsequent author self-citation.

Interpretation: Nearly one-fifth of all citations to articles about diabetes mellitus in clinical journals in the year 2000 were author self-citations. The frequency of self-citation was not associated with the quality of publications. These findings are likely applicable to the general clinical medicine literature and may have important implications for the assessment of journal or publication importance and the process of scientific discovery.

CMAJ 2004;170(13):1925-7

Author self-citation refers to citing one's previous publications in a new publication. Author self-citation exists when the citing and the cited papers have at least 1 author in common. This practice is distinct from journal self-citation, in which publications in a journal cite previous publications in the same journal. Critics of the impact factor as a metric of journal importance have noted the bias that results from journal self-citation,^{1,2} but little is known about the impact of author self-citation.

The scientific community uses bibliometric data, including citation counts of articles and impact factors of

the journals in which the articles were published, to judge the importance of articles.³ Academic promotion committees similarly use these data to assess the productivity of faculty members and the scientific merit of their work.⁴ Author self-citations are not removed from citation counts or from the calculation of impact factors. As a result, author self-citations may misrepresent the importance of individual articles, skew the calculation of journal impact factors and bias perceptions of the importance of a publication.⁵ The effects of author self-citation on the process of research and discovery are unknown and potentially important.

Because the extent of author self-citation is unknown, we designed a study to identify the extent to which this practice occurs in the literature on diabetes mellitus, a major clinical field whose literature would represent the general medical literature, and to determine how self-citation relates to selected types of clinical articles and the quality of the reported research.

Methods

We identified 170 clinical journals by accessing the impact factors of Science Citation Index (Thomson ISI, The Thomson Corporation, Stamford, Conn.), obtaining recommendations from clinicians and librarians, and evaluating journals' yields of publications of scientific merit and clinical relevance. For these journals, 6 research associates collected bibliographies from every article in each issue published in 2000 and applied methodologic criteria. The final database included 49 028 articles. Further details on the creation of the database are available elsewhere.⁶ Owing to the large number of articles, we limited our study to a single major clinical field (diabetes mellitus) representative of general medicine. We classified all publications about diabetes by type (original article, narrative review, systematic review or meta-analysis) and assessed each publication for methodologic rigour according to explicit criteria for clinical research.⁶ A review was defined as any publication in which the title or a section heading identified it as a review, overview or meta-analysis or the text claimed to review or summarize the medical literature on a specific topic.⁶ A systematic review was defined as any review in which the authors identified a specific topic, stated how and from what sources data were retrieved, and provided explicit study inclusion and exclusion criteria. All other reviews were classified as methodologically limited. A meta-analysis was defined as any review that combined data from previous publications. The research

associates were trained to assess the methodology of articles, and the level of interrater agreement was very high (Cohen's kappa value > 0.8).⁶

On Apr. 1, 2003, the database contained 295 articles about diabetes therapy, prognosis, diagnosis, harm or cause. For each such publication, we retrieved the complete bibliographic citation, the article type and the methodologic grade. During the period Apr. 1–7, 2003, we queried the ISI Web of Knowledge (www.isiknowledge.com [a subscription is required]), which reports citation counts, impact factors and other bibliometric data. For 6 of the 295 articles, the ISI Web of Science did not track citations; these articles were excluded from further consideration. For each of the remaining 289 articles, we collected the citation count — the number of times the publication had been cited in subsequent publications. We examined every such citation, compared the lists of authors of the cited and citing publications, and identified the number of self-citations. Despite the large number of articles and citations, there was no instance in which article or author names caused confusion.

We created frequency tables of counts of author self-citations and all citations for the publications, categorized by type and methodologic rigour. Because the proportion of self-citations among all citations had a skewed distribution, we estimated the median. We also estimated the 95% confidence intervals (CIs) for the medians and for differences of medians (by study type and by methodologic rigour) using the bootstrapping technique with 10 000 iterations.

