Citation bias and other determinants of citation in biomedical research

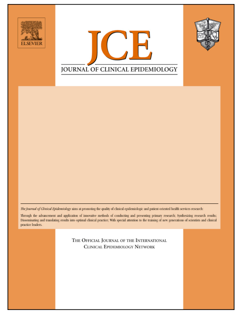
Urlings, M J E; Duyx, B; Swaen, G M H; Bouter, L M; Zeegers, M P

published in
Journal of clinical epidemiology
2021

DOI (link to publisher)
10.1016/j.jclinepi.2020.11.019

citation for published version (APA)
Urlings, M. J. E., Duyx, B., Swaen, G. M. H., Bouter, L. M., & Zeegers, M. P. (2021). Citation bias and other determinants of citation in biomedical research: findings from six citation networks. *Journal of clinical epidemiology*, *132*, 71-78. https://doi.org/10.1016/j.jclinepi.2020.11.019

Journal Pre-proof



Citation bias and other determinants of citation in biomedical research: findings from six citation networks

M.J.E. Urlings, B. Duyx, G.M.H. Swaen, L.M. Bouter, M.P. Zeegers

PII: S0895-4356(20)31197-5

DOI: https://doi.org/10.1016/j.jclinepi.2020.11.019

Reference: JCE 10354

To appear in: *Journal of Clinical Epidemiology*

Received Date: 1 July 2020

Revised Date: 5 October 2020

Accepted Date: 2 November 2020

Please cite this article as: Urlings MJE, Duyx B, Swaen GMH, Bouter LM, Zeegers MP, Citation bias and other determinants of citation in biomedical research: findings from six citation networks, *Journal of Clinical Epidemiology* (2021), doi: https://doi.org/10.1016/j.jclinepi.2020.11.019.

Citation bias and other determinants of citation in biomedical research: findings from six citation networks

M.J.E. Urlings[1*], B. Duyx[1], G.M.H. Swaen[1], L.M. Bouter[2,3], M.P. Zeegers[1,4]

[1] Nutrition and Translational Research in Metabolism (School NUTRIM), Maastricht University, Maastricht, The Netherlands.

[2] Department of Epidemiology and Data Science, Amsterdam University Medical Centers, Amsterdam, The Netherlands

[3] Department of Philosophy, Faculty of Humanities, Vrije Universiteit, Amsterdam, The Netherlands

[4] Care and Public Health Research Institute (School CAPHRI), Maastricht University, Maastricht, The Netherlands.

* Corresponding author, mje.urlings@maastrichtuniversity.nl

Conflict of Interest

All authors declare that they have no conflict of interest (other than the project funding received for this study; see Funding statement).

Funding

This work was supported by the Long-range Research Initiative (LRI) from the European Chemical Industry Council (CEFIC), grant no. LRI-Q3-UM. LRI had no role in study design, data collection and analysis, preparation of the manuscript, or decision to publish.

Author contributions

All authors have been involved in the development of the citation analyses methodology. Funding was obtained by GMHS and MPZ. Data collection and analysis on each of the citation networks was performed in collaboration by M.J.E.U., B.D., and G.M.H.S. All co-authors were actively involved in the interpretation of the data. Drafting this article was performed by M.J.E.U.; the article was critically revised by all co-authors. All co-authors approved the final version.

Abstract

Objective: When the probability of being cited depends on the outcome of that study, this is called citation bias. Aim of this study is to assess the determinants of citation and how these compare across six different biomedical research fields.

Study design: Citation network analyses were performed for six biomedical research questions. After identifying all relevant publications, all potential citations were mapped together with the actually performed citations in each network. As determinants of citation we assessed: study outcome, study design, sample size, journal impact factor, gender, affiliation, authority and continent of the corresponding author, funding source, title of the publication, number of references and self-citation. Random effect logistic regression analysis was used to assess these factors.

Results: Four out of six networks showed evidence for citation bias. Self-citation, authority of the author and journal impact factor were also positively associated with the probability of citation in all networks.

Conclusion: The probability of being cited seems associated with positive study outcomes, the authority of its authors and the journal in which that article is published. Additionally, each network showed specific characteristics that impact the citation dynamics and that need to be considered when performing and interpreting citation analyses.

Key words:

Citation bias, citation network analyses, questionable research practices, biomedical research, H-index

Running title: Citation bias and determinants of selective citation in six biomedical research fields

What is new?

Key Findings:

- Citation bias might be quite common and amplifies the well-documented effects of publication bias
- Authority of authors and journals as well as self-citation were also found to be generic determinants of citation

What this adds to what was known?

- Previously, citation analyses have been performed with a variation of methodologies. By using the same method on six networks, we were able to compare citation patterns across six research questions.

What is the implication and what should change now?

- Authors, reviewers and editors should be aware that citation bias is probably common and can hamper a balanced interpretation of study findings.
- The probability of being cited is often not related to the quality of a study. This provides support for the growing concern about using the number of citations as metric in the assessment of researchers and institutes.

