Applied Medical Informatics
Editorial **Vol. 46, No. 1 /2024, pp: 1-4**

Citation Statements Practices and the Role of an Assistive Tool

Sorana D. BOLBOACĂ

Department of Medical Informatics, "Iuliu Haţieganu" University of Medicine and Pharmacy Cluj-Napoca, Louis Pasteur Str., no. 6, 400349 Cluj-Napoca, Romania
E-mail: sbolboaca@umfcluj.ro

* Author to whom correspondence should be addressed;

Published online: 31 March 2024

Citations play a pivotal role in scientific writing, giving credit and acknowledge other's work, supporting an argument, presenting similar or dissimilar results, explaining specific results, and not ultimately directing the readers to the original source of information. By fostering a culture of transparency and accountability, citation practices support the credibility and trustworthiness of research. After a brief introduction of the concept of citation, its use and misuse, this editorial introduces one assisting tool and highlights the need towards quality, not quantity, in citation behavior in scientific writing.

Citation is a mandatory practice in scientific writing used to "*properly referring to others' ideas, thoughts, or concepts*" [1]. We used citations to give credit and acknowledge other's work, to support an argument, to present similar or dissimilar results, and to explain our results by directing the readers to the original source of information. Citation is a good practice in scientific writing and allows the authors to avoid plagiarism. Citations could take the form of paraphrasing (a summary of other's work written in the researcher's words by presenting the main points of the original work) or quotation (quotation from others' works). Regardless of form, the original source's bibliographic details accompany a citation. Bibliographic details respect the Vancouver style in health and biomedical sciences [2]. As Lipson stated, "*Good citations should reveal your sources, not conceal them. They should honestly reflect the research you conducted.*" [3].

Eugene Garfield (1925–2017) is considered the father of citation analysis for scientific literature. Garfield proposed in 1955 [4] the first science citation index and envisioned the automation of the process [5]. Development of citation indexes simplifying the scientific literature navigating process [6]. Scientific communities misuse citations [7] as the key measurement for impact, either at the journal-level (Impact Factor, journal classifications), the article-level (AIS – article influence score), or the author-level (Hirsh index [8]) [9]. Critics have been questioning the use of citation analysis as a measure to trace intellectual influence and impact since the 1950s, arguing that "*impact is not the same as importance or significance*" [10]. None of the metrics used to measure the impact consider an individual author's contribution to a paper, with possible high recognition for all authors, regardless of whether all have or not significant contributions to research [11]. Furthermore, co-citation increased the citations and correlates with patterns of interaction between scientists [12], while self-citation (journal, institution, language, filed or country) artificially increases citation-related metrics also in the absence of an increase in scientific production (e.g., Colombia, Egypt, Indonesia, Iran, Italy, Malaysia, Pakistan, Romania, Russian Federation, Saudi Arabia, Thailand, and Ukraine) [13]. Despite its limitations, citations are used to evaluate the scientific performance of research groups, departments, and institutions [14], evaluate the research proposals [15,16], allocate research funds [17], hire researchers and academic staff [18], and rankings institutions (e.g., (THE) World University Rankings give a weight of 30% for research quality and 29% for research environment [19]) [20]. Different metrics are derived from citations to correct its disadvantages and differences between specific research fields. One example is the standardized citation metric described by Ioannidis et al. [21,22,23]. Considering that a single metric does not capture the performance of an academic scholar, new

metrics were proposed and have been evaluated (e.g., Figure-of-Merit (FOM), Enhanced Research Quality Index (ERQI) [24]). Altmetrics [25], alternative metrics to measure the impact of a scientific paper, had been proposed to consider social media events related to scholarly communication (https://www.altmetric.com/), considering the translation of the research results to the general population readability. Different publishers adopted altmetrics (e.g., BioMed Central, Public Library of Science (PLOS) Nature Publishing Group, and Elsevier) but paid service (altmetrics to individual articles).

All the metrics based on citations suffer from a fundamental problem: they are limited to the form (number of citations) and do not consider the content. Smart Citations by scite.ai (https://scite.ai/) implements a classification of citations into three different categories (contrasting, supporting, mentioning). The developers of scite.ai use machine learning to classify and contextualize scientific research by classification of citations based on their intent and context [26]. Josh Nicholson and Yuri Lazebnik developed the tool under the frame of National Science Foundation (NSF) and National Institute on Drug Abuse (NIDA) financial support. Nicholson et al. introduced a new paradigm in analyzing citations, as citation context analysis [26]. The following are the features of scite.ai:

- ✓ Display the number of publications citing a specific manuscript, and classify the citation statements as contrasting, supporting, or mentioning. Indicate the paper section (e.g., Introduction, Methods, Results, Discussion) where the citation statement was given, and classify the publication type (article, preprint, book, etc.) and relationship (as self-citation or independent citation).
- ✓ Offer the possibility of collecting bibliographic data of a specific manuscript in seven standard styles (APA, MLA, Chicago, Hardvard, Vancouver, IEEE, and BibTex).
- ✓ List the citation statistics for a specific author. The list includes the total number of publications, the number of citation statements received by the author, and number of citation statements given by the author. The classification of citation statements includes contrasting, supporting, mentioning, or unclassified.
- ✓ ***Assistant by scite*** is an automated tool developed to assist scientists in finding the answers to specific questions and queries. The input is given by the user and the output by the ***assistant by scite.*** Scientific references support the response, and a search strategy, consulted publications and the list of references used to generate the text underlying the information used are provided. The user can either copy the generated response or regenerate the response. A dedicated section is displayed and lists the paragraphs from the references that served as the source of the information, along with the corresponding sections in the manuscript where the paragraphs were located. The question is *How accurate are the citations*?

