DOCTORAL THESIS

Shifting Evidence: Factors Affecting the Reliability of Systematic Reviews in Biomedicine

Bastian, Hilda

Award date:
2021

BOND UNIVERSITY

Shifting Evidence:

Factors Affecting the Reliability of

Systematic Reviews in Biomedicine

Hilda Bastian

Submitted in total fulfillment of the requirements of the degree of Doctor of Philosophy (PhD)

October 2020

Faculty of Health Sciences and Medicine

Professor Paul Glasziou and Professor Chris Del Mar

This research was supported by an Australian Government Research Training Program Fellowship

Abstract

Background

Each systematic review or systematic review version is a static snapshot. However, the evidence shift for several reasons, with new additions or problems identified in the synthesised evidence. The aim of this thesis is to explore the ways shifts in the underlying evidence could affect the reliability of systematic reviews, and consider the implications.

Methods

The growth of evidence and practices used in systematic reviews to deal with it was assessed in a series of descriptive studies, including a longitudinal study of updating in Cochrane reviews from 2003 to 2018, and a study of Cochrane reviews designated as having enough evidence. As a precursor to developing research integrity filters, a comprehensive search strategy was developed, the prevalence of retracted publications in PubMed and PubMed Central was estimated, and their discoverability assessed. Research integrity filters for post-publication events that can compromise publications were developed and piloted in assessing 36,462 trials included in Cochrane reviews and 83,302 non-Cochrane systematic reviews. The prevalence of potentially compromised included trials affecting Cochrane reviews was estimated.

Results

The number of trials and systematic reviews continues to increase dramatically, with ongoing and published clinical trials rising more quickly than systematic reviews. The number of Cochrane reviews has plateaued, and they are increasingly out of date. The median time to update of a cohort of reviews updated in 2003 was three years, but by the end of 2018, the median time since their last update was seven years.

It is uncommon for systematic reviews to resolve questions about health care, as results often give rise to further questions about the optimal use of care, and which groups of people could benefit. Other areas of concern to clinicians and patients, however, may not attract research activity. Systematic reviews may be more likely to end because of a lack of expected research than because there is enough evidence.

Although prioritising scarce systematic reviewing resources is critical, there is no consensus on methods for determining when there is enough evidence to cease updating. Fewer than 5% of protocols for Cochrane reviews specify how they will determine when there is enough evidence.

Emerging evidence of compromised trials could have a similar effect on the reliability of a systematic review's results as new studies but monitoring of these events by systematic review groups appears inadequate. System problems and journal practice contribute: a substantial proportion (over 30%) of retracted publications in PubMed/PMC are not indexed as such, and journal compliance with recommended practice is low. Using pilot research filters, 28% of a group of 2,025 Cochrane reviews with above average numbers of included studies were found to be affected by at least one trial that had been retracted or had an erratum or expression of concern published.

Conclusions

The increase in evidence and problems identified in trials is outpacing the methods, infrastructure, and collaborations needed to enable systematic reviewers to keep up with shifting evidence. Research integrity filters were shown to improve identification of potentially problematic trials. Centralised efforts to enable timely and efficient responses to these trials should be undertaken.

Keywords: systematic reviews, clinical trials, updating, search filters, research integrity, retractions, errata, post-publication peer review.

Declaration and authorship

This thesis is submitted to Bond University in fulfilment of the requirements of the degree of *Doctor of Philosophy*.

This thesis represents my own original work towards this research degree and contains no material that has previously been submitted for a degree or diploma at this University or any other institution.

Hilda Bastian is the sole author on the Introduction and Discussion chapters, 1 and 7, and lead author on all other chapters, one of which is a multi-author published paper with updated data and figures, and the others being multi-author manuscripts being posted as preprints and submitted for publication.

The original research work underpinning all chapters was driven in every case primarily by Hilda Bastian, who also produced initial drafts of each manuscript, and managed the collaborative research projects. Specific contributorship is detailed at the end of each multi-author chapter. One of the projects, chapter 3, developed some unpublished pilot work into a longitudinal study, substantially expanded in concept, with most data collection and analysis carried out during the PhD candidature. The other work submitted in this thesis was carried out during the PhD candidature.

All figures and graphs were prepared for the PhD projects, and all are open access. A substantial proportion of the work in chapter 5 was undertaken when the author and co-authors of the chapter worked for a U.S. government agency, and that work cannot be copyrighted. In that capacity, the author received funding from the Intramural Research Program of the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM), National Institutes of Health (NIH).

The Vancouver style is used for referencing. Where original references had currently non-functioning links, these have been updated.

Research outputs

Peer-reviewed journal articles arising from this thesis.

- Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLOS Medicine*. 2010 Sep 21;7(9):e1000326. *[Chapter 2]*

- Bastian H. A stronger post-publication culture is needed for better science. *PLOS Med.* 2014; 11(12):e1001772. *[Chapter 6, 7]*

Manuscripts being posted as preprints and submitted for publication arising from this thesis.

- Bastian H, Doust J, Clarke M, Glasziou P. The epidemiology of systematic review updates: a longitudinal study of updating of Cochrane reviews, 2003 to 2018. *[Chapter 3]*

- Bastian H, Hemkens L. Enough evidence and other endings: a descriptive study of stable Cochrane systematic reviews in 2019. *[Chapter 4]*

- Bastian H, Jordan DC, Vaught M. The prevalence of retractions and their identification in PubMed and PubMed Central: a descriptive study. *[Chapter 5]*

- Bastian H, Sampson M, Doust J, Vaught M. Systematic reviews and the perpetuation of error in the biomedical literature: a descriptive study and pilot of research integrity filters for PubMed. *[Chapter 6]*

Conference presentations arising from this thesis.

- Bastian H, Clarke M, Doust J, Glasziou P. From Barcelona to Madrid: History and quality of update reporting of Cochrane Reviews flagged as updates in 2003 and analysed for the Barcelona Colloquium. 19th Cochrane Colloquium; 2011, Madrid. *[Chapter 3]*

- Bastian H, Hemkens L. 'No more updates': a descriptive study of the 126 'stable' Cochrane reviews in March 2012. 20th Cochrane Colloquium; 2012, Auckland. *[Chapter 4]*

- Bastian H, Hemkens L. Reaching certainty: a descriptive study of 'stable' Cochrane reviews and coming to firm conclusions. 21st Cochrane Colloquium; 2013, Quebec. *[Chapter 4]*

- Bastian H. Post-publication evaluation: what can we expect? PLOS Medicine 10th Anniversary Symposium; 2014, San Francisco. *[Chapter 6]*

- Bastian H. Tackling the twin beasts of information overload and reviewer workload. 22nd Cochrane Colloquium; 2015, Vienna. *[Chapter 2]*

- Bastian H. Reproducibility and the prevention of bias: before, during, and after research. Priscilla M. Mayden Lecture, Research Reproducibility 2016; 2016, Utah. *[Chapter 6]*

- Bastian H. Trials, systematic reviews, and the perpetuation of error in the biomedical literature. Synthesis Seminar Series; 2017, Center for Clinical Trials and Evidence, Johns Hopkins University, Baltimore. *[Chapter 6]*

- Bastian H. Accessibility hurdles for meta-research and science correction. Association for Interdisciplinary Meta-Research and Open Science 2019; the University of Melbourne. *[Chapter 6]*

Acknowledgments

I am deeply grateful for the knowledge, support, patience, and encouragement of my supervisors, Paul Glasziou and Chris Del Mar, and the support of all the team at the Institute for Evidence-Based Healthcare at Bond University. One of the best parts of this exercise has been the opportunity to work with some of my favourite colleagues: I cannot thank all my co-authors enough for their generosity, effort, and expertise: Iain Chalmers, Jenny Doust, Paul Glasziou, Lars Hemkens, Diana Jordan, Mike Clarke, Margaret Sampson, and Melissa Vaught.

I am also deeply grateful for the support and inspiration of former colleagues at the U.S. National Library of Medicine, in particular the former Director, Dr David Lipman, and Melissa Vaught and Diana Jordan. This degree could not have been undertaken without their support and stimulating collegiality. Working with this group was a privilege and a joy. I also owe thanks to the many people at the U.S. National Library of Medicine who endured endless questions from me about PubMed and PubMed Central.

I am also greatly appreciative to members of the Cochrane Editorial Unit and UK Cochrane Centre who provided information, data, and permission or guidance on the use of Cochrane data, Anne Eisinga, Harriet MacLeHose, Toby Lasserson, Chris Mavergames, Karla Soares-Weiser, as well as Herm Lamberink for generously sharing data from another project on Cochrane reviews.

Thanks and appreciation are not enough to express my gratitude to Iain Chalmers, for recruiting me into a lifelong interest in systematic reviews, and following that up with decades of inspiration, intellectual challenge, and friendship – and irresistible tasks to tackle. A giant thank you to Jenny Doust, who has also provided an endless supply of support, fun, and stimulating ideas and conversation over the years, as well as indispensable encouragement to see this work through to the end.

My heartfelt gratitude to my sons and their families, for their understanding and encouragement in the demanding last leg of this marathon.

TABLE OF CONTENTS

List of Tables and Boxes

List of Figures

Abbreviations

API	Application programming interface
CDSR	*Cochrane Database of Systematic Reviews* (journal component of The Cochrane Library)
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
COPE	Committee on Publication Ethics
CRDT	PubMed publication created date
CRG	Cochrane review group
CSE	Council of Science Editors
CSV	Comma separated values
DOI	Digital object identifier
EOC	(Editorial) expression(s) of concern
Epub	e-publication ahead of print issue
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
ICMJE	International Committee of Medical Journal Editors
IQR	Interquartile range
MeSH	NLM's Medical Subject Headings
NCBI	National Center for Biotechnology Information
NICE	National Institute for Health and Care Evidence
NLM	U.S. National Library of Medicine
ORI	Office of Research Integrity of the U.S. Department of Health and Human Services
PMC	PubMed Central
PMID	PubMed identification number
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SoF	GRADE Summary of Findings table in Cochrane reviews
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology statement
XML	eXtensible Markup Language (for encoding human- and machine-readable files)

Chapter 1

Introduction

Summary

This chapter describes the emergence of systematic reviews, outlining the research aims and questions of this thesis. The chapter concludes with brief introductions to thesis research projects presented in following chapters.

Hilda Bastian

Background

From the start of their broadening adoption from the 1950s, (1) randomised clinical trials became more influential on healthcare decision-making. In 1962, the US Food and Drug Administration (FDA) began requiring evidence of efficacy for new drugs, which led to a pivotal role for randomised trials in healthcare evidence and spurred more widespread adoption of them internationally. (2)

The number of clinical trials quickly outpaced people's ability to keep up with them and weigh their often-disparate results. By the 1970s, efforts to gather clinical trials relevant to a specific question, and synthesise their results to generate reliable evidence, were gaining pace. (3,4) In 1979 at a symposium on medicines, epidemiologist Archie Cochrane estimated that there was then an average of 14 clinical trials appearing each day, a number too high for people to digest. He criticized the medical profession for not producing a "critical summary, by speciality or subspeciality, adapted periodically, of all relevant randomised controlled trials". (5) Cochrane would later use the term *systematic review* to describe this work, which was beginning to gather pace in medicine by the 1980s. (3) New techniques were being developed to support the increasing sophistication of systematic reviews, such as the forest plot to display the results of pooled studies in 1982, (6) and the funnel plot to explore publication and other biases in 1984. (7)

Awareness of the risks of making decisions based on non-systematic reviews was increased by a study by Cindy Mulrow in 1987, where she assessed 50 reviews published in a year in four major medical journals. (8) Applying eight criteria, Mulrow determined that none met all the criteria, and only one met six of them. Reviews in medicine at the time, she concluded, "do not routinely use scientific methods to identify, assess, and synthesize information".

In contrast, the systematic review community aimed to provide a more reliable basis for decision-making by being more rigorous, reducing opportunities for bias, and increasing transparency. David Moher and colleagues described this approach to

research review: “What distinguished systematic reviews was the use of formal explicit methods, in other words pre-specification, of what exactly was the question to be answered, how evidence was searched for and assessed, and how it was synthesized in order to reach the conclusion. Importantly, these formal methods were described as part of the review itself in a Methods section”. (9)

Although systematic review can be applied with studies of any type, the burgeoning systematic review community in biomedicine in the 1990s focused on reviewing randomised trials, as the least biased approach to evaluating the effects of interventions. The impact this approach could have on patient care was demonstrated by an influential paper in 1992, in which Elliott Antman and colleagues established that more beneficial care could have been introduced over a decade sooner if trials had been systematically reviewed. (10)

Undertaking periodically updated systematic reviews of randomised trials on a large scale was made feasible by developments in information technology and publishing. Searching the biomedical literature was made increasingly accessible to researchers in institutions from 1971, when the US National Library of Medicine’s database of literature, MEDLARS, went online (called MEDLINE). (11) Secondly, the potential for accomplishing dissemination and updating of data-heavy systematic reviews “increased dramatically” with the emergence of electronic publishing by the 1990s. (3)

The systematic review community in biomedicine was given impetus by the establishment of the international Cochrane Collaboration in 1993. The organisation evolved out of extensive pioneering work in the area of perinatal care spearheaded by Iain Chalmers, and was fueled by the support provided by the NHS Research and Development Programme to establish a Cochrane Centre to produce systematic reviews in Oxford, and a Centre for Reviews and Dissemination (CRD) in York. (12)

The Cochrane Collaboration published systematic reviews of the effects of healthcare interventions, and it did so while also growing a community of practice and use, advocating internationally for the support and adoption of systematic

reviews as a fundamental pillar of evidence-based health care. The Collaboration established methodological standards, developed infrastructure and training, and became a rallying site for people to develop methodology and meta-research of clinical research. (12–14)

A second key track in the evolution of systematic reviews and their methodology came with their adoption by health technology assessment (HTA) agencies. HTA agencies began to appear in the 1970s, with the establishment in the US of the Office of Technology Assessment in 1972 and the National Centre for Health Care Technology in 1978, both now defunct. (15) In the 1980s, HTA agencies were established in more countries, mostly undertaking and/or commissioning systematic reviews as part of their work. (15–17) HTAs do not always include systematic reviews, or they may go about them in somewhat different ways. (18) Members of HTA agencies and researchers supported by them contributed to methodological development of both HTA and systematic reviewing. (19–21)

A third key track was the increasing interest in systematic reviewing and updating in the clinical practice guideline (CPG) development community. Developing from the beginning of the 1990s, (22) many organisations producing CPGs began adopting more systematic methods, exemplified by the consensus development of an instrument to appraise the methodological quality of CPGs in 2003 by the Appraisal of Guidelines, Research and Evaluation (AGREE) Collaboration. (23)

Formal linkages and individuals in common typically strengthen and influence the development of reforming groups and movements. (24) That kind of cross-fertilization was a strong feature of the development of each of the three communities of interest or movements described here – the Cochrane Collaboration, HTA agencies, and CPG producers. They had significant overlaps in membership, and funding from HTA commissions supported groups of systematic reviewers and methodology developers. (25)

The methodological cross-fertilisation resulting from these inter-relationships is exemplified by the methodology for CPG producers developed by the Grading of

Recommendations, Assessment, Development and Evaluation (GRADE) Working Group. The Working Group began in 2000, (26) with membership including overlap with HTA and the Cochrane Collaboration: one of its leaders, Gordon Guyatt, who coined the term “evidence-based medicine”, was co-convenor of the Cochrane Applicability and Recommendations Methods Group. (25,27,28) The GRADE Working Group developed methods for grading the strength of a body of evidence rather than an individual study, and for systematically making recommendations based on that body of evidence, with judgments made explicit and transparent. (29) The GRADE Working Group incorporated and built on the concepts about risk of bias developed by the Cochrane Collaboration, (30) and the Cochrane Collaboration, in turn, incorporated GRADE methodology for strength of evidence and summarising findings into Cochrane reviews. (31) This dynamic can enable new practices and methods to diffuse through far-reaching networks of influence.

While that results in some consolidation of methods, there has also been an extensive proliferation of methods across the groups working to assess and or/keep up with the results of trials and other studies. From the origins of systematic reviewing, where there was a coordinated effort to zero in on increasingly strenuous gold standards for the process, the field has moved into a phase of forking and branching. This has been in part due to developing ways to overcome limitations of the standard processes, and in part extending the process to address different questions than effectiveness of healthcare interventions. However, most of the proliferation of reviewing methods has been to take shortcuts in the onerous process. (9)

In 2009, Maria Grant and Andrew Booth developed a typology of 14 review types and methodologies, including systematic reviews, rapid reviews, qualitative evidence syntheses, and umbrella reviews (reviews of systematic reviews). (32) Since then, there has been further diversification, and some types not included in that typology have gained traction, notably realist reviews, (33) living systematic reviews, (34) and network meta-analyses. (35,36)

Consensus and disagreements about the comparative rigour, value, and efficiency of review methods have increasingly been supported or addressed by meta-research – research on research, (37) also called methodological research. By the 1980s, empirical studies were addressing key methodological challenges in the conduct, reporting, and interpretation of randomised trials, for example, exploring the impact of randomisation in 1983, (38) and testing the effects of blinding people assessing the quality of clinical reports in 1996. (39)

Systematic reviewing shines a spotlight on the general standard of quality and reporting of the types of study they analyse. (40) Increasingly, meta-research established that methods used in systematic reviews were often not rigorously developed or evaluated. For example, reviews assessing the tools used to critically assess non-randomised studies found that most of them did not have a robust method of development. (41) The GRADE Group authors summarised the overall state of knowledge in 2011 of the pivotal issue of rating the quality of studies this way: "attempts to show systematic difference between studies that meet and do not meet specific criteria have shown inconsistent results". (30)

However, meta-research on systematic review methods, as well as the methods themselves, has been overwhelmingly forward-looking: they address ways of finding and processing new studies or addressing new questions, either for a de novo review or for an update. Systematic review methods are therefore implicitly predicated on the primary studies that have already been found and incorporated being static. Typically, though, methods guidance does not direct researchers to look backwards to identify shifts in the evidence already assembled. The Cochrane Collaboration's Handbook has some provision for this, although it is not raised in the section on updating: the search strategy chapter advises downloading retraction and errata fields when gathering citations from MEDLINE, and "When updating a review, it is important to search MEDLINE for the latest version of the citations to the records for the included studies" (section 6.4.10). (31) However, the updating section (3.4.2.2) is explicitly forward-looking: "When no new studies meeting the selection criteria are found, the review update will simply require that this finding be recorded in the relevant sections of the review". (31) There appear to be few other exceptions of

authors that have advocated for reviewing the current status of previously reviewed papers to ascertain corrections, for example. (42,43)

This thesis focuses on factors involving primary studies that can potentially shift bodies of evidence and make the conclusions of systematic reviews unreliable, whether they are new or overlooked studies, or emerging information that affects previous publications.

Research aim and questions

Each systematic review or systematic review version is a static snapshot. However, the evidence that is eligible for the review is not likely to be static. Evidence can shift for several reasons. Relevant new primary studies may be done, and new data can become available for a primary study that has already been analyzed, rendering a review's conclusions out-of-date. (44) In addition, errors, overlooked studies, or other problems in a primary study can be identified, which can have serious implications for the reliability of the review. (42,45–48)

The aim of this thesis is to explore the ways shifts in the underlying evidence could affect the reliability of systematic reviews, and consider the implications for the practice of systematic reviewing and its infrastructure needs. Central to this consideration is assessing the magnitude of the issues. Understanding the size of these problems is critical for determining the feasibility and potential value of strategies to address them, especially in the context of escalating methodological demands for undertaking and updating systematic reviews.

My focus on the broad topic of shifting evidence grew from an interest in keeping up with new evidence and using systematic reviews as the basis for the production of reliable, up-to-date consumer health information. (49,50) In grappling with the challenges of managing that task over a number of years, I relied on, and undertook meta-research to develop the work, with a particular interest in shifting and conflicting

evidence. While subsequently curating the systematic review and methods literature at PubMed, (51) I began addressing another aspect of changing evidence: post-publication events such as retractions and errata. When I developed a plan for a PubMed tool to identify post-publication events associated with biomedical publications, I identified a range of barriers, including the under-ascertainment of these events in PubMed, as well as a pattern of research gaps. The following chapters report on a series of research projects undertaken for this thesis addressing these broad questions from various angles:

1. To what extent are systematic reviews keeping up with new clinical trials?
2. How frequently are Cochrane reviews of healthcare interventions updated, and what growth is there in the inclusion of trials?
3. What other shifts in the underlying evidence could affect the reliability of systematic reviews, and how common are they?
4. To what extent could systematic reviews be reducing or perpetuating error in the biomedical literature?
5. What are practical options that could be implemented by systematic reviewers or communities to enable systematic reviews to be responsive to critical shifts in evidence?

Following chapters

A key goal of this thesis is to keep the implications of shifts in evidence in perspective. Chapter 2 analyses the epidemiology of systematic reviews in relation to the growth of clinical trials, and considers our ability to keep up with growing bodies of evidence. This is done by charting key historical events as well as the rate of publication of systematic reviews and clinical trials, using data from PubMed, the U.S. National Library of Medicine's database of biomedical literature, (52) the Cochrane Clinical Trials Register (CCRT), (53) Epistemonikos, a database including an extensive collection of systematic reviews in health, (54) and data on Cochrane reviews provided by the Cochrane Collaboration.

Chapter 3 describes the updating history over 15 years of a cohort of systematic reviews from the *Cochrane Database of Systematic Reviews* that were first updated in 2003. In this journal, the Cochrane Collaboration has produced and maintained a body of reviews that have the goal of updating to keep up with evidence, usually from clinical trials. (55) Studying this unique large-scale effort to stay systematically up-to-date with clinical trials provides an opportunity to consider the implications of evidence shifts arising from new studies.

In chapter 4, the questions of when and why evidence stops shifting are explored. Here analysis of the Cochrane systematic reviews designated as “stable” and therefore not requiring updating, provide an opportunity to consider how much evidence is enough, and the reasons for changing directions in systematic reviews.

Chapter 5 turns the attention to post-publication events. The prevalence of retracted studies in PubMed is estimated by a systematic analysis of multiple sources, and search filters to identify retracted publications are developed and tested. This study identifies the extent of limitations in the biomedical literature’s responsiveness to post-publication changes, and develops capacity to capture the post-publication record of a publication. After this study began, an analysis of a much smaller category of post-publication event, editorial expressions of concern, was also done. (56)

The techniques developed in the study of retracted publications are applied to systematic reviews and key post-publication events for included studies in them identifiable in PubMed in Chapter 6. The prevalence of these events is estimated, along with the potential for systematic reviews to perpetuate error because of compromised included studies, and ways this problem could be reduced. Methods that could be used in the production and updating of systematic reviews are offered, and the implications for biomedical literature collections are discussed.

