The Royal Society of Medicine

Journal of the Royal Society of Medicine; 2022, Vol. 115(1) 31–35
DOI: 10.1177/01410768221075881

# Citation bias: questionable research practice or scientific misconduct?

**Peter C Gøtzsche**
Institute for Scientific Freedom, 2970, Copenhagen, Denmark
**Corresponding author:** Peter C Gøtzsche. Email pcg@scientificfreedom.dk

## Citation bias

Citation bias occurs when authors preferentially cite research that supports their own findings or claims, or research that showed what they had hoped to find but didn't find in their research. In research articles, citation bias may occur in the Introduction section, where the researchers argue why their own research is important, and in the Discussion section, where they put their findings into context and perspective.

The first use of the term 'citation bias', of which I am aware, was not in biomedicine. In 1985, researchers in physics referred to 'a citation bias against Eastern-bloc [*particle*] accelerators'.[1]

## Demonstration of citation bias in a systematic review of trials of non-steroidal anti-inflammatory drugs for rheumatoid arthritis

In 1987, I reported, in an article published in the *British Medical Journal*, what seems likely to be the first demonstration of citation bias in a systematic review of trials in healthcare.[2] For my doctoral thesis, I used exhaustive search strategies, read the references of the reports identified, and wrote to manufacturers in an attempt to assemble all published and unpublished reports of double-blind trials that had compared two or more non-steroidal anti-inflammatory drugs in patients with rheumatoid arthritis.

I found a wide variety of biases favouring sponsors' drugs and disfavouring comparator drugs. When bias in the Conclusions or Abstracts consistently favoured one of the drugs, it favoured the control drug in only one report and the new drug in the remaining 81 reports ($P = 3.4 \times 10^{-23}$).[3] This observation was why I became interested in assessing whether trial reports were also biased when they cited earlier, similar trials.

I examined the reference lists for studies that had compared the same two drugs as those in the index trial report. For each article, I noted whether the proportion of references with a positive outcome for the new drug was the same, lower, or higher than the proportion among all articles assumed to have been available to the authors (those published more than two years earlier than the index article). Ten articles had a neutral selection of references, 22 a negative selection and 44 a positive selection ($P < 0.01$; sign test). The bias was not caused by overrepresentation of highly cited journals among the articles with positive selection of references, or by better methodological quality of the cited articles. And the trials that were least cited were not published in journals or books that are difficult to identify in a search, or to obtain through a library.

I concluded my report as follows:

> "The reference bias shown in this study seems to be real. Such a finding has important implications, since there is no reason to believe that rheumatologists are more biased than others in selecting references. A reader tracing the literature on any new drug using the reference lists given in the articles might risk obtaining a biased sample. Reference bias has another serious implication: it may render the conclusion of the individual article less reliable. Is this also true for review articles, and for other disciplines in medicine?"[2]

I called the bias 'reference bias,' but 'citation bias' is a better term. When I did a PubMed search on these terms in the title of articles (in quotation marks; 27 August 2021), I retrieved 16 articles using 'reference bias' and 21 articles using 'citation bias.' The only relevant records for 'reference bias' were two of my own papers,[2,4] a letter to the editor about the first one and my translation of the first one into Danish. The other 12 articles with 'reference bias' in the title were about problems related to the use of a 'reference genome' or 'reference tomography,' or to the use of self-reports of health or quality of life, for example when patients used

themselves as the reference. In contrast, all 21 records retrieved using 'citation bias' as the search term were relevant, although three were comments on other papers, one was about handsearching literature, one was about gender bias, and one was authored by Italian researchers claiming that they were being cited less frequently than they should have been.

## Cholesterol and coronary heart disease

In 1992, Uffe Ravnskov published an analysis based on 14 cholesterol-lowering trials regarded by the trial directors as supportive of beneficial effects, and 10 were considered unsupportive.[5] Ravnskov found that the trials deemed to be supportive were cited almost six times more often than other trials, and that unsupportive trials were not cited at all after 1970, even though they were similar in number to those considered supportive.

Criticism of the role of a low-fat diet for preventing heart disease is often met with the assertion that consensus committees have settled the issue unanimously. Using three authoritative reviews, Ravnskov studied the work of such committees.[6] As he found that fundamental parts of the hypothesis seemed to be based on biased quotation, he used the term 'quotation bias' in the title of his article.