Introduction

The functioning of the scientific and scholarly publication system and its impact on correct development of knowledge are increasingly being scrutinized. An important aspect in this regard is publication bias, meaning selective publication of studies with welcome, spectacular or statistically significant results, refraining from publishing studies not belonging to these categories (1). Publication bias is well studied and widely recognized as a problem in scientific and scholarly research (2). Secondly, outcome reporting bias refers to selective reporting of certain findings *within* a publication (3). The current study focuses on the next step in knowledge development, namely citation of earlier research. By selectively citing certain publications, additional to publication and outcome reporting bias, knowledge development can be subtly driven into a certain direction (4). In case the citation or non-citation of publications is based on the nature and direction of the results, is called citation bias (5).

Citations assign priority or ownership of a claim or finding to authors of the cited publication and are used to integrate findings in the existing consensus (6). The number of received citations is often used as metric to evaluate performance of academics (7), institutes and journals (8). Because the reward system focuses on the number of received citations, being cited has become a goal in itself, relating to high publication pressure (9). Consequently, researchers are tempted to cut corners and violate basic methodological principles, to get results that might be cited more often (10, 11). Multiple surveys showed that selectively citing publications to enhance one's own findings and please editors and reviewers is a frequent misbehavior in multiple disciplines (12, 13). Recently the idea that number of publications and citations are a suitable indicator of study quality has been challenged (14). Recommendations to get away from the focus on these metrics are formulated in the San Francisco Declaration on Research Assessment (DORA) and the Hong Kong Principles (15, 16). By evaluating and valuing also other types of scientific output, publication pressure may be reduced and more focus can go to actual scientific quality of studies and researchers. However, time is needed for this research culture to actually be altered.

For correct knowledge development, citations should give a balanced overview of available supportive and non-supportive evidence with regard to a given hypothesis. Apart from study outcome, many other determinants can drive this selection. We realize that not all selective citation leads to biased knowledge development. Study quality is, apart from the subject of the publication, the only clearly justified determinant of citation. Although it is difficult to measure study quality, checklists are available for a number of study designs (17, 18). Proxies for study quality can be found in study design and sample size. Additionally, journal impact factor and authority of the authors might affect the probability of citation. However the journal impact factor is measured at the journal level, which does not necessarily translate to a single publication (7). Research showed that 80% of the journal impact factor is attributable to 20% of the publications (7, 19). Author-related factors that might relate to the probability of citation are gender, affiliation and geographical location of the authors. With regard to gender, contradicting theories exist. Bornmann (2006) hypothesized that the probability of being cited might be higher for men compared to women (8). Contradicting research has suggested that women publish less (20), but make up for that in terms of citations (21). Regarding affiliation, for-profit affiliations are expected to be cited less compared to universities, following lower perceived trust in industry-funded research (22). Regarding the continent where studies are performed, research showed North American authors tend to cite each other, while leaving out others (23). The effect that all these factors have on citation and on knowledge development is expected to vary across research fields.

Consequences of selective citation are difficult to quantify and can occur on different levels. By disregarding counter-evidence, unfounded consensus (24) or polarisation (25) can develop. Citation bias might lead to ill-advised research programmes and contribute to research waste (24, 26). Citation bias can also spread outside the research community, by contributing to distorted information in the media (27), and lead to misplaced medical or societal policy decisions (28).

In a systematic review we demonstrated that citation bias has been reported in several academic fields, with varying magnitudes. Great variability in the methods was observed, contributing to high heterogeneity in the accompanying meta-analysis (29). Learning from these earlier studies, we have optimized our method to assess citation bias and other determinants of citation. The current publication describes the findings of citation analyses we performed for six biomedical research fields.

Methods

Citation analyses were performed for the following six biomedical topics: the association between trans fat intake and serum cholesterol; swimming in chlorinated water in association with childhood asthma; epidemiological research on health effects of bisphenol A; the validity of the hygiene hypothesis for the etiology of asthma; epidemiological research on health effects of phthalates; the association between diesel emission and lung cancer in humans. These topics were selected for several reasons. First, they all related to epidemiological research, so that the authors could understand the methods described in the publications. Second, the topics were selected so that none of the authors of this publication has a conflict of interest, by being active in the topics under study. Third, each network described a clearly delineated research questions, so that each publication had a realistic probability of being cited by the later publications in the network. Finally, we aimed for the network to comprise between 100 and 200 publications. This decision was made for pragmatic reasons, since article selection and data extraction needed to be done manually and in duplo. Each citation analysis was based on a preregistered research protocol, which are available as online supplements to this publication. The data set for each citation analysis can be requested via the corresponding author and via https://dataverse.nl/dataverse/SoundScience.