In the final chapter, the variety of ways in which evidence can shift and affect the reliability of systematic reviews are discussed and placed in perspective. The implications for individual systematic review authors and users are addressed, with

recommendations for systematic review communities as well as biomedical literature publishers and curators. Finally, the need for a stronger post-publication culture in biomedical research is emphasised (Chapter 7).

References

1. Hill AB. Controlled clinical trials: papers delivered at the conference convened by the Council for International Organizations of Medical Sciences established under the joint auspices of UNESCO and WHO. Oxford: Blackwell Scientific Publications; 1960.

2. Barron BA, Bukantz SC. The Evaluation of New Drugs: Current Food and Drug Administration Regulations and Statistical Aspects of Clinical Trials. Arch Intern Med. 1967 Jun 1;119(6):547–56.

3. Chalmers I, Hedges LV, Cooper H. A Brief History of Research Synthesis. Eval Health Prof. 2002 Mar 1;25(1):12–37.

4. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLOS Med. 2010 Sep 21;7(9):e1000326.

5. Cochrane AL. 1931–1971: a critical review, with particular reference to the medical profession. In: Medicines for the year 2000. London: Office of Health Economics; 1979. p. 1–11.

6. Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. BMJ. 2001 Jun 16;322(7300):1479–80.

7. Light R, Pillemer D. Summing up: The Science of Reviewing Research. Cambridge, Massachusetts: Harvard University Press; 1984.

8. Mulrow CD. The medical review article: state of the science. Ann Intern Med. 1987 Mar;106(3):485–8.

9. Moher D, Stewart L, Shekelle P. All in the Family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. Syst Rev. 2015 Dec 22;4(1):183.

10. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA. 1992 Jul 8;268(2):240–8.

11. Miles WD. A history of the National Library of Medicine: the nation's treasury of medical knowledge [Internet]. Bethesda: US National Library of Medicine, NIH; 1982. Available from: https://collections.nlm.nih.gov/catalog/nlm:nlmuid-8218545-bk

12. Chalmers I. The Cochrane Collaboration: Preparing, Maintaining, and Disseminating Systematic Reviews of the Effects of Health Care. Ann N Y Acad Sci. 1993;703(1):156–65.

13. Daly J. Evidence-Based Medicine and the Search for a Science of Clinical Care. Berkeley: University of California Press; 2005.

14. Chandler J, Hopewell S. Cochrane methods - twenty years experience in developing systematic review methods. Syst Rev. 2013 Sep 20;2(1):76.

15. Banta D. The development of health technology assessment. Health Policy Amst Neth. 2003 Feb;63(2):121–32.

16. Jonsson E. History of health technology assessment in Sweden. Int J Technol Assess Health Care. 2009 Jul;25 Suppl 1:42–52.

17. Hailey D. The history of health technology assessment in Australia. Int J Technol Assess Health Care. 2009 Jul;25(S1):61–7.

18. Rotstein D, Laupacis A. Differences between systematic reviews and health technology assessments: a trade-off between the ideals of scientific rigor and the realities of policy making. Int J Technol Assess Health Care. 2004;20(2):177–83.

19. McDonagh MS, Jonas DE, Gartlehner G, Little A, Peterson K, Carson S, et al. Methods for the drug effectiveness review project. BMC Med Res Methodol. 2012 Sep 12;12:140.

20. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 1: introduction. Med Decis Mak Int J Soc Med Decis Mak. 2013;33(5):597–606.

21. Wong J. The History of Technology Assessment and Comparative Effectiveness Research for Drugs and Medical Devices and the Role of the Federal Government. Biotechnol Law Rep. 2014 Dec 1;33(6):221–48.

22. Bhaumik S. Use of evidence for clinical practice guideline development. Trop Parasitol. 2017 Dec;7(2):65–71.

23. AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care. 2003 Feb;12(1):18–23.

24. Aveni AF. Organizational Linkages and Resource Mobilization: The Significance of Linkage Strength and Breadth. Sociol Q. 1978 Mar 1;19(2):185–202.

25. Atkins D, Fink K, Slutsky J, Agency for Healthcare Research and Quality, North American Evidence-based Practice Centers. Better information for better health care: the Evidence-based Practice Center program and the Agency for Healthcare Research and Quality. Ann Intern Med. 2005 Jun 21;142(12 Pt 2):1035–41.

26. GRADE Working Group. GRADE home [Internet]. GRADE. [cited 2019 Aug 3]. Available from: http://www.gradeworkinggroup.org/

27. Cochrane Collaboration. The Cochrane Collaboration Methods Groups Newsletter [Internet]. UK Cochrane Centre; 2001. Available from: https://web.archive.org/web/20040228124031/http://www.cochrane.org/newslett/MGNews_2001.pdf

28. Smith R, Rennie D. Evidence based medicine--an oral history. BMJ. 2014 Jan 21;348:g371.

29. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924–6.

30. Guyatt GH, Oxman AD, Vist G, Kunz R, Brożek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol. 2011 Apr;64(4):407–15.

31. Higgins JP, Green SE. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 5.1.0. Oxford: Cochrane Collaboration; 2011. Available from: http://handbook-5-1.cochrane.org/

32. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Inf Libr J. 2009 Jun;26(2):91–108.

33. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy. 2005 Jul;10 Suppl 1:21–34.

34. Elliott JH, Turner T, Clavisi O, Thomas J, Higgins JPT, Mavergames C, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. PLoS Med. 2014 Feb;11(2):e1001603.

35. Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med. 2002 Aug 30;21(16):2313–24.

36. Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J Clin Epidemiol. 2011 Feb;64(2):163–71.

37. Ioannidis JPA, Fanelli D, Dunne DD, Goodman SN. Meta-research: Evaluation and Improvement of Research Methods and Practices. PLoS Biol. 2015 Oct;13(10):e1002264.

38. Chalmers TC, Celano P, Sacks HS, Smith H. Bias in Treatment Assignment in Controlled Clinical Trials. N Engl J Med. 1983 Dec 1;309(22):1358–61.

39. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996 Feb;17(1):1–12.

40. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. Nature. 2018 07;555(7695):175–82.

41. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. Int J Epidemiol. 2007 Jun 1;36(3):666–76.

42. Royle P, Waugh N. Should systematic reviews include searches for published errata? Health Inf Libr J. 2004 Mar;21(1):14–20.

43. Sampson M. Updating searches for systematic reviews [Internet]. Aberystwyth: University of Wales Aberystwyth; 2009. Available from: https://pure.aber.ac.uk/portal/files/10374343/Sampson_Updating_Searches_for_Systematic_Reviews_PhD.pdf

44. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007 Aug 21;147(4):224–33.

45. Farrah K, Rabb D. Errata for trial publications are not uncommon, are frequently not trivial, and can be challenging to access: a retrospective review. J Med Libr Assoc JMLA. 2019;107(2):187–93.

46. Egger M, Smith GD. Bias in location and selection of studies. BMJ. 1998 Jan 3;316(7124):61–6.

47. Roberts I, Smith R, Evans S. Doubts over head injury studies. BMJ. 2007 Feb 24;334(7590):392–4.

48. Garmendia CA, Nassar Gorra L, Rodriguez AL, Trepka MJ, Veledar E, Madhivanan P. Evaluation of the Inclusion of Studies Identified by the FDA as Having Falsified Data in the Results of Meta-analyses: The Example of the Apixaban Trials. JAMA Intern Med. 2019 Mar 4;

49. Bastian H. Health literacy and patient information: developing the methodology for a national evidence-based health website. Patient Educ Couns. 2008 Dec;73(3):551–6.

50. Bastian H, Scheibler F, Knelangen M, Zschorlich B, Nasser M, Waltering A. Choosing health technology assessment and systematic review topics: the development of priority-setting criteria for patients' and consumers' interests. Int J Technol Assess Health Care. 2011 Oct;27(4):348–56.

51. Bastian H. PubMed review methods filter and methods research support from PubMed Health. Cochrane Database Syst Rev. 2016 Oct 23;(10 Suppl 1):27–8.

52. U.S. National Library of Medicine. MEDLINE, PubMed, and PMC (PubMed Central): How are they different? [Internet]. National Library of Medicine NIH. 2018 [cited 2019 Aug 12]. Available from: https://www.nlm.nih.gov/bsd/difference.html

53. Cochrane Controlled Register of Trials (CENTRAL) | Cochrane Library [Internet]. [cited 2019 Nov 14]. Available from: https://www.cochranelibrary.com/central/about-central

54. Rada G, Pérez D, Capurro D. Epistemonikos: a free, relational, collaborative, multilingual database of health evidence. Stud Health Technol Inform. 2013;192:486–90.

55. Starr M, Chalmers I, Clarke M, Oxman AD. The origins, evolution, and future of The Cochrane Database of Systematic Reviews. Int J Technol Assess Health Care. 2009 Jul;25 Suppl 1:182–95.

56. Vaught M, Jordan DC, Bastian H. Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central. Res Integr Peer Rev [Internet]. 2017;2. Available from: https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-017-0030-2

Chapter 2

The growth of systematic reviews and their place in the biomedical literature.

Summary

This chapter charts the rise of systematic reviews in relation to the rise of trials and other biomedical literature, considering the implications for keeping up with the evidence. The original version was published in 2010. (1) The chapter includes updates of the data and methods on the growth of systematic reviews and trials in the context of other biomedical literature.

Publication, with updated data for figures:

(1) Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLOS Medicine*. 2010 Sep 21;7(9):e1000326.

Hilda Bastian

Paul Glasziou

Iain Chalmers

Summary points

- When Archie Cochrane reproached the medical profession for not having critical summaries of all randomised controlled trials, about 14 reports of trials were being published per day. There are now 75 trials, and 11 systematic reviews of trials, per day and a plateau in growth has not yet been reached.

- Although trials, reviews, and health technology assessments have undoubtedly had major impacts, the staple of medical literature synthesis remains the non-systematic narrative review. Only a small minority of trial reports are being analysed in up-to-date systematic reviews. Given the constraints, Archie Cochrane's vision will not be achieved without some serious changes in course.

- To meet the needs of patients, clinicians, and policymakers, unnecessary trials need to be reduced, and systematic reviews need to be prioritised. Streamlining and innovation in methods of systematic reviewing are necessary to enable valid answers to be found for most patient questions. Finally, clinicians and patients require open access to these important resources.

Thirty years ago, and a quarter of a century after randomised trials had become widely accepted, Archie Cochrane reproached the medical profession for not having managed to organise a “critical summary, by speciality or subspeciality, adapted periodically, of all relevant randomised controlled trials”. (2) Thirty years after Cochrane's reproach we feel it is timely to consider the extent to which health professionals, the public and policymakers could now use “critical summaries” of trials for their decision-making.

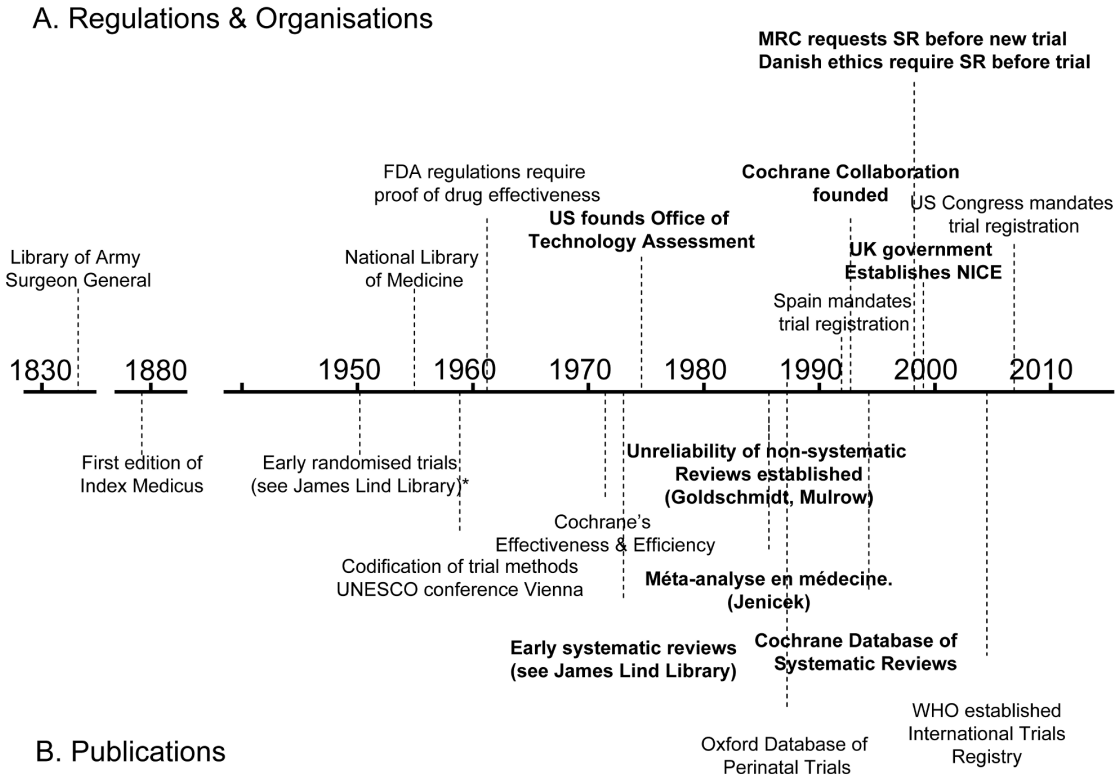
The landscape

Keeping up with information in health care has never been easy. Even in 1753, when James Lind published his landmark review of what was then known about scurvy, he needed to point out that "… before the subject could be set in a clear and proper light, it was necessary to remove a great deal of rubbish". (3) And 20 years later, Andrew Duncan launched a publication summarising research for clinicians, lamenting that critical information "…is scattered through a great number of volumes, many of which are so expensive, that they can be purchased for the libraries of public societies only, or of very wealthy individuals". (4) We continue to live with these two problems—an overload of unfiltered information and lack of open access to information relevant to the well-being of patients.

A century later, the precursor of the US National Library of Medicine (NLM) began indexing the medical literature. Between 1865 and 2006, the index grew from 1,600 references to nearly 10 million. (5) Even with the assistance of electronic databases such as NLM's MEDLINE, the problem of having to trawl through and sift vast amounts of data has grown. As mountains of unsynthesised research evidence accumulate, we need to keep improving our methods for gathering, filtering, and synthesising it. Some of the key events in the story so far are shown on the timeline in Figure 1.

A legal regulatory framework overseen by the US Food and Drug Administration (FDA) requiring proof of efficacy of new drugs was introduced in 1962, and other countries followed suit. These developments made it inevitable that randomised trials would increasingly become an important component of the evidence base. (6) Government health technology assessment agencies were also established as policymakers sought to have more reliable evidence of the effects of other forms of health care interventions. (7)

Figure 1. Policy and academic milestones in the development of trials and the science of reviewing trials.



As the number of clinical trials grew, so too did the science of reviewing trials. Systematic reviews and meta-analyses endeavouring to make sense of multiple trials began to appear in a variety of health fields in the 1970s and 1980s (see Box 1). An important early example showed that postoperative radiotherapy after surgical treatment of breast cancer was associated with a previously unrecognised increased risk of death. (8) Another challenged beliefs about vitamin C and the common cold. (9) A third suggested a previously unrecognised advantage of some forms of fetal monitoring during labour in reducing neonatal seizures. (10)

Box 1. Early systematic reviews of the effects of health care interventions.

- Stjernswärd J (1974) Decreased survival related to irradiation postoperatively in early breast cancer. Lancet 304: 1285-1286.
- Chalmers TC (1975) Effects of ascorbic acid on the common cold. An evaluation of the evidence. Am J Med 58: 532-536.
- Cochran WG, Diaconis P, Donner AP, Hoaglin DC, O'Connor NE, Peterson OL, Rosenoer VM (1977) Experiments in surgical treatments of duodenal ulcer. In: Bunker JP, Barnes BA, Mosteller F, eds. Costs, risks and benefits of surgery. Oxford: Oxford University Press. pp 176-197.
- Smith ML, Glass GV (1977) Meta-analysis of psychotherapy outcome studies. Am Psychol 32: 752-760.
- Hemminki E, Starfield B (1978) Routine administration of iron and vitamins during pregnancy: Review of controlled clinical trials. Br J Obstet Gynaecol 85: 404-410.
- Hemminki E, Starfield B (1978) Prevention and treatment of premature labour by drugs: Review of controlled clinical trials. Br J Obstet Gynaecol 85: 411-417.
- Chalmers I (1979) Randomized controlled trials of fetal monitoring, 1973–1977. In: Thalhammer O, Baumgarten K, Pollak A, eds. Perinatal medicine. Stuttgart: Georg Thieme. pp 260-265.
- Policy Research Incorporated (1979) Medical Practice Information Demonstration Project. Bipolar disorder, a state of the science report. Baltimore: Policy Research Incorporated.
- Editorial (1980) Aspirin after myocardial infarction. Lancet 1:1172-1173. [Published anonymously but written by Richard Peto.]
- Baum ML, Anish DS, Chalmers TC, Sacks HS, Smith H, Fagerstrom RM (1981) A survey of clinical trials of antibiotic prophylaxis in colon surgery: Evidence against further use of no-treatment controls. N Engl J Med 305:795-799.
- Hampton JR (1982) Should every survivor of a heart attack be given a beta-blocker? Part I: Evidence from clinical trials. BMJ 285:33-36.
- Stampfer MJ, Goldhaber SZ, Yusuf S, Peto R, Hennekens CH (1982) Effect of intravenous streptokinase on acute myocardial infarction: Pooled results from randomized trials. N Engl J Med 307: 1180-1182.
- Sacks HS, Chalmers TC, Berk AA, Reitman D (1985) Should mild hypertension be treated? An attempted meta-analysis of the clinical trials. Mt Sinai J Med 52: 265-270.
- Yusuf S, Peto R, Lewis J, Collins R, Sleight P (1985) Beta blockade during and after myocardial infarction: An overview of the randomized trials. Prog Cardiovasc Dis 27: 335-371.

By the mid-1980s, the need to minimise the likelihood of being misled by the effects of biases and the play of chance in reviews of research evidence was being made evident in articles (11–15) and textbooks. (16) In 1988, regularly updated electronic publication of systematic reviews and meta-analyses, along with bibliographies of randomised trials, began in the perinatal field. (17,18) This provided a model for the inauguration of the international Cochrane Collaboration in 1993 to prepare, maintain, and disseminate systematic reviews of the effects of health care interventions.

Where are we now?

Despite this progress, the task keeps increasing in size and complexity. We still do not know exactly how many trials have been done. For a variety of reasons, a large proportion of trials have remained unpublished. (19,20) Furthermore, many trials have been published in journals without being electronically indexed as trials, which makes them difficult to find. One of the first steps in being able to adequately review literature is that scientific contributions which predate digitalised information systems and trial indexing need to be “rediscovered and inserted into the memory system”. (21) Through the 1990s, to identify possible reports of controlled trials, the Cochrane Collaboration mobilised thousands of volunteers around the globe to comb the major databases, and to hand-search nondigitalised health literature, unpublished conference proceedings, and books. The result of this collaborative effort is the Cochrane Controlled Trials Register (CCTR) (now called the Cochrane Central Register of Controlled Trials).

The differences between the numbers of trial records in MEDLINE and CCTR (see Figure 2) have multiple causes. Both CCTR and MEDLINE often contain more than one record from a single study, and there are lags in adding new records to both databases. The NLM filters are probably not as efficient at excluding non-trials as are the methods used to compile CCTR. Furthermore, MEDLINE has more language restrictions than CCTR. In brief, there is still no single repository reliably showing the true number of randomised trials. Similar difficulties apply to trying to estimate the number of systematic reviews and health technology assessments (HTAs).

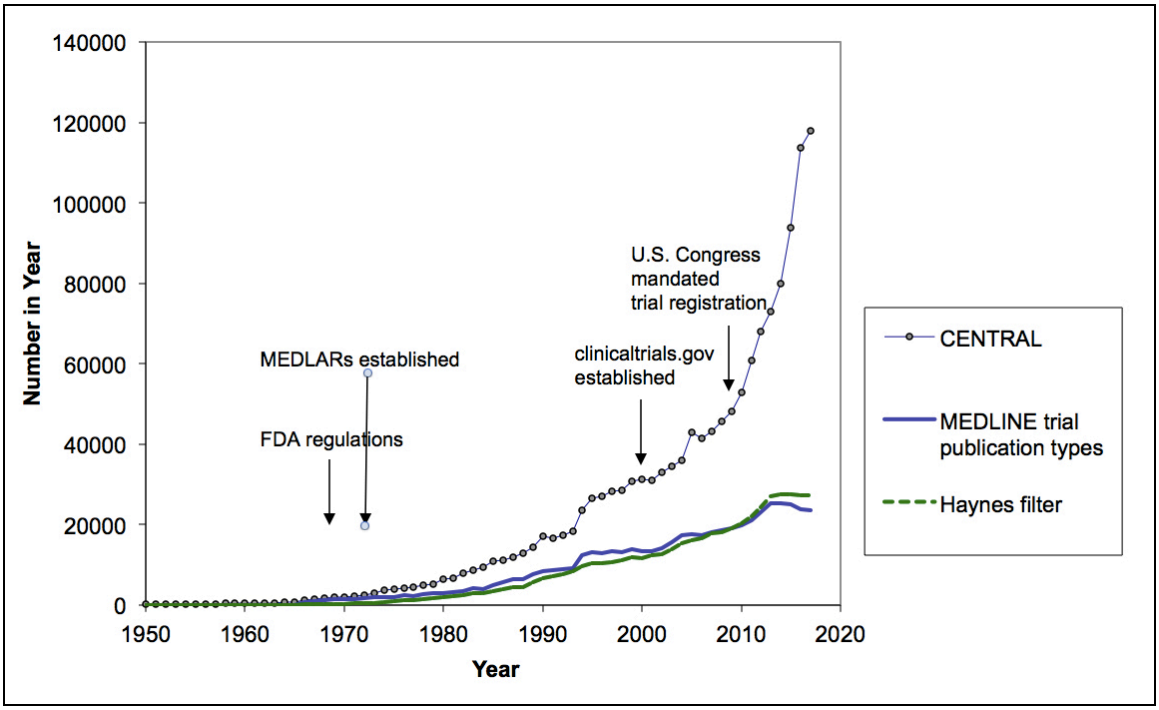


Figure 2 (updated). The number of trials, 1950 to 2017.

CENTRAL (formerly CCTR) includes prospective trial registry entries, half of which may be completed and not necessarily published: both published and unpublished trials can be included and analysed in systematic reviews. Other measures include only published trials. The Haynes filter uses the "narrow" version of the Therapy filter in PubMed's Clinical Queries; see Appendix 1. In Figures 2 and 3, we use a variety of data sources to estimate the numbers of trials and systematic reviews published from 1950. (See Appendix 1 for methods.)