## Overcitation of unsupportive studies

Rarely, citation bias goes in the opposite direction. A 1995 study found that five reviews cited more unsupportive than supportive trials of the effectiveness of pneumococcal vaccines; two cited more supportive trials, and one cited an equal proportion of supportive and unsupportive trials.[7] Overall, unsupportive trials were twice as likely to be cited as supportive trials, but there was also a time issue. Results of all seven trials in adults published before 1980 were supportive, whereas six of the seven trials published in 1980 or later were unsupportive. The citation pattern might therefore reflect a tendency of authors to preferentially cite recent studies.

## A review of studies designed to detect citation bias

During the next 20 years, an average of about two studies designed to detect citation bias were published annually. In 2017, a Dutch research group[8] published a systematic review of 52 studies, from across scientific disciplines, designed to detect citation bias. The authors claimed to have done 'the first systematic review of citation bias' and that they had taken account of 'all available evidence'. Because they had not included my 1987 report,[2] I downloaded the protocol for their review[9] to explore how this could have happened. One of their search terms was 'referenc* bias*.' When I used that term to search PubMed (17 August 2021), I retrieved my 1987 paper and another systematic review from my research group which the Dutch researchers had also failed to include,[4] even though both papers had 'reference bias' in their titles. Although the Dutch researchers claimed to have checked the reference lists of the reports they had included, 3 of the 36 biomedical studies of citation bias which they had included (their references 18, 22 and 29) had quoted my 1987 report.

Duyx et al. found that positive articles were cited 'about 1.3 to 3.7 times more often' than negative articles and that statistically significant articles were cited 1.6 times (95% confidence interval [CI] 1.3–1.8) as often as statistically nonsignificant articles. When they looked at the direction of the results and whether they supported the investigators' hypotheses, they found pooled ratios of 2.1 (95% CI 1.3–3.6) and 1.8 (95% CI 1.4–2.4), respectively. Articles with a positive conclusion that supported the investigators' conclusion were cited 2.7 times (2.0–3.7) more often than others. Research quality was not related to the number of citations, whereas journal impact factor was.

One of the reports included by the Dutch researchers had assembled 458 eligible articles.[10] It also found that statistically significant studies were cited twice as often as statistically non-significant studies, but that the association disappeared after adjustment for journal impact factor. Jannot et al. were aware that the journal impact factor may be considered an intermediary causal variable between statistical significance and citation frequency and that it is therefore wrong to adjust for journal impact factor. If the association is strong, an adjustment for an intermediary causal factor along the causal pathway could remove completely any true association between the primary causal factor and the outcome. The authors of another report included in the Dutch review had not understood this. They declared that 'The prestige and visibility of journals is a potential confounding factor, which should be adjusted for', which they then did.[11]

The Dutch team went on to conduct a citation network analysis based on 108 articles.[12] They judged that disproportionate attention had been paid to articles suggesting a harmful effect of trans fat on cholesterol. Reporting statistically significant results was a strong predictor of citation, together with sample size, journal impact factor and the 'authority' of the authors.

Most recently, the Dutch team assessed citation bias in six research areas.[13] They concluded that '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.' They illustrated themselves that 'authority' plays a role, as 12 of their 39 references were to papers co-authored by the group's senior researcher, Lex Bouter. Self-citation cannot be avoided if one has done most of the relevant research in an area, but this is not the case here. Several of their self-citations were of questionable relevance to the points being made.

## Case study: 'Of mites and men'

In 1998, my research group published a systematic review of chemical and physical interventions intended to reduce antigens against house dust mites.[14] Our findings, that the interventions did not have an effect on patients with asthma, were very robust but unwelcome everywhere, including within the Cochrane Collaboration.[15]

The citation bias in this field is extreme. We documented this in a 2005 systematic review of narrative review articles in which opinions had been expressed about the clinical effects of physical or chemical interventions.[4] We titled our article 'Of Mites and Men', inspired by John Steinbeck's novel *Of Mice and Men.*

Narrative review articles on this topic usually asserted that several methods were effective. We judged positive bias to have been present if the reference list contained a higher proportion of references to trials with statistically significant results favouring the interventions than the proportion among all trials judged to have been available to the authors (published two years or more prior to the review). Of the 38 narrative reviews that recommended physical interventions, 10 had neutral selection of trial references, 1 negative selection and 27 positive selection ($P = 2 \times 10^{-8}$, sign test).

