Dubious data and contamination of the research literature on pain

Amanda C de C Williams

Professor of Clinical Health Psychology, University College London
Consultant Clinical Psychologist, University College Hospitals NHS Foundation Trust

With several colleagues, I have recently stumbled into investigating what we call "untrustworthy" data in pain. The story started when we were updating a systematic review and meta-analysis of psychological interventions for chronic pain[1]. Three of the 70+ eligible papers had results that were staggeringly better than anyone else's, by an order of magnitude. The same team had produced all three papers. Either they had discovered spectacularly effective ways of delivering CBT and exercise to people with musculoskeletal (spinal) pain, in which case it was urgent that we all learned from the trials, or there was a problem with their data.

Our questions to the authors about their treatments, even when answered, did not elicit useful information, and the author team itself seemed rather less expert than we expected. Eventually we decided to exclude the three trials from the meta-analysis, but we had become curious about the author group and the number of papers – many of them large RCTs – they had published in pain and that had found their way into meta-analyses and guidelines.

We systematically searched for their RCTs on physical and/or psychological interventions for spinal pain and found 10 trials. We ran these through a risk-of-bias tool, which turned up little, mostly because information was missing. Then we applied the Cochrane Pregnancy and Childbirth review group's Trustworthiness Screening Tool, developed for routine use by this group on trials eligible for meta-analysis. This tool checks for features of good practice, such as trial preregistration, and publicly available ethics application, and also examines feasibility and distributions of data, from baselines and from tests. This generated concerns about 8 of the 10 trials, such as identical data at baseline across trials, zero attrition, and all changes extraordinarily large. We published our findings[2].

We then approached the authors of the six journals that had published these trials (see [3]) with a copy of the published paper[2], expressing concern. Three of the journals instigated investigations consistent with the COPE (Committee on Publishing Ethics) guidelines they endorsed (as does the *British Journal of Pain*). This resulted in two retractions by journals and one by the trial authors. Of the other three, one (which had published four of the papers concerned) wrote to the first author, were told he was unavailable, and decided to take it no further; the two others appeared to find it distasteful that we had raised the subject, implying that we were behaving unprofessionally, and took the first author's assurances at face value. One of those has since reconsidered and retracted the paper; the other (though fully signed up to COPE) preferred resolution by 'academic debate', as if authenticity of data is a matter of personal preference. We declined.

We remained concerned about the number of systematic reviews, meta-analyses, and guidelines that these trials had been included in and whether, as in our meta-analysis, they would have inflated the effect sizes. Using citation tracking, it turned out[4] that 32 reviews or guidelines, with 55 comparisons, had included the studies, and that removing them on average halved the effect sizes and improved precision. None of the reviews or guidelines had raised concerns about these trials.

This problem is not limited to psychological and physical interventions for pain. In the pain field overall[5], 389 papers have so far been retracted; one author accounts for 33 of them! First authors were most commonly from China, publishing preclinical research, followed by mainly clinical papers

from Japan, the USA, South Korea, and Germany. The UK accounted for 1.8% of total retractions. Authors of both preclinical and clinical papers tended to be in clinical, rather than academic, posts. In some countries, doctors' careers will stagnate unless they publish, or large bonuses are offered for publishing and being cited. This has generated a market in authorship of ready-written fake papers, and citation cartels.

Such practices distort the field and potentially (when the falsified papers are in rare cancers where few papers are published) lead to harmful treatments. Retraction is intended to correct the scientific record, and Retraction Watch https://retractionwatch.com/ is a highly informative and reliable source of information. Of course, errors requiring retraction can arise for completely innocuous reasons, although correction may be a better option. However, most retractions are for research misconduct, including lack of adherence to ethical guidelines, but reasons are often poorly specified, particularly when authors themselves retract when investigations begin. Problematically, there is a lag of several years before papers are retracted and, even after retraction, they find their way into reviews and clinical guidelines[6,7] that may never be corrected. For instance, two systematic reviews purported to show that ivermectin reduced COVID deaths, with massive publicity, but several of the trials were found to have blocks of repeated data and other problems indicating fabrication: when excluded, there were no benefits to giving ivermectin.

There is a fightback. Apart from Retraction Watch, Elisabeth Bik is a microbiologist with a real gift for spotting manipulated (rotated, cropped, etc.) images in scientific papers, from microscopy to genetic sequences. She runs the Science Integrity Digest, https://scienceintegritydigest.com/, which makes for sobering reading (and some humour: human photosynthesis, anyone?). Academic groups in Manchester and Sydney are developing tools to detect problematic RCTs. If you want to help develop the tool, contact Jack Wilkinson (Jack.wilkinson@manchester.ac.uk).

In a thoughtful editorial about the need for such tools to identify untrustworthy data, Tugwell and Knottnerus[8] emphasise the disproportionate damage done by these trials and their inclusion in reviews and guidelines. But, misconduct apart, there are various sleights of hand that are common in reporting clinical practice but that also produce untrustworthy data. It is worth reading about Open Science practices to understand this better, or you could follow Bauer's tongue-in-cheek checklist for highlighting the "strengths" of weak studies[9], including options such as 'Report only hypotheses with most interesting findings', 'Report rounded-down p values', and 'Report only the most obvious limitations of the research'. It is increasingly hard to do research as a clinician: time and other resources are scarce, and the complex documentation for ethical and R&D approval is voluminous and daunting. But what research is done needs to be reported as transparently as possible, whatever the findings, for us to progress from the current massive overload of poor quality research[10].

References

1 Williams ACdeC, Fisher E, Hearn L, et al. Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane Database Syst Rev* 2020, Issue 8. Art. No.: CD007407. doi: 10.1002/14651858.CD007407.pub4

2 O'Connell NE, Moore RA, Stewart G, et al. Investigating the veracity of a sample of divergent published trial data in spinal pain. *PAIN* 2023;164:72–83

3 Williams ACdeC, Hearn L, Moore RA, et al. Effective quality control in the medical literature: investigation and retraction vs inaction. *J Clin Epidemiol* 2023 in press. DOI:10.1016/j.jclinepi.2023.02.022

4 O'Connell NE, Moore RA, Stewart G, et al. Trials we cannot trust: investigating their impact on systematic reviews and clinical guidelines in spinal pain. *PAIN* 2023 in press.

5 Ferraro MC, Moore RA, Williams ACdeC, et al. Characteristics of retracted publications related to pain research: a systematic review. *PAIN* 2023 in press

6 Avenell A, Stewart F, Grey A, et al. An investigation into the impact and implications of published papers from retracted research: systematic search of affected literature. *BMJ Open* 2019;9:e031909. doi:10.1136/bmjopen-2019-031909

7 Fanelli D, Wong J, Moher D. What difference might retractions make? An estimate of the potential epistemic cost of retractions on meta-analyses, *Accountability in Research* 2022;29(7):452-9. DOI: 10.1080/08989621.2021.1947810

8 Tugwell P, Knottnerus JA. How does one detect scientific fraud – but avoid false accusations. *J Clin Epidemiol* 2017;87:1-3. Doi:10.1016/j.jclinepi.2017.08.013

9 Bauer GE. Biased Outcome reporting Guidelines for Underwhelming Studies (BOGUS) statement and checklist. *BMJ* 2021;375:e067350 http://dx.doi.org/10.1136/bmj-2021-067350

10 Moore A, Fisher E, Eccleston C. Flawed, futile, and fabricated—features that limit confidence in clinical research in pain and anaesthesia: a narrative review. *Brit J Anaesthesia* 2022; Nov 8. doi: 10.1016/j.bja.2022.09.030