Even though these figures must be seen as more illustrative than precise, multiple data sources tell the same story: astonishing growth has occurred in the number of reports of clinical trials since the middle of the 20th century, and in reports of systematic reviews since the 1980s – and a plateau in growth has not yet been reached. With a median of perhaps 80 participants per trial, the number of people being enrolled in trials is likely to be more than 2,000,000 per year. (22) Prospective trial registration establishes a new genre of evidence repository: trials are registered

in these databases at inception, theoretically enabling an overview of all published and unpublished trials.

In 2004, the International Committee of Medical Journal Editors (ICMJE, http://www.icmje.org/) announced that their journals would no longer publish trials that had not been prospectively registered. (23) Before this announcement, an average of 30 trials a week were being prospectively registered around the world. Once the journal editors' deadline came into force, more than 200 ongoing trials per week were being registered. (24) In 2007, the US Congress made detailed prospective trial registration legally mandatory. (25) As WHO's international clinical trials platform develops, it will become possible to generate a more realistic picture of how many trials are being done. This registry draws together standardised core data from all the trial registries meeting specified quality criteria. Registering full protocols and reporting trial results in these registries are the next frontiers.

How close are we to Archie Cochrane's goal?

In 1986 and 1987, Goldschmidt and Mulrow showed how great the potential is for error in reviews of health literature that were not conducted systematically. (11,12) Looking at data such as those in Figure 3 could provide the comforting illusion that systematic reviews have displaced other less reliable forms of information. However, as Figure 4 shows, this is far from the case. The growth has been even more remarkable in non-systematic ("narrative") reviews and case reports. Journal publishing of non-systematic reviews, and the emergence of many journals whose sole product is non-systematic reviews, has far outstripped the growth of systematic reviews and HTAs, as impressive as the latter has been. And the number of case reports—which can also provide important new information such as adverse effects—is far higher than the number of trials or systematic reviews. Trials, systematic reviews, and HTAs have undoubtedly had major impacts, including on clinical guidelines: they are more likely to be cited and read than other study types. (26) However, the staple of medical literature synthesis remains the non-systematic narrative review.

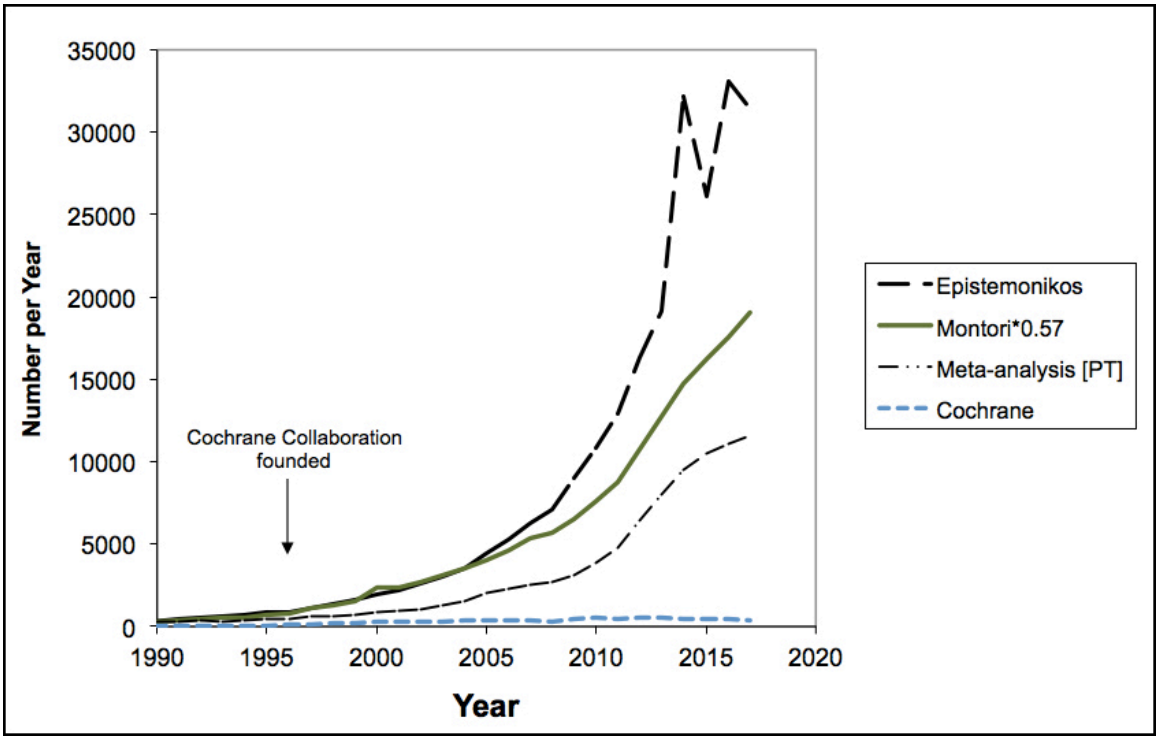


Figure 3 (updated). The number of systematic reviews in health care, 1990 to 2017 estimated by searches in Epistemonikos, PubMed (Montori and meta-analysis filters), and systematic reviews in the Cochrane Database of Systematic Reviews.

Epistemonikos is a database that includes machine-learning search results. (27) The Montori systematic review filter is detailed in Appendix 1. (This figure previously contained data from a health technology assessment database that is no longer up-to-date.)

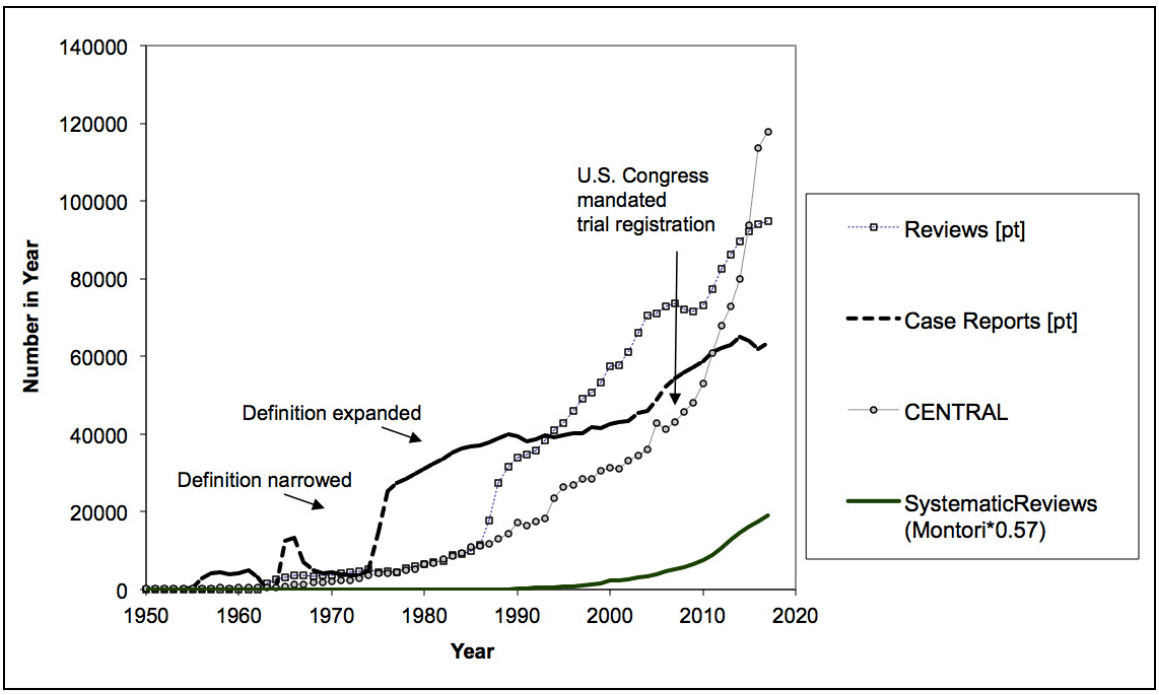


Figure 4 (updated). The rise in non-systematic reviews, case reports, trials (including unpublished), and systematic reviews, 1950 to 2017, based on filtered searches in PubMed and the annual totals in Cochrane's CENTRAL register of trials.

Furthermore, we are a long way from having all relevant trials incorporated into good systematic reviews. The workload involved in producing reviews is increasing, and the bulk of systematic reviews are now many years out of date. (28) The median number of trials contained within individual systematic reviews has been variously estimated at between six and 16 (Cochrane reviews now include an average of over 12 trials per review; (29,30) M Clarke, personal communication), but many reviews have covered much the same territory. Thus, in the 30 years since systematic reviews began in earnest, with around 15 years of intensified and large-scale reviewing effort, only a minority of trials have been assessed in systematic reviews. Given the triple constraint posed by the growth in trials, the increasing complexity of review methods, and current resources, Archie Cochrane's vision will not be achieved without some serious changes in course—in particular, with a greater concentration on Cochrane's use of the word "relevant".

Where to now?

First, we need to prioritise effectively and reduce avoidable waste in the production and reporting of research evidence. (31) This has implications for trials as well as systematic reviews. Some funders and others will now not consider supporting a trial unless a systematic review has shown the trial to be necessary. (32) It is essential that this requirement be more widely adopted. And it is essential that reviews address questions that are relevant to patients, clinicians and policymakers.

Second, we may need to choose between elaborate reviews of a quarter of the questions clinicians and patients have or “leaner” reviews of most of what we want to know. The methodological standards for systematic reviewing have been increasing over time, (30) and the evolution of standards in the Cochrane Collaboration and in HTA has been remarkable. The increase in steps and reporting required is reflected in the length of reviews. Early Cochrane reviews could typically be printed out in 10 or 20 pages, even when they incorporated several trials. Today, it is not unusual for a review by a health technology agency to run to several hundred pages. Often the reviews are longer than the combined length of the reports of all the included trials.

A contributing factor here is the increasing expectation for reviews to include study types other than randomised trials. This will often be essential for detecting less common adverse effects. However, the inclusion of all study types to answer all questions about the effects of treatments would not necessarily provide better quality information in every instance – while it would unquestionably increase the time and resource requirement for reviews. While it is vital that reviews are scientifically defensible, burdening those preparing them with excessive requirements could result in having valid answers to relatively few questions.

In particular, we need leaner and more efficient methods of staying up-to-date with the evidence. Using current methods, the Cochrane Collaboration has not been able to keep even half of its reviews up-to-date, (33) and other organisations are in a similar predicament. (34) We need to develop innovative methods to reduce the

labour of updating, and provide what clinicians and patients need: an assurance that a conclusion is not out of date, even if not every later trial is included within every analysis. It is also the responsibility of reviewer authors and journal editors to ensure that every new systematic review places itself clearly in context of other systematic reviews and HTAs. It will be to little avail to the average clinician, patient, and information provider, however, if the resulting knowledge is not comprehensible and openly accessible.

Finally, although more funding for evaluative clinical research internationally remains a priority, more international collaboration could result in better use being made of resources for systematic reviewing and HTAs. While multiple reviews on topics can provide a rounded picture of an area as well as a de facto form of updating when the reviews are conducted several years apart, there is also considerable duplication of review effort.

In November 2009, an international meeting in Cologne formed a new collaboration called “KEEP Up,” which will aim to harmonise updating standards and aggregate updating results. This should reduce the workload and enable organisations to be alerted when there are important shifts in evidence. Initiated and coordinated by the German Institute for Quality and Efficiency in Health Care (IQWiG) and involving key systematic reviewing and guidelines organisations such as the Cochrane Collaboration, Duodecim, the Scottish Intercollegiate Guidelines Network (SIGN), and the National Institute for Health and Clinical Excellence (NICE), this effort will provide a platform for tackling practical and methodological issues involved in keeping up-to-date.

There is nevertheless a risk that the increasing burdens placed on the methods of systematic reviewing could make the goal of keeping up-to-date with the knowledge won from trials recede ever more quickly into the distance. Perhaps one of the first questions we should ask whenever an additional process or more demanding methodology for systematic reviewing is proposed is this: Will this development serve or hinder our ability to better understand and communicate enough results from trials? In 1979, when Archie Cochrane argued that we needed critical summaries to keep up with the crucial knowledge those trials were generating, there were perhaps

14 trials a day being published. Thirty years later, it would be just as hard to keep up with the systematic reviews. Every day there are now 11 systematic reviews and 75 trials, and there are no signs of this slowing down: but there are still only 24 hours in a day.

Coda: The number of published trials and systematic reviews per day has more than quadrupled between 2007 and 2017. Including the full CENTRAL database (published and unpublished trials), there were 323 trials and 55 systematic reviews a day in 2017.

Methods for data and update in Appendix 1. Data available: https://osf.io/awktv

Acknowledgments

We are grateful to Sigrid Droste from the German Institute for Quality and Efficiency in Health Care (IQWiG) for her guidance on National Library of Medicine documentation on changes in reporting of publication types across time.

Author contributions (original)

ICMJE criteria for authorship read and met: HB PG IC. Agree with the manuscript's results and conclusions: HB PG IC. Designed the experiments/the study: HB. Analyzed the data: HB. Collected data/did experiments for the study: HB. Wrote the first draft of the paper: HB. Contributed to the writing of the paper: PG IC. Conceived the idea, participated in searching, data collection, and analysis for both the historical and publication trends: HB. Guarantor for the article: HB. Searched and analyzed publication trends: PG. Participated in the searching and analysis of historical trends: IC.

Update: Searched and analysed publication trends: HB.

References

1. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLOS Med. 2010 Sep 21;7(9):e1000326.

2. Cochrane AL. 1931–1971: a critical review, with particular reference to the medical profession. In: Medicines for the year 2000. London: Office of Health Economics; 1979. p. 1–11.

3. Lind J. A treatise of the scurvy. In three parts. Containing an inquiry into the nature, causes and cure, of that disease. Together with a critical and chronological view of what has been published on the subject. [Internet]. Edinburgh: Printed by Sands, Murray and Cochran for A Kincaid and A Donaldson; 1753 [cited 2009 Apr 26]. Available from: http://www.jameslindlibrary.org/trial_records/17th_18th_Century/lind/lind_kp.html

4. Duncan A. Introduction. Med Philos Comment. First(Part 1):6–7.

5. Cummings M. The National Library of Medicine. In: Coping with the biomedical literature: A primer for the scientist and the clinician. New York: Praeger; p. 161–173.

6. Barron BA, Bukantz SC. The Evaluation of New Drugs: Current Food and Drug Administration Regulations and Statistical Aspects of Clinical Trials. Arch Intern Med. 1967 Jun 1;119(6):547–56.

7. Banta D. The development of health technology assessment. Health Policy Amst Neth. 2003 Feb;63(2):121–32.

8. Stjernswärd J. Decreased survival related to irradiation postoperatively in early breast cancer. Lancet. 1974;304:1285–1286.

9. Chalmers TC. Effects of ascorbic acid on the common cold. An evaluation of the evidence. Am J Med. 1975;58(4):532–6.

10. Chalmers I. Randomized controlled trials of fetal monitoring 1973–1977. In: Perinatal Medicine. Stuttgart: Georg Thieme; 1979. p. 260–5.

11. Goldschmidt P. Information synthesis: a practical guide. Health Serv Res. 1986;21:215–36.

12. Mulrow CD. The medical review article: state of the science. Ann Intern Med. 1987 Mar;106(3):485–8.

13. L'Abbé K, Detsky A, O'Rourke K. Meta-analysis in clinical research. Ann Intern Med. 107:224–32.

14. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. N Engl J Med. 1987 Feb 19;316(8):450–5.

15. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. CMAJ Can Med Assoc J J Assoc Medicale Can. 1988 Apr 15;138(8):697–703.

16. Jenicek M. Méta-analyse en médecine. Évaluation et synthèse de l'information clinique et épidémiologique. St. Hyacinthe and Paris: EDISEM and Maloine Éditeurs; 1987.

17. Chalmers I. The work of the National Perinatal Epidemiology Unit. One example of technology assessment in perinatal care. Int J Technol Assess Health Care. 1991;7(4):430–59.

18. Starr M, Chalmers I, Clarke M, Oxman AD. The origins, evolution, and future of The Cochrane Database of Systematic Reviews. Int J Technol Assess Health Care. 2009 Jul;25 Suppl 1:182–95.

19. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009 Jan 21;(1):MR000006.

20. Lee K, Bacchetti P, Sim I. Publication of clinical trials supporting successful new drug applications: a literature analysis. PLoS Med. 2008 Sep 23;5(9):e191.

21. Kass E. Reviewing reviews. In: Coping with the biomedical literature: a primer for the scientist and the clinician. New York: Praeger; 1981. p. 79–91.

22. Chan A-W, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. Lancet Lond Engl. 2005 Apr 26;365(9465):1159–62.

23. International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals [Internet]. International Committee of Medical Journal Editors (ICMJE); 2018. Available from: http://www.icmje.org/icmje-recommendations.pdf

24. Zarin DA, Ide NC, Tse T, Harlan WR, West JC, Lindberg DAB. Issues in the registration of clinical trials. JAMA. 2007 May 16;297(19):2112–20.

25. One Hundred Tenth Congress of the United States of America. H.R.3580 - Food and Drug Administration Amendments Act of 2007 [Internet]. U.S. Congress; 2007 [cited 2019 Nov 9]. Available from: https://www.congress.gov/bill/110th-congress/house-bill/3580

26. Dijkers MPJM, Task Force on Systematic Reviews and Guidelines. The value of traditional reviews in the era of systematic reviewing. Am J Phys Med Rehabil. 2009 May;88(5):423–30.

27. Rada G, Pérez D, Capurro D. Epistemonikos: a free, relational, collaborative, multilingual database of health evidence. Stud Health Technol Inform. 2013;192:486–90.

28. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007 Aug 21;147(4):224–33.

29. Mallett S, Clarke M. THE TYPICAL COCHRANE REVIEW: How Many Trials? How Many Participants? Int J Technol Assess Health Care. 2002 Dec;18(4):820–3.

30. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and Reporting Characteristics of Systematic Reviews. PLOS Med. 2007 Mar 27;4(3):e78.

31. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet Lond Engl. 2009 Jul 4;374(9683):86–9.

32. Danish Research Ethics Committee System. Recommendation No. 20: Controlled clinical trials - the influence of existing and newly acquired scientific results on the research ethical evaluation. Copenhagen: Danish Research Ethics Committee System; 1997.

33. Koch G. No improvement - still less than half of the Cochrane reviews are up to date. In Dublin: Cochrane Collaboration; 2006. Available from: https://cmr.cochrane.org/?CRGReportID=10066

34. Garritty C, Tsertsvadze A, Tricco AC, Sampson M, Moher D. Updating systematic reviews: an international survey. PloS One. 2010 Apr 1;5(4):e9914.

Chapter 3

The epidemiology of systematic review updates: a longitudinal study of updating of Cochrane reviews, 2003 to 2018.

ABSTRACT

Background:

The Cochrane Collaboration has been publishing systematic reviews in the *Cochrane Database of Systematic Reviews* (*CDSR*) since 1995, with the intention that these be updated periodically.

Objectives:

To chart the long-term updating history of a cohort of Cochrane reviews and the impact on the number of included studies.

Methods:

The status of a cohort of Cochrane reviews updated in 2003 was assessed at three time points: 2003, 2011, and 2018. We assessed their subject scope, compiled their publication history using PubMed and *CDSR*, and compared them to all Cochrane reviews available in 2002 and 2017/18.

Results:

Of the 1,532 Cochrane reviews available in 2002, 11.3% were updated in 2003, with 16.6% not updated between 2003 and 2011. The reviews updated in 2003 were not markedly different to other reviews available in 2002, but more were retracted or declared stable by 2011 (13.3% versus 6.3%). The 2003 update led to a major change of the conclusions of 2.8% of the 177 updated reviews. The cohort had a

median time since publication of the first full version of the review of 18 years and a median of three updates by 2018 (range 1–11). The median time to update was three years (range 0–14 years). By the end of 2018, the median time since the last update was seven years (range 0–15). The median number of included studies rose from eight in the version of the review before the 2003 update, to 10 in that update and 14 in 2018 (range 0–347).

Conclusions:

Most Cochrane reviews get updated, however they are becoming more out-of-date over time. Updates have resulted in an overall rise in the number of included studies, although they only rarely lead to major changes in conclusion.

Hilda Bastian
Jenny Doust
Mike Clarke
Paul Glasziou

Background

Systematic reviews use explicit formal methods to identify and analyse studies on specific research questions, and synthesise the findings of these studies. (1) Reviews' findings can become outdated if they are overtaken by new studies or data, missing studies or errors in the review or included studies are identified, or methodology improves in critical ways. In a 2007 paper, Shojania et al reported on the need for updates in 100 systematic reviews from 1995 to 2005. (2) They concluded that around half may have been out-of-date, with a signal for required updating within two years of the evidence search in 11% of reviews. The median period without a signal suggesting a need to update was 5.5 years (95% confidence interval (CI), 4.6 to 7.6 years).

The intention to keep reviews up-to-date was a cornerstone of the Cochrane Collaboration. Named for Archie Cochrane, this international network aimed to achieve a goal he articulated by filling the need for: "a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomized controlled trials". (3) The Cochrane Collaboration began publishing systematic reviews of the effects of healthcare interventions based on clinical trials in the *Cochrane Database of Systematic Reviews* (*CDSR*) in 1995. (4) Its organisers wrote in an introductory brochure in the 1990s that evidence reviews "must be prepared systematically and they must be kept up-to-date to take account of new evidence". (5) A Cochrane review, they wrote, "is updated and amended as new evidence becomes available and errors are identified".

The original expectation was that reviews would be updated at least annually, but in 2000 it was agreed that this was unsustainable, (6) and the expected update interval was changed to every two years, unless a reason was given for a different schedule. (7) However, this also proved to be unsustainable, and in 2016 the organisation moved towards updating by perceived need or priority, preparing a consensus document on updating systematic reviews. (8) With the publication of a revision of the *Cochrane Handbook for Systematic Reviews of Interventions* (*Cochrane Handbook*) in 2019, there is no longer any default interval at which Cochrane reviews

are judged to be out of date, and the possibility of retracting reviews deemed outdated will cease. (9)

Cochrane reviews are the result of a unique large-scale and long-term effort to stay systematically up-to-date with health and social care evidence across a broad spectrum of topics. This collection of reviews also provides an opportunity to study the practice of systematic reviewing and growth of evidence across time.

We assembled a cohort of Cochrane reviews, comprising all the reviews flagged as updated in the *CDSR* in 2003, and assessed these across 15 years at three time points: 2003, 2011, and 2018. The aims of our study were to describe the cohort of updated reviews in relation to all other Cochrane reviews that were available at the end of 2002 and, in 2011 and around 2018. We also charted and followed up on the updating history of the 2003 update cohort. We assessed how often the 2003 updates led to a major change in reviews' conclusions. In addition, we charted the growth in the number of included studies in the reviews over time.