Our citation analysis method consisted of several steps, which were similar for each of the six subjects. First an extensive literature search was conducted in Web of Science – Core Collection. Although this database is not sufficient to identify all available publications, it is the only database that enabled us to download the references that are needed for the citation networks. Broad search strategies were developed, in consultation with content experts for each of the subjects, to include as many relevant publications as possible. Reference lists of the identified publications were not checked for missing publications, since this would interfere with the research question at hand. Table 1 provides an overview of the in- and exclusion criteria used in each of the subjects, to create networks of publications that can be expected to cite each other. Each network was restricted to English language publications with human participants. Article selection was based on title, abstract and full text, and was done by two researchers independently. Frequent consensus meetings were held between the researchers and in case of disagreement, a third researcher was consulted.

- Insert table 1 -

After the eligible publications had been identified, data were extracted for potential determinants of citation. The following determinants were identified on the basis of previous research: study outcome, sample size, study design, gender, affiliation and continent of the corresponding author, journal impact factor, number of affiliations involved on the publication, authority of the author and citations in the network under study and occurrence of self-citation. Some variations existed in the operationalisation of these variables between the six networks due to topic-specific differences. The code book for data extraction in each of the networks was part of the pre-registered study protocols. Data extraction was done independently by two researchers, with consensus meetings that solved all disagreements.

The networks of publications and citations were created by specialized software, called CitNetExplorer (30), by downloading the publications and their reference lists from Web of Science. To assess which factors influenced the probability of citation, we required an overview of all *potential* citation pathways and of the *performed* citations. A potential citation pathway exist when the online publication date of the cited publication was before the submission date of the citing publication. Only a selection of all potential citation pathways have actually taken place. The actual citations, as mentioned in the reference lists, were therefore addressed in this study as *performed* citations.

In terms of statistical analysis, random effect logistic regression was performed. The outcome variable was measured dichotomously as a performed or not performed citation. Univariate analyses were performed for each of the aforementioned determinants. The following continuous determinants were reduced to three categories: sample size, number of affiliations, journal impact factor, number of references and authority of the author. Categorisation was done based on the tertiles of the continuous variables and was done to deal with their large continous variation and non-parametric distribution. Each univariate analysis was adjusted for study design, as a categorical variable. Systematic reviews were thereby considered of the highest quality, followed by intervention studies, observational studies and narrative reviews. A random effect model was used to deal with the hierarchical data structure of the citation network. Namely, citation networks have a tree-like structure, consisting of the level of the publication and the level of the citation. Multiple citations are bundled in one publication. Because of this structure, the citations are not completely independent of each other, since they belong in groups to the same publication, sharing the same characteristics. Additional to this random effect analyses, we assessed whether authors are more likely to cite publications that share the same characteristics. We describe this as concordance between the cited and citing publication. If the cited and citing publication have the same score on a determinant, e.g. if both publications have a supportive outcome, this was scored as concordant. By means of fixed effect logistic regression, we assessed whether concordance between the cited and citing publication influenced the probability of citation in comparison to the non-concordant situation in which the cited and citing publication did not share the same characteristic. All statistical analyses have been performed in Stata 13.

Results

Table 2 displays basic characteristics of the six citation networks under study. The results of each citation analysis have been published earlier (31-35). Five out of six networks contain around 100 publications or more, with a citation prevalence between 6% and 15%. Number of citations per publication is very skewly distributed, as displayed by the median and IQR in table 2. The supportive publications outnumbered the non-supportive publications in each network, suggesting presence of publication bias. A publication was defined supportive when the authors' conclusion was in line with the prevailing hypothesis in that field.

- insert table 2 -

An overview of the descriptive statistics for all determinants in the six networks can be found in supplementary table 1. Table 3 provides an overview of odds ratios, adjusted for study design, for each of the assessed determinants of citation, over the six networks. With this table, we aim to create a schematic overview of the magnitude and signifance level of all tested associations, across the six networks. Four networks found a significant, positive association between studies reporting supportive findings and the probability of citation. Five networks found a positive correlation between sample size and the probability of citation. Journal impact factor and the authority of the author were positively correlated to citation in all six networks. To avoid interference between these determinants and the probability of being cited in our analyses, journal impact factor and the authority of all authors were measured at the moment the cited article was published, so its citations were not included in the calculation of the impact factor and authority. Other determinants under study showed substantial variation in their association to the probability of citation across these networks.

- insert table 3 -

Additionally we looked at the concordance between the cited and citing publication in relation to the probability of citation. Table 4, reports the results of these analyses.

- insert table 4 -

Interestingly, we did not find that authors are more likely to cite publications with the same study conclusion. This means both supportive and non-supportive studies are likely to cite supportive publications more often. Self-citation, defined as at least one common author on the citing and cited publication, was significantly associated with the probability of citation in all networks. Authors were three to six times more likely to cite their own publications compared to the publications of others. Finally, authors are more likely to cite studies performed on the same continent.

