TITLE: SPIN, DOCTOR: THE UNDEREXAMINED (OR OVERLOOKED) ISSUE OF EXAGGERATION IN MEDICAL RESEARCH

AUTHOR: MAARIA KASMANI

For a trial of the concept of evidence-based medicine, the COVID-19 pandemic is where the rubber hits the road. The race to find effective interventions for COVID-19 infections is this decade's first iteration of a gold rush, bearing all the frenzied hallmarks of an industry with everything to gain and a clock ticking down. Peer review processes have been abbreviated in response to the emerging crisis, and the resulting volumes of data serve to highlight an issue that has been growing at an alarming rate for many years [1]. A caveat to the increased accessibility to information provided by the internet is the decreased oversight in what is being published, resulting in an exponentially increasing number of medical publications and an onslaught of information. Distinguishing fact from artefact and correlation from coincidence is key in evaluating which potential therapies stand the best chance of success, and benefits greatly from an evidence-based approach [2]. Yet, many of the steps in distinguishing good science from bad have been suspended due to the urgency of the situation, so that research can be published as quickly as possible.

Evidence-based medicine has been defined as "the conscientious, explicit, judicious use of current best evidence in making decisions about the care of individual patients" [2]. Since the term entered popular scientific thought in the mid-1990s, medical schools have introduced the concept into their curricula, with the aim of preparing future clinicians for the novel responsibilities of an evidence-based medical practice.

On the surface, carrying out responsible literature review is not that complicated. Much of the common advice will likely sound familiar: use articles from prestigious, peer-reviewed journals, pay mind to the study type, look for statistical significance, and make note of biases within the study. After all, systems such as peer-review specifically exist to assure the quality of the studies being published, and the highest impact journals are deeply respected in the scientific community. The dense lettering and scientific jargon, complicated methodologies and voluminous data inherent to medical research seem to imbue the work with a sense of sacrosanctity, especially when presented in a respected journal.

The illusion crumbles upon closer examination. Possibly one of the most damaging research articles to ever be published — Andrew Wakefield's 1998 paper that gave credence to damaging anti-vaccination rhetoric through its claims of a causal relationship between the Measles, Mumps and Rubella (MMR) vaccine and autism development— was published by The Lancet and was thereby endorsed by one of the most reputable journals in medical science for twelve years before it was retracted in 2010 [3]. And this is hardly uncommon. An online database of retracted articles can be accessed at retractiondatabase.org, featuring thousands of articles in hundreds of journals, including bastions of medical research such as The British Medical Journal and The Journal of the American Medical Association. Some research is quickly retracted, but other papers remain in publication for years, sometimes decades before they are identified as misleading, plagiarized, or inaccurate. This is to be expected, with the rate of biomedical research increasing as it is. While peer review and journal standards decrease the amount of misinformation floating around, each retracted study on the online database demonstrates the insufficiency of these systematic quality checks. Additionally, systematic quality checks may be unable to prevent intentional publication of fraudulent or misleading information. Regardless, the existing checks are greatly disarmed in the present situation. One study that examined the basic characteristics of peer-reviewed original articles related to COVID-19 found that, depending on study type, anywhere from 47% (case-control studies) to 83% (cross-sectional studies) were at high risk of bias [1]. The ability to critically evaluate medical research is therefore inherent to the successful implementation of evidence-based medicine.

While the benefits of an evidence-based approach are clear, who has the time? Time limitations consistently rank as one of the greatest barriers to an evidence-based practice across different health-care systems, economies, and fields within medicine [4]. An evidence-based practice requires clinicians to pore over journals and studies, increasing the mental and temporal demands of an already exhausting career. As a result, article abstracts take on a deeper importance, as these summaries can drastically reduce the time and energy required

to understand the core conclusions of a study. Additionally, article abstracts are often the only part of an article to which access is completely unimpeded. Recent information is hard to come by, but a study in 2000 found that clinicians tended to read only the abstract for 63% of the articles they encounter [5]. This is a reasonable solution to a problem that cannot be ignored: according to the universal Consolidated Standards of Reporting Trials (CONSORT) guidelines on medical publishing standards, a well-written abstract should include all of the information the reader would require to critically evaluate a randomized-controlled trial [6]. This includes disclosure of study population, randomization and blinding protocol, and primary outcome measures. Additionally, taking a data-first approach to the information through use of infographics and graphical abstracts can help avoid the risk of misleading readers from the facts of the study.

Surprisingly, abstracts often fall short of CONSORT guidelines. A 2016 study found that, even in the five highest-impact medical journals, adherence is inconsistent at best. The highest adherence rate to the CONSORT guidelines was found in randomized controlled trials published in the Lancet at 78%, while those published in The New England Journal of Medicine only had a shocking 55% adherence rate [7]. An alarmingly common shortcoming is the phenomenon known as "spin", which appears to saturate even the most highly regarded journals. Spin, or interpretive bias, can be defined as a conscious or unconscious misrepresentation or overstatement of a study's results [8]. A 2019 study examined the presence of interpretive bias in randomized controlled trials published in six high-impact journals and identified its presence in 57% of abstracts [9]. A similar study from 2017 examined the type of spin typically present in article abstracts, to find that impacted abstracts are often worded to suggest therapeutic efficacy of their experimental treatment even in cases where there is little or no evidence of a statistically or clinically significant effect [8]. Numerous studies have parroted these findings across the various fields of medicine: 86% of articles published regarding Periodontal therapy and Cardiovascular Disease outcomes and 40% of articles published in high-impact surgical journals featured some degree of spin in their abstract, with 64% and 23% of abstracts, respectively, suggesting therapeutic potential unsubstantiated by the study's results [10,11].