Methods

Study aim 1: Describe the cohort of reviews in the CDSR *updated in 2003 in relation to all Cochrane reviews available at the end of 2002, 2011, and around 2018.*

The cohort was established in 2003 by searching each of the four issues of the *CDSR* published that year for reviews flagged as updated. The issue(s) in which each review was flagged as updated was recorded.

(a) Subject scope.

Systematic reviews in different subject areas might become outdated more quickly than others. (2) Cochrane reviews are produced and maintained by editorial groups called Cochrane Review Groups (CRGs). These cover one or more subject areas, and have their own editorial policies. Differences in editorial practice could also affect

updating. The number of CRGs grew throughout the early years of the Collaboration, with some merging in later years.

In order to gauge how similar in subject and editorial scope the 2003 cohort reviews might be to Cochrane reviews overall, we compared the spread of reviews among CRGS. A list of CRGs in 2019 was collected from the Cochrane website. (10) The names of the CRGs in each year from 1995 to 2002 were collected using the Internet Archive, (11) from a combination of archived issues of *Cochrane News* (12) and the Cochrane website. (13) The original CRG names were retained, and also normalised to the current CRG name, and mergers of groups were noted.

The CRG base for each of the reviews in the cohort was collected, noting where CRGs which have since been merged into another CRG. We compared the number of CRGs represented in the cohort with the number of CRGs in 2002 and 2019.

To compare the subject mix of the 2003 cohort with current Cochrane review production and maintenance, we gathered the number of new and updated reviews by CRG from the online Cochrane Library, collecting data for both 2017 and 2018 to reduce the impact of annual fluctuations. As the search function does not enable separation of new and updated reviews, the totals of new and updated reviews from 2017 to 2018 were collected from the dashboards on Cochrane's website. (14)

(b) Likelihood of being updated.

In 2011, data was collected to enable comparison of updating history between the 2003 cohort and other Cochrane reviews. In the time since 2003, the *CDSR* had moved from publication on CD-ROM to online publication, and had a change of publisher. (4,15) It became apparent that many previous versions of Cochrane reviews were no longer published in the *CDSR* and so we used PubMed to establish a cohort of reviews that had been published in the *CDSR* by the end of 2002, supplemented by a search of *CDSR* for reviews still dated pre-2003. The number of reviews that had been published by the end of 2002 was available from another project, (16) but that did not provide identifiers for those reviews.

In 2011, the status of each review that had been available in 2002 was identified using both PubMed and the latest version in the *CDSR*. There were three possible status designations of Cochrane reviews in 2011 at that time: ongoing, designated stable (no further update required), or withdrawn. Which of these had been applied to each review was recorded. As some of the reviews were no longer included in the CDSR without any withdrawal notice, we assigned that as a fourth category. The categories withdrawn and no longer included without notice were classified together as retracted. Both PubMed and *CDSR* were used to establish whether the reviews had been updated between 2003 and 2011. Dates of updates were defined as the date of publication of a new citation version of the review, and date of search or incorporation of new data if no new cited version was published. When a version of a review involved only a software update, we did not count this as an update.

To compare updating and status between the reviews available in 2002 which had, and had not, been updated in 2003, we analysed the rate of ongoing, stable, and retracted reviews. We also analysed the proportions of ongoing reviews that had not been updated since 2002, those that were updated in 2003 only, and those that had been updated at least once between 2004 and 2011.

Study aim 2: Describe the updating history of the cohort, major changes in conclusions in 2003, and the growth of included studies over time.

All reviews flagged as updated in 2003 were collected. To assess the changes to the reviews' conclusions, one author (HB) assessed all of these updates, and another (JD) assessed those published as updates in two of the four issues of the *CDSR*. Both authors agreed on a final group of reviews that had a major change in their conclusions following the update.

The first follow-up of the 2003 cohort of reviews was done in 2011. Using the first issue of *CDSR* in that year (published in April), the number of years from the review's last reported search for eligible studies to April 2011 was recorded. After publication of the fourth and final issue of the *CDSR* for 2011, as complete as possible an

updating history of the 2003 cohort was assembled. For each review in the cohort, PubMed was searched for previous records of the review. The version of the review published in issue 4 of 2011 was reviewed for information about updates in the three parts of the review that were expected to report the review's history: What's New, history, and notes. (17) As practice in what constituted an update had changed since 2003, two types of events were recorded as updates: the year of publication of a new citation of a version of the review in PubMed, and the year reported in the review for an update search or incorporation of new data even if no new cited version was published.

In March 2019, we added the year for each update from 2012 to the end of 2018. This was again based on PubMed searches and the information recorded in the reviews' What's New, history, and notes sections. In addition, we added the year the review was first published and the most recent PubMed identifier for the review.

This final data collection for the 2003 cohort alone included an assessment of the status of these reviews at the end of 2018, using the same categories as previously. However, we added a category for 2018: republished after previous retraction. For reviews that were stable, retracted, or republished after previous retraction, we recorded the years of these events.

When we originally identified reviews as updated in 2003, we recorded the number of studies included in that update, as well as the number included in the prior issue of the review. For one review, the updated version of the review was the only one available for assessment. We subsequently recorded the number of included studies in the versions of each review at the end of 2011 and the end of 2018.

Data management and analysis

One author (HB) undertook all data collection, curation, and visualisation. Data for the first collection in 2003 was originally recorded in document tables, and added to an Excel spreadsheet in 2019. All other collections were recorded in Excel. Data was analysed using RStudio 1.1463 running R 3.5.2. (9) Packages tidyverse and (20)

reshape2 (21) were used for analyses and data visualisation. Summary statistics were used to describe the cohort. Data for this project, including analytic code, will be deposited at Github. (22)

Results

2003 cohort in perspective: subject scope.

We identified 1,532 Cochrane reviews via PubMed and the *CDSR* that were available at the end of 2002. This was fewer than the number reported by the Cochrane Collaboration since the beginning of the *CDSR* in 1995 (1,558). (16) Some reviews had been indexed in PubMed after 2002 (n = 13), and this is likely to be the case for more of the shortfall of 26 reviews. Other reviews may have been retracted prior to the indexing of the *CDSR* in PubMed in 2000. (15) A total of 177 reviews were flagged as updated in the *CDSR* in 2003. As the first version of four of these was also published in 2003, the update rate in 2003 of 1,532 reviews indexed in PubMed to the end of 2002 was 11.3%.

There were 49 Cochrane Review Groups (CRGs) with editorial responsibility for Cochrane reviews in 2002, compared with 53 in 2019. Table 1 shows the rise in the number of CRGs between 1995 and 2002, as well as annual review production based on Cochrane-supplied data and the years of first publication of the 173 reviews in the update cohort reviews that had been published by the end of 2002. CRGs had a median of two reviews updated in 2003, ranging from none to 27 (interquartile range (IQR): 3).

Table 1. Annual number of CRGs and new Cochrane reviews, 1995–2002, with year of first publication for the 2003 update cohort.

Year	Number of CRGs	Number of reviews		Year of first publication of the reviews in the 2003 update cohort	
		New	Cumulative total	Number	% of all reviews[1]
1995	18	65	65	10	15.4
1996	22	105	170	13	7.7
1997	30	139	309	15	4.9
1998	45	208	517	21	4.1
1999	46	184	701	20	2.9
2000	48	264	965	37	3.8
2001	49	309	1274	36	2.8
2002	49	284	1558	21	1.3
Total	*n.a.*	*1,558*	*n.a.*	*173*[2]	*n.a.*

[1] Percentage of total number of Cochrane reviews available in that year.

[2] This excludes four reviews in the 2003 update cohort that were also published for the first time in 2003.

Only four CRGs had been in existence for less than four years in 2002, and 19 were no more than five years old. At that time annual updating of Cochrane reviews was expected, but only 42 of the 49 CRGs (85.7%) flagged reviews as updated in 2003, and only 11.3% of the reviews published by the end of 2002 were updated during 2003.

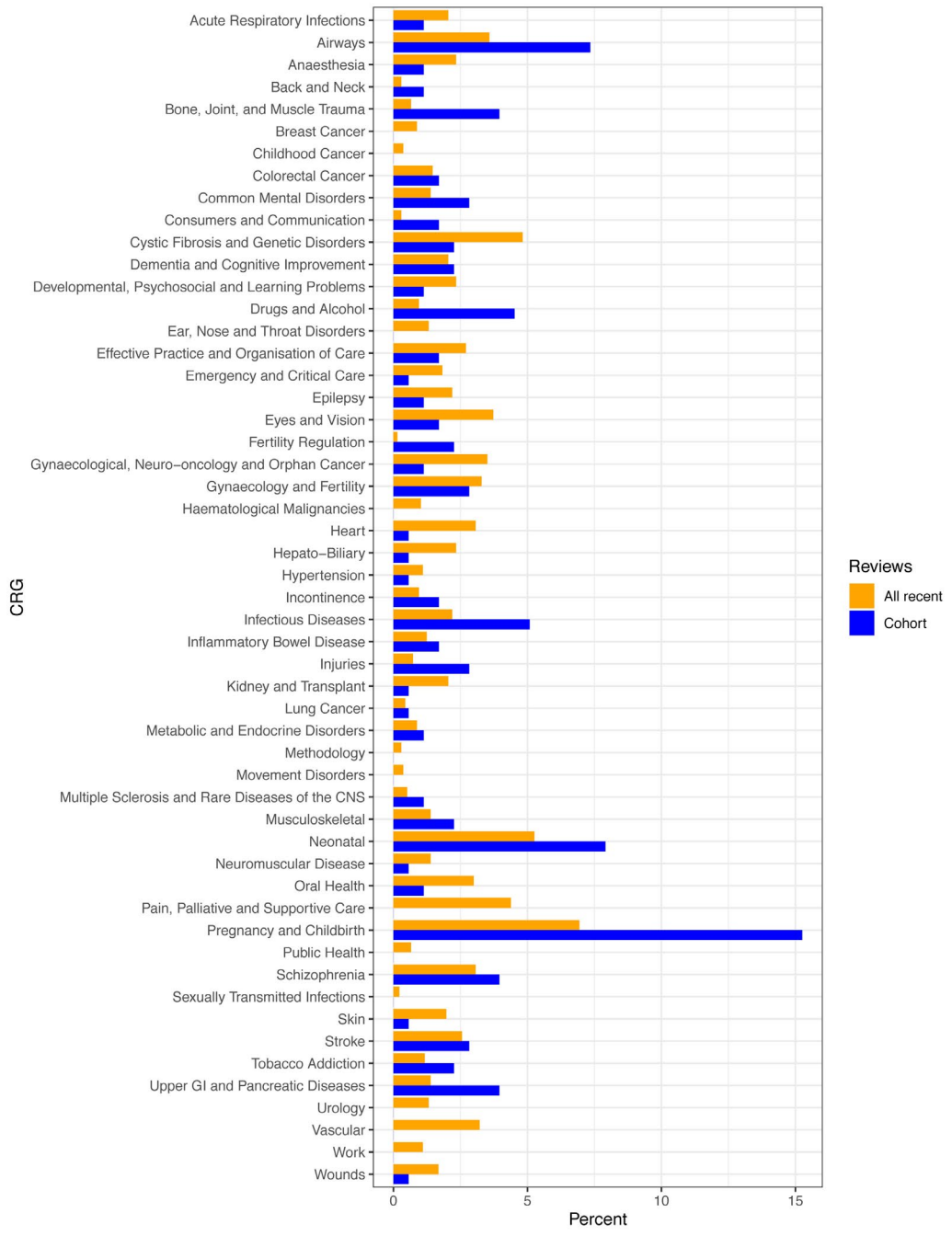
To gauge how similar the subject scope of reviews in the 2003 update cohort are to the current subject scope of Cochrane reviews, we compared the spread of reviews among CRGs in the 2003 update cohort with the spread of new and updated reviews published in 2017/18 (Table 2). The *CDSR* does not enable breaking these down into new and updated reviews but the 2017/2018 totals reported in Cochrane dashboards online show that 44.8% of these reviews were updates. (14) (The dashboards report fewer total new and updated reviews for 2017/2018: 1,353 versus 1,369 returned by *CDSR* search).

Table 2. Proportion of reviews per CRG: 2003 update cohort and all new and updated reviews in 2017/18.

Individual CRGs	2003 update cohort (n = 177)	All reviews 2017/2018 (n = 1,369)
Number of reviews per CRG (median)	3	19
Number of reviews per CRG (range)	1 to 27 (IQR 3)	2 to 95 (IQR 23)
Proportion of the CRG's reviews (median)	1.7%	1.4%
Proportion of the CRG's reviews (range)	0.6% to 15.3%	0.1% to 6.9%

Cochrane reviews in the 2003 update cohort were distributed across a narrower range of CRGs, and therefore topic areas, than the new and updated reviews in 2017/18 (Table 2). The extent of this shift is illustrated in the case of the oldest CRGs. Of the 18 CRGs that had formed by 1995, two later merged. Those 17 CRGs were responsible for 57.6% of the 2003 update cohort (102 of 177 reviews), and 48.0% of the new and updated reviews in 2017/18 (657 of 1,369 reviews). The individual CRGs, based on their 2019 names and grouping, are shown with their proportion of reviews in both time periods in Figure 5.

Figure 5. Proportion of reviews per CRG: 2003 update cohort and new and updated reviews in 2017/18.



Note: Two individual CRGs from 2003 had merged by 2019 and are treated as merged in 2003 for this comparison.

2003 cohort in perspective: likelihood of being updated.

The updating status of the 1,532 reviews published by the end of 2002 was categorised in 2011 to show whether each review was ongoing, stable (no longer being updated), or retracted at that time, and whether or not they had been updated at least once between 2003 and 2011. The retracted reviews included those with a withdrawal notice as well as six reviews that were no longer in the *CDSR*, with no record to explain their absence.

We compared the 173 reviews in the 2003 update cohort for which the original version had been published before 2003 with the other reviews that were available at the end of 2002 (Table 3).

Table 3. Comparison of 2002 reviews updated in 2003 (*n* = 173) with update status of other reviews available in 2002 (*n* = 1532).

Review status by the end of 2011	**Updated in 2003**		**No 2003 update**		**Total**	
	No.	*%*	*No.*	*%*	*No.*	*%*
Ongoing	150	86.7	1,274	93.7	1,424	93.0
Stable	11	6.4	26	1.9	37	2.4
Retracted	12	6.9	59	4.3	71	4.6
Total reviews	***173***	***100***	***1,359***	***100***	***1,532***	***100***
Ongoing with no update since 2002	–	–	237	18.6	237	16.6
Ongoing and updated in 2003 only	22	14.7	–	–	22	1.5
Ongoing and updated at least once between 2004 and 2011	128	85.3	1,037	81.4	1,165	81.8
Total ongoing reviews	***150***	***100***	***1,274***	***100***	***1,424***	***100***

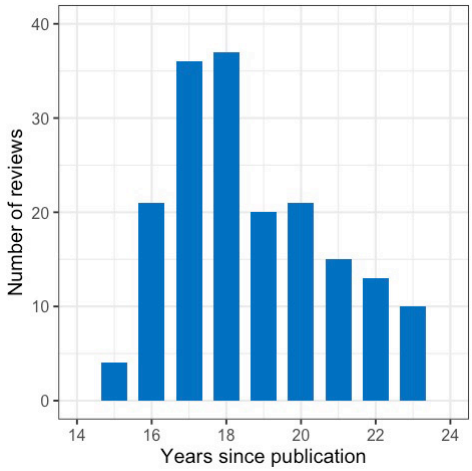
The proportion of 2002 Cochrane reviews that were ongoing in 2011 and had been updated at least once between 2004 and 2011 was high, whether they had been updated in 2003 (85.3%) or not (81.4%), suggesting that the 2003 update cohort is not markedly different to the other 2002 reviews in regard to subsequent updating

over the following eight years. However, a higher proportion of the 2003 cohort was retracted or declared stable than other reviews from 2002 by 2011 (13.3% versus 6.3%). By 2019, the overall proportion of non-retracted Cochrane reviews being declared stable was 6.6% *[Chapter 4]*, compared with 6.8% for our 2003 update cohort (11 of 161 non-retracted reviews). In regard to the number of included studies, the 2003 update cohort include two reviews with no included studies in 2011 (1.1%), which was a lower proportion than Yaffe et al found for all Cochrane reviews in 2010 (8.7%). (23)

Updating history of the 2003 update cohort from publication to 2018.

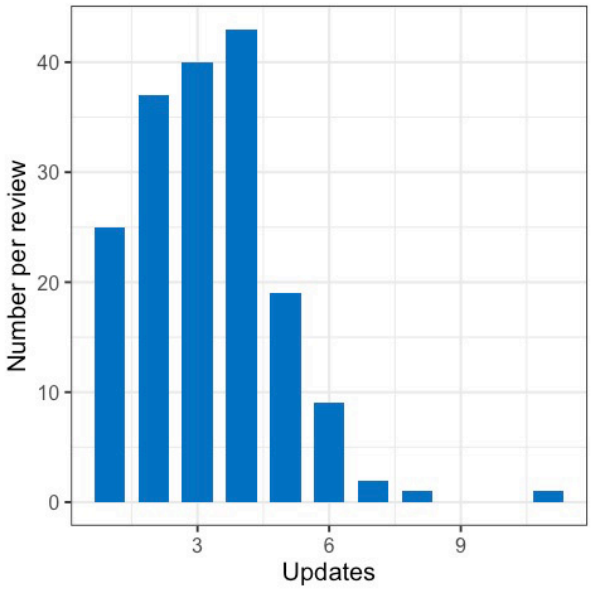
The reviews in the 2003 cohort were first published as full Cochrane reviews between 1995 and 2003, with a median time since first publication of 18 years by the end of 2018 (range: 15 to 23 years; IQR 3) (Figure 6).

Figure 6. Years since first publication of reviews to the end of 2018 (n = 177).



From first publication until 2018, the median number of updates per review was three (range: 1 to 11, IQR 2) (Figure 7).

Figure 7. Number of updates per review to 2018 (n = 177).



Among the 177 reviews in the 2003 update cohort, 150 were ongoing in 2018: the other 27 were either retracted (n = 15) or designated stable (n = 12). One review had been retracted in 2010, but was updated and republished the following year, and is included among the 150 ongoing reviews.

The median time to each update of these reviews was three years (Table 4). The shortest was zero in the case of the first update, as some reviews were updated during the year they were published. The longest interval between updates was 14 years between the first and second updates of one review.

Table 4. Median time in years to first and subsequent updates for the 150 ongoing reviews from the 2003 update cohort.

No. (%)	Time to update in years – median (range)										
	1st	*2nd*	*3rd*	*4th*	*5th*	*6th*	*7th*	*8th*	*9th*	*10th*	*11th*
150 (100)	3 (0-8)										
135 (90.0)		3 (1-14)									
107 (71.3)			3 (1-8)								
72 (48.0)				3 (1-10)							
32 (21.3)					3 (1-8)						
13 (8.7)						3 (1-9)					
4 (2.7)							3 (1-5)				
4 (2.7)								3 (1-5)			
2 (1.3)									2		
1 (0.7)										2	
1 (0.7)											2

Figure 8 charts the ongoing reviews in 2018 by the years since their last update. A total of 36 (24.0%) reviews had not been updated in the previous 10 years. Further analyses are included in Table 5, together with other analyses of the currency of the 2003 update cohort.

Figure 8. Ongoing reviews by years since last update, 2018 (n = 150).

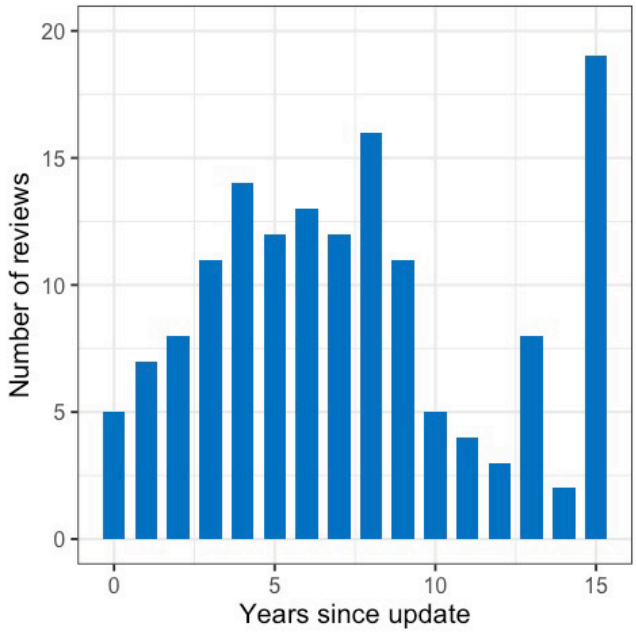


Table 5. Additional indicators of currency of the cohort of reviews.

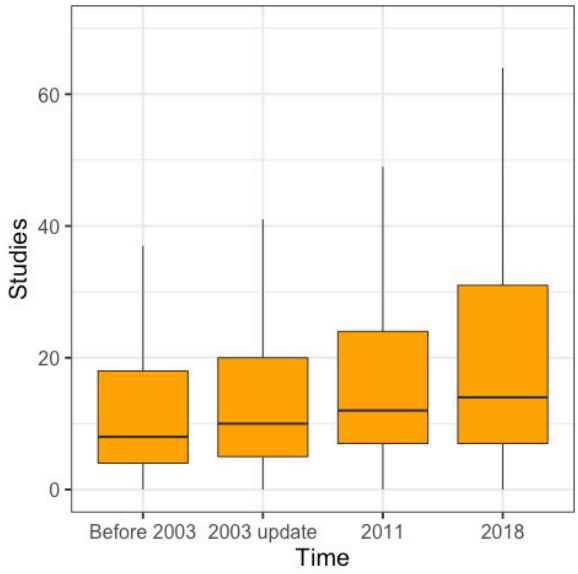
Ongoing reviews at the end of 2018 (n = 150)	
Years since last update *(Figure 4)*	Median 7 years (range: 0 – 15; IQR 6)
Number and proportion with an update within the last two years	20 (13.3%)
Number and proportion with an update within the last six years[1]	70 (46.7%)
Proportion of reviews' publication life spent as "up-to-date" based on 2-year updating interval	Median 43.8% (range: 13.3% – 94.4%; IQR 22.3)
Date of last search (2011 data)	
Years since date of last search (as of April 2011)[2]	Median 3 years (range: 0 – 13; IQR 5)
Years since date of last search for reviews that have not been updated since 2011 (n = 106, 59.9%)[2]	Median 11 years (range: 7 – 15; IQR 6)

[1] The median "survival time" of a systematic review based on Shojania et al (2) is 5.5 years.
[2] Search date was unknown for 13 reviews.

Impact of updating over time.