Discussion

Aim of this study was to describe determinants of citation across six biomedical fields. Four out of six networks showed a statistically significant higher probability of being cited for supportive publications compared to non-supportive publications. The magnitude of this association varied among the networks, finding a 40 to 300% higher probability of being cited for a positive study in comparison to a negative study.. Authority of the author and the journal impact factor have also been found significant determinants of citation across the six fields, together with self-citation. It is difficult to determine the potential harm of self-citation on knowledge development. When the citation is done in the methods section, to refer to an earlier used methodology, this may not lead to bias. If self-citations are used in the introduction or discussion, to stress one's own conviction, this could lead to bias. Looking at the content of studies, five networks showed a significantly higher probability of being cited for publications with a large sample size compared to publications with less participants. This is good news, given that a higher sample size could be seen as a proxy of the quality of a study. With regard to study design, results were different than expected. We expected that reviews would have a higher probability of being cited compared to empirical studies, following the idea that empirical studies will be substantially less cited once they have been taken up in a (systematic) review. Nevertheless, in four networks empirical studies actually had a significantly higher probability of being cited compared to reviews. Potentially these findings can be explained by the high number of *narrative* reviews in the networks compared to *systematic* reviews. Due to the low number of systematic reviews in each network, it was not possible to analyse their effect seperately.

In interpreting and comparing the findings of the six citation networks we learned that each research topic had very specific characteristics. The field of trans fats and cholesterol has a consensus that trans fats increase LDL-cholesterol and decrease HDL-cholesterol (36). This was strongly driven by the first human publication in this network (37), which attracted many citations within the network. The network on chlorinated swimming water and childhood asthma was known for two opposing research groups. These two groups do cite each other, showing no evidence for citation bias, but probably with the goal to refute each other's findings. The networks on bisphenol A and phthalates included multiple health outcomes. Therefore multiple subnetworks exist within the studied pool of publications. In the public debate, BPA and phthalates are labeled as endocrine disruptors, being harmful for human health, without making reference to specific health outcomes (38). In the network on BPA 60 out of 169 publications did not receive any citations, indicating research waste. In the network on diesel emission and lung cancer, industry involvement played a large role. The diesel industry criticized findings from studies that showed harmful effects of diesel exhaust exposure, referring to methodological flaws in these studies. We learned that these network characteristics impact the citation dynamics in a field and need to be taken into account in performing and interpreting citation analyses.

Limitations

We recognise our study has several limitations. Firstly, we have created the networks only on the basis of Web of Science, making it likely that not all available publications have been identified. However, this database was the only one who enabled us to set up a citation network, by downloading all the underlying reference lists. Nevertheless, we think that the networks as created give a good representation of the available literature and we have no reason to believe that the determinants of citation would have been different if literature from other sources had been included. Second, we only adjusted for study design as potential confounder. Initially we planned to include all determinants that were statistically significant at univariate level in the multivariate analysis. However, this led to very wide confidence intervals, making its interpretation very difficult. Based on several tests with study quality checklists, we learned that most variation in quality exists between study designs instead of within designs. Another limitation lies in the fact that continuous variables were reduced to three categories. This was done to improve the interpretation of the results, since there was a large spread in the continuous variables. We recognise this might lead to a reduction of statistical power. On the other hand, literature suggests that the number of type I and type II errors are not impacted when no multicollinearity is expected (39), which is not the case in our analyses.

Conclusion

This research showed that the probability of being cited was associated with study outcome in several research fields, indicating evidence for citation bias. Additionally, the odds of being cited was higher with increased journal impact factor, authority of the author and self-citation, across all six fields under study. With regard to other determinants of citation, many differences between the six biomedical topics were identified. We have learned that each citation network has specific characteristics that impact the data extraction and interpretation of the results, such as relevant confounders and influential publications. Concluding, attention should be paid to balanced citations of publications, to make sure knowledge development takes place in a balanced manner. In this publication we studied citations in a quantitative way. Further, qualitative, studies are needed to also acquire insight in the content and correctness of citations in their effect on knowledge development.

References

1. Rosenthal R. The file drawer problem and tolerance for null results. Psychological bulletin. 1979;86(3):638.
2. Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan A-W, Cronin E, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PloS one. 2008;3(8):e3081.
3. Swaen GM, Urlings MJ, Zeegers MP. Outcome reporting bias in observational epidemiology studies on phthalates. Annals of epidemiology. 2016;26(8):597-9. e4.