Results

Of the 170 journals, 50 published the 289 publications about diabetes mellitus for which citation information was available (see Table 1, available online at www.cmaj.ca/cgi

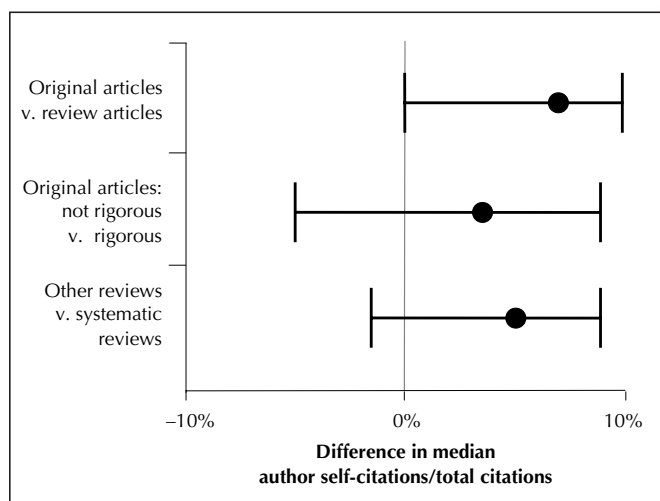


Fig. 1: Differences in median proportions of author self-citations among all citations of publications about diabetes mellitus in 2000 that were cited at least once before April 2003. Horizontal bars represent the 95% bootstrap confidence interval (CI) for the differences. Original articles (7%) v. reviews (0%): 7% (95% CI 0% to 10%). Original articles that were not rigorous (8%) v. those that were rigorous (4.5%): 3.5% (95% CI -5% to 9%). Other reviews (8%) v. systematic reviews (3%): 5% (95% CI -1.5% to 9%).

/content/full/170/13/1925/DC1); there were 256 original articles and 33 reviews. Only 61 of the original articles had rigorous methods. Of the 10 systematic reviews, 8 included meta-analysis. Citation counts of individual publications ranged from 0 to 347 (median 6, interquartile range [IQR] 2–12). Author self-citation counts per publication ranged from 0 to 16 (median 1, IQR 0–2). Author self-citations accounted for an average of 18% (95% CI 15%–21%) and a median of 7% (95% CI 5%–11%) of all citations of each publication that was cited at least once ($n = 266$).

Original articles had double the mean proportion of author self-citations compared with review articles (19% v. 9%; median 7% v. 0%, difference 7%, 95% CI 0–10%). Methodologic rigour and review type were not significantly associated with subsequent author self-citation (Fig. 1).

Interpretation

The important finding of this study is that author self-citations comprise nearly one-fifth of all citations to articles about diabetes published in clinical journals in 2000. In general, however, self-citations had little relation with the quality of an article, and authors did not preferentially cite rigorously performed research or reviews. These findings are likely applicable to the general clinical medicine literature and may have important implications for the assessment of the “importance” of journals and publications and for the process of scientific discovery.

Author self-citation serves necessary functions in medical literature. It allows an author or group to expand on previous hypotheses, refer to established study designs and methods, and justify further investigations on the basis of prior results. Author self-citation may be inevitable when the published data in a specific field are solely the work of 1 investigator or research group. For example, we cited our own work once in this paper in order to refer the reader to relevant methodologic issues.

The greatest risk of author self-citation may be its effects on the process of scientific discovery. Citations build connectivity between publications, and this is essential to the growth and progression of medical knowledge.⁴ Repeated self-citation accentuates one's credibility or expertise⁷ and may perpetuate one's interpretations or opinions of specific research findings or general constructs. Self-citations, when pervasive, might falsely validate the conclusions of an author or group and could even limit scientific discovery if other investigators do not challenge what might be perceived as developing or accepted concepts.

