# Identifying spin bias of nonsignificant findings in biomedical studies

Renée O'Leary[1*], Giusy Rita Maria La Rosa[2], Robin Vernooij[3,4] and Riccardo Polosa[1,2]

## Abstract

**Objective** The purpose of this research note is to share a technique for the identification of spin bias that we developed as part of a living systematic review on the cardiovascular testing of e-cigarette substitution for cigarette smoking. While some researchers have remarked on the subjective nature of ascertaining spin bias, our technique objectively documents forms of spin bias arising from the misrepresentation of nonsignificant findings and from the omission of data.

**Results** We offer a two-step process for the identification of spin bias consisting of tracking data and findings and recording of data discrepancies by describing how the spin bias was produced in the text. In this research note, we give an example of the documentation of spin bias from our systematic review. Our experience was that nonsignificant results were presented as causal or even as significant in the Discussion of studies. Spin bias distorts scientific research and misleads readers; therefore it behooves peer reviewers and journal editors to make the effort to detect and correct it.

**Keywords** Spin bias, Data discrepancies, Bias assessment, Peer review, Nonsignificant results, Systematic literature reviews, Quality assessment tools

## Introduction

Our review team conducted an extensive quality assessment on the clinical studies we included in our living systematic review of the cardiovascular effects of e-cigarette substitution for cigarettes [1]. When we scrutinized the studies for many types of bias (drawn from the Oxford Centre for Evidence-Based Medicine *Catalogue of Bias* [2]), we were particularly concerned with spin bias. Spin bias is the distortion of data that misleads readers [3]. Spin bias is "a misrepresentation of study results, regardless of motive" [4]. One type of spin bias occurs when statistically nonsignificant results are reported as "showing an effect" [5] or where unjustified claims are made for results with *p-values* > 0.05 [6].

The reason for our vigilance for spin bias is because the use of e-cigarettes is a controversial and divisive field [7–9]. We, along with other researchers, are deeply concerned that "polarized stances on e-cigarettes will threaten the integrity of research" [9]. Calling out spin is critical because biased conclusions from studies garner media reporting [10] that becomes a source of misinformation, influencing public and clinicians' opinions and actions. Senior members of the research team reported reading numerous instances of spin bias in e-cigarette

*Correspondence:
Renée O'Leary
renee.oleary@eclatrbc.it
[1]Center of Excellence for the Acceleration of Harm Reduction, University of Catania, Via Santa Sofia, 89 Torre Biologica 11 piano, 95123 Catania, Italy
[2]Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy
[3]Department of Nephrology and Hypertension, University Medical Center Utrecht, Utrecht, The Netherlands
[4]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

studies. Would we find spin bias in the set of clinical studies included in our systematic review?

Yes, we did. Seven of 26 studies exhibited spin bias of nonsignificant findings (See Table 1). Spin bias may spring from "a strong position that relies more on their opinion than on the study results" [10], or it may be prompted as defense against publication bias against null results [11]. Whatever the motive, we developed a technique for an objective identification of spin bias for nonsignificant and misreported findings.

**Table 1** Data discrepancies and nonsignificant data spin

| Data | Spin Bias: Section and Text | Study |
|---|---|---|
| Table 3 differences between TC and ENDS on blood pressure not significant | Discussion "blood pressure are less impacted by…[ENDS] than by TC suggests that they might be less detrimental" | [13] |
| Figure 7 no significant difference in PCO$_2$ for ENDS with or without nicotine | Discussion "increased PCO$_2$ after vaping without nicotine" | [14] |
| Table S3 LDL cholesterol changes not significant for ENDS or TC, not significant between study groups Results "Mean LDL cholesterol levels… according to the RMANCOVA analysis, they were not different between the two study groups" | Discussion "[ENDS] users also appeared to have greater decreases in… LDL cholesterol than [TC] smokers" | [15] |
| Results cSBP "no significant difference at any time point in the [ENDS] group" | Discussion "We demonstrate acute changes in central blood pressure for a short period of time after vaping a liquid with nicotine" | [16] |
| Results "lower systolic blood pressure among arms…was not statistically significant" | Results "greater reduction in systolic blood pressure for [ENDS] group than in the TC group" | [17] |
| Figure 3a differences between interventions are nonsignificant | Results "Statistically significant differences between interventions…Figs. 2b and 3a" | [18] |
| Table 4 no significant difference in HR minimum between control group and ENDS group. | Results "significant difference between C [control] and [ENDS]… for min[imum] HR" | [19] |

