

In our Section 4, parentheses should be included as Tedeschi indicates, but in his letter, he has reversed the subscripts 1 and 2.

Fred M. Hoppe
*Department of Mathematics and Statistics
 McMaster University
 Hamilton, ON L8S4K1, Canada*

Daniel J. Hoppe
*University of Toronto Orthopaedic Sports Medicine
 76 Grenville St., Toronto, ON M5S1B2, Canada*

Stephen D. Walter
*Department of Clinical Epidemiology and Biostatistics
 McMaster University
 Hamilton, ON L8N3Z5, Canada*

*Corresponding author. Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S4K1, Canada.
 Tel.: +1-905-525-9140x24688; fax: +1 905-522-0935.
 E-mail address: hoppe@mcmaster.ca (F.M. Hoppe)

References

- [1] Tedeschi F. A clarification on odds ratios as conditional risk ratios. *J Clin Epidemiol* 2018;97:133–4.
- [2] Hoppe FM, Hoppe DJ, Walter SD. Odds ratios deconstructed: a new way to understand and explain odds ratios as conditional risk ratios. *J Clin Epidemiol* 2017;82:87–93.

<http://dx.doi.org/10.1016/j.jclinepi.2017.10.009>

Citation analysis is also useful to assess the eligibility of biomedical research works for inclusion in living systematic reviews



I have read with much interest the JCE series advocating the use of human efforts and machine automation to create and update living systematic reviews (LSRs) [1]. I recognize that the series provides important information on how biomedical research works are verified as eligible for inclusion in LSRs using semantic classification and crowdsourcing techniques [1]. However, this paper has not dealt with another technique that has been recently shown to be useful (when jointly used with semantic classification and crowdsourcing techniques) in assessing the eligibility of papers for inclusion in LSRs: This important technique is citation analysis [2–6].

It has been shown that evidence identified by traditional methods that are neither cited by nor citing other initial search results (not included in the citation network of the initially selected papers) do not relate to the topic of the LSR and had been included in the initial set of search results by error. Such papers can be easily

eliminated without having to semantically analyze them [2,4,5].

Furthermore, citation analysis based on two easy-to-do snowballing techniques can be useful to add the papers that are eligible for the LSR but that were missed by semantic classification and crowdsourcing techniques¹.

Backward snowballing consists of analyzing the references of the initial search results and of all the papers that were cited by the initial search results and then include the papers that cite or are cited by a significant number of initial search results in the LSR. Forward snowballing consists of analyzing the references of all the papers indexed in the database and include in the LSR the papers that cite a significant number of initial search results [3,6].

Moreover, papers that are co-cited with the initial search results in the analyzed database (and that are not included in the initial set of considered papers for the LSR) are very likely to be related to the topic of the LSR [2,7]. These papers are easy to find and to include in the LSR using the analysis of the references of all the papers that cite initial search results.

For these reasons, citation analysis should be used for creating and updating LSRs.

Houcemeddine Turki
*Faculty of Medicine of Sfax, University of Sfax
 Sfax, Tunisia
 Tel.: +21629499418.
 E-mail address: turkiabelwaheb@hotmail.fr*

References

- [1] Thomas J, Noel-Storr A, Marshall I, Wallace B, McDonald S, Mavergames C, et al. Living systematic reviews: 2. Combining human and machine effort. *J Clin Epidemiol* 2017;91:31–7.
- [2] Chen TT. The development and empirical study of a literature review aiding system. *Scientometrics* 2012;92(1):105–16.
- [3] Choong MK, Galgani F, Dunn AG, Tsafnat G. Automatic evidence retrieval for systematic reviews. *J Med Internet Res* 2014;16(10):e223.
- [4] Octaviano FR, Felizardo KR, Maldonado JC, Fabbri SC. Semi-automatic selection of primary studies in systematic literature reviews: is it reasonable? *Empirical Softw Eng* 2015;20(6):1898–917.
- [5] Octaviano F, Silva C, Fabbri S. Using the SCAS strategy to perform the initial selection of studies in systematic reviews: an experimental study. Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering 2016;25. ACM. Available at <https://dl.acm.org/citation.cfm?doi=2915970.2916000>.
- [6] Wohlin C. Guidelines for snowballing in systematic literature studies and a replication in software engineering. Proceedings of the 18th international conference on evaluation and assessment in software engineering 2014;38. ACM. Available at <https://dl.acm.org/citation.cfm?id=2601268>.

¹ Less than 1% of the total number of papers that should be considered for the LSR according to [1]. This fact can be explained by the nonindexation of these few papers in the analyzed reference database or by the use of a language other than English when writing them.

- [7] Small H. Co-citation in the scientific literature: a new measure of the relationship between two documents. *J Assoc Inf Sci Technology* 1973;24(4):265–9.

<http://dx.doi.org/10.1016/j.jclinepi.2017.11.002>

Citation analysis may well have a role to play in study identification, but more evaluation and system development are required



I would like to thank Houcemeddine Turki [1] for his interest in the paper that I and my colleagues recently published [2], and for drawing our attention to the potential utility of citation networks for identifying eligible studies for inclusion in systematic reviews.

I agree that citation analysis may well have a role to play in the future along the lines suggested but feel that further research and development is needed. Most significantly, no study has evaluated these approaches at scale (eg, by looking across all systematic reviews in the Cochrane Library), so we do not know yet how generalizable these techniques are. One of the reasons for this is that it has been quite difficult to generate citation networks automatically. This factor has not only limited the evaluation work that has hitherto been possible but also points to a weakness in our information architecture more broadly: that, even if citation networks *might* be useful, we cannot generate them in a reliable and automated way. (This last limitation may be changing with the emergence of the new Microsoft Academic Graph API, which is designed to support the creation of citation networks automatically.)

When using citation data, the reviewer needs to be always cautious that they are not simply reinforcing the effect of publication bias. We know that researchers tend to operate as communities, publishing with one another, attending the same conferences, and citing each other's papers. An over-reliance on snowballing techniques can miss important 'clusters' of studies in situations where they are not connected to the papers already identified. I agree with Mr Turki, that citation data might be considered useful when combined with semantic classification, rather than on its own, and looking forward to see more evaluative studies in this area, using services which enable citation networks to be created quickly, reliably, and at scale.

James Thomas

EPPI-Centre

University College London

18 Woburn Square

London, WC1H 0NR, United Kingdom

Tel.: +44 207 612 6844; fax: +44 207 612 6400.

E-mail address: james.thomas@ucl.ac.uk

Reference

- [1] Turki H. Citation analysis is also useful to assess the eligibility of biomedical research works for inclusion in living systematic reviews. *J Clin Epidemiol* 2018;97:135–6.
- [2] Thomas J, Noel-Storr A, Marshall I, Wallace B, McDonald S, Mavergames C, et al. Living systematic reviews: 2. Combining human and machine effort. *J Clin Epidemiol* 2017;91:31–7.

<http://dx.doi.org/10.1016/j.jclinepi.2017.11.001>