The SSSPIN study—spin in studies of spin: meta-research analysis

Lisa Bero professor¹, Kellia Chiu graduate student¹, Quinn Grundy assistant professor¹ ²

¹The University of Sydney, Charles Perkins Centre, School of Pharmacy, Faculty of Medicine and Health, Sydney, NSW 2006, Australia; ²University of Toronto, Faculty of Nursing, Toronto, ON, Canada

Abstract

Objectives — To identify and calculate the prevalence of spin in studies of spin.

Design — Meta-research analysis (research on research).

Setting — 35 studies of spin in the scientific literature.

Main outcome measures — Spin, categorised as: reporting practices that distort the presentation and interpretation of results, creating misleading conclusions; discordance between results and their interpretation, with presentation of favourable conclusions that are not supported by the data or results; attribution of causality when study design does not support it; and over-interpretation or inappropriate extrapolation of results.

Results — Five (14%) of 35 spin studies contained spin categorised as reporting practices that distort the presentation and interpretation of results (n=2) or categorised as over-interpretation or inappropriate extrapolation of results (n=3).

Conclusion — Spin occurs in research on spin. Although researchers on this topic should be sensitive to spinning their findings, our study does not undermine the need for rigorous interventions to reduce spin across various research fields.

Conclusion with spin — Our hypothesis that spin will be less prevalent in spin studies than in studies on other topics has been proven. Spin scholars are less likely to spin their conclusions than other researchers, and they should receive substantial resources to launch and test interventions to reduce spin and research waste in reporting.

Introduction

Spin in scientific papers refers to practices that distort the interpretation of results and mislead readers to view the results in a more favourable light. Identification and characterisation of spin is an active area of research with a systematic review of spin studies published in 2017 and additional recent studies. The findings of these studies suggest that the prevalence of spin is high across a variety of research fields and that interventions to reduce spin are needed. Scholars on spin studies (doctorate prepared or biomedical researchers) conduct meta-research (research on research) to study the nature and extent of spin in the scientific literature. Spin scholars use various observational study designs to examine spin, including reviews and cross sectional studies. A systematic review of studies of spin in the scientific literature found that the nature and prevalence of spin varied by the study designs included in the spin studies, with the highest and most variable level of spin present in trials.

Spin has been defined in various ways, ranging from misleading conclusions to conclusions being more favourable than the results, inappropriate attribution of causality, and over-reaching extrapolation of results. This variability in definitions and subjectivity in coding spin provides an opportunity for spin scholars to come to different conclusions about the nature and prevalence of spin. Thus, there is a risk that studies of spin will be spun. We expect that spin scholars will be hyperaware and super vigilant about detecting spin in their own studies. On the other hand, they could have a tendency to produce conclusions that favour support for further research on spin and interventions to reduce spin. We hypothesise that spin will be less prevalent in spin studies than studies on other topics when matching by study design.

Methods

To test our hypothesis, we conducted a systematic search for spin in the main text of 35 studies examining the nature and prevalence of spin that were included in our previously published methodological systematic review. The 35 studies included in this review investigated spin in clinical trials, observational studies, and systematic reviews on a variety of
topics including pharmaceuticals, obesity, rheumatology, and oncology. The highest prevalence of spin, but also the greatest variability, was found in trials. The median prevalence of spin was 56% in trials, compared with 26% in systematic reviews. We also found considerable variation in how spin scholars defined spin.

**Identification and prevalence of spin**

All authors independently categorised instances of spin and reached consensus. We categorised instances of spin into the four categories derived from our systematic review:

- Reporting practices that distort the presentation and interpretation of results, creating misleading conclusions
- Discordance between results and their interpretation, with presentation of favourable conclusions that are not supported by the data or results
- Attribution of causality when study design does not support this
- Over-interpretation or inappropriate extrapolation of results.

We compared the prevalence of spin detected in the 35 spin studies with the prevalence of spin reported in the meta-research studies included in our systematic review. The meta-research studies of spin examined spin in various bodies of evidence, across a range of biomedical fields of research. When possible, we conducted subgroup analyses by the study designs used to examine spin (review, cross sectional, case study, and retrospective cohort designs).