4. De Vries Y, Roest A, de Jonge P, Cuijpers P, Munafò M, Bastiaansen J. The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: the case of depression. Psychological medicine. 2018;48(15):2453-5.
5. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions: John Wiley & Sons; 2019.
6. Nigel Gilbert G. Referencing as persuasion. Social studies of science. 1977;7(1):113-22.
7. Neylon C, Wu S. level metrics and the evolution of scientific impact. PLoS biology. 2009;7(11):e1000242.
8. Bornmann L, Daniel H. What do citations measure? A review of studies on citing behavior. Journal of Documentation. 2006;64(1):45-80.
9. Haven TL, Bouter LM, Smulders YM, Tijdink JK. Perceived publication pressure in Amsterdam: Survey of all disciplinary fields and academic ranks. PloS one. 2019;14(6).
10. Bouter LM. Commentary: Perverse incentives or rotten apples? Accountability in research. 2015;22(3):148-61.
11. Anderson MS, Ronning EA, De Vries R, Martinson BC. The perverse effects of competition on scientists' work and relationships. Science and engineering ethics. 2007;13(4):437-61.
12. Bouter LM, Tijdink J, Axelsen N, Martinson BC, ter Riet G. Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. Research Integrity and Peer Review. 2016;1(1):17.
13. Haven T, Tijdink JK, Pasman HR, Widdershoven G, Bouter L. Do research misbehaviours differ between disciplinary fields? A mixed methods study among academic researchers in Amsterdam. 2019.
14. Seglen PO. Citations and journal impact factors: questionable indicators of research quality. Allergy. 1997;52(11):1050-6.
15. Cagan R. The San Francisco declaration on research assessment. The Company of Biologists Ltd; 2013.
16. Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. The Hong Kong principles for assessing researchers: Fostering research integrity. OSF Preprints. 2019;17.
17. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. European journal of epidemiology. 2010;25(9):603-5.
18. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. Journal of clinical epidemiology. 1998;51(12):1235-41.
19. Campbell P. Escape from the impact factor. Ethics in science and environmental politics. 2008;8(1):5-7.
20. West JD, Jacquet J, King MM, Correll SJ, Bergstrom CT. The role of gender in scholarly authorship. PloS one. 2013;8(7):e66212.
21. Eisenberg N, Martin CL, Fabes RA. Gender development and gender effects. 1996.
22. DeAngelis CD. Conflict of interest and the public trust. Jama. 2000;284(17):2237-8.
23. Allik J. Bibliometric analysis of the journal of cross-cultural psychology during the first ten years of the new millennium. Journal of Cross-Cultural Psychology. 2013;44(4):657-67.

24. Greenberg SA. How citation distortions create unfounded authority: Analysis of a citation network. BMJ. 2009 Jul;339. PubMed PMID: WOS:000268351400004.
25. Trinquart L, Johns DM, Galea S. Why do we think we know what we know? A metaknowledge analysis of the salt controversy. Int J Epidemiol. 2016 Feb;45(1):251-60. PubMed PMID: WOS:000374230100032.
26. Andrade NS, Flynn JP, Bartanusz V. Twenty-year perspective of randomized controlled trials for surgery of chronic nonspecific low back pain: Citation bias and tangential knowledge. Spine J. 2013 Nov;13(11):1698-704. PubMed PMID: WOS:000327431800043.
27. Chapman S, Ragg M, McGeechan K. Citation bias in reported smoking prevalence in people with schizophrenia. Aust NZ J Psychiat. 2009;43(3):277-82. PubMed PMID: WOS:000263366700011.
28. Koren G, Nickel C. Perpetuating fears: bias against the null hypothesis in fetal safety of drugs as expressed in scientific citations. J Popul Ther Clin Pharmacol. 2011 2011 (Epub 2011 Jan;18(1). PubMed PMID: MEDLINE:21289377. English.
29. Duyx B, Urlings MJ, Swaen GM, Bouter LM, Zeegers MP. Scientific citations favor positive results: a systematic review and meta-analysis. Journal of clinical epidemiology. 2017;88:92-101.
30. Van Eck NJ, Waltman L. CitNetExplorer: A new software tool for analyzing and visualizing citation networks. Journal of Informetrics. 2014;8(4):802-23.
31. Urlings MJ, Duyx B, Swaen GM, Bouter LM, Zeegers MP. Citation bias in the literature on dietary trans fatty acids and serum cholesterol. Journal of clinical epidemiology. 2019;106:88-97.
32. Duyx B, Urlings MJ, Swaen GM, Bouter LM, Zeegers MP. Selective citation in the literature on swimming in chlorinated water and childhood asthma: a network analysis. Research integrity and peer review. 2017;2(1):17.
33. Duyx B, Urlings MJ, Swaen GM, Bouter LM, Zeegers MP. Selective citation in the literature on the hygiene hypothesis: a citation analysis on the association between infections and rhinitis. BMJ open. 2019;9(2):bmjopen-2018-026518.
34. Urlings M, Duyx B, Swaen G, Bouter L, Zeegers M. Selective citation in scientific literature on the human health effects of bisphenol A. Research integrity and peer review. 2019;4(1):6.
35. Urlings MJ, Duyx B, Swaen GM, Bouter LM, Zeegers MP. Determinants of Citation in Epidemiological Studies on Phthalates: A Citation Analysis. Science and Engineering Ethics. 2020:1-15.
36. Hall JE. Guyton and Hall textbook of medical physiology e-Book: Elsevier Health Sciences; 2015.
37. Mensink RP, Katan MB. Effect of dietary trans fatty acids on high-density and low-density lipoprotein cholesterol levels in healthy subjects. New England Journal of Medicine. 1990;323(7):439-45.
38. Rochester JR. Bisphenol A and human health: a review of the literature. Reproductive toxicology. 2013;42:132-55.
39. Iacobucci D, Posavac SS, Kardes FR, Schneider MJ, Popovich DL. The median split: Robust, refined, and revived. Journal of Consumer Psychology. 2015;25(4):690-704.