In 2003, the updating of a review resulted in major changes to its conclusions for five reviews (2.8%). We compared the number of included studies in the updated review with the number in its prior version (excluding one review had been retracted without leaving a copy of the original review in *CDSR*, leading to a single missing baseline value). Some updates did not lead to the inclusion of further studies, but the number of included studies grew substantially over time (median: 8 from the version before the 2003 update to 14 in the version of the review at the end of 2018) (Figure 9).

Figure 9. Included studies in reviews: baseline, 2003, 2011, and 2018.



Notes: Outliers not displayed; 1 missing value at baseline (before 2003).

Time	*Median*	*Range*	*IQR*
Before 2003	8	0–176	14
2003 update	10	0–176	15
2011	12	0–347	17
2018	14	0–347	24

Discussion

Our study describes the updating history over 15 years of a cohort of 177 Cochrane reviews that were first updated in 2003. For most of this period, the Cochrane updating policy was a recommended two-year interval until a review was regarded as out of date. By that measure, reviews in this cohort could be considered out of date for more than half of their publication life. However, in another study Shojania et al concluded that the median life of a systematic review before it became outsided was 5.5 years (CI, 4.6 to 7.6 years), based on a sample of 100 systematic reviews. As the median time between updates for the Cochrane reviews in our 2003 update cohort was three years, most of this cohort were likely to be up-to-date most of the time.

However, this cohort has been getting more out of date over time. The median time to first update was three years, but by the end of 2018, more than half the reviews were more than seven years since their last update (range 0–15 years). This is reflective of Cochrane reviews more generally, as more new reviews are published than updated and the legacy of existing reviews now exceeds 8,000. If indeed the interval between review updates continues to increase in the long term, or more and more reviews are never updated, up-to-date Cochrane reviews will become the exception, not the rule.

Until recently, Cochrane policy allowed for withdrawing a review from publication when it was seriously out of date. This resulted in a higher proportion of retraction for our cohort (6.9%), relative to its fellow Cochrane reviews that were not updated in 2003 (4.3%). With the introduction of a new update classification system that aims to phase out the withdrawal of outdated reviews, (24) the retraction rate in future is likely to be considerably lower. The proportion of reviews showing as apparently current in the *CDSR* but which are seriously outdated will be correspondingly higher.

Our cohort had a higher number of updates than other Cochrane reviews from the early 2000s. In addition, they included a particularly low proportion of “empty” reviews in 2011 (1.1% without included studies) compared to the 8.7% Yaffe et al reported for all Cochrane reviews in 2010. (23) In our study, the median number of included

studies prior to the 2003 update was eight. Mallett and Clarke reported that the median number of included studies in all Cochrane reviews at the beginning of 2001 was six. (25) The 2019 *Cochrane Handbook* advises authors of reviews to consider, among other things, whether new eligible studies are likely to be found before deciding to update. (9) Our results are consistent with Cochrane groups already concentrating updating effort on research questions that were generating new studies.

Our results for the earlier years are similar to the results from studies with shorter follow-up than ours. We identified 13 other studies of updating in Cochrane reviews or protocols that addressed some similar outcomes (overview in Appendix). (26–38) Mean or median times to updates were comparable to those in our study. Our finding that a major change in conclusion after update is rare (2.8%) is also consistent with others' results. Jaidee et al found a 2.0% rate of major changes in their conclusions at the first update of 101 Cochrane reviews. (35) Bashir et al found a 3.9% change in their conclusions in 8 out of 204 reviews with a meta-analysis for a primary outcome. (38) The highest rate was found by French et al, (30) who found a 9.1% change in review conclusions in 23 out of 254 reviews, but these changes were not necessarily major.

There have been two studies with similar analyses of updating non-Cochrane evidence syntheses. The National Institute for Health and Care Excellence (NICE) reported on updating of their systematic review-based clinical practice guidelines. (39) For 11 guidelines, the median time to update was 5.3 years (range: 3.3 to 6.5 years), with major changes in recommendations in six of them. Peterson et al studied 41 comparative effectiveness reviews of drugs, finding a median time to update of just over two years. (40)

Our study has several important strengths, particularly its long-term follow-up and open data that could enable others to extend this longitudinal study. The collection of cross-sections of data in 2003, 2011, and 2017/2018 was valuable. By maintaining a historical collection of data and using PubMed as well as the *CDSR*, we established that the updating and version history of Cochrane reviews in the *CDSR* has important

gaps. While most older versions of reviews are online in the *CDSR*, some complete reviews and some update versions of reviews have been removed from the *CDSR* without leaving any public record other than PubMed. Some updated versions of Cochrane reviews were never submitted to PubMed. In some cases this may be in error, but it may also be a matter of policy. (7) This has left no complete public record of all Cochrane reviews, their updates, versions, and retractions. Not submitting updates for indexing at PubMed has critical implications for searches: reviews with misleadingly older dates will not be retrieved by searches that are limited to more recent records.

The inadequate historical record was one of the limitations of our study. We also relied on manual collection of data and quality assurance by a single author. In assessing the growth of included studies, we did not collect data on eligible studies awaiting inclusion and we do not know if this would have affected our results.

Our study raises some issues for Cochrane and others to consider. We could find the date of last search clearly reported in 92.7% of the reviews in 2011. In a study published in 2013, Beller et al found that only 90.0% of a sample of systematic reviews from PubMed reported the date of last search. (41) The availability of that date is critical for users to assess the currency of evidence and the adequacy of overlapping periods in update searches.

In 2002, Koch drew attention to the poor quality of update reporting in Cochrane reviews, (27) and we found this to be an ongoing problem. Contributing factors include missing versions of Cochrane reviews in the *CDSR*, and reviews not carrying forward events from “What’s New” and other notes sections into the history of each subsequent version of the review as Cochrane advises. (7) Automation of that process could improve this situation.

A further critical gap is the lack of reporting of searches undertaken before it is determined that an update is needed. Although searches that found no new studies were generally reported as updates to the reviews in early years when annual or two-yearly updates were expected, these are no longer as visible. (8,9) Transparency of

this element of updating activity is important for users, including other systematic reviewers and producers of clinical practice guidelines and health information generally. A survey of health organisations producing systematic reviews (and often clinical practice guidelines) by Garritty et al in 2010 had 114 respondents (30% of them from Cochrane). Only eight of the organisations did not have any updating activity, but resources for updating for those that did were stretched, and the groups regarded too many of their reviews as too far out of date. Sharing the results of searches is critical in this context to prevent large numbers of groups using resources on the same futile searches.

In 2010, two of us advocated that major changes were needed to keep up with the evidence, given the massive rise in clinical trials and increasing complexity of systematic review methods, with constrained resources. (16) Prioritisation of systematic reviews, reduction of avoidable waste, and “leaner and more efficient methods of staying up-to-date” were stressed. Since then, the rise in clinical trials has escalated: ClinicalTrials.gov amassed over 96,000 trials in its first decade to 2010, and the total nearly doubled in the next five years. (44,45) It now stands at over 320,000 (as of November 2019). Progress in streamlining methods and avoiding waste are not moving as fast. Methodological expectations for Cochrane reviews have increased, (9,42) and trial registry entries have been added to Cochrane’s trial database as the organization grapples with the implications of incorporating unpublished trials and data. (46) There has, however, been a concerted effort to develop and implement methods for prioritisation of updating for Cochrane reviews. (8) With this transition to targeted updating, new questions about processes and impact need to be tackled. Are the right systematic reviews being updated? Are people being harmed by reliance on outdated systematic reviews?

Author contributions

HB initiated the study, undertook the data collections in 2003, 2011, and 2018, curated, analysed and visualised data, and drafted the manuscript. HB and JD jointly conceptualised the establishment of the cohort in 2003 and agreed on reviews with major changes in conclusions. HB, JD, PCG, and MC participated in conceptualising the 2011 follow-up. The 2018 follow-up was initiated and conceptualised by HB. All authors contributed to interpretation of the data and critical revision of the manuscript.

Disclosures

Hilda Bastian was a member of the Cochrane Collaboration's governing body from the organisation's founding in 1993 until 2001, the coordinating editor of a CRG (Consumers and Communication) from 1997 to 2001, and a member of the GRADE working group from 2006 to 2012. She received support from an Australian Government Research Training Program Scholarship. Jenny Doust is an editor for the Cochrane Acute Respiratory Infectious Group. Mike Clarke was a member of the Cochrane Collaboration's governing body from 1998 until 2004 (and its Chair from 2002 to 2004), the Director of the UK Cochrane Centre from 2002 to 2011, and since 2007 has been the coordinating editor of a CRG (Methodology Review Group).

Acknowledgments

We are grateful to Robyn Brown, who participated in the first assessment of the cohort in 2003. We are also grateful to Claire Allen, who in 2009 provided the data underlying the chart of Cochrane reviews from 1995 to 2008 then published on the Cochrane website.

References

1. Moher D, Stewart L, Shekelle P. All in the Family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. Syst Rev. 2015 Dec 22;4(1):183.

2. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007 Aug 21;147(4):224–33.

3. Cochrane AL. 1931–1971: a critical review, with particular reference to the medical profession. In: Medicines for the year 2000. London: Office of Health Economics; 1979. p. 1–11.

4. Starr M, Chalmers I, Clarke M, Oxman AD. The origins, evolution, and future of The Cochrane Database of Systematic Reviews. Int J Technol Assess Health Care. 2009 Jul;25 Suppl 1:182–95.

5. Cochrane Collaboration. The Cochrane Collaboration [Internet]. Cochrane Collaboration; 1998. Available from: https://web.archive.org/web/19980115184957/http://som.flinders.edu.au/FUSA/COCHRANE/cochrane/cc-broch.htm

6. Cochrane Collaboration. Minutes of meeting of the Cochrane Collaboration Steering Group, Cape Town, South Africa, 24 and 29 October 2000 [Internet]. Cochrane Collaboration; 2000. Available from: https://community.cochrane.org/sites/default/files/uploads/inline-files/minutes_reports/ccsg/2000.10.24.htm

7. Higgins JP, Green SE. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 5.1.0. Oxford: Cochrane Collaboration; 2011. Available from: http://handbook-5-1.cochrane.org/

8. Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ. 2016 20;354:i3507.

9. Higgins JPT, Thomas J. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 6.0. London: Cochrane Collaboration; 2019 [cited 2019 Oct 2]. Available from: https://training.cochrane.org/handbook/current

10. Cochrane Collaboration. Cochrane Review Groups and Networks [Internet]. Cochrane Library. [cited 2019 Oct 10]. Available from: https://www.cochranelibrary.com/about/cochrane-review-groups

11. Internet Archive. Wayback Machine [Internet]. Internet Archive Wayback Machine. [cited 2019 Oct 18]. Available from: https://archive.org/web/

12. Cochrane Collaboration. Cochrane News (Internet Archive) [Internet]. The Cochrane Collaboration. 2012. Available from: https://web.archive.org/web/20130214130017/http://www.cochrane.org/news/newsletters/cochrane-news

13. Cochrane Collaboration. Cochrane website (Internet Archive) [Internet]. The Cochrane Collaboration. [cited 2019 Nov 13]. Available from: https://web.archive.org/web/*/cochrane.org

14. Cochrane Collaboration. Dashboard [Internet]. Cochrane Community. [cited 2019 Oct 30]. Available from: https://community.cochrane.org/organizational-info/resources/strategy-2020/dashboard

15. U.S. National Library of Medicine. The Cochrane database of systematic reviews. In: NLM Catalog [Internet]. Bethesda: US National Library of Medicine, NIH; [cited 2019 Jul 29]. Available from: https://www.ncbi.nlm.nih.gov/nlmcatalog/100909747

16. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLOS Med. 2010 Sep 21;7(9):e1000326.

17. Higgins JPT, Green SE. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 5.0.0. Oxford: Cochrane Collaboration; 2008. Available from: https://handbook-5-1.cochrane.org/v5.0.0/

18. RStudio Team. RStudio: integrated development environment for R. [Internet]. Boston: RStudio Inc; 2015 [cited 2019 Aug 12]. Available from: https://www.rstudio.com/

19. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing; 2016. Available from: https://www.r-project.org/

20. Wickham H. tidyverse [Internet]. RStudio Inc; 2017. Available from: https://CRAN.R-project.org/package=tidyverse

21. Wickham H. reshape2 [Internet]. 2017. Available from: https://CRAN.R-project.org/package=reshape2

22. Bastian H. Cochrane updating. GitHub [Internet]. 2019 Nov; Available from: https://github.com/hildabast/cochrane-updating

23. Yaffe J, Montgomery P, Hopewell S, Shepard LD. Empty reviews: a description and consideration of Cochrane systematic reviews with no included studies. PloS One. 2012;7(5):e36626.

24. Cochrane Editorial Unit. Updating Classification System: guide to applying to Cochrane Reviews [Internet]. Cochrane Collaboration; 2019. Available from: https://community.cochrane.org/sites/default/files/uploads/inline-files/Cochrane_UCS-

Guide_8Jul16_0.pdf#targetText=The%20Updating%20Classification%20System%20(UCS,decisions%20for%20individual%20Cochrane%20Reviews.

25. Mallett S, Clarke M. THE TYPICAL COCHRANE REVIEW: How Many Trials? How Many Participants? Int J Technol Assess Health Care. 2002 Dec;18(4):820–3.

26. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al. Methodology and Reports of Systematic Reviews and Meta-analyses: A Comparison of Cochrane Reviews With Articles Published in Paper-Based Journals. JAMA. 1998 Jul 15;280(3):278–80.

27. Koch G. Are Cochrane Reviews Updated Regularly or Not? In Stavanger: Cochrane Collaboration; 2002. Available from: https://cmr.cochrane.org/?CRGReportID=4572

28. Koch G. Less than half of the Cochrane Reviews are current. In Barcelona: Cochrane Collaboration; 2003. Available from: https://cmr.cochrane.org/?CRGReportID=5113

29. González de Dios J. [Bibliometric analysis of systematic reviews in the Neonatal Cochrane Collaboration. Its role in evidence-based decision making in neonatology]. An Pediatr Barc Spain 2003. 2004 May;60(5):417–27.

30. French SD, McDonald S, McKenzie JE, Green SE. Investing in updating: how do conclusions change when Cochrane systematic reviews are updated? BMC Med Res Methodol. 2005 Oct 14;5:33.

31. Henderson S, Hampson L, Atherton D, Neilson J. Ten years of updating Cochrane reviews: the experiences of the Pregnancy and Childbirth Group. In Dublin: Cochrane Collaboration; 2006 [cited 2019 Jul 29]. Available from: https://cmr.cochrane.org/?CRGReportID=9981

32. Koch G. No improvement - still less than half of the Cochrane reviews are up to date. In Dublin: Cochrane Collaboration; 2006. Available from: https://cmr.cochrane.org/?CRGReportID=10066

33. Shea B, Boers M, Grimshaw JM, Hamel C, Bouter LM. Does updating improve the methodological and reporting quality of systematic reviews? BMC Med Res Methodol. 2006 Jun 13;6:27.

34. Tricco AC, Brehaut J, Chen MH, Moher D. Following 411 Cochrane protocols to completion: a retrospective cohort study. PloS One. 2008;3(11):e3684.

35. Jaidee W, Moher D, Laopaiboon M. Time to update and quantitative changes in the results of cochrane pregnancy and childbirth reviews. PloS One. 2010 Jul 13;5(7):e11553.

36. Ngamjarus C, Jaidee W, Laopaiboon M. Progress in time to update of Cochrane Systematic Reviews. In: Cochrane Database of Systematic Reviews [Internet].

Cochrane Collaboration; 2011. p. 159–60. Available from: https://cmr.cochrane.org/?CRGReportID=16807

37. Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. BMJ. 2013 Dec 13;347:f7191.

38. Bashir R, Surian D, Dunn AG. Time-to-update of systematic reviews relative to the availability of new evidence. Syst Rev. 2018 17;7(1):195.

39. Lyratzopoulos G, Barnes S, Stegenga H, Peden S, Campbell B. UPDATING CLINICAL PRACTICE RECOMMENDATIONS: IS IT WORTHWHILE AND WHEN? Int J Technol Assess Health Care. 2012 Jan;28(1):29–35.

40. Peterson K, McDonagh MS, Fu R. Decisions to update comparative drug effectiveness reviews vary based on type of new evidence. J Clin Epidemiol. 2011 Sep 1;64(9):977–84.

41. Beller EM, Chen JK-H, Wang UL-H, Glasziou PP. Are systematic reviews up-to-date at the time of publication? Syst Rev. 2013 May 28;2(1):36.

42. Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. Methodological Expectations of Cochrane Intervention Reviews (MECIR) [Internet]. London: Cochrane Collaboration; 2019. Available from: https://community.cochrane.org/mecir-manual

43. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009 Oct;62(10):e1-34.

44. Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of clinical trials registered in ClinicalTrials.gov, 2007-2010. JAMA. 2012 May 2;307(17):1838–47.

45. van Valkenhoef G, Loane RF, Zarin DA. Previously unidentified duplicate registrations of clinical trials: an exploratory analysis of registry data worldwide. Syst Rev. 2016 15;5(1):116.

46. Cochrane Controlled Register of Trials (CENTRAL) | Cochrane Library [Internet]. [cited 2019 Nov 14]. Available from: https://www.cochranelibrary.com/central/about-central

Chapter 4

Enough evidence and other endings: a descriptive study of stable Cochrane systematic reviews in 2019.

Abstract

Background

From 2006 to 2019, Cochrane reviews could be designated “stable” if they were not being updated but highly likely to be current. This provides an opportunity to observe practice in ending systematic reviewing and what is regarded as enough evidence.

Methods

We identified Cochrane reviews designated stable in 2013 and 2019 and reasons for this designation. For those with conclusions stated to be so firm that new evidence is unlikely to change them, we assessed conclusions, strength of evidence ratings, and recommendations for further research. We assessed the fate of the 2013 stable reviews. We also estimated usage of formal analytic methods to determine when there is enough evidence in protocols for Cochrane reviews.

Results

Cochrane reviews were rarely designated stable. In 2019, there were 507 stable Cochrane reviews (6.6% of 7,645 non-withdrawn reviews). The most common reasons related to no, little, or infrequent research activity expected (331 of 505; 65.5%). Only 39 reviews were stable because of firm conclusions unlikely to be

changed by new evidence (7.7%), but that declaration was mostly not supported by judgments made in the review about strength of evidence and implications for research. Among the 180 reviews stable in 2013, 16 reverted to normal status (8.9%), with 2 of those changing conclusions because of new studies. Few Cochrane protocols specified an analytic method for determining when there was enough evidence to stop updating the review (116 of 2,415; 4.8%).

Conclusion

Cochrane reviews were more likely to end because important future primary research activity was believed to be unlikely, than because there was enough evidence. Judgments about the strength of evidence and need for research were often inconsistent with the declaration that conclusions were unlikely to change. The inconsistencies underscore the need for reliable analytic methods to support decision-making about the conclusiveness of evidence.

Hilda Bastian
Lars Hemkens

Background

The question of when there is enough evidence to be certain about effects in health is a critical one: both premature and overdue certainty can do damage. In 1992, Antman and colleagues demonstrated that harmful clinical advice and initiation of redundant clinical trials could continue long past the time a question has been definitively answered. (1) They argued that keeping on top of trial results with systematic reviews could help reduce this problem, a point of view reiterated by Chalmers and Glasziou in a 2009 paper on avoidable research waste. (2)

Systematic reviews involve searching for studies on a question and synthesising the findings, using explicit formal methods. (3) They can be outdated by subsequent studies, sometimes quite quickly. (4) Systematic reviews of clinical trials therefore need to be monitored and updated as long as a critical question about effectiveness or safety remains. This was made feasible on a large scale by developments in information technology and publishing, and enabled the establishment of the Cochrane Collaboration in 1993. (5) The Cochrane Collaboration developed a reviewer and methodologist community around the production and updating of systematic reviews, underpinned by a clinical trials register and dissemination through its own journal, the *Cochrane Database of Systematic Reviews* (CDSR). (6,7) The Cochrane community also developed a range of technical and social mechanisms, and played an often-pioneering role in establishing and evaluating scientific methodology. (8)

The Collaboration had an initial goal of continual updating and at least one update a year, moving to a default update interval of two years in 2000, (9) and dropping the routine updating expectation by 2019. (10) Attempting to keep up with evidence in health like this raises a variety of complex issues. (11) Methods for updating systematic reviews became a widespread concern among systematic reviewers and users of systematic reviews, as well as an area of study. (3,12–14) At some point, updating a systematic review can become redundant, but there is a risk in misjudging this.

In the wake of a mega-trial that overturned the results of a meta-analysis in 1995, Egger and Smith argued “several medium sized trials of high quality seem necessary to render results trustworthy”. (15) Pogue and Yusuf developed a method using optimal information size and cumulative meta-analysis, advocating for its use for prospective determination and monitoring of what would be enough evidence in systematic review protocols, just as for clinical trials. (16,17) Egger et al identified limitations in this method, and it never crossed over into practice. (18,19)

Wetterslev et al built on the Pogue/Yusuf method in 2008, incorporating consideration of heterogeneity, with a method they called trial sequential analysis with cumulative meta-analysis. (20–22) Cochrane’s 2019 guidance to reviewers discourages the use of these methods for updating reviews, unless it is prospectively included in the protocol and used as a secondary analysis only, or in a prospective meta-analysis of a defined group of trials. (10) A 2016 consensus report shows that there is still no well-validated methodology in routine use in systematic reviewing, however, that can reliably show when we have reached, or passed, the point of “enough evidence”. (14)

Up until 2006, there were two status options for Cochrane reviews: “normal” (active reviews, ordinarily with an interval of two years until update was due), and withdrawn (the review is retracted). In 2006, the Collaboration’s governing body decided to add a third option, designating a review as “stable”, in the following software release for Cochrane reviews. (23) This status was codified in the 2008 version of the organisation’s handbook for systematic reviewers, (24) defining a stable review as one that is no longer updated but “highly likely to maintain its current relevance for the foreseeable future (measured in years rather than months)”. This status was to be reviewed periodically, and two uses were specified:

- “The intervention is superseded (bearing in mind that Cochrane reviews should be internationally relevant);
- The conclusion is so certain that the addition of new information will not change it, and there are no foreseeable adverse effects of the intervention”.

In 2008, the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) Group introduced a categorisation of “high quality” in assessing the strength of evidence, defined as “Further research is very unlikely to change our confidence in the estimate of effect”, and this was soon incorporated into Cochrane review summary of findings tables. (24,25) However, the GRADE Group moved away from this conceptualisation by 2017, in favour of “high certainty of evidence”. (26) The GRADE Handbook also recommends using an optimal or required information size calculation to assess the adequacy of precision of an estimate of effect, and that was incorporated in the 2019 edition of the Cochrane Handbook. (10,27)

Cochrane retired the status “stable” in 2019, and a new updating status of “no longer being updated” adopted, with different criteria for use. (10,28) The pool of systematic reviews that have carried the designation “stable” provides an opportunity to study a critical stage in the life cycle of systematic reviews, and when some systematic reviewers believe there is enough evidence. The aims of this study were to assess the extent of usage of stable status, review the fate of reviews designated stable in 2013, and categorise reported reasons for cessation of systematic review updating. We also aimed to describe the reviews with conclusions stated to be unlikely to change with results of new studies, and estimate the extent to which Cochrane reviews were using analytic methods over and above meta-analysis to determine when there was enough evidence.