Note: the studies used multiple terms for e-cigarettes and tobacco cigarettes. E-cigarettes terms are translated as ENDS (electronic nicotine delivery systems) and tobacco cigarettes as TC.

cSBP – central systolic blood pressure

ENDS – e-cigarettes

HR – heart rate

LDL – low-density lipoprotein

PCO$_2$ – partial pressure of carbon dioxide

RMANCOVA – repeated measures analysis of covariance

TC – tobacco cigarette

## Technique for the identification of spin bias

Our technique for identifying spin bias within an article has two steps.

First, the data and findings from the results, including in tables and figures and in appendices or supplementary materials, are tracked throughout their mentions in the study text. Data discrepancies and "pairs of statements that cannot both be true" [12] point to potential instances of spin bias. Our technique examines the misreporting or misrepresentation of nonsignificant data in the Discussion, but additionally this process can detect data discrepancies between the abstract and the conclusion of a study. Tracking can also reveal the omission of specific findings from the Discussion, another type of spin bias.

Second, the discrepancies identified in the data are reported with exact quotes (see Table 1) for objective identification. In our systematic review, many of the data discrepancies were between the data presented in a table or figure and what was stated in the text.

For non-significant findings in some instances, the spin is made with causal language, with a claim or by stating that an effect occurred where the finding was not significant [5, 10]. For example, in Table 1, there was no significant difference between e-cigarette and cigarette use on blood pressure, but in the Discussion the claim was that e-cigarettes had a lower impact than cigarettes. In some cases, the authors in their Discussion flatly stated that a finding was significant when it was reported as nonsignificant in the Results section. We observed this misreporting in two studies in our review.

For accuracy and completeness, two reviewers independently should check for data discrepancies and misreporting of nonsignificant findings. Differences in their assessments and observations should be resolved by discussion; this was our procedure. A third team member (in our review, the Project Leader) should verify all evidence of spin bias.

## Our systematic review: nonsignificant findings and spin bias in the studies

Table 1 displays the occurrences of spin bias we found in our systematic review with this technique and how we documented the evidence of spin bias.

## Critical discussion

It could be argued that authors could legitimately frame their nonsignificant findings by stating that "the findings *could be…*but not sufficient evidence." This linguistic turn obscures the results of the actual data collected. The authors' *could be* is likely their preferred hypothesis, yet nonsignificant data could be construed in any number of ways. The accurate statement would be that it was not a significant finding. Our living systematic review reported

over 66% nonsignificant cardiovascular test results. This was important data indicating that e-cigarettes had no difference in cardiovascular effects than cigarettes. Nonsignificant findings provide evidence that should not be drowned out by the noise of speculations.

In the broader biomedical literature, misreporting and misinterpretation of study findings are evidently common practices that produce spin bias [10, 20, 21]. A comparison of 896 abstracts with their full text conclusions observed that 15–35% were "inconsistent" [22]. Another study documented that 22% of trials with nonsignificant results (75 of 346 studies) had high levels of spin bias in their Conclusions [5]. In our systematic review, 27% of the studies exhibited spin bias with nonsignificant results. Nonsignificant results appear to be the most prone to spin bias: "the only factor that seems consistently associated with spin is non-statistically significant results" [23, see also 5].

## Limitations

Certainly our technique for identifying spin bias requires further testing, evaluation, and validation. As far as we know, this is the only report on the identification of nonsignificant results and spin bias in the Discussion sections of published articles. Spin bias has been examined between an abstract and the text in randomized controlled trials [20, 22, 24, 25] and in systematic reviews [21, 26, 27]. Spin bias in abstracts is especially serious because many readers look only at the abstract.