Table 1: Overview of the inclusion and exclusion criteria for each of the six citation networks

Citation networks	Inclusion criteria	Exclusion criteria

Trans fatty acids and cholesterol	- The effect of industrially produced trans fatty acids on LDL- and/or HDL-cholesterol is reported as one of the study outcomes. - Study design is observational (cross-sectional or cohort), intervention study, review or other type of synthesis paper (editorial, commentary etc.) in humans	- Publications solely including ruminant trans fatty acid, dairy products or conjugated linoleic acid as determinant - Study outcomes not directly related to cholesterol (e.g. sudden cardiac death, gut inflammation, cancer etc.) - Determinants other than dietary intake of fatty acids (e.g. erythrocyte, plasma lipid, circulating lipids, etc.)
Swimming in chlorinated water and childhood asthma	- Publications assessing the relation between swimming in indoor chlorinated pools and asthma - Study design is observational (cross-sectional, cohort, case control, ecological), experimental, review or other type of synthesis paper in humans - Asthma can be assessed in several ways: by doctor's diagnosis, self-assessment (or parents' assessments), asthma-related symptoms, lung tests, and blood biomarkers	- Publications not including indoor chlorinated swimming pools, or in which indoor chlorinated pools can not be clearly distuinguished from outdoor pools or non-chlorinated pools. - Publications based on non-swimmers, e.g. swimming pool workers. - Publications not including 18- swimmers, or in which 18- swimmers can not be clearly distuinguished from 18+ swimmers
Epidemiological research on bisphenol A	- Epidemiological studies on human subjects studying the association between bisphenol A and any health effect - Study designs is cohort study, cross sectional study, case control study, experimental study, narrative review and systematic review with or without meta-analysis	- In vitro or animal studies on the potential health effects of bisphenol A - Ecological studies on exposure rates of bisphenol A - Policy documents on the regulation of availability of bisphenol A in the environment
Hygiene hypothesis and etiology of asthma	- Articles assessing evidence for (or against) the hygiene hypothesis. One or both of the following health outcomes must be measured: asthma, or hay fever/allergic rhinitis. Determinant is exposure to bacteria and/or allergens. This exposure can be measured by proxy. - Study design is observational (cross-sectional, cohort, case-control, ecological), experimental, review, or editorial in humans.	- Basic studies that focus on the mechanism by which hygiene has an impact on allergy - Meeting abstracts, news items, no fulltext available, not peer-reviewed articles, book chapters - Articles mainly focused on auto-immune diseases (diabetes) or inflammatory bowel disease.
Epidemiological research on phthalates	- Studies on human health effects of MEHP - Study designs is cohort study, cross sectional study, case control study, experimental study, narrative review and systematic review with or without meta-analysis	- *In vitro* or animal studies on the potential health effects of phthalates - Ecological studies describing exposure rates of phthalates - Policy documents on the regulation of availability of phthalates in the environment
Diesel emission and lung cancer	- Publications providing or assessing evidence on the association between diesel exposure and cancer. At least one of the following health outcomes must be evaluated: bladder cancer or lung cancer. Determinant is diesel exposure. This exposure can be measured by proxy or job occupation. - Study design is cross-sectional study,	- Meeting abstracts, news items, no fulltext available, non-peer-reviewed publications, book chapters - Basic studies that focus on the mechanism by which diesel exposure has an impact on bladder cancer or lung cancer.

	cohort study, case-control study, experimental study, or review (narrative, systematic, meta-analysis) in humans	

Table 2. Overview of basic characteristics of six citation networks

	Trans fatty acid and cholesterol	Chlorinated water and asthma	Bisphenol A and health	Hygiene hypothesis	Phthalates and health	Diesel exposure and lung cancer
Number of publications in each network	108	36	169	110	112	96
Number of potential citation pathways among the publications in the network	5041	570	12432	5551	5684	4317
Percentage performed citations in the network	13%	34%	6%	7%	10%	16%
Median number (IQR) of citations per publication within the network	2 (0-7)	4 (0-8)	1 (0-5)	1 (0-4)	2 (0-7)	5 (1-10)
Highest number of citations received by one publication	73	26	64	35	33	34
Time period of publication years	1990-2015	2002-2015	2002-2017	1995-2017	2000-2018	1988-2017
Number of supportive vs non-supportive publications*	86 vs 16	16 vs 10	92 vs 28	41 vs 35	35 vs 30	51 vs 34