**Reflexivity**

Meta-research and methodological systematic reviews involve interpretation and multiple judgements. Thus, understanding how the researchers have inevitably influenced the research process is important. The reflexivity section reports how our preconceptions, personal beliefs and values, assumptions, and theoretical and disciplinary positions shape our research on spin. Firstly, we are members of the community we study; all the authors consider themselves spin scholars who have published a systematic review of spin studies and have made presentations on the topic of spin. Thus, we are motivated to offer this constructive critique in the hopes that it works towards a more robust science.

The authors bring different methodological perspectives to the study of spin. LB is primarily a quantitative researcher, QG had training in advanced qualitative methods, and KC is a doctoral student receiving training in quantitative and qualitative methods. The result of this combination of perspectives is that we treat spin as a social and political construct. Whether interpretation becomes defined as spin will depend on the context and on time and place. For example, in another scientific era, spin might not be considered a problem or might be defined differently. LB has thought about the concept of spin for decades in the context of her research on bias, conflicts of interest, and peer review in biomedicine and public health. LB is also an avid knitter, but does not spin her own wool, and thus, values the contributions of spinners in other contexts.

**Patient and public involvement**

No patients were involved in the design or conduct of this study.

**Results**

**Identification of spin**

Five (14%) of the 35 spin studies contained spin. Two of the spin studies were reviews of studies of quality improvement interventions or psychological treatments. These reviews contained instances of spin that we categorised as discordance between results and their interpretation, with presentation of favourable conclusions that are not supported by the data or results. In these two spin studies, authors reported statistically non-significant results as a trend to suggest that their primary outcome, the presence of spin, was related to the study design or to authors’ non-financial conflicts of interest. In one spin study, the authors reported a statistically non-significant result as a trend to suggest that in article abstracts, ”authors of nonrandomized studies did not use more cautious language when describing the causal inference in their studies compared with authors of RCTs.” The second spin study also reported statistically non-significant results, stating in the abstract that “spin was related in trend to the inclusion of own primary studies in the systematic review and researcher allegiance.”

Three meta-research spin studies (two reviews and one cross sectional study) contained instances of spin that we categorised as over-interpretation or inappropriate extrapolation of results. One spin study classified any observational study that made a recommendation as containing spin if it did not also provide a statement that “randomized controlled trials should be conducted.” Two spin studies classified use of causal language in observational studies as spin, even if accompanied by a discussion or disclaimer about study design limitations. In one study, the words classified as implying causal language included “modify, increase, decrease, improve, influence or impact.”

The other study found that observational studies of PEBO (proposed effect of breakfast on obesity) were acceptable only if they used the word “associated” or other words limited to associations.

**Prevalence of spin**

We detected spin in five of 35 spin studies (14%, 95% confidence interval 4.8% to 30.3%). Of the 35 spin studies, four (19%) of 21 reviews had spin and one (14%) of seven cross sectional studies had spin. We compared this prevalence of spin to the median prevalence of spin that has been previously reported for studies of similar design, but on different topics (such as clinical topics; table 1). The prevalence of spin in spin studies was lower than the prevalence of spin in studies on other topics, even when matching by study design. The limited number of spin studies with spin did not allow for any investigation of author characteristics or other factors that might contribute to spin.

**Discussion**

**Principal findings**

We found that spin occurs in studies of spin, but at a lower proportion than in studies on other topics. The spin manifested as discordance between results and their interpretation, with presentation of favourable conclusions that are not supported by the data or results; and over-interpretation or inappropriate extrapolation of results. A so-called trend in significance, as detected in two studies, does not justify a positive conclusion. Interpretations that blur the boundary between a trend in significance and statistical significance could be reproduced within conclusions, abstracts, conference proceedings, media, and all forms of science communication. Investigations of
factors contributing to spin have focused on individual author characteristics, but should be expanded to study the structures and cultures of research that could incentivise spin. Our findings suggest that even spin scholars are not immune to the multiple pressures that contribute to spin, which are believed to be associated with publication bias, confirmation bias, conflicts of interest, pressures from funders, academic incentives, and placating peer reviewers and editors.