Two recent checklists could incorporate our two-step process for identifying spin bias with nonsignificant findings. The recently published Quality Output Checklist and Content Assessment tool (QuOCCA) includes one item on the spin of nonsignificant results, but it purposefully excludes checking the Discussion because spin in that section of a study was "difficult to identify" [6]. This was not our experience. The Discussion section is a key section for reporting and interpreting results, and should be checked with the QuOCCA, not excluded. PRIOR (Preferred Reporting Items for Overviews of Reviews) [28] is another recently published tool with checks for reporting bias and data discrepancies. Both tools would be enriched with our technique to identify the spin of nonsignificant data, the most common reporting bias.

Our technique does not identify all instances of spin bias. It cannot document where particular datapoints, such as secondary outcomes or subgroup findings, have been over-emphasized over primary outcomes, although our technique does document findings which are omitted from the Discussion. Identifying spin bias from overgeneralizations and overstatements entails analyzing rhetoric and checking for ascertainment bias. Our technique will require adaptation to be useful for uncovering the spin bias from undocumented deviations from a clinical

trial registry or protocol and the published study [29]. A rating for the intensity of spin bias could be based on the number of instances identified, with multiple occurrences a flag for potential researcher bias.

Finally, we can only wonder out loud about how much of an appetite there is among editors and peer reviewers to routinely take on an additional check. Spin bias distorts scientific research and misleads readers. Editors and peer reviewers should be vigilant for spin bias [5]. and "in theory, peer-reviewers and editors should determine whether the conclusions match the results" [10]. But it does not have to be all or nothing. Knowing that nonsignificant findings are the most common source of reporting bias, peer reviewers should spot check for how nonsignificant findings are presented in the discussion and abstract, without it being too onerous or time-consuming a task. The editorial team should routinely check for reporting bias of all kinds to prevent errors that result in retractions. Hopefully our technique can assist.

## Declarations

**Ethics approval and consent to participate**
Not applicable.

**Competing interests**
"RO did not receive any funding for this study. Within the past five years RO was supported by a contract with ECLAT, Srl, and ECLAT has received funding from the Foundation for a Smoke-Free World. RP is full tenured professor of Internal Medicine at the University of Catania (Italy) and Medical Director of the Institute for Internal Medicine and Clinical Immunology at the same University. He has received grants from U-BIOPRED and AIR-PROM, Integral Rheumatology & Immunology Specialists Network (IRIS), Foundation for a Smoke Free World, Pfizer, GlaxoSmithKline, CV Therapeutics, NeuroSearch A/S, Sandoz, Merk Sharp & Dohme, Boehringer Ingelheim, Novartis, Arbi Group Srl., Duska Therapeutics, Forest Laboratories and Ministero dell' Università e della Ricerca (MUR) Bando PNRR 3277/2021 (CUP E63C22000900006) and 341/2022 (CUP E63C22002080006), funded by NextGenerationEU, the European Union (EU) economic recovery package. He is founder of the Center for Tobacco Prevention and Treatment (CPCT) at the University of Catania and of the Center of Excellence for the Acceleration of Harm Reduction at the same university. He receives consultancy fees from Pfizer, Boehringer Ingelheim, Duska Therapeutics, Forest Laboratories, CV Therapeutics, and Sermo Inc. He is being paid textbook royalties from Elsevier. He is also involved in a patent application for ECLAT Srl. He is a pro bono scientific advisor for Lega Italiana Anti Fumo (LIAF) and the International Network of Nicotine Consumers Organizations (INNCO); and he is Chair of the European Technical Committee

for Standardization on "Requirements and test methods for emissions of electronic cigarettes" (CEN/TC 437; WG4).