Overall, most citation statements classified by scite.ai are "mentioning", showing the current citation behavior (if classification is accurate) by listing citations without an appropriate assessment of their content.

Researchers should appropriately assess how scite.ai performs. Such evaluations are limited. Bakker et al. [27] reported a low accuracy of scite.ai as compared to human classification when investigated 324 citations out of which 98 citations classified by scite.ai, with supporting and contrasting citations classified as mentioning (40 supporting and 17 contrasting).

Norman Kaplan listed in 1965 [28] several questions regarding citation behavior, questions that still apply: "*How often are the works of others cited without having read them carefully*?", "*How often are citations simply lifted from the bibliography in someone else's work without either reading or giving credit to the man who did the original search of the literature*?" (*secondary referencing* or *secondhand citation*).

Extending the context underlined by Norman Kaplan, other questions regarding citation behavior are of interest: *How often the cited document is critically summarized?*, *How often does the cited manuscript support the information*?, *How could the citation of a retracted manuscript without acknowledgments of retraction be explained*?, *How would the automated research assistive tools impact the scientific writing*?, *How can automated research assistant tools be used by respecting research integrity*?.

What to expect next? As the amount of scientific literature grows exponentially, we can expect an increased use of assistive technology and algorithms. We also expect the development of new research assistive tools. But how can assistive research tools help in increasing the quality of citations? First, automatic checking could ensure the accuracy of the bibliographic data of citations. This feature would be useful for the authors and publishers, ensuring that all bibliometric data are correct and accurate. Second, before the reviewing process, the statements with secondhand citations could be identified and the manuscript automatically returned to the authors, a

process that is currently left on the shoulders of the reviewer and/or editors. Third, research assistive tools could automatically detect citation inaccuracy and misinterpretations by comparing the cited text with the original source. Such a feature can help researchers to evaluate the reliability and validity of cited information critically, shifting the citation towards content. Fourth, assistive research tools could become valuable tools for research evaluation, review and identification of works related to a specific research topic, facilitating a more effective navigation in the scientific research landscape. Finally, researchers must appropriately evaluate the validity of the assistive research tools and develop guidelines for their use that respect research integrity. Researchers must undergo training on how to use such tools and the implementation of the guidelines must be evaluated.

To sum up, citation statements should reflect the content of the cited documents, must be accurate and specific. Listing a series of references that report a specific subject is uninformative, so a citation statement must move from presenting what other researchers did towards a critical appraisal or previous reported studies. Advanced technology and innovative approaches could empower researchers to navigate scientific literature accurately and help them increase accuracy and specificity in scientific communication. It remains essential to uphold principles of integrity and transparency in scientific communication, ensuring that the referred sources appropriately support citation statements as reliable information guiding the advancement of medical knowledge based on accurate evidence. Digital innovations are expected to contribute to reproducibility and reshape the way researchers cite while respecting research integrity.

List of Abbreviations: Not applicable.

Funding: This research received no funding.

Ethics Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflict of Interest: The author declares no conflict of interest.

References

1. Bahadoran Z, Mirmiran P, Kashfi K, Ghasemi A. The Principles of Biomedical Scientific Writing: Citation. Int J Endocrinol Metab. 2020 Apr 27;18(2):e102622. https://doi.org/10.5812/ijem.102622.
2. Fletcher S, Fletcher R, Thomas L, Lock S, Lundberg GD, Fox R, Nylenna M, Colaianni LA, Relman AS, Angell M, Riis P, Robinson RG, Squires BP, Clever L, Medicine International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical Journals. BMJ 1991;302:338-341. https://doi.org/10.1136/bmj.302.6772.338.
3. Lipson C. "Why Cite?" Cite Right: A Quick Guide to Citation Styles--MLA, APA, Chicago, the Sciences, Professions, and More. Chicago: U of Chicago, 2006.
4. Garfield E. Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. Science 1955;122(3159):108-111. https://doi.org/10.1126/science.122.3159.108
5. Garfield E. Can Citation Indexing be Automated? Stat Assoc Methods Mech Doc Symp Proc. 1965;269:189–192.
6. Tagliacozzo R. Citations and citation indexes: a review. Methods of Information in Medicine 1967;6(03):136-142. https://doi.org/10.1055/s-0038-1636363
7. Margolis J. Citation indexing and evaluation of scientific papers. Science. 1967 Mar 10;155(3767):1213-9. https://doi.org/10.1126/science.155.3767.1213.
8. Hirsch J. An index to quantify an individual's scientific research output. Proceedings of the National academy of Sciences 2005;102(46):16569–16572. https://doi.org/10.1073/pnas.0507655102.
9. Roldan-Valadez E, Salazar-Ruiz SY, Ibarra-Contreras R, Rios C. Current concepts on bibliometrics: a brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised

Impact per Paper, H-index, and alternative metrics. Ir J Med Sci. 2019 Aug;188(3):939-951. https://doi.org/10.1007/s11845-018-1936-5.
10. Garfield E. Citation indexes in sociological and historical research. Arherican Documentation 1963;14:289-291. https://doi.org/10.1002/asi.5090140405
11. Fazel S, Wolf A. What is the impact of a research publication? Evid Based Ment Health. 2017 May;20(2):33-34. https://doi.org/10.1136/eb-2017-102668.
12. Small HG. A co-citation model of a scientific specialty: A longitudinal study of collagen/research, Social Studies of Science 1977;7(2):139-166. https://doi.org/10.1177/030631277700700202.
13. Baccini A, Petrovich E. A global exploratory comparison of country self-citations 1996-2019. PLoS One. 2023 Dec 29;18(12):e0294669. https://doi.org/10.1371/journal.pone.0294669.
14. Moed HF. Citation analysis in research evaluation. Dordrecht, The Netherlands: Springer, 2005.
15. Cabezas-Clavijo A, Robinson-García N, Escabias M, Jiménez-Contreras E. Reviewers' ratings and bibliometric indicators: hand in hand when assessing over research proposals? PLoS One. 2013 Jun 28;8(6):e68258. https://doi.org/10.1371/journal.pone.0068258. PMID: 23840840; PMCID: PMC3695904.
16. Venable GT, Khan NR, Taylor DR, Thompson CJ, Michael LM, Klimo P Jr. A correlation between National Institutes of Health funding and bibliometrics in neurosurgery. World Neurosurg. 2014 Mar-Apr;81(3-4):468-72. https://doi.org/10.1016/j.wneu.2013.11.013. Epub 2013 Nov 13. PMID: 24239737.
17. Carlsson H. (2009). Allocation of research funds using bibliometric indicators—Asset and challenge to Swedish higher education sector. *InfoTrend*, 64(4), 82-88.
18. Holden G, Rosenberg G, Barker K.. Bibliometrics: A potential decision making aid in hiring, reappointment, tenure and promotion decisions. In Holden G., Rosenberg G., Barker K. (Eds.), Bibliometrics in social work, 2005, pp. 67-92. New York, NY: Routledge.
19. World University Rnaking 2024 methodology. [Internet] [accessed on 30 March 2024]. Available from: https://www.timeshighereducation.com/sites/default/files/the_2024_world_university_rankings_methodology.pdf
20. Szluka P, Csajbók E, Győrffy B. Relationship between bibliometric indicators and university ranking positions. Sci Rep. 2023 Aug 30;13(1):14193. https://doi.org/10.1038/s41598-023-35306-1.
21. Ioannidis JPA, Baas J, Klavans R, Boyack KW. A standardized citation metrics author database annotated for scientific field. PLoS Biol. 2019 Aug 12;17(8):e3000384. https://doi.org/10.1371/journal.pbio.3000384.
22. Ioannidis JPA, Boyack KW, Baas J. Updated science-wide author databases of standardized citation indicators. PLoS Biol. 2020 Oct 16;18(10):e3000918. https://doi.org/10.1371/journal.pbio.3000918.
23. Ioannidis JPA, Boyack KW, Collins TA, Baas J (2023) Gender imbalances among top-cited scientists across scientific disciplines over time through the analysis of nearly 5.8 million authors. PLoS Biol 21(11): e3002385. https://doi.org/10.1371/journal.pbio.3002385
24. Mina R, Homsi F. A figure of merit that includes 5 distinct performance indicators to improve research evaluation of academic scholars'. Heliyon. 2024 Feb 10;10(4):e26235. https://doi.org/10.1016/j.heliyon.2024.e26235.
25. Priem J, Taraborelli D, Groth P, Neylon C. [26 October 2010]. "Altmetrics: A manifesto (v 1.01)". Altmetrics. [accessed 30 March 2024]. Available from: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1187&context=scholcom
26. Nicholson JM, Mordaunt M, Lopez P, Uppala A, Rosati D, Rodrigues NP, Grabitz P, Rife SC. Scite: a smart citation index that displays the context of citations and classifies their intent using deep learning. Quantitative Science Studies 2021;2(3):882-898. https://doi.org/10.1162/qss_a_00146
27. Bakker C, Theis-Mahon N, Brown SJ. Evaluating the Accuracy of scite, a Smart Citation Index. Hypothesis [Internet]. 2023 Sep. 13 [cited on 2024 Mar. 30];35(2). Available from: https://journals.iupui.edu/index.php/hypothesis/article/view/26528
28. Kaplan N. The norms of citation behavior: Prolegomena to the footnote. American Documentation, 1965;16(3):179-184. https://doi.org/10.1002/asi.5090160305