** The number of supportive and non-supportive publications do not add up to the number of total publications in each network. Reason for this discrepancy is that a number of publications did not report one clear supportive or non-supportive conclusion. These inconclusive or mixed publications were excluded from the last row of the table.*

Tabel 3: Overview of odds ratios, adjusted for study design, of the association between various determinants and the probability of being cited in six different networks

Determinant	Categories	Trans fatty acids and cholesterol	Swimming in chlorinated water and childhood asthma	Bisphenol A and health	Hygiene hypothesis	Phthalates and health	Diesel emmission and lung cancer
Statistical significance	Significant vs non-significant	**3.2 (2.5-4.2)**	1.3 (0.8-2.3)	**1.5 (1.2-1.8)**	-*	1.0 (0.7-1.3)	-*
Authors'	supportive vs.	**2.4 (1.9-3.1)**	1.4 (0.9-2.3)	**1.7 (1.3-2.0)**	**3.1 (2.2-4.3)**	0.8 (0.7-0.9)	**1.4 (1.1-1.7)**

conclusion	Non-supportive						
Study design**	Empirical vs review	**3.9 (3.2-4.8)**	**3.6 (2.3-5.7)**	**1.6 (1.3-1.9)**	**4.3 (3.2-5.7)**	1.1 (0.9-1.4)	1.1 (0.9-1.4)
Sample size**	High vs low	**7.7 (4.6-13)**	**6.4 (3.1-13)**	**1.4 (1.1-1.7)**	**1.9 (1.2-3.0)**	0.7 (0.6-0.8)	**3.3 (2.4-4.4)**
Title of publication	Conclusive vs non-conclusive	0.9 (0.7-1.2)	1.1 (0.7-1.7)	**1.2 (1.0-1.4)**	0.3 (0.2-0.4)	**1.3 (1.1-1.5)**	0.8 (0.5-1.1)
Number of affiliations	High vs low	1.1 (0.9-1.4)	1.4 (0.9-2.3)	**1.3 (1.1-1.7)**	**1.6 (1.1-2.2)**	1.1 (0.9-1.3)	**3.4 (2.6-4.3)**
Journal impact factor	High vs low	**5.4 (3.7-7.8)**	**1.7 (1.1-2.8)**	**1.2 (1.1-1.4)**	**4.9 (3.2-7.6)**	**1.5 (1.3-1.8)**	**4.0 (3.0-5.5)**
Funding source	Profit vs not-for-profit	1.2 (0.9-1.6)	**2.0 (1.2-3.5)**	-*	0.8 (0.6-1.2)	0.9 (0.7-1.2)	**1.4 (1.0-1.9)**
Number of references	High vs low	0.8 (0.6-1.0)	**6.2 (3.1-12)**	0.8 (0.7-0.9)	0.8 (0.5-1.3)	0.8 (0.6-1.0)	**1.7 (1.3-2.2)**
Gender	Male vs female	**1.5 (1.3-1.9)**	**1.8 (1.1-2.8)**	1.0 (0.9-1.1)	0.7 (0.6-1.0)	**1.4 (1.2-1.7)**	1.0 (0.8-1.2)
Affilation	all other vs university	1.2 (0.8-2.0)	0.8 (0.5-1.3)	1.3 (0.6-3.0)	**2.0 (1.5-2.5)**	0.7 (0.6-0.8)	1.3 (1.0-1.5)
Continent	N.-America vs Europe	-*	1.2 (0.7-2.2)	**1.6 (1.3-2.0)**	0.9 (0.5-1.4)	1.2 (1.0-1.4)	**1.8 (1.5-2.2)**
Authority	High vs low	**6.4 (4.7-8.9)**	**3.6 (1.9-7.1)**	**3.3 (2.6-4.2)**	**2.7 (2.0-3.7)**	**1.6 (1.3-2.1)**	**4.9 (3.7-6.5)**

** Missing values, due to insufficient data to perform the respective analysis. ** For study design and sample size, only the crude OR are reported.*

Bold figures indicate statistically significant findings in the expected direction (p<0.05)