Methods

Study aim 1: assessing the extent of usage of stable status

The identification numbers of all Cochrane systematic reviews designated as stable in the March 2012 and February 2013 issues of the *CDSR* had been identified from the journal’s encoded version (XML markup language), as well as the number of non-withdrawn published Cochrane reviews. Each review’s identification number was recorded, as well as the year it was designated stable according to the published “What’s New” section of the review.

In 2019, we identified the *CDSR* reviews designated stable up to 20 August via the advanced search option in Archie (the internal Cochrane contributors' database on Cochrane reviews and other documents). (29,30) Identification number, title, and status, were collected, as well as the Cochrane Review Group (CRG) responsible for the review. CRGs are the editorial groups responsible for reviews in specific topic areas. We also collected two fields related to the review's updating status ("rationale" and "explanation"), which had been introduced in a new Updating Classification System in 2016. (28) Cases where a review had both stable and retracted ("withdrawn") status were excluded. Reviews that were stable in 2013 and still described as stable in the *CDSR* were included in the study, even if they were not among the reviews declared stable downloaded from Archie.

The total number of Cochrane reviews on 20 August, excluding those that were withdrawn, was obtained from the Cochrane Editorial Unit. A full listing of the CRGs in 2019 was compiled from The Cochrane Library website in March 2019. (31) The reviews of a CRG that no longer exists (HIV/AIDS) were merged with those of the CRG now responsible for that subject area (Infectious Diseases).

Study aim 2: reviewing the fate of reviews designated stable in 2013

We compared the list of stable reviews from 2013 with those in 2019, identifying those that were no longer designated stable. Data on the current status of those that were no longer designated stable were collected from the *CDSR*, and events reported in the "What's New" table since 2013 were summarized by one author (HB). Each review's status was categorised as normal, stable, or withdrawn. These cases were evaluated by both authors.

Study aim 3: categorising reported reasons for cessation of systematic reviewing

In 2012 and 2013, both authors had reviewed the reasons for the designation given in the "What's New" section, assigning categories which were developed and agreed on iteratively. Where reasons were not given in the "What's New" section, the

abstract, discussion, and conclusion sections were reviewed. Differences in category assignment were resolved by discussion.

We modified and added to our categories from 2013 when reviews did not fit an existing category. We also used the information in the unpublished fields for update classification in Archie to supplement the information published in the reviews for our categorisation.

We assessed the currency of our 2013 categorisations by checking them against the unpublished "rationale". When these matched, we retained our original category without further review.

Where the internal "rationale" and "explanation" for a review were unambiguously consistent, and unambiguously matched one of our categories, we assigned that category and undertook no further analysis, unless it was a review we had categorised as having a firm conclusion in 2013. All other cases were initially reviewed by one author (HB), who extracted data on reasons for the designation from the "What's New" section of that review on the *CDSR* and assigned a category to the review. A random sample of 70 of these were independently assigned a category by the second author (LGH), and differences were resolved by discussion. The second author also reviewed all cases assigned as reaching firm conclusions, and differences were resolved by discussion. Of the final sample of included stable reviews, 42% were assessed by both authors, 27% by one author (HB), and 31% were based on Cochrane classifications alone (S5 File).

Study aim 4: describing reviews with firm conclusions unlikely to change with new studies

To explore this category of reviews, we categorised the main conclusions of these reviews, resolving differences by discussion. Authors' judgments in two other parts of Cochrane reviews are directly relevant to aspects of a firm conclusion, and could be expected to support the review's conclusiveness. The first is the authors' conclusion on the implications of their findings for research, and the other is the judgment on

strength of evidence. We categorised the reviews' section on implications for future research, resolving differences by discussion.

In addition, one author (HB) collected the highest GRADE rating for certainty of evidence in the summary of findings (SoF) table. Where there was no summary of findings table, the description of evidence quality or certainty in the abstract, results, or discussion sections of the review was collected. When the only evidence rating was at the individual study level, the rating for the best-rated study was collected. To determine whether these reviews reported used specific formal analytic method in addition to meta-analysis to reach their determination of enough evidence, the methods, results, and discussion sections were also reviewed by one author (HB).

Study aim 5: estimating the extent of usage of formal analytic methods to determine when there is enough evidence in protocols

To estimate the potential usage of analytic methods for analysing whether there is enough evidence above and beyond meta-analysis, a full text search of all protocols in *CDSR* was done for the phrases "trial sequential analysis", "value of information", "optimal information size", or "required information size" on 29 September 2019. Protocols were screened by one author (HB) to identify those that included an analytic method for determining when there would be enough evidence. The number of protocols in the *CDSR* was recorded.

Data management and analysis

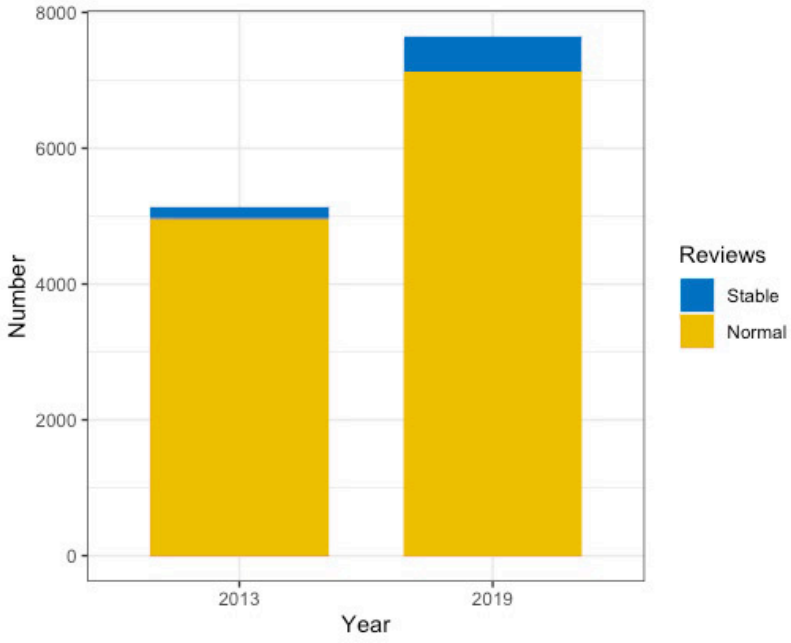
Data were collected in Excel and analysed using RStudio 1.1463 running R 3.5.2, (32,33) using tidyverse and reshape2 packages. (34,35) Summary statistics were used to describe the cohort. Analytic code (S1 Code, S1 File) and meta-data (S2 File), as well as data, are included in supporting information. Data for this project, including analytic code, are also deposited at the Open Science Framework. (36)

Results

Usage of stable status

We identified 507 reviews classified stable among 7,645 non-withdrawn reviews (6.6%) in August 2017 (Figure 10). In February 2013 there had been 180 Cochrane reviews classified stable, which were 3.5% of all 5,137 non-withdrawn Cochrane reviews.

Figure 10. The number of stable reviews among non-withdrawn Cochrane reviews, 2013 and 2019.



There was an increase in the proportion of reviews designated stable between 2013 and 2019. This was in large part due to a single Cochrane Review Group (CRG). Table 6 shows the breakdown of stable reviews across CRGs in 2019, showing that they are not distributed evenly across the 53 groups and most are designated stable by a few groups.

Table 6. Number and proportion of stable reviews across Cochrane Review Groups (CRGs).

	0 stable reviews		1-10 stable reviews		11-20 stable reviews		>20 stable reviews*	
	No.	%	No.	%	No.	%	No.	%
Number and proportion of CRGs by level of volume of stable reviews (n = 53)	20	37.7	23	43.4	4	7.5	6	11.3
Number and proportion of stable reviews in CRGs by level of volume of stable reviews (n = 507)	-	-	79	15.6	52	10.3	376	74.2

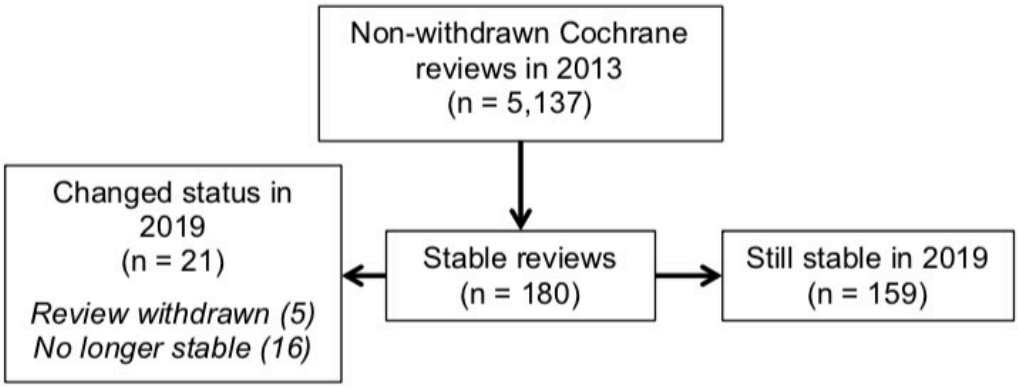
* Range of 24 to 219

The number of stable reviews per CRG ranged from 0 to 219, with a median of 1 (IQR 7). Of the six CRGs responsible for over 20 stable reviews, three CRGs were responsible for 59.2% of all stable reviews, with a single CRG designating 219 reviews stable (43.2% of all stable reviews).

Fate of reviews designated stable in 2013

Most of the 180 reviews with stable status in 2013 were still designated stable in 2019 (n = 159, 88.3%), but 16 reverted to normal status (8.9%) and five were withdrawn (2.8%) (Figure 11). One was withdrawn because it was no longer a priority for the editorial group, and the others were being replaced by one or more new reviews or protocols.

Figure 11. Status in 2019 of stable reviews from February 2013.



Most stable reviews that reverted to normal status did not have new included studies (11 of 16; 68.8%) (Table 7). Two of the reviews had been declared stable in 2013 because of firm conclusions judged unlikely to be changed by new evidence. In one case, however, it was because new trials had been found and an update of the review was in progress. In the other, it was because of a reader's criticism that the firm conclusion was unjustified.

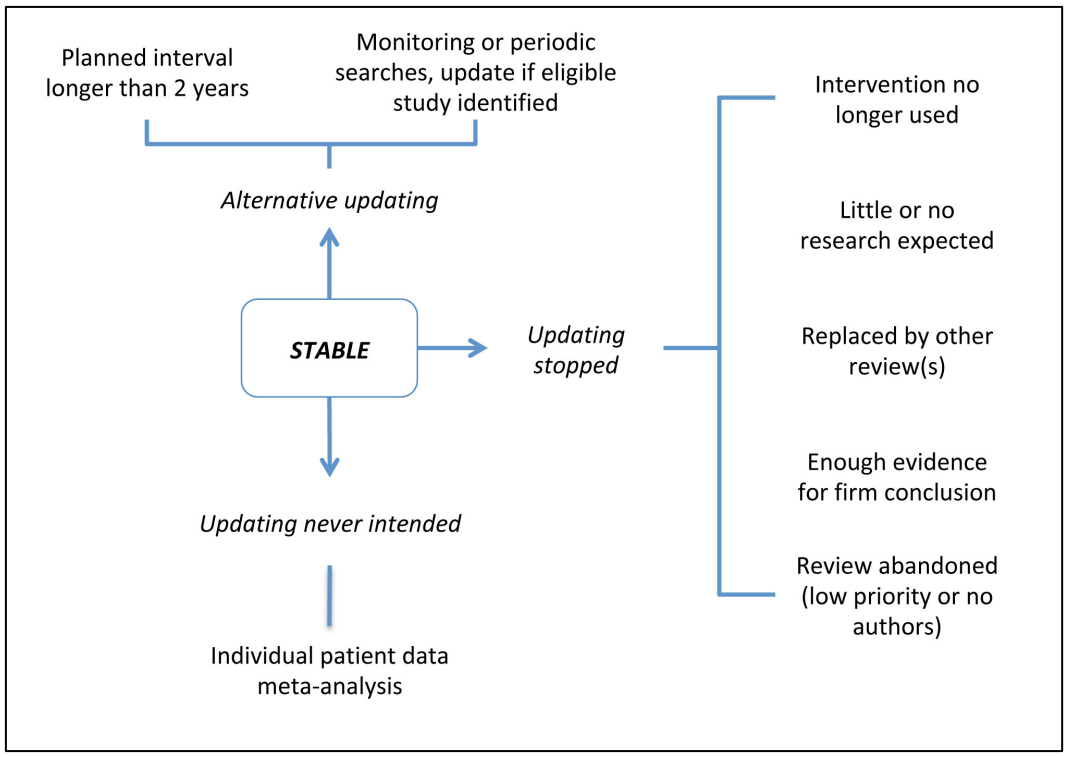
Table 7. Reasons formerly stable reviews reverted to normal status (n = 16).

	Number	Percent
Editorial action or update with no new included studies, stable status not renewed	10	62.5
New included studies without change of conclusion	3	18.8
Changed conclusions because of new included studies	2	12.5
Change of status in response to criticism, no new studies	1	6.2
Total	*16*	*100*

Reasons for reviews being designated stable

We agreed on eight categories for the reasons reported for declaring reviews stable. In Figure 12, we summarise these reasons and group them into those with an alternative schedule for updating, those where updates have ceased, and those where it appears that updating was never intended.

Figure 12. Categories of reasons reported for stable status of reviews.



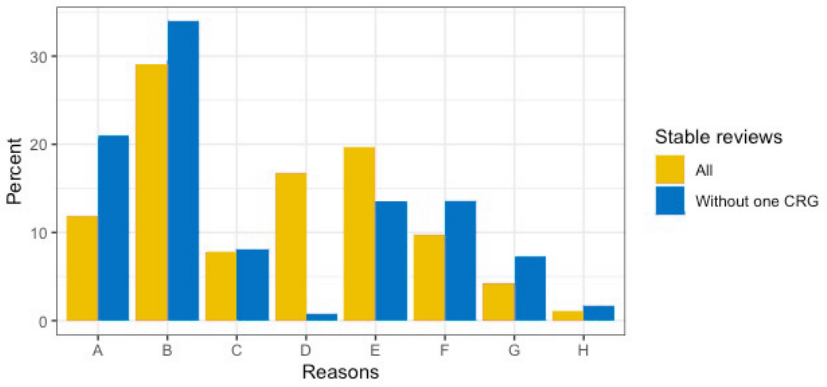
Although each category is distinct, they are not mutually exclusive: a review could be assigned to ongoing monitoring rather than scheduled updating because little further research is expected, for example. Table 8 shows the proportion of reviews assigned to each of the eight categories, in order of frequency. For two reviews, no reasons were reported for declaring the review stable.

Table 8. Reported reasons for declaring Cochrane reviews stable (n = 505).

Reason (ID)	Number	Percent
Little or no further evidence expected (B)	147	29.1
Monitoring new trials or periodic searches, to be updated only if new eligible study identified or standards change (E)	99	19.6
Update required, but in more than 2 years (D)	85	16.8
Intervention not in general use, superseded, or withdrawn from market (A)	60	11.9
Review has been or will be superseded by a new review (F)	49	9.7
Firm conclusions; new studies would be unlikely to change conclusions (C)	39	7.7
Abandoned; low priority and/or authors unavailable for updating (G)	21	4.2
Individual patient data review (H)	5	1.0
Total	505	100

The three most common reasons for declaring a review stable all relate to no, little, or infrequent research activity expected (331 of 505; 65.5%). As such a large proportion of these reviews were from one CRG, Figure 13 illustrates the proportion of reviews in each of the eight categories as in Table 8, with and without its reviews. This suggests that editorial practices were variable between CRGs that applied stable status.

Figure 13. Reported reasons for declaring Cochrane reviews stable, with and without 219 reviews from a single CRG.



Notes: There were 505 stable reviews with reported reasons, and 286 without the 219 reviews from a single CRG.

Reviews with firm conclusions reported to be unlikely to change

There were 39 reviews designated stable with firm conclusions reported to be unlikely to change with further evidence, which was 0.5% of all 7,645 non-withdrawn Cochrane reviews in 2019. They came from 11 CRGs, including 16 from the single CRG with the most stable reviews (41.0%, approximately the same proportion as the group has for stable reviews overall). Two of the reviews coming to firm conclusions reported an analytic method for this decision in their methods (5.1%). In one, it was cumulative meta-analysis, and in the other, it was a sample size calculation based on data from the larger included trials.

In 20 of the 39 reviews, the firm conclusion was that there was a benefit (51.3%). The firm conclusion in the 19 others was an absence of evidence of benefit or superiority, with five of those concluding there was evidence of adverse effect(s) (12.8%).

The majority of the reviews concluded there were still open questions. The authors in 20 of the 39 reviews with firm conclusions wrote that further research was needed on the subject for which their conclusion was firm (59.0%). Of the 16 reviews where authors concluded no further research on that question was needed, six recommended research on other questions related to the review's subject (a further 15.4% of the 39 reviews).

Table 9 shows the evidence rating and future research recommendation for the 14 reviews that had GRADE SoF tables, broken down by the type of firm conclusion.

Table 9. Stable reviews coming to firm conclusions, with a GRADE Summary of Findings (SoF) table (n = 14).

Firm conclusion on effect(s)	More research needed on that question		Highest quality of evidence rating in SoF[1]		
	No	*Yes*	*High*[2]	*Moderate*[3]	*Low*[4]
Benefit (n = 8)	2	6	3	4	0
Some evidence of benefit (n = 1)	0	1	0	0	1
Some evidence of benefit, adverse effects (n = 1)	0	1	1	0	0
No evidence of superiority (n = 1)	0	1	0	1	0
No evidence of benefit, adverse effects (n = 3)	2	1	2	1	0
Total	*4*	*10*	*6*	*6*	*1*

Note: The shaded columns indicate possible results that could be incongruent with a firm conclusion unlikely to be changed by future research.

1 One review with an SoF table did not rate the quality of the evidence.
2 Defined in the review as "High: further research is very unlikely to change our confidence in the estimate of effect".
3 Defined in the review as "Moderate: further research is likely to have an important effect on our confidence in the estimate of effect and may change the estimate".
4 Defined in the review as "Low: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect".

In those 14 reviews with evidence rated using GRADE, we saw little consistency between the firm conclusion/stable status, the quality of evidence, and the recommendation about future research. Only three of the 14 reviews (21.4%) rated the quality of the evidence as high, with no further research recommended. The others are not universally contradictory – for example, in one review, the lack of good evidence of effectiveness was coupled with reference to literature on the biological implausibility of possible benefit from the intervention. (37) That is exceptional, however, and the apparent internal contradictions in conclusions were typically not explained. Some of the discrepancy could be related to conflating the question of whether a strong enough study will be done, with what the impact of one would be.

For the group of 25 reviews without a GRADE SoF table, 12 concluded there was no need for further research on the subject of their firm conclusion (48.0%). However, we could not assess the consistency with the firm conclusion and the authors' judgment of the strength of the evidence. Many reported no overall "rating" of the quality of the included evidence (14 of 25; 56.0%), and of those that did, there was no other consistent method for reaching that judgment.

Use of analytic methods for determining the evidence is enough in protocols for Cochrane reviews

We identified 116 out of 2,415 review protocols that reported some planned use of one of the formal analytic methods we searched for (Table 10).

Table 10. Use of analytic methods in protocols of Cochrane reviews identified by text search (n = 116).

Method	**Number**	**Percent**	**Percent of all protocols** *(n = 2,415)*
Trial sequential analysis	107	92.2	4.4
Value of information	0	0	0
Optimal information size	5	4.3	0.2
Required information size	4	3.4	0.2
Total	116	100	4.8

The method proposed was mostly trial sequential analysis, and most uses of it came from the CRG associated with the development of the method that advises its authors to use it (79 of 107, 73.8%). (38) Each use of optimal information size was in relation to GRADE assessment.

In all, 19 CRGs in this search had at least one protocol using one of these methods (35.8% of all CRGs). The range of protocols per CRG was 1 to 79 (median 1, IQR 2).

Discussion

The group of reviews declared stable by Cochrane authors or editorial groups demonstrate several ways in which a “live” systematic review can reach its end. It was usually not because there was enough reliable evidence. There were some relatively common reasons for declaring a Cochrane review stable. One was that the research focus shifted or the review needed to be split or merged with another or others. Another was that the clinical question had lapsed because the interventions involved are no longer available or have been superseded by other forms of care.

However, the overwhelming reason for declaring reviews stable related to the perceived likelihood of there being further eligible studies. Two-thirds of the reviews were declared stable because of some variation of infrequent or no research activity around the intervention(s). This is also a critical factor to others. For example, for the National Institute for Health and Care Evidence (NICE), not identifying any major ongoing studies is key to retiring a question. (39)

Low likely research yield is a logical criterion for the use of scarce updating resources. However, following active research areas only could tilt the systematic review agenda towards the agendas of those who invest in trials, rather than clinical and consumer relevance. “Continuing importance of the review question to decision makers” is explicitly a key consideration about updating in Cochrane guidance. (10) That goal might be at risk from considering whether there are new studies that could impact results before deciding to update. The latest issue of the Cochrane Handbook also recommends considering metrics such as review clicks and citations in deciding whether or not to update. (10) That may lead in the same direction, perhaps channeling resources issues of less health value because they can go viral.

This situation becomes critical when the question still concerns consumers and clinicians and an updated systematic review could have encouraged further studies, or the judgment about the likelihood of important evidence proves to be wrong. In our sample, authors of at least 5% of stable reviews reversed their decision between 2013 and 2019, with two of those reviews having changed conclusions because of

new evidence. Although that is reassuringly low, we do not know whether important new evidence or other developments affected further reviews that did not search for them.

The 2016 consensus statement on updating systematic reviews (14) argues that decisions not to update a systematic review need to be made in a context where new studies are under surveillance, because "… it is still important to assess new studies that might meet the inclusion criteria. New studies can show unexpected effects (eg, attenuation of efficacy) or provide new information about the effects seen in different circumstances (eg, groups of patients or locations)". The remit of Cochrane's review groups includes maintaining a register of trials within their subject scope, but we do not know the extent to which they all systematically assess incoming studies.