## Consent for publication
Not applicable.

## Abbreviations
No abbreviations in text. Abbreviations for Table 1 in Legend.

## References

1. La Rosa G, Vernooij R, Qureshi M, Polosa R, O'Leary R. Clinical testing of the cardiovascular effects of e-cigarette substitution for smoking: a living systematic review. Intern Emerg Med. 2023. https://doi.org/10.1007/s11739-022-03161-z.
2. Catalogue of Bias Collaboration. Catalogue of Bias. Oxford: Centre for Evidence Based Medicine. ; 2022. Available from: https://catalogofbias.org/.
3. Catalogue of Bias Collaboration, Mahtani KR, Chalmers I, Nunan D. Spin Bias. In: Catalogue of Bias. 2019. Available from: https://catalogofbias.org/biases/spin-bias/. Accessed 27 July 2022.
4. Bradley SH, DeVito NJ, Lloyd KE, Logullo P, Butler JE. Improving medical research in the United Kingdom. BMC Res Notes. 2022. https://doi.org/10.1186/s13104-022-06050-y.
5. Chiu K, Grundy Q, Bero L. Spin' in published biomedical literature: a methodological systematic review. PLoS Biol. 2017. https://doi.org/10.1371/journal.pbio.2002173.
6. Héroux ME, Butler AA, Cashin AG, McCaughey EJ, Affleck AJ, Green MA, et al. Quality output Checklist and Content Assessment (QuOCCA): a new tool for assessing research quality and reproducibility. BMJ Open. 2022. https://doi.org/10.1136/bmjopen-2022-060976.
7. Klein DE, Hawkins B, Schwartz R. Understanding experts' conflicting perspectives on tobacco harm reduction and e-cigarettes: an interpretive policy analysis. SSM - Qualitative Research in Health. 2022. https://doi.org/10.1016/j.ssmqr.2022.100197.
8. Balfour DJK, Benowitz NL, Colby SM, Hatsukami DK, Lando HA, Leischow SJ, et al. Balancing consideration of the risks and benefits of e-cigarettes. Am J Public Health. 2021. https://doi.org/10.2105/AJPH.2021.306416.
9. Carroll DM, Denlinger-Apte RL, Dermody SS, King JL, Mercincavage M, Pacek LR, et al. Polarization within the field of tobacco and nicotine science and its potential impact on trainees. Nicotine Tob Res. 2021. https://doi.org/10.1093/ntr/ntaa148.
10. Boutron I, Ravaud P. Misrepresentation and distortion of research in biomedical literature. Proc Natl Acad Sci USA. 2018. https://doi.org/10.1073/pnas.1710755115.
11. Meerpohl JJ, Schell LK, Bassler D, Gallus S, Kleijnen J, Kulig M, et al. Evidence-informed recommendations to reduce dissemination bias in clinical research: conclusions from the OPEN (overcome failure to publish nEgative fiNdings) project based on an international consensus meeting. BMJ open. 2015. https://doi.org/10.1136/bmjopen-2014-006666.
12. Puljak L, Riva N, Parmelli E, González-Lorenzo M, Moja L, Pieper D. Data extraction methods: an analysis of internal reporting discrepancies in single manuscripts and practical advice. J Clin Epidemiol. 2020. https://doi.org/10.1016/j.jclinepi.2019.09.003.
13. Biondi-Zoccai G, Sciarretta S, Bullen C, Nocella C, Violi F, Loffredo L, et al. Acute effects of heat-not-burn, electronic vaping, and traditional tobacco combustion cigarettes: the Sapienza University of Rome-Vascular Assessment of Proatherosclerotic Effects of Smoking (SUR - VAPES) 2 randomized trial. J Am Heart Assoc. 2019. https://doi.org/10.1161/JAHA.118.010455.
14. Chaumont M, van de Borne P, Bernard A, Van Muylem A, Deprez G, Ullmo J, et al. Fourth generation e-cigarette vaping induces transient lung inflammation and gas exchange disturbances: results from two randomized clinical trials. Am J Physiol Lung Cell Mol Physiol. 2019. https://doi.org/10.1152/ajplung.00492.2018.