Table 4. Concordance analyses for six citation networks

	Trans fatty acids and serum cholesterol	**Swimming in chlorinated water and childhood asthma**	**Bisphenol A and health**	**Hygiene hypothesis**	**Phthalates and health**	**Diesel exposure and lung cancer**
Author conclusion	**1.9 (1.6-2.4)**	1.6 (1.0-2.6)	1.1 (0.8-1.4)	**3.4 (1.6-7.1)**	0.9 (0.7-1.1)	1.2 (1.0-1.4)
Study design	**1.3 (1.0-1.5)**	1.4 (1.0-1.9)	**1.4 (1.2-1.6)**	1.1 (0.9-1.3)	**1.8 (1.4-2.2)**	0.4 (0.3-0.5)
Journal impact factor	0.5 (0.1-2.2)	0.9 (0.6-1.3)	**1.3 (1.1-1.6)**	1.0 (0.8-1.3)	0.8 (0.7-1.0)	1.2 (0.9-1.5)
Funding source	0.8 (0.6-1.0)	1.1 (0.6-2.0)	**1.5 (1.2-1.9)**	1.2 (0.8-1.8)	1.0 (0.8-1.2)	0.7 (0.5-0.9)
Gender	1.1 (0.9-1.3)	1.1 (0.8-1.6)	1.1 (0.9-1.3)	1.0 (0.8-1.3)	1.2 (1.0-1.5)	1.1 (0.9-1.3)
Affiliation of corresponding author	1.1 (0.9-1.3)	**2.1 (1.0-4.4)**	0.9 (0.7-1.1)	0.8 (0.7-1.0)	1.0 (0.9-1.2)	1.0 (0.8-1.2)
Continent	-*	-*	-*	**2.0 (1.6-2.5)**	**1.4 (1.2-1.7)**	**1.8 (1.5-2.1)**
Self-citation	-*	**5.4 (3.2-9.2)**	**5.2 (3.8-7.0)**	**6.1 (3.7-9.9)**	**3.2 (2.5-4.1)**	**4.1 (2.9-5.7)**

** Missing values, due to insufficient data to perform the respective analysis.*

Bold figures are significant in the expected direction (p<0.05)

Citation bias and other determinants of citation in biomedical research: findings from six citation networks

M.J.E. Urlings[1*], B. Duyx[1], G.M.H. Swaen[1], L.M. Bouter[2,3], M.P. Zeegers[1,4]

[1] Nutrition and Translational Research in Metabolism (School NUTRIM), Maastricht University, Maastricht, The Netherlands.

[2] Department of Epidemiology and Data Science, Amsterdam University Medical Centers, Amsterdam, The Netherlands

[3] Department of Philosophy, Faculty of Humanities, Vrije Universiteit, Amsterdam, The Netherlands

[4] Care and Public Health Research Institute (School CAPHRI), Maastricht University, Maastricht, The Netherlands.

* Corresponding author, mje.urlings@maastrichtuniversity.nl

Conflict of Interest

All authors declare that they have no conflict of interest (other than the project funding received for this study; see Funding statement).

Funding

This work was supported by the Long-range Research Initiative (LRI) from the European Chemical Industry Council (CEFIC), grant no. LRI-Q3-UM. LRI had no role in study design, data collection and analysis, preparation of the manuscript, or decision to publish.

Author contributions

All authors have been involved in the development of the citation analyses methodology. Funding was obtained by GMHS and MPZ. Data collection and analysis on each of the citation networks was performed in collaboration by M.J.E.U., B.D., and G.M.H.S. All co-authors were actively involved in the interpretation of the data. Drafting this article was performed by M.J.E.U.; the article was critically revised by all co-authors. All co-authors approved the final version.

Citation bias and other determinants of citation in biomedical research: findings from six citation networks

M.J.E. Urlings[1*], B. Duyx[1], G.M.H. Swaen[1], L.M. Bouter[2,3], M.P. Zeegers[1,4]

- Key points -

What is new?

Key Findings:

- Citation bias might be quite common and amplifies the well-documented effects of publication bias
- Authority of authors and journals as well as self-citation were also found to be generic determinants of citation

What this adds to what was known?

- Previously, citation analyses have been performed with a variation of methodologies. By using the same method on six networks, we were able to compare citation patterns across six research questions.

What is the implication and what should change now?

- Authors, reviewers and editors should be aware that citation bias is probably common and can hamper a balanced interpretation of study findings.
- The likelihood of being cited is often not related to the quality of a study. This provides support for the growing concern about using the number of citations as metric in the assessment of researchers and institutes.

Citation bias and other determinants of citation in biomedical research: findings from six citation networks

M.J.E. Urlings[1*], B. Duyx[1], G.M.H. Swaen[1], L.M. Bouter[2,3], M.P. Zeegers[1,4]

- Conflict of interest statement -

Conflict of Interest

All authors declare that they have no conflict of interest (other than the project funding received for this study; see Funding statement).

Funding

This work was supported by the Long-range Research Initiative (LRI) from the European Chemical Industry Council (CEFIC), grant no. LRI-Q3-UM. LRI had no role in study design, data collection and analysis, preparation of the manuscript, or decision to publish.

Author contributions

All authors have been involved in the development of the citation analyses methodology. Funding was obtained by GMHS and MPZ. Data collection and analysis on each of the citation networks was performed in collaboration by M.J.E.U., B.D., and G.M.H.S. All co-authors were actively involved in the interpretation of the data. Drafting this article was performed by M.J.E.U.; the article was critically revised by all co-authors. All co-authors approved the final version.