New studies are not the only development that could affect or compromise a systematic review's conclusions and estimates. Identification of major error in an included study, for example, could reduce the effect size in a meta-analysis, (40) or could invalidate a review's conclusions. The 2019 *Cochrane Handbook for Systematic Reviews of Interventions* notes studies retracted for data fabrication should be removed from Cochrane reviews. (10)

The designation "stable" has been replaced in 2019 by the status, "no update planned". Given the problems our study shows with the implementation of the stable status, its end is justified. The replacement status has five options, none of which cover having enough evidence: (a) the intervention or (b) review is superseded, (c) the research area is no longer active, (d) the review is of low priority, and (e) "other". (28)

"No update planned" is one of three potential statuses from 2019, the others being "up to date" and "update pending". These categories aim to "provide readers with a guide to the status of the Cochrane review, and the likely future plans for the Cochrane Review with respect to updating". (28) The usefulness of the "up to date" categorisation to readers depends on how current and accurate that judgment is.

“Update pending” may be inherently misleading to readers, however. It could mean an update is around the corner, but it could be a euphemism for “out of date”.

Our study’s results have only limited application to Cochrane’s new system. Analysing these reviews was useful for charting part of the life cycle of systematic reviews and exploring some practices, but it had had major limitations in assessing prevalence of those practices. A substantial number of Cochrane’s editorial groups never applied the “stable” status, and practices among those that did appears to be highly variable. The data therefore do not reflect the proportion of Cochrane reviews to which the status could apply. Reasons for declaring a review stable were often poorly reported or ambiguously stated. A substantial proportion of our category assignments were based on internal Cochrane information without assessing the text of the review, or were made by a single author.

However, several of our findings are relevant for Cochrane reviews and for the development of methods in systematic reviews generally. We found that judgments about the conclusiveness of evidence and potential importance of future research to current findings were often inconsistent within the small group of reviews coming to firm conclusions. The potential for an error in judgment is high for Cochrane reviews, (21,41–43) and there can be considerable differences in authors’ GRADE-based assessments in systematic reviews in general. (44) Differing interpretations at the additional level suggested by this study underscore the value that improved methodology could offer.

Although the *Cochrane Handbook* discourages the use of methods such as trial sequential analysis, (10) value of information and related methods have been advocated or are in use for determining future research needs based on meta-analysis, (45,46) as well as methods for determining priorities in updating them. (14,47,48) We found that a small proportion of Cochrane protocols are incorporating similar methods to determine when updating is no longer required. Our search terms would not have identified all the protocols using a methodology to pre-specify when a review could be closed.

Optimal information size is now explicitly recommended in the *Cochrane Handbook* for considering the imprecision of trial results, (10) and that may be more widely used in future. The impact of that should be assessed, both for systematic reviewers and users of reviews. The inconsistencies we identified in this study underscore the need for reliable analytic methods to support decision-making about the conclusiveness of evidence. Those decisions include health care choices and recommendations, as well as conducting, funding, approving, and participating in clinical trials. Being able to decide when there is enough evidence, with reasonable reliability, is both a practical and ethical necessity.

Author contributions

HB initiated the study, curated, analysed and visualised data, and drafted the manuscript. Both authors contributed to the design of the study, study assessments, and interpretation of the data. HB acquired the data from the *CDSR* and Cochrane website, and LGH acquired the data from Archie. LGH revised the manuscript critically.

Disclosures

Hilda Bastian was a member of the Cochrane Collaboration's governing body from the organisation's founding in 1993 until 2001, the coordinating editor of a CRG (Consumers and Communication) from 1997 to 2001, and a member of the GRADE working group from 2006 to 2012. She received support from an Australian Government Research Training Program Scholarship. Lars G. Hemkens is a member of the Cochrane Collaboration's Adverse Events Methods Group and a CRG (Heart), and is co-author of several Cochrane reviews. Both authors previously worked for another organisation that undertakes systematic reviews, the German Institute for Quality and Efficiency in Healthcare (IQWiG).

Acknowledgments

We are grateful to Toby Lasserson from Cochrane's Editorial Unit for assistance with data on non-withdrawn reviews, and Anne Eisinga from the UK Cochrane Centre for assistance with a source on Cochrane's original updating policy. We also thank Harriet MacLeHose from the Cochrane Editorial Unit for guidance on the use of data from Archie.

References

1. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA. 1992 Jul 8;268(2):240–8.

2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet Lond Engl. 2009 Jul 4;374(9683):86–9.

3. Moher D, Stewart L, Shekelle P. All in the Family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. Syst Rev. 2015 Dec 22;4(1):183.

4. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007 Aug 21;147(4):224–33.

5. Chalmers I, Hedges LV, Cooper H. A Brief History of Research Synthesis. Eval Health Prof. 2002 Mar 1;25(1):12–37.

6. Chalmers I. The Cochrane Collaboration: Preparing, Maintaining, and Disseminating Systematic Reviews of the Effects of Health Care. Ann N Y Acad Sci. 1993;703(1):156–65.

7. Starr M, Chalmers I, Clarke M, Oxman AD. The origins, evolution, and future of The Cochrane Database of Systematic Reviews. Int J Technol Assess Health Care. 2009 Jul;25 Suppl 1:182–95.

8. Chandler J, Hopewell S. Cochrane methods - twenty years experience in developing systematic review methods. Syst Rev. 2013 Sep 20;2(1):76.

9. Cochrane Collaboration. Minutes of meeting of the Cochrane Collaboration Steering Group, Cape Town, South Africa, 24 and 29 October 2000 [Internet].

Cochrane Collaboration; 2000. Available from: https://community.cochrane.org/sites/default/files/uploads/inline-files/minutes_reports/ccsg/2000.10.24.htm

10. Higgins JPT, Thomas J. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 6.0. London: Cochrane Collaboration; 2019 [cited 2019 Oct 2]. Available from: https://training.cochrane.org/handbook/current

11. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLOS Med. 2010 Sep 21;7(9):e1000326.

12. Garritty C, Tsertsvadze A, Tricco AC, Sampson M, Moher D. Updating systematic reviews: an international survey. PloS One. 2010 Apr 1;5(4):e9914.

13. Tsertsvadze A, Maglione M, Chou R, Garritty C, Coleman C, Lux L, et al. Updating Comparative Effectiveness Reviews: Current Efforts in AHRQ's Effective Health Care Program. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2011 [cited 2019 Sep 16]. (AHRQ Methods for Effective Health Care). Available from: http://www.ncbi.nlm.nih.gov/books/NBK66066/

14. Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ. 2016 20;354:i3507.

15. Egger M, Smith GD. Misleading meta-analysis. BMJ. 1995 Mar 25;310(6982):752–4.

16. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Control Clin Trials. 1997 Dec;18(6):580–93; discussion 661-666.

17. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. Lancet Lond Engl. 1998 Jan 3;351(9095):47–52.

18. Egger M, Smith GD, Sterne JA. Meta-analysis: is moving the goal post the answer? Lancet Lond Engl. 1998 May 16;351(9114):1517.

19. Nüesch E, Jüni P. Commentary: Which meta-analyses are conclusive? Int J Epidemiol. 2009 Feb;38(1):298–303.

20. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. J Clin Epidemiol. 2008 Jan;61(1):64–75.

21. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive--Trial sequential analysis adjustment of random error risk

due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. Int J Epidemiol. 2009 Feb;38(1):287–98.

22. Imberger G, Thorlund K, Gluud C, Wetterslev J. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review. BMJ Open. 2016 12;6(8):e011890.

23. Cochrane Collaboration. Minutes of meetings of the Cochrane Collaboration Steering Group, Khon Kaen, Thailand 24-26 April 2006 [Internet]. Cochrane Collaboration; 2006 [cited 2019 Sep 20]. Available from: https://community.cochrane.org/sites/default/files/uploads/inline-files/minutes_reports/ccsg/1MinutesofKhonKaenCCSGmeetingwithposthoccorrections.htm

24. Higgins JPT, Green SE. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 5.0.0. Oxford: Cochrane Collaboration; 2008. Available from: https://handbook-5-1.cochrane.org/v5.0.0/

25. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924–6.

26. Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, et al. The GRADE Working Group clarifies the construct of certainty of evidence. J Clin Epidemiol. 2017 Jul;87:4–13.

27. Schünemann H, Brożek J, Guyatt G, Oxman AD. GRADE Handbook [Internet]. Hamilton: GRADE; 2013. Available from: https://gdt.gradepro.org/app/handbook/handbook.html

28. Cochrane Editorial Unit. Updating Classification System: guide to applying to Cochrane Reviews [Internet]. Cochrane Collaboration; 2019. Available from: https://community.cochrane.org/sites/default/files/uploads/inline-files/Cochrane_UCS-Guide_8Jul16_0.pdf#targetText=The%20Updating%20Classification%20System%20(UCS,decisions%20for%20individual%20Cochrane%20Reviews.

29. Ahtirschi O, van Valkenhoef G. Advanced document search options - Archie Help [Internet]. Documentation.cochrane.org. 2019 [cited 2019 Sep 23]. Available from: https://documentation.cochrane.org/display/Archie/Advanced+document+search+options

30. Archie [Internet]. London: Cochrane Collaboration; 2019. Available from: https://community.cochrane.org/help/tools-and-software/archie

31. Cochrane Collaboration. Cochrane Review Groups and Networks [Internet]. Cochrane Library. [cited 2019 Oct 10]. Available from: https://www.cochranelibrary.com/about/cochrane-review-groups

32. RStudio Team. RStudio: integrated development environment for R. [Internet]. Boston: RStudio Inc; 2015 [cited 2019 Aug 12]. Available from: https://www.rstudio.com/

33. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing; 2016. Available from: https://www.r-project.org/

34. Wickham H. tidyverse [Internet]. RStudio Inc; 2017. Available from: https://CRAN.R-project.org/package=tidyverse

35. Wickham H. reshape2 [Internet]. 2017. Available from: https://CRAN.R-project.org/package=reshape2

36. Bastian H. Shifting evidence and the reliability of systematic reviews. 2019 Sep 23 [cited 2019 Sep 23]; Available from: https://osf.io/awktv/

37. Enayati AA, Hemingway J, Garner P. Electronic mosquito repellents for preventing mosquito bites and malaria infection. Cochrane Database Syst Rev. 2007 Apr 18;(2):CD005434.

38. Cochrane Hepato-Biliary. Information for authors [Internet]. Cochrane Hepato-Biliary. [cited 2019 Sep 29]. Available from: /information-authors

39. National Institute for Health and Care Excellence. Developing NICE Guidelines: The Manual [Internet]. London: National Institute for Health and Care Excellence (NICE); 2015 [cited 2019 Aug 24]. (NICE Process and Methods Guides). Available from: http://www.ncbi.nlm.nih.gov/books/NBK310375/

40. Fanelli D, Moher D. What difference do retractions make? An estimate of the epistemic impact of retractions on recent meta-analyses [Internet]. bioRxiv; 2019. Available from: https://www.biorxiv.org/content/10.1101/734137v1

41. Fleming PS, Koletsi D, Ioannidis JPA, Pandis N. High quality of the evidence for medical and other health-related interventions was uncommon in Cochrane systematic reviews. J Clin Epidemiol. 2016;78:34–42.

42. Garcia-Alamino JM, Bankhead C, Heneghan C, Pidduck N, Perera R. Impact of heterogeneity and effect size on the estimation of the optimal information size: analysis of recently published meta-analyses. BMJ Open. 2017 Nov 8;7(11):e015888.

43. Gartlehner G, Nussbaumer-Streit B, Wagner G, Patel S, Swinson-Evans T, Dobrescu A, et al. Increased risks for random errors are common in outcomes graded as high certainty of evidence. J Clin Epidemiol. 2019 Feb;106:50–9.

44. Berkman ND, Lohr KN, Morgan LC, Kuo T-M, Morton SC. Interrater reliability of grading strength of evidence varies with the complexity of the evidence in systematic reviews. J Clin Epidemiol. 2013 Oct;66(10):1105-1117.e1.

45. Myers E, Sanders GD, Ravi D, Matchar D, Havrilesky L, Samsa G, et al. Evaluating the Potential Use of Modeling and Value-of-Information Analysis for Future Research Prioritization Within the Evidence-Based Practice Center Program [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2011 [cited 2019 Sep 22]. (AHRQ Methods for Effective Health Care). Available from: http://www.ncbi.nlm.nih.gov/books/NBK62134/

46. Roloff V, Higgins JPT, Sutton AJ. Planning future studies based on the conditional power of a meta-analysis. Stat Med. 2013 Jan 15;32(1):11–24.

47. Sampson M, Shojania KG, McGowan J, Daniel R, Rader T, Iansavichene AE, et al. Surveillance search techniques identified the need to update systematic reviews. J Clin Epidemiol. 2008 Aug;61(8):755–62.

48. Sutton AJ, Donegan S, Takwoingi Y, Garner P, Gamble C, Donald A. An encouraging assessment of methods to inform priorities for updating systematic reviews. J Clin Epidemiol. 2009 Mar;62(3):241–51.

Chapter 5

The prevalence of retractions and their identification in PubMed and PubMed Central: a descriptive study.

Abstract

Background

The visibility and discoverability of retractions may be inadequate at journals and bibliographic databases. This could be contributing to the continuing influence of invalid data and publications despite their retraction.

Objectives

We aimed to estimate the prevalence of retracted publications over time and their identification in PubMed and PubMed Central (PMC), as well as adherence to aspects of guidelines on retracting publications.

Methods

We developed extended search strategies for PubMed and PMC, searching for retracted publications to the end of 2017, as well as the Retraction Watch database, Web of Science, CrossRef, and major publisher websites. We searched references of studies of retractions, and PubMed records that were still “e-published ahead of print” status up to the end of 2012. We analyzed the proportion of retractions identified in PubMed/PMC, as well as the retractions by year, time to retraction, and compliance with practices intended to improve identification of retractions.

Results

We screened 24,437 records and identified 8,828 retracted publications in PubMed/PMC to the end of 2017, below 0.1% of new publications in PubMed from 1985 to 2015. A total of 2,705 (30.6%) of these publications did not have retraction status in PubMed/PMC. A third of those had been “e-published ahead of print” since before 2013, a category of record that is not indexed for PubMed by the National Library of Medicine (NLM). A high proportion of those records were retracted (over 40%). Journal non-compliance with recommended practice was high, with separate notices of retraction not submitted to PubMed for 29.4% of retracted publications.

Conclusions

Retracted publications are rare, and under-identified in PubMed/PMC. There are multiple causes, some of which could be resolved by the NLM and others that are the responsibility of journals. All but a small minority of retracted publications could be identified by filtered searches.

Hilda Bastian
Diana C. Jordan
Melissa Vaught

Background

Invalid data and publications continue to influence research and decision-making, either directly, or second-hand via publications which have relied on them. Although retraction tends to reduce citation of a publication, these publications continue to be used without apparent awareness of their status. (1–8) As well as their importance to the integrity of literature, retractions are relevant to research integrity investigations of individuals and analysis of journal performance, and as an indicator of research integrity and evolving publishing practice. (9,10) The visibility and discoverability of retractions are, therefore, vital elements of scholarship infrastructure and the historical record, but both may be inadequate at journals and bibliographic databases. (11–14)

As a major bibliographic database used by millions every day that can gather and standardize post-publication events across journals, PubMed is an important resource for authors, investigators, and developers of publication-related resources. PubMed Central (PMC) is its accompanying full-text archive, which is also used independently of PubMed. Both are free public databases produced by the U.S. National Library of Medicine (NLM) at the National Institutes of Health (NIH), providing application program interfaces (APIs) that enable data on publications and their retracted status to be incorporated in other literature resources. PubMed is also a common basis for studies of retractions. (10) Ensuring more rapid and comprehensive coverage of retractions with higher accuracy, visibility, and discoverability in PubMed and PMC could play a critical role in reducing error and improving accountability in biomedical science.

PubMed includes publications from selected journals, books, and documents, predominantly related to biomedicine. Publications funded by NIH and several other funders and government agencies are also included, whether or not the journal is one selected for inclusion in PubMed or PMC. (15) As well as processing the data provided by publishers and authors of eligible manuscripts, the NLM creates linkages between records. The NLM adds indexing terms and categories to MEDLINE-

indexed journals, some of which can be added to non-MEDLINE publications in PubMed. MEDLINE-indexed journals are the primary component of PubMed. (16)

The NLM added the indexing category “Retraction of publication” for retraction notices in MEDLINE-indexed journals in 1984. (17,18) This enabled NLM indexers to generate a record for this event, noting it in the citation record of the affected publication. Retraction was only used when a publication “was based on fraudulent research”. (18) Journal retractions issued for other reasons were not originally classified by NLM as retracted, for example a series of retractions by the editors of the American Journal of Cardiology in 1987. (19,18) The categorization was applied retrospectively to all research classified as fraudulent that could be identified by NLM in 1984. (18)

The International Committee of Medical Journal Editors (ICMJE) added retraction of research findings to its Uniform Requirements for medical journals in 1987, also with a view to identifying only publications determined to be fraudulent. (20,21)

The NLM added an indexing category for the retracted publication itself in 1989. (22) In 1991, “retraction of publication” and “retracted publication” became formal publication types, a new feature at NLM, covering the retraction or withdrawal of a publication. (23) A retraction notice in PubMed can retract multiple publications or a publication from before that journal was indexed by the NLM. By 1992 the NLM’s definition had expanded beyond fraudulent research and continued to be refined through the 1990s. (23–25) An indexing category for “partial retraction” was added in 2007. (26) It was rarely used, and was phased out in 2016, with the status of these events changed to erratum. (27) As of 2016, publishers have also been able to edit and change the status and most content of their PubMed records, as well as link them, but not delete them. (28,29) PMC contributors can include XML tagging for retraction notices, with a citation for the retracted article linking it as a related article, but only the National Center for Biotechnology Information (NCBI) can add, remove, or change PMC records. (30) Retracted publications and retractions in PMC are searchable via PMC filters.

Current definitions of retracted or withdrawn publications by the NLM, (31) Committee on Publication Ethics (COPE), (32) Council of Science Editors (CSE), (33) and the International Committee of Medical Journal Editors (ICMJE) (34) specify several categories of reasons for retracting an article, including:

- Scientific misconduct (NLM, COPE, CSE, ICMJE);
- Invalid or unreliable data (NLM, COPE, ICMJE);
- Plagiarism (NLM, COPE, ICMJE);
- Pervasive ("honest") error (NLM, COPE, CSE);
- Irreproducible research/replication failure (NLM, CSE);
- Redundant publication (findings already published elsewhere) (COPE, CSE, ICMJE);
- Unethical research (COPE, CSE);
- Research compromised by inappropriate methodology (ICMJE);
- Failure to disclose competing interests likely to influence interpretations or recommendations (COPE).

Publishing practice and policies continue to evolve. Studies of retracted publications identify further reasons, including editorial misconduct and peer review fraud, (35,36) as well as publisher or administrative error, data ownership or copyright issues, and authorship disputes. (35,37,38)

For this study, we regarded a retracted publication as one that has been retracted, removed, or withdrawn from the literature for any reason. The aims of the study were to estimate the prevalence of retracted publications and their identification in PubMed and PMC over time, and estimate rates of adherence to aspects of NLM, COPE, and *ICMJE* guidance on retracting publications.

This was the second study in the PubMed Commons program of evaluation of post-publication activity at PubMed/PMC, aiming to identify opportunities to improve coverage, visibility, discoverability, and time lag in identification of events. The first PubMed Commons study identified and analyzed editorial expressions of concern.

(39) This resulted in the addition of a MEDLINE publication type for expressions of concern in 2018. (40)

Methods

This study is reported according to the STROBE statement for reporting observational studies, (41) and the UK InterTASC Information Specialists' Sub-Group (ISSG) Search Filter Appraisal Checklist. (42) Details of search strategies for the prevalence of retracted publications are reported in accordance with the PRISMA reporting guidelines for systematic reviews. (43)

Inclusion criteria for retracted publications

Publications were classified as retracted for this study without any restriction on reasons for the designation. Publications included articles, letters, and any other item that can generate a PubMed or PMC record. A retraction notice is the statement issued by the journal advising that the publication is withdrawn. This status could be indicated by any variant of the word retracted, removed, withdrawn, or similar language to indicate that the publication has been rescinded. Retracted publications and retraction notices in any language were eligible for inclusion.

We included publications as retracted if at least two authors confirmed retraction status at the journal website or printed publication, or if there was an annotation to the PubMed or PMC record, unless we judged the annotation in PubMed or PMC to be in error (for example, assigned to the wrong publication).

Partial retractions were regarded as errata per current NLM policy, not retractions. (31) When it was unclear from a journal's retraction notice whether or not it was a retraction of the full article, NLM indexing of the record as retracted was accepted, as additional communication between NLM and the publisher may have taken place that is not publicly available. Publications that are retracted can also be republished, with the republished version indexed as such. (31) We did not collect and analyze

republications. Articles can also be corrected and republished: we did not classify these as retractions.

To be eligible for inclusion in analyses as a retracted publication, a publication had to have a record in PubMed and/or PMC, whether or not the notification of its retraction was also in PubMed/PMC. When PubMed records were deleted after we had classified the record as a retracted publication, the record remained in our dataset.

To be eligible for inclusion in this study as a report of retraction, the notice could be a separate publication, an annotation to the retracted publication at the journal and/or in PubMed/PMC, or a note replacing the retracted publication. However, when the retracted publications were not indexed in PubMed/PMC, the notices and publications were excluded as the full denominator of non-PubMed-indexed publications is unknown. This includes, for example, retractions in a MEDLINE-indexed journal for publications that predate that journal's acceptance into MEDLINE.

Retractions of retractions were treated as "de-retractions", so the originally retracted publication was no longer designated as retracted. However, the original retraction notice was itself classified as a retracted publication. When an annotation "temporary removal" was revoked, we did not record this as a de-retraction, although we classified those publications with this annotation at our final search as retracted, if the publication had the same status displayed at the journal.

Abstracts within conference proceedings can have separate PubMed records, or a single PubMed record for the full set. PMC sometimes issues separate records for abstracts in conference proceedings, where PubMed has issued only one. We did not attempt to disaggregate the total number of abstracts retracted and followed the PubMed practice for categorization at the proceedings only. Thus, if PubMed issued a single record for a conference where multiple abstracts were retracted, this remained a single retracted publication even if PMC retracted multiple records.

There were cases where a journal published a summary version of a publication as well as a full-length version, each with a separate PubMed record. When the full-

length version was retracted, we included both as retracted publications, whether or not the summary was explicitly mentioned in the retraction notice.

The classifications in this study are those of the authors only, and may not reflect the final decisions for PubMed/PMC.

Assessing the prevalence of retracted publications

Although there have been multiple studies of retractions in the literature, we did not identify a formally validated search strategy for retracted publications. Our searches of PubMed and PMC began with the databases' specific indexing terms for retracted publications and retraction notices, and text terms to capture versions of "retracted" and "withdrawn" added to titles or abstracts. As retracted publications were identified from other sources that would not be detected by our PubMed/PMC search, we developed our strategy iteratively. Final extensions of the search strategy were agreed on after discussion by all authors, taking into consideration the additional time required to screen records. Our search strategies were not peer reviewed.