15. Cravo AS, Bush J, Sharma G, Savioz R, Martin C, Craige S, et al. A randomised, parallel group study to evaluate the safety profile of an electronic vapour product over 12 weeks. Regul Toxicol Pharmacol. 2016. https://doi.org/10.1016/j.yrtph.2016.10.003.
16. Franzen KF, Willig J, Cayo Talavera S, Meusel M, Sayk F, Reppel M, et al. E-cigarettes and cigarettes worsen peripheral and central hemodynamics as well as arterial stiffness: a randomized, double-blinded pilot study. Vasc Med. 2018. https://doi.org/10.1177/1358863X18779694.
17. George J, Hussain M, Vadiveloo T, Ireland S, Hopkinson P, Struthers AD, et al. Cardiovascular effects of switching from tobacco cigarettes to electronic cigarettes. J Am Coll Cardiol. 2019. https://doi.org/10.1016/j.jacc.2019.09.067.
18. Kerr DMI, Brooksbank KJM, Taylor RG, Pinel K, Rios FJ, Touyz RM, et al. Acute effects of electronic and tobacco cigarettes on vascular and respiratory function in healthy volunteers: a cross-over study. J Hypertens. 2019. https://doi.org/10.1097/HJH.0000000000001890.
19. Sumartiningsih S, Lin HF, Lin JC. Cigarette smoking blunts exercise-induced heart rate response among young adult male smokers. Int J Environ Res Public Health. 2019. https://doi.org/10.3390/ijerph16061032.
20. Shaqman M, Al-Abedalla K, Wagner J, Swede H, Gunsolley JC, Ioannidou E. Reporting quality and spin in abstracts of randomized clinical trials of periodontal therapy and cardiovascular disease outcomes. PLoS ONE. 2020. https://doi.org/10.1371/journal.pone.0230843.
21. Corcoran A, Neale M, Arthur W, Ottwell R, Roberts W, Hartwell M, et al. Evaluating spin in the abstracts of systematic reviews and meta-analyses on cannabis use disorder. Subst Abus. 2022. https://doi.org/10.1080/08897077.2021.1944953.
22. Li G, Abbade LPF, Nwosu I, Jin Y, Leenus A, Maaz M, et al. A scoping review of comparisons between abstracts and full reports in primary biomedical research. BMC Med Res Methodol. 2017. https://doi.org/10.1186/s12874-017-0459-5.
23. Boutron I. Spin in scientific publications: a frequent detrimental research practice. Ann Emerg Med. 2020. https://doi.org/10.1016/j.annemergmed.2019.11.002.
24. Tosatto D, Bonacina D, Signori A, Pellicciari L, Cecchi F, Cornaggia CM, et al. Spin of information and inconsistency between abstract and full text in RCTs investigating upper limb rehabilitation after stroke: an overview study. Restor Neurol Neurosci. 2022. https://doi.org/10.3233/RNN-211247.
25. Nascimento DP, Costa LO, Gonzalez GZ, Maher CG, Moseley AM. Abstracts of low back pain trials are poorly reported, contain spin of information, and are inconsistent with the full text: an overview study. Arch Phys Med Rehabil. 2019. https://doi.org/10.1016/j.apmr.2019.03.024.
26. Balcerak G, Shepard S, Ottwell R, Arthur W, Hartwell M, Beaman J, et al. Evaluation of spin in the abstracts of systematic reviews and meta-analyses of studies on opioid use disorder. Subst Abus. 2021. https://doi.org/10.1080/08897077.2021.1904092.
27. Yavchitz A, Ravaud P, Altman DG, Moher D, Hrobjartsson A, Lasserson T, et al. A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity. J Clin Epidemiol. 2016. https://doi.org/10.1016/j.jclinepi.2016.01.020.
28. Gates M, Gates A, Pieper D, Fernandez RM, Tricco AC, Moher D. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. BMJ. 2022. https://doi.org/10.1136/bmj-2022-070849.
29. Lancee M, Lemmens CMC, Kahn RS, Vinkers CH, Luykx JJ. Outcome reporting bias in randomized-controlled trials investigating antipsychotic drugs. Transl Psychiatry. 2017. https://doi.org/10.1038/tp.2017.203.

## Publisher's Note