So, we have data demonstrating that even articles published in high-profile journals may misrepresent their findings. We also have evidence that such misrepresentation does impact the way that doctors interpret the studies they read: the SPIIN randomized controlled trial found that experimental cancer treatments in abstracts containing spin were significantly more likely to be interpreted as beneficial by clinicians than the same abstracts with the interpretive bias edited out [12]. The implications are scary, but the results of a similar study conducted on members of the Japanese Primary Care Association (JPCA), which regularly holds Evidence Based Medicine (EBM) workshops, should provide some solace. The DOCTOR trial found that, as long as the necessary information concerning the primary objective was provided, primary care physicians were not swayed by spin in abstracts, and in fact rated conclusions in the spin-free articles as more valid [13]. Another study involving the JPCA found that doctors who have been out of education for longer are more likely to overstate treatment benefit in abstracts with spin [14]. This lines up with the relative nascency of EBM: physicians who received their training prior to the widespread introduction of the concept were likely never taught the skills necessary to identify spin in medical school.

At the present moment, successful critical appraisal of the literature could inform the choice between two experimental treatments. The traditional EBM approach would be to consult the body of evidence for each topic: when the data is compiled as such, the biases found in individual studies tend to cancel each other out. Differences in outcome due to differences in demographics, methods of data collection, and so on can be identified in individual studies when they are systematically compared against one another, allowing for a glimpse at what the true effect of a given intervention may be. But this takes time and the bodies of evidence for COVID-19-related research have not had the years they require to accumulate, meaning that it becomes even more important to critically evaluate each study's strengths and weaknesses. Luckily, a standardized

and systematic approach to this evaluation already exists in the form of CONSORT (Table 1). The limitation is therefore not in the need for such a framework to be developed, but rather for this framework to become a universally understood and applied component of both medical research and practice.

References

1. Raynaud, M., Zhang, H., Louis, K. et al. COVID-19-related medical research: a meta-research and critical appraisal. BMC Med Res Methodol. 2021; 21(1).
2. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-2.
3. Eggertson L. Lancet retracts 12-year-old article linking autism to MMR vaccines. CMAJ. 2010;182(4):E199-E200.
4. Barzkar F, Baradaran HR, Koohpayehzadeh J. Knowledge, attitudes and practice of physicians toward evidence-based medicine: A systematic review. J Evid Based Med. 2018;11(4):246-51.
5. Saint S, Christakis DA, Saha S, et al. Journal reading habits of internists. Journal of General Internal Medicine. 2000;15(12):881-4.
6. Schulz K, Altman D, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. Trials 2010;11(1).
7. Hays M, Andrews M, Wilson R, Callender D, O'Malley PG, Douglas K. Reporting quality of randomised controlled trial abstracts among high-impact general medical journals: a review and analysis. BMJ Open. 2016;6(7):e011082.
8. Turrentine M. It's All How You "Spin" It: Interpretive Bias in Research Findings in the Obstetrics and Gynecology Literature. Obstet Gynecol. 2017;129(2):239-42.
9. 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(5):e192622.
10. 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;15(4):e0230843.
11. Arunachalam L, Hunter IA, Killeen S. Reporting of Randomized Controlled Trials With Statistically Nonsignificant Primary Outcomes Published in High-impact Surgical Journals. Ann Surg. 2017;265(6):1141-5.
12. Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. J Clin Oncol. 2014;32(36):4120-6.
13. Shinohara K, Aoki T, So R, et al. Influence of overstated abstract conclusions on clinicians: a web-based randomised controlled trial. BMJ Open. 2017;7(12):e018355.
14. sujimoto Y, Aoki T, Shinohara K, et al. Physician characteristics associated with proper assessment of overstated conclusions in research abstracts: A secondary analysis of a randomized controlled trial. PLoS One. 2019;14(1):e0211206.

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item
Title and abstract		
	1a	Identification as a randomised trial in the title
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts[21 31])
Introduction		
Background and objectives	2a	Scientific background and explanation of rationale
	2b	Specific objectives or hypotheses
Methods		
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons
Participants	4a	Eligibility criteria for participants
	4b	Settings and locations where the data were collected
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed
	6b	Any changes to trial outcomes after the trial commenced, with reasons
Sample size	7a	How sample size was determined
	7b	When applicable, explanation of any interim analyses and stopping guidelines
Randomisation:		
Sequence generation	8a	Method used to generate the random allocation sequence
	8b	Type of randomisation; details of any restriction (such as blocking and block size)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how
	11b	If relevant, description of the similarity of interventions
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses
Results		
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome
	13b	For each group, losses and exclusions after randomisation, together with reasons
Recruitment	14a	Dates defining the periods of recruitment and follow-up
	14b	Why the trial ended or was stopped
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[28])
Discussion		
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses
Generalisability	21	Generalisability (external validity, applicability) of the trial findings
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence
Other information		
Registration	23	Registration number and name of trial registry
Protocol	24	Where the full trial protocol can be accessed, if available
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration[13] for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials,[11] non-inferiority and equivalence trials,[12] non-pharmacological treatments,[32] herbal interventions,[33] and pragmatic trials.[34] Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

Table 1: CONSORT 2010 checklist of information to include when reporting a randomized trial[6]