

Red for danger in systematic reviews?

Andrew Moore

How do you feel about red traffic lights—stop immediately, or the basis for a negotiation? Most people stop most of the time, which is why traffic works—because we obey the rules and we know red means danger. We may not have the statistics at our fingertips but we know that running a red light is a risky business. The USA Federal Highway Administration claims that roughly 45% of car collisions take place at a road intersection and the cause is usually related to ignoring a stop sign or running a red stop light. In 2018, 846 people were killed in the USA and 139 000 injured in crashes involving crossing a red light.¹

Initiated in Cochrane reviews, many systematic reviews now use a red–amber–green system in the assessment of the risk of bias relating to the methods used in included studies. Green indicates no known risk of bias, amber (used when an item is not mentioned in a study) an unknown level of risk of bias, and red when a methodological issue is known to be associated with a high risk of bias.

Assessment of risk of bias is good, but how that risk of bias assessment is used is crucial in determining the overall quality and reliability of evidence, and the confidence we have in it. Trustworthy evidence should be based on studies that are green across the board—with no likelihood of bias. That is a rare event. Most systematic reviews use small and flawed studies where red for high risk of bias occurs with almost every included study, often for more than one item.

The fact of this high risk of bias in individual studies is all too often ignored when making conclusions or recommendations, but ignoring the red signals is a dangerous business. For example, experience shows that, for trials of efficacy of interventions, methodological issues such as lack of randomisation, lack of blinding, small study size, and inappropriate imputation methods to handle data from dropouts can all result in a positive assessment of efficacy where none exists. They give the wrong answer: ignore these red lights at your peril. A systematic review pronouncing on efficacy when each

included study has one high risk of bias is like running a succession of red lights.

Ideally, a systematic review will highlight these problems and come to appropriate conclusions. For example, a comprehensive review of cannabinoids for pain included 37 randomised, double blind trials.² None was without some risk of bias with green across the board, and 28 had at least one red light. The appropriate conclusion was that ‘Studies ... have unclear or high risk of bias, and outcomes had GRADE rating of low-quality or very low-quality evidence. We have little confidence in the estimates of effect.’

Trials may have inadequate method reporting, but so do systematic reviews. For cannabinoids, there were more systematic reviews (57) than randomised trials (37). Confidence in the results in the systematic reviews using AMSTAR-2 definitions was critically low (41), low (8), moderate (6) or high (2); 86% were inadequate.^{3,4} Low or critically low confidence in the results of systematic reviews when assessed using AMSTAR-2 is all too common.

Another growing concern is that of research misconduct, to the extent of recognising that fraudulent studies pollute our literature. About 2% of scientists have fabricated, falsified or modified data or results at least once, and 34% admitted other questionable research practices, those numbers increase to 14% and 72% when asked about the behaviour of colleagues.⁵ Retraction Watch (<https://retractionwatch.com/>) is a website dedicated to highlighting fraud and research misconduct, now with over 20 000 retracted papers on its database. It came about because of a Japanese anaesthetist who faked—and had retracted—183 randomised trials, mainly concerning anti-vomiting medications. This was known for years before it came to light, and it is an interesting story.^{6–8} A blog at the journal Science provides a fascinating insight on retractions,⁹ and tools are evolving better to detect fraud.¹⁰

The scandal of poor medical research is well known.¹¹ We are becoming aware of the poor quality of most systematic reviews. Between 1986 and 2016, annual publication of systematic reviews increased by over 2700%, 18 times greater than for all studies. The 1400 systematic reviews published in 2000 grew to 29 000

by 2019, about 80 published each day.¹² A superb critical analysis confirms a growing opinion that many systematic reviews are redundant and misleading; only 3% may be trustworthy and clinically useful.¹³ China is the largest producer of meta-analyses and second largest of systematic reviews,¹⁴ with considerable concerns about their conclusions.¹³

These considerations, plus many others, place a huge burden on journal editors and peer reviewers. Plucking the useful 3% from the many systematic review submissions is hard work, but rewarding. One of the top cited articles for the EJHP over the past 3 years was an overview review on medication adherence.¹⁵ It used evidence from eight good systematic reviews; evidence from a further 24 was excluded because of limited evidence, poor review methods, or both, meaning that 75% of our literature on this topic had to be discarded. But it was cited because quality was valued.

Journals are now beginning to demand more than manuscripts—for clinical trials it is the raw data for independent analysis. For systematic reviews and meta-analyses, we also need raw data for independent analysis, together with the confidence to reject the inadequate. It is a big ask, with a big bill. The question is who should pay the insurance cost of running red lights.

For readers of systematic reviews, just bear in mind that red means stop.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Commissioned; internally peer reviewed.

© European Association of Hospital Pharmacists 2021. No commercial re-use. See rights and permissions. Published by BMJ.



To cite Moore A. *Eur J Hosp Pharm* 2021;**28**:299–300.

Published Online First 6 October 2021

Eur J Hosp Pharm 2021;**28**:299–300.
doi:10.1136/ejhp-2021-003080

REFERENCES

- Peterman DR. Safety impact of speed and red light cameras. congressional research service R46552. Available: <https://www.itskrs.its.dot.gov/sites/default/files/docs/R46552.pdf>
- Fisher E, Moore RA, Fogarty AE, *et al.* Cannabinoids, cannabis, and cannabis-based medicine for pain management: a systematic review of randomised controlled trials. *Pain* 2021;**162**:S45–66.
- Shea BJ, Reeves BC, Wells G, *et al.* AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised

Correspondence to Dr Andrew Moore, Pain Research, Newton Ferrers, Plymouth PL8 1DA, UK; andrew.moore@omkltid.org

- studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
- 4 Moore RA, Fisher E, Finn DP, *et al.* Cannabinoids, cannabis, and cannabis-based medicines for pain management: an overview of systematic reviews. *Pain* 2021;162:S67–79.
 - 5 Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One* 2009;4:e5738.
 - 6 Moore RA, Derry S, McQuay HJ. Fraud or flawed: adverse impact of fabricated or poor quality research. *Anaesthesia* 2010;65:327–30.
 - 7 Carlisle JB. The analysis of 168 randomised controlled trials to test data integrity. *Anaesthesia* 2012;67:521–37.
 - 8 Tramèr MR. The Fujii story: a chronicle of naive disbelief. *Eur J Anaesthesiol* 2013;30:195–8.
 - 9 Brainard J. Rethinking retractions. *Science* 2018;362:390–3.
 - 10 Li W, Bordewijk EM, Mol BW. Assessing research misconduct in randomized controlled trials. *Obstet Gynecol* 2021;138:338–47.
 - 11 Altman DG. The scandal of poor medical research. *BMJ* 1994;308:283–4.
 - 12 Hoffmann F, Allers K, Rombey T, *et al.* Nearly 80 systematic reviews were published each day: observational study on trends in epidemiology and reporting over the years 2000–2019. *J Clin Epidemiol* 2021;138:1–11.
 - 13 Ioannidis JPA. The mass production of redundant, misleading, and Conflicted systematic reviews and meta-analyses. *Milbank Q* 2016;94:485–514.
 - 14 Fontelo P, Liu F. A review of recent publication trends from top publishing countries. *Syst Rev* 2018;7:147.
 - 15 Wilhelmsen NC, Eriksson T. Medication adherence interventions and outcomes: an overview of systematic reviews. *Eur J Hosp Pharm* 2019;26:187–92.