Spin manifesting as inappropriate extrapolation of results occurred in studies that examined spin in studies of observational design. Such studies are commonly used to answer public health questions assessing potential harms, for example, from tobacco or chemical exposures. Observational studies are also used to assess the effects of exposures to which people cannot or should not be randomised, such as living near a toxic waste dump or adhering to a lifelong diet. While inappropriate causal statements or a lack of evidence from randomised controlled trials was considered spin in the context of these observational study designs, not all conclusions from observational studies should be labelled as spin, otherwise it could undermine public health research that necessarily relies on these studies to provide evidence for practice and policy. Although choosing the most rigorous research design possible for a research question is very important, we caution researchers against being a bit too keen to identify spin in observational studies.

**Strengths and limitations**

A strength of our study was that in the reflexivity methods section we sought to make explicit our disciplinary, theoretical, and personal perspectives in order to account for their influence on the research. In this way, reflexivity might help mitigate against spin by allowing researchers, reviewers, editors, and readers to judge the rigour of interpretation. Reflexivity is also helpful in distinguishing between interpretation (which will produce a range of acceptable positions shaped by theoretical, disciplinary, intellectual, and personal perspectives) and spin (which crosses a line into over-interpretation or misrepresentation of findings). We believe that interpretation is inherent to the scientific process and that it can be conducted transparently and rigorously. However, the boundary between robust interpretation, grounded in evidence, and spin must be carefully maintained.

Our study had limitations. We compared the prevalence of spin in spin studies with historical data on spin prevalence calculated in our previous systematic review. However, recent studies of spin suggest that the prevalence of spin has remained stable over the past two years. Different study designs are prone to different biases. For example, randomised controlled trials minimise bias due to confounding better than studies of observational designs. The nature and prevalence of spin might also vary by study design, but we recommend caution in making direct comparisons between study designs, owing to the variability in definitions of spin used for different study designs.

Our findings support previous recommendations for alerting authors, peer reviewers, and editors to the manifestations of spin. Our findings also support a need for more rigorous research on spin, including the social, cultural, and political factors that incentivise spin. Specifically, randomised controlled trials of context specific interventions to reduce spin across different research areas and study designs are needed. Spin research should continue and be repeated in the future because the findings might change over time and in different contexts. Spin can influence how the public understands science and, thus, spin studies should be conducted by not only researchers but also members of communities who use the research.

**Conclusion**

Our findings suggest that spin occurs in research on spin. Although researchers on this topic should be sensitive to spinning their findings, our study does not undermine the need for rigorous interventions to reduce spin across various research fields.

**Conclusion with spin**

Our hypothesis that spin will be less prevalent in spin studies than studies on other topics has been proven. Spin doctors are less likely to spin their conclusions than other researchers. They should receive substantial resources to launch and test interventions to reduce spin and research waste in reporting.

**What is already known on this topic**

Spin scholars—PhD prepared researchers or biomedical researchers—conduct meta-research to study the nature and extent of spin in the scientific literature.

Spin in the scientific literature is defined in various ways that include practices distorting the interpretation of results and misleading readers to view the results as more favourable than they really are; spin occurs across different scientific disciplines and study designs.

A systematic review of spin studies found that the median prevalence of spin was 56% in trials, compared with 26% in systematic reviews.

**What this study adds**

Spin occurs in studies of spin.

Spin scholars appear to be more cautious about spinning their findings than other researchers.

We thank Matthew Page for inspiring us to use an acronym for our study in order to enhance its chance of publication.

**Contributors:** LB devised the study, supervised and participated in analysis, and wrote the first draft of the manuscript; KC participated in data collection and analysis and critically revised manuscript drafts; and OG participated in analysis and commented on the draft. LB is the guarantor of the study. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Funding:** There was no funding for this work.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

**Ethical approval:** This study was exempt from ethical review according to the guidelines of the University of Sydney’s human research ethics committee.

**Data sharing:** The full data are publicly available at The Sydney eScholarship Repository at http://hdl.handle.net/2123/16826 (doi:10.4227/11/59362060b5d76).