The most quoted trial had only seven patients per group. The statistically significant benefit claimed for this trial was erroneous and was not based on a clinical outcome. The recommendations were often based on non-randomised studies, of which the most quoted study had only 10 patients per group, yet still claimed very positive results. In contrast, the most recent version of our review has 55 randomised trials and a total of 3121 patients.[16]

## Questionable research practice or scientific misconduct?

Citation bias is a questionable research practice, and it is sometimes so gross that it amounts to scientific misconduct. According to the *British Medical Journal*, scientific misconduct includes deceptive selective reporting of findings and omission of conflicting data; wilful suppression or distortion of data; serious deviation from accepted practices in proposing or carrying out research; and improper reporting of results (https://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/scientific-misconduct; accessed 18 August 2021). The *British Medical Journal* also considers failures of transparency to be forms of scientific misconduct. Reporting is deceptive when citations do not support what is claimed or when the available evidence contradicts what is claimed, but these references are not quoted, or they are misrepresented.

The studies that have been performed since my 1987 *British Medical Journal* article show that citation bias is common. This was confirmed in a 2016 survey conducted by Lex Bouter. People attending international research integrity conferences ranked selective citation to enhance one's own findings or convictions, or to please editors, reviewers or colleagues, as the most frequently occurring form of research misbehaviour.[17]

A review of 1523 trial reports published from 1963 to 2004 found that fewer than a quarter of preceding trials had been cited, comprising fewer than a quarter of the participants enrolled in all relevant prior trials.[18] Potential implications of this include ethically unjustifiable trials, wasted resources, incorrect conclusions and unnecessary risks for trial participants.

A cumulative meta-analysis of trials of intravenous streptokinase for myocardial infarction showed that a consistent, statistically significant reduction in total mortality was achieved in 1973 after only eight trials involving 2432 patients had been reported.[19] The results of the 25 subsequent trials, which enrolled an additional 34,542 patients through 1988, had little or no effect on the odds ratio, but exposed the enrolled patients to an increased risk of death.

Another cumulative meta-analysis indicated that aprotinin greatly decreased the need for perioperative blood transfusion, stabilising at an odds ratio of 0.25 by the 12th study.[20] Citation of previous trials was extremely low, with a median of 20% of prior trials cited. Only 7 of 44 (15%) subsequent reports referenced the largest trial (N = 1784), which was 28 times larger than the median trial size.

As these examples illustrate, it is very important to discuss other trials comprehensively and in an unbiased way. The CONSORT guideline states that this can best be achieved by including a formal systematic review in the Results or Discussion sections of the report.[21] When it is impractical, it is often possible to quote a systematic review of similar

trials. At a minimum, the discussion should be as systematic as possible and be based on a comprehensive search, rather than being limited to studies with results that are concordant with those of the current trial.

The Introduction in the report of a clinical trial and in protocols for trials is also important and should also be systematic, preferably including reference to systematic reviews of relevant existing evidence.[22] An early example of this was the 1986 report of the first International Study of Infarct Survival trial, which assessed the effects of beta-blockade in myocardial infarction.[23] In the Discussion section, the result of the trial was combined with a recent review of similar drugs. A decade later, we mentioned in the Introduction of our 1995 report of a trial of somatostatin for bleeding oesophageal varices the two trials that had been carried out previously and provided a meta-analysis of the data from the three trials in the Results section.[24]

## Conclusion

Some of the illustrative instances I have described above amount to scientific misconduct, as it seriously distorts readers' perceptions of what the best available evidence tells them, or constitutes misleading justification for conducting additional placebo-controlled trials. More attention to citation bias is needed to reduce potentially harmful consequences for patients, both directly, by distortion of evidence, and indirectly, by eroding trust in science.

### Declarations

**Competing Interests:** None declared.

**Funding:** Institute for Scientific Freedom.

**Ethics approval:** Not applicable

**Guarantor:** PCG

**Contributorship:** Sole authorship.

**Acknowledgements:** I thank Iain Chalmers for encouraging me to write this article and for his comments on earlier drafts; and Dr Hamish Chalmers, for identifying a 1985 use of the term 'citation bias.'