To the extent that assessors of the importance of a publication rely on bibliometric indices based on citation counts (perhaps for lack of a better metric),^{4,5} author self-citation may artificially inflate an article's importance to the general scientific community. Author self-citation may even be performed knowingly for this purpose alone,⁷ a practice that has been satirized.⁸ Apart from mere ego-

tism, this practice may be promoted by the fact that citations are “a hallmark of academic achievement for authors and journals” and “correlate highly with the opinions of peers as to a scientist’s contributions to his/her field and are used by medical school deans for promotion reviews.”⁴ One journal published an open call for authors to cite more recent articles from journals with a high impact factor to raise the former’s impact factor.⁹ Although this would affect counts of journal self-citation (as opposed to author self-citation), it makes apparent the motivation and means by which citation counts may be manipulated. Our findings add to the debate regarding the role that bibliometric data should play in judgements about the scientific importance of published work.

This article has been peer reviewed.

From the Department of Internal Medicine, Mayo Clinic College of Medicine, Rochester, Minn. (Gami, Montori), the Department of Clinical Epidemiology and Biostatistics (Montori, Wilczynski, Haynes) and the Department of Medicine (Haynes), McMaster University, Hamilton, Ont.

Competing interests: None declared.

Contributors: Apoor Gami and Victor Montori conceived the study and acquired and analyzed the data. All of the authors were involved in study design and data interpretation. Victor Montori provided statistical expertise and study supervision. Apoor Gami drafted the manuscript. All of the authors provided critical revisions and approved the version to be published.

Acknowledgements: We thank other members of the research team, including Angela Eady, Susan Marks, Ann McKibbin, Doug Morgan, Cindy Walker-Dilks, Stephen Walter, Stephen Werre and Sharon Wong, for their contributions.

This study was funded by the US National Library of Medicine.

References

1. Garfield E. How can impact factors be improved? *BMJ* 1996;313:411-3.
2. Fassoulaki A, Paraskeva A, Papilas K, Karabinis G. Self-citations in six anaesthesia journals and their significance in determining the impact factor. *Br J Anaesth* 2000;84:266-9.
3. Garfield E, Welljams-Dorof A. Citation data: their use as quantitative indicators for science and technology evaluation and policy-making. *Sci Public Policy* 1992;19:321-7.
4. Callaham M, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA* 2002;287:2847-50.
5. DeMaria AN. A report card for journals. *J Am Coll Cardiol* 2003;42:952-3.
6. Montori VM, Wilczynski NL, Morgan D, Haynes RB. Systematic reviews: a cross-sectional study of location and citation counts. *BMC Med* 2003;1(1):2.
7. Hyland K. Self-citation and self-reference: credibility and promotion in academic publication. *J Am Soc Inf Sci Technol* 2003;54:251-9.
8. Craddock N, O'Donovan MC, Owen MJ. Introducing Selfcite 2.0-career enhancing software. *BMJ* 1996;313:1659-60.
9. Kapoor VK, Aggarwal R. A comparison of citations in the *Indian Journal of Gastroenterology* with other journals. *Indian J Gastroenterol* 1993;12(Suppl 1):S12-6.

Correspondence to: Dr. Apoor S. Gami, Division of Cardiovascular Diseases, Mayo Clinic, 200 First St. SW, Rochester MN 55905, USA; fax 507 255-7070; gami.apoor@mayo.edu



CMA CENTRE FOR
PHYSICIAN HEALTH AND WELL-BEING
HEALTHY PHYSICIANS, VIBRANT PROFESSION

**If you are feeling
overwhelmed by
professional and personal
demands, you are not alone.
Research has found
that 45.7% of
Canadian physicians are
in an advanced
state of burnout.***

CMA believes in a culture of medicine that values the health of its practitioners and does not stigmatize those who seek personal assistance. CMA's new Centre for Physician Health and Well-being provides national leadership and advocacy on issues affecting the health and morale of Canadian physicians.

To learn more about the work of the centre or to access information and resources, visit the centre's new Web section at cma.ca or call 1 877 CMA-4-YOU (1 877 262-4968).

The centre supports and complements the work of the provincial and territorial physician health programs but does not provide individual assessment, support or treatment advice. Contact numbers for the programs are available at cma.ca

*2003 CMA Physician Resource Questionnaire.

ASSOCIATION
MÉDICALE
CANADIENNE



CANADIAN
MEDICAL
ASSOCIATION

cma.ca