The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

1. Bouton I, Dutton S, Ravued P, Atman DG. Reporting and interpretation of randomised controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 2010;303:2058-64. 10.1001/jama.2010.651/20501928
2. Chiu K, Grundy G, Bere L. Spirit in published biomedical literature: A methodological systematic review. *PloS Biol* 2017;15:e2002173. 10.1371/journal.pbio.2002173.2889480
3. Cooper CM, Gray HM, Ross AE, et al. Evaluation of spin in the abstracts of otolaryngology randomized controlled trials. *Laryngoscope* 2018;2036-40.30578543
4. Kempf E, de Beyer JA, Cook J, et al. Over-interpretation and misreporting of prognostic factor studies in oncology: a systematic review. *Br J Cancer* 2018;119:1288-96. 10.1038/s41416-018-0305-5.3053050
5 Khan MS, Lateef N, Siddiqi TJ, et al. Level and prevalence of spin in published cardiovascular randomized clinical trial reports with statistically nonsignificant primary outcomes: a systematic review. *JAMA Netw Open* 2019;2:e192622. 10.1001/jamanetworkopen.2019.2622 31050775
6 Motosko CC, Ault AK, Kimberly LL, et al. Analysis of spin in the reporting of studies of topical treatments of photoaged skin. *J Am Acad Dermatol* 2019;80:516-522.e12. 10.1016/j.jaad.2018.04.034 29689326
7 Patel SV, Zhang L, Elsohly B, Yu D, Chadi SA. Spin in articles about minimally invasive transanal total mesorectal excision: an assessment of the current literature. *Colorectal Dis* 2019;21:8-14. 10.1111/codi.14451 30341922
8 Li LC, Moja L, Romero A, Sayre EC, Grimshaw JM. Nonrandomized quality improvement intervention trials might overstate the strength of causal inference of their findings. *J Clin Epidemiol* 2009;62:959-66. 10.1016/j.jclinepi.2008.10.008 19211223
9 Lieb K, von der Osten-Sacken J, Stoffers-Winterling J, Reiss N, Barth J. Conflicts of interest and spin in reviews of psychological therapies: a systematic review. *BMJ Open* 2016;6:e010606. 10.1136/bmjopen-2015-010606 2718267
10 Prasad V, Jorgenson J, Ioannidis JPA, Cifu A. Observational studies often make clinical practice recommendations: an empirical evaluation of authors’ attitudes. *J Clin Epidemiol* 2013;66:361-366.e4. 10.1016/j.jclinepi.2012.11.005 23394591
11 Cofield SS, Corona RV, Allison DB. Use of causal language in observational studies of obesity and nutrition. *Obes Facts* 2010;3:353-6. 10.1159/000322940 21196788
12 Brown AW, Bohan Brown MM, Allison DB. Belief beyond the evidence: using the proposed effect of breakfast on obesity to show 2 practices that distort scientific evidence. *Am J Clin Nutr* 2013;98:1298-308. 10.3945/ajcn.113.064410 24004890
13 Gibbs NM, Gibbs SV. Misuse of ‘trend’ to describe ‘almost significant’ differences in anaesthesia research. *Br J Anaeth* 2015;115:337-9. 10.1093/bja/aev149 26025239

**Accepted:** 27 08 2019

© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ. http://creativecommons.org/licenses/by-nc/4.0/ This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.
## Table

| Study design                  | Proportion of studies with spin | Studies on other topics* (median (%), range); No | Studies on other topics* (median (%), range); No |
|------------------------------|---------------------------------|-----------------------------------------------|-----------------------------------------------|
| Reviews                      | Spin studies (%; 95% CI); No   | 19.0 (5.4 to 41.9); n=2 spin studies including 219 reviews | 26.3 (24.2-28.4); n=2 spin studies including 219 reviews |
| Cross sectional studies      | 14.3 (0.4 to 57.9); n=7        | 85.6 (85.6-85.6); n=1 spin study including 167 observational studies | 85.6 (85.6-85.6); n=1 spin study including 167 observational studies |

* Data from Chiu et al. Other topics included drug studies, obesity and nutrition, surgical studies, psychological treatments, quality improvement interventions, diagnostic tests, or unrestricted studies across fields such as oncology, obesity or rheumatology.
† Data for observational studies which include cross sectional studies.