**Provenance:** Invited article from the James Lind Library.

### References

1. Irvine J and Martin BR. Basic research in the East and West: a comparison of the scientific performance of high energy physics accelerators. *Soc Stud Sci* 1985; 15: 293–341. In: Science Policy Study – Hearings, Volume 13. British Science Evaluation Methods: Hearing Before the Task Force on Science Policy. Committee on Science and Technology No. 59, US Government Printing Office, Washington.
2. Gøtzsche PC. Reference bias in reports of drug trials. *Br Med J* 1987; 295: 654–656.
3. Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal, antiinflammatory drugs in rheumatoid arthritis. *Controlled Clin Trials* 1989; 10: 31–56 (erratum: 356).
4. Schmidt LM and Gøtzsche PC. Of mites and men: reference bias in narrative review articles: a systematic review. *J Fam Pract* 2005; 54: 334–338.
5. Ravnskov U. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. BMJ 1992; 305: 15–19.
6. Ravnskov U. Quotation bias in reviews of the diet-heart idea. *J Clin Epidemiol* 1995; 48: 713–719.
7. Hutchison BG, Oxman AD and Lloyd S. Comprehensiveness and bias in reporting clinical trials. Study of reviews of pneumococcal vaccine effectiveness. *Can Fam Physician* 1995; 41: 1356–1360.
8. Duyx B, Urlings MJE, Swaen GMH, Bouter LM and Zeegers MP. Scientific citations favor positive results: a systematic review and meta-analysis. *J Clin Epidemiol* 2017; 88: 92–101.
9. Duyx B, Urlings MJE, Swaen G, Bouter LM and Zeegers MP. Protocol for systematic review of citation bias. See http://hdl.handle.net/10411/20710 (2015, accessed 17 August 2021).
10. Jannot AS, Agoritsas T, Gayet-Ageron A and Perneger TV. Citation bias favoring statistically significant studies was present in medical research. *J Clin Epidemiol* 2013; 66: 296–301.
11. Nieminen P, Rucker G, Miettunen J, Carpenter J and Schumacher M. Statistically significant papers in psychiatry were cited more often than others. *J Clin Epidemiol* 2007; 60: 939–946.
12. Urlings MJE, Duyx B, Swaen GMH, Bouter LM and Zeegers MPA. Citation bias in the literature on dietary trans fatty acids and serum cholesterol. *J Clin Epidemiol* 2019; 106: 88–97.
13. Urlings MJE, Duyx B, Swaen GMH, Bouter LM and Zeegers MP. Citation bias and other determinants of citation in biomedical research: findings from six citation networks. *J Clin Epidemiol* 2021; 132: 71–78.
14. Gøtzsche PC, Hammarquist C, Burr M. House dust mite control measures in the management of asthma: meta-analysis. *BMJ* 1998; 317: 1105–1110.
15. Gøtzsche PC. *Survival in an Overmedicated World: Look up the Evidence Yourself*. Copenhagen: People's Press, 2019.
16. Gøtzsche PC and Johansen HK. House dust mite control measures for asthma: systematic review. *Allergy* 2008; 63: 646–659.
17. Bouter LM, Tijdink J, Axelsen N, Martinson BC and Ter Riet G. Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Res Integr Peer Rev* 2016; 1: 17.
18. Robinson KA and Goodman SN. A systematic examination of the citation of prior research in reports

of randomized, controlled trials. *Ann Intern Med* 2011; 154: 50–55.
19. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F and Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; 327: 248–254.
20. Fergusson D, Glass KC, Hutton B and Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? *Clin Trials* 2005; 2: 218–229.
21. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c869.
22. Clarke M and Hopewell S. Many reports of randomised trials still don't begin or end with a systematic review of the relevant evidence. *J Bahrain Med Soc* 2013; 24: 145–148.
23. ISIS-1. Randomised trial of intravenous atenolol among 16 027 cases of suspected acute myocardial infarction: ISIS-1. First International Study of Infarct Survival Collaborative Group. *Lancet* 1986; 2: 57–66.
24. Gøtzsche PC, Gjørup I, Bonnén H, Brahe NEB, Becker U and Burcharth F. Somatostatin v placebo in bleeding oesophageal varices: randomised trial and meta-analysis. *BMJ* 1995; 310: 1495–1498.