We searched PubMed and PMC, downloading abstracts and MEDLINE (44) or PMC data associated with the record. We also searched the Retraction Watch database, Web of Science, CrossRef (via API), and several publisher websites (Bentham, Cell Press, Elsevier's Science Direct, IEEE, Springer Link, SagePub, Taylor and Francis, and Wiley Online Library). Additional searches included comments on PubMed Commons, a commenting platform previously available on PubMed, (45) two reports of mass retractions, and one publisher of guidelines (American College of Obstetricians and Gynecologists, ACOG). Details of search strategies and numbers of records retrieved are reported in Appendix 3.

Initial screening of results was by two authors for PubMed, PMC, Retraction Watch, and Web of Science, and one author for the others (HB or MV). All potentially retracted publications identified by these searches were reviewed by at least two authors, with differences resolved by consensus of at least two authors. Final searches of PubMed, PMC, and Retraction Watch incorporated retractions to the end

of 2017, but other searches were only up to the end of August 2017. Our first search was run in November 2015 and the final update search for PubMed/PMC using the finalized search strategy reported here was run on 20 February 2018. Publications and retraction notices dated in 2018 were excluded. However, it is possible that some records retracted only by changes to the record itself in 2018 without a separate retraction notice were included.

In addition, we searched all PubMed records up to the end of 2012 that still had “epublished ahead of print” (epub) status. While the number of retracted publications in this set was relatively high, the volume of them after 2012 was too high for us to screen. There were more than 125,000 records still with this status published between 2013 and 2017 at that time. Epub records up to the end of 2012 were screened by 2 authors (DJ, MV).

When a publication is included in a journal issue, the already-created epub record should be updated to reflect the change in status. However, this can fail in practice. Therefore, we looked for potential duplicates for the epub record, to identify cases where a new record for the same publication had been created instead of updating the epub. For publications listed as epubs at PubMed, but which could not be found on journal websites, Google Scholar and Google were searched to identify possible accessible versions.

We identified a further indicator for possible retraction: retraction notices with a title that was only a precise repetition of the title of the retracted publication, and empty abstracts. However, the volume of records with identical repeated titles or empty abstracts within a journal was too high for us to screen.

To identify available collections of retracted publications from other studies, and to inform our search strategies, we conducted searches for studies of retractions in the biomedical literature, with data on retractions from 1990 onwards. PubMed and reference lists for seed studies were manually searched by one author (HB), including the studies found in a 2017 review of articles on retractions, (10) with reference lists of the further identified studies also hand-searched. In addition, all

citations in Google Scholar for a pivotal study (38) up to 21 August 2017 were screened. The articles with reference lists that were hand-searched, and those with available identified retractions, are included in Appendix 3. The search strategy and list of 117 primary publications and one review is Appendix 4, including an update in 2019 that did not contribute to our reference searching for retracted publications.

Up to 2017, a total of 53 data collections of retractions reported in 63 publications were identified. These studies could cover publications that are not included in PubMed or PMC. Reliance on PubMed/MEDLINE as a data source was common, and for 45% of studies, this was the sole data source (n = 24). Only 10 of the studies included identification of some or all of the retracted publications the authors had analyzed. (35,36,46–53)

One person searched PubMed for the retracted publications cited in 9 of those 10 studies (HB), and those not indexed as retractions were added to the database for further screening. The tenth study published an Excel sheet of PubMed IDs, and this was de-duplicated with the database before adding records for further screening. (47)

Data collection and analysis

Searches of and data retrieval from NLM's Catalog, PubMed, and PMC were conducted via the NCBI e-utilities, (54) using the Rstats package rentrez to build and submit queries. (55) Queries to CrossRef (56) were done using rcrossref. (57) Analyses were performed using RStudio 1.04 running R 3.3.1, (58,59) and we used tidyverse, survminer, and survival packages for R. (60–62) Data for this project will be deposited at the Open Science Framework. (63) Rayyan was used for screening PubMed downloads, (64) with Microsoft Excel used for smaller batches. Our database of retracted publications and retraction notices was originally maintained in Microsoft Access, and migrated to R.

Summary statistics were used to describe the cohort, with Kaplan-Meier survival analysis for time to retraction.

Assessing prevalence

Each included retracted publication remained paired with an identifier for the retraction, which could be a PubMed or PMC ID, a DOI, or a short citation. Where the form of retraction was only an annotation or replacement within the PubMed or PMC record, the ID for the retracted publication served as the ID for the retraction notice. Separate retraction notices in PubMed/PMC were defined as those where the ID was not identical to that of the retracted publication, and in journals where there was a separate DOI or citation.

Data from Web of Science, CrossRef, the publisher IEEE, and the batch of records provided by Retraction Watch were processed first by the journal of publication. The list of journal titles across these four sources was de-duplicated. ISSNs or ISBNs for a given journal in data sources were collapsed into a single query and searched against the NLM Catalog, and NLM title abbreviations retrieved. If multiple NLM title abbreviations were retrieved for a given journal, all title abbreviations were searched simultaneously (using OR Boolean).

For MEDLINE-indexed journals, one author (MV) searched for all publications noted as retracted in these sources. For other journals, the number of total PubMed records for that journal was reviewed. After discussion, data from journals with no publications, or few publications and a non-biomedical focus, were excluded. These were deemed unlikely to be in PubMed.

Where a DOI was provided for a publication in an eligible journal, we searched the PubMed identifier fields for the DOI. If a PMID was not returned, then PMC was searched for the DOI, restricting search to the DOI field. If no DOI was provided or the DOI did not return a PMID, then the PubMed was searched for journal title abbreviation(s) with: volume and first page number; author last names and publication year. If no hits were returned from these searches, then PubMed was searched for the journal title abbreviation(s) with the full article title, and if not retrieved, then search again for major content terms from the title in the PubMed title field.

Data from Retraction Watch were processed in three batches (June and August 2017, and January 2018). (65) The first two were scraped from the website and structured into citations, with unique records searched for in PubMed. The first of these were searched for individually by DOI or title. The second batch was included in the group processing described above. A final update of records identified by Retraction Watch after 1 September 2017 was provided by Retraction Watch. These were screened by one author (HB), who manually searched by title in PubMed, with the exception of IEEE journals, and economic and business journals which did not have records in PubMed. The websites for Elsevier (ScienceDirect), Springer (SpringerLink), Taylor & Francis, and Wiley (including The Cochrane Library separately) were searched online by one author (HB). The websites for Bentham, Cell Press, and Sage Publications were searched online by another author (MV).

For all included records of retracted publications and retraction notices as of February 2018 in PubMed XML (66) and the following data was retrieved for our analyses:

- PubMed identification number (and PMC identification number if the publication was also in PMC);
- The year from the publication, Entrez, Epub date, and PubMed dates;
- Titles and abstracts;
- Retraction status fields in PubMed; and
- Publication status “epub ahead of print” in PubMed.

Similar identifying and date data were downloaded for PMC records, including retraction status fields.

For retraction notices, we noted how many publications each notice retracted. Those counts were confirmed by at least two authors.

Each retracted publication and retraction notice was assigned a publication year based on the earliest date associated with that record. For example, an article

epublished in 2009, for which a PubMed record was created in 2010, was dated 2009.

Where a PubMed record had been deleted as of July 2018, the latest data we had for that record was used.

As well as assessing prevalence of retractions in PubMed and PMC, we assessed concordance between PubMed and PMC on the retraction status of publications, including whether or not retraction-related records in PMC had PubMed records.

Time trends

We chose 1984 as the starting point for time trend analyses, as this was the year that NLM began prospectively classifying retractions. We calculated the proportion of retracted publications by year of their publication using only retracted publications with records in PubMed, with the total number of new journal publication records in PubMed by year. (67)

Our dataset only included dates of retraction for publications in PubMed with separate retraction notices in PubMed. We did not attempt to account for biases in the data resulting from the missing data being potentially non-random. After choosing the earliest year for both retracted publications and retraction notices, we excluded cases where the date of retraction occurred before the date of publication. We then calculated the proportion of all publications retracted in the year they were published. We assessed length of time from publication to retraction via Kaplan-Meier analysis and calculated the median and interquartile range. We also calculated the cumulative percentage of retracted publications for new publications in PubMed by year from 1984.

Non-compliance with selected NLM, COPE, and ICMJE recommendations

Several recommendations aim to ensure that publications are clearly identified as retracted, and to aid in their identification. In its current policy, the NLM requires publishers to state clearly that a publication is retracted in a citable, separate retraction notice. (31) It also encourages following guidelines such as those by COPE and ICMJE. (31) We used PubMed records as proxies for publisher compliance with several recommendations for ensuring visibility of retracted status to users. As we noted in the introduction, journal publishers are able to do this themselves for PubMed records. We did not assess the visibility of retractions at journal websites.

The first practice analyzed was issuing a separate retraction notice, which is stipulated by NLM, (31) COPE, (32) and ICMJE (34). Our proxy measure for non-compliance with this requirement was how many retracted publications did not have a separate PubMed record for a retraction notice.

The second practice was including the title of the affected publication in the title of the retraction notice. This should be done according to the ICMJE, (34) and is mentioned by COPE as one way of associating a retraction notice with its retracted publication(s). (32) Our proxy measure for non-compliance with this practice was how many titles of retraction notices were shorter than six words (strings of letters/numbers separated by spaces).

We also assessed compliance with the NLM policy on publications withdrawn while still with epub status, a policy in force since at least 2006. (31,68,69) When an epub is removed at the journal or replaced by a retraction notice, the NLM policy states that a replacement record, also with epub status, should be submitted for PubMed with the word “Withdrawn” as the first word of the title, and recommended wording incorporated into the abstract. Our proxy measure for non-compliance with this policy was how many titles of epub status retracted publications did not begin with the word “withdrawn” or “withdrawal”.

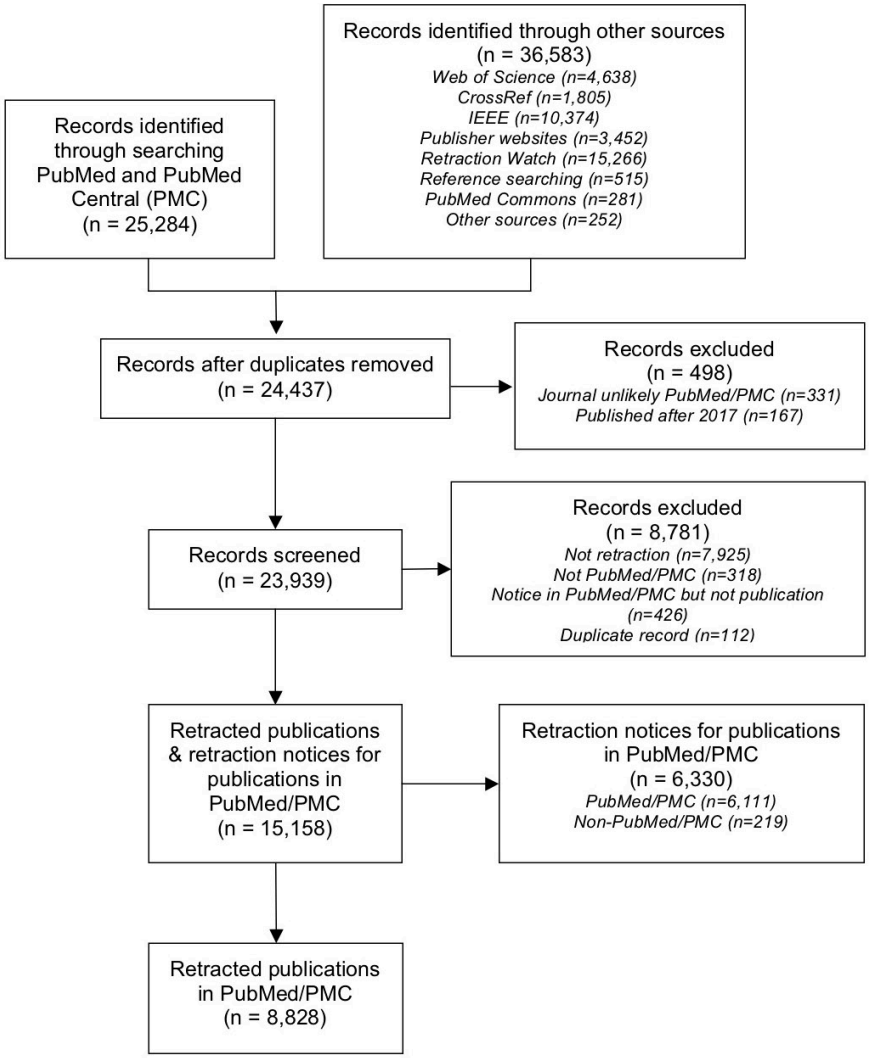
Updating all PubMed records, epub or not, to indicate retracted status in the title of the retracted publication ensures that retracted status is clear to users promptly, whether or not the publication type retracted has been assigned. Although the ICMJE does not refer to PubMed records, it recommends that a retracted article "should be clearly labeled in all its forms". (34) We calculated the proportion of retracted publications in PubMed for which a version of "retracted" or "withdrawn" had not been added as the first word in the title.

Results

Table 11. Overview of retraction terminology and eligibility for inclusion.

Term	Description	Eligibility
Retracted publication	A publication retracted from a journal for any for any reason (with or without separate notification of retraction).	Included, if the record was in PubMed or PMC.
Retraction notice	Separate PubMed or PMC record for a notice issued in a journal notifying the retraction of one or more publications.	Included, if the retraction applies to one or more publications in PubMed or PMC.
Partial retraction	A former indexing option at NLM; now classified as an erratum.	Excluded.
De-retraction	Reversal of retraction, resulting in re-instatement of the original publication in the journal. There is no separate indexing option or term for this at NLM.	Included as a retraction notice. The former retraction notice included as a retracted publication. The original publication, now re-instated, was excluded.
Temporary removal	Publication removed at the journal, with a notice that the publication's removal is temporary, pending final decision on status. The PubMed record may or may not be amended with this information.	Included as a retracted publication if still removed at study end.

Figure 14. Flow diagram for retraction search results.



Note: Retraction notices included 264 additional separate notices in PMC for conference abstracts indexed as single conference proceeding in PubMed, which were handled as single retracted publications throughout.

Search results

Our searches and online screening identified a total of 61,867 records, which included 24,437 unique records (Figure 14). Of these, 498 were excluded as it was determined the journals in which they were published would be very unlikely to yield retracted PubMed-indexed publications. This left 23,939 records for full screening.

We finally included 8,828 retracted publications in PubMed/PMC up to the end of 2017, and excluded 744 retracted publications that were not in PubMed/PMC, whether or not a retraction notice for them was in PubMed/PMC.

Prevalence of retracted publications and retraction notices

Table 12 details numbers of retracted publications and retraction notices. Of the 8,828 retracted publications, 6,123 were given some kind of retraction status in PubMed or PMC indicating it was either a retracted publication or retraction notice (69.4%), and 2,705 were not (30.6%). There were 14 retracted publications only in PMC, with no PubMed record: 9 of them were assigned PMC retracted publication status. For 66 retracted publications, the PubMed record had been deleted at the time of our last data update. These remain in our analyses.

Table 12: Total numbers of retracted publications and retraction notices identified in PubMed and PMC, and proportions indexed as retractions.

	Number	*Percent (n = 8,828)*
Retracted publications with separate retraction notice in PubMed and/or PMC:		
With retraction-related indexing	5,605	63.5
No retraction-related indexing	739	8.4
Total	***6,344***	***71.9***
Retracted publications with no separate retraction notices in PubMed/PMC		
With retraction-related indexing	518	5.9
No retraction-related indexing	1,966	22.3
Total	***2,484***	***28.1***
Separate retraction notices in PubMed or at journals:		
Retracting a single publication	5,838	66.1
Retracting multiple publications*	228	2.6
Total	***6,066***	***68.7***

* Retraction notices retracting more than one publication retracted a total of 964 publications (range: 2 to 107 publications per notice).

A publication in a MEDLINE-indexed journal in PubMed can have a PMC record if a full text version is deposited there, and a PMC-only publication should have a PubMed record. However, PubMed and PMC processes are independent, as is indexing and tagging of retracted publications and retraction notices. We assessed concordance in status between PubMed and PMC for retracted publications (Table 13). There were some discordant records.

Table 13. Discordant retraction indexing status between PubMed and PMC for records tagged as retracted publications or retraction notices (n = 15,041).

Discordant status	Number
Records with retraction status in PubMed but not in associated PMC records	60
Records with retraction status in PMC but not in associated PubMed records	43
Records with retraction status in PMC that had no PubMed record*	80

* Excludes conference abstracts where proceedings have a single PubMed record.

Our inclusion criteria did not differentiate reasons for withdrawal or retraction. We therefore included 433 systematic reviews in PubMed from the *Cochrane Database of Systematic Reviews* (*CDSR*) as retractions, although we excluded one of their retracted reviews because it was not indexed in PubMed. Only one *CDSR* publication was assigned retracted status in PubMed. The other 432 account for 16.0% of those we included that PubMed has not assigned any kind of retracted status. *CDSR* is unusual in having a large number of records withdrawn that could have either the publication is out of date or that it “contains a major error”, and the reason for withdrawal is not routinely stated. (70) Another publication also had a high proportion of retraction for out-of-date publications, but the number was not as large. (71) Opinions about whether *CDSR* withdrawals constitute retractions vary, with some apparently treating them all as retracted as we have done, (13,52) and others, like Retraction Watch, (65) doing so selectively. As of 2019, withdrawals of *CDSR*

reviews will signify serious errors or types of misconduct and can be expected to reduce. (72)

We found six cases of “de-retraction” – retraction of retraction. We classified the original retraction notice as the retracted publication, not the original article. For two of them, neither the original retraction notice nor its retraction had a record in PubMed, so they are not included in these analyses. The four we included were two where the original articles were assigned as retracted publications in PubMed, one where the retraction notices was still assigned as such although it had itself been retracted, and one where both the first retraction notice and its retraction were assigned as errata.

E-publication ahead of print

When journals provide records for “e-publication ahead of print”, these are created with the intention of updating the record when the publication is formally included in a journal issue. Those ahead-of-print records are not routinely indexed for MEDLINE. (69) As indexing is the process by which the publication types for retraction are assigned, this can contribute to under-ascertainment of retraction status, if neither the retracted publication nor the retraction notice is ever included in an issue of the journal.

To assess the potential impact of this practice when publications remain as epubs for a prolonged time, we assessed the status of all PubMed records up to the end of 2012 that were epub ahead of print as of June 2017 (n = 2,265). The cut-off date was chosen because of time constraints. The number of these records rose steeply in 2013: there were 663 epub records dated 2012, and 2,125 dated 2013.

Table 14. Publications with e-published ahead of print status in PubMed to the end of 2012 (n = 2,265).

	Number	*Percent*
Confirmed retracted publication	926	40.9
Confirmed retraction notice	15	0.7
Records which may include retractions:		
Likely retracted publication or retraction notice	19	0.8
Possible retracted publication or retraction notice	124	5.5
Unlikely retracted publication or retraction notice	455	20.1
Records confirmed as not related to retraction	726	32.1
Total	***2,265***	***100***

The oldest epub was dated 1999 and the median year was 2010. A very high proportion of records up to 2012 that were still “epub ahead of print” in 2017 were confirmed as retracted (40.9%), with a further 0.7% being epub retraction notices (Table 14). We believe this is an under-estimate. A further 0.8% appeared likely to be retraction-related (either retracted publications or retraction notices), but we could not confirm this as the items appeared to have been removed from the journals without enough information for us to confirm status. For a further 5.5%, it was possible that they were retraction-related, but we could not confirm with certainty that the item had in fact been published. These could have been publications retracted without issuing a notice, or publications provided to PubMed when publication was planned, but the journal decided not to publish in the interim. Finally, we categorized 20.1% as most likely not retraction-related, although there may be some records in there that reflect publications that were withdrawn and later republished in corrected form without reference to the epub version.

Of the 941 confirmed retraction-related records in this analysis, only 14 were assigned NLM status as such (1.5%), including 11 of the 926 retracted publications (1.2%). In addition, seven of the publications retracted by the epub notices were not assigned retracted status in PubMed. These 915 publications represent 33.8% of the

2,705 retracted publications we identified without retracted status assigned in PubMed/PMC. As we only assessed these to the end of 2012, a potentially important percentage of retracted publications remains unrecognized among the pool of epub records that we did not investigate. In March 2018, there were more than 97,000 epub records from 2013 to 2017.

Temporary removals

“Temporary removal” is not a recognized formal status at PubMed/PMC. Out of searches from November 2015 onwards, we identified 114 PubMed records annotated “temporary removal”, none of which were assigned retraction status by the NLM at that time. We classified records with “temporary removal” or retracted status at their journals in February 2018 as retracted (n = 64). For 51 publications, this was because they were still labeled removed at both the journal and PubMed. Six had been retracted at the journal (5.3% of the original group), although they were still annotated “temporary removal” at PubMed. For three, the label was still at PubMed, although it had been removed at the journal, and for four, the journal still labeled it removed although the label was gone in PubMed. In the other 50 cases, the label was gone from both the journal and PubMed (43.9% of the original group).

Time trends

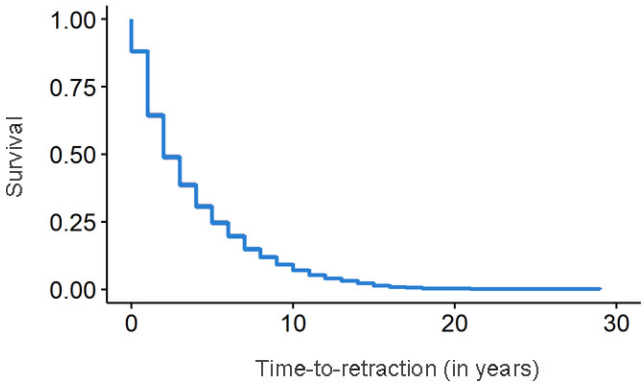
Retractions often occur years after an article was published, leaving important levels of uncertainty around retraction rates in recent years. In addition, when there is no separate retraction notice, it can be unclear when retraction took place, leading to high levels of missing data.

We only had dates for both publication and retraction for those publications with retraction notices in PubMed and PMC (Table 12). We used PubMed data only for the time trend calculation, to ensure standard date fields. Data for 16 of these were excluded as the date for the retraction record preceded the publication date for the retracted publication, resulting in a high rate of missing data for time to retraction (29.6%). Historical trends in editorial practice could have affected both time to

retraction and issuing indexed retraction notices, particularly the increasing trend to e-publication ahead of print. Our ascertainment of retractions among ahead of print records was likely to be higher up to the end of 2012. The missing data are not likely to be random. However, the retracted publications in PubMed were published in 2,356 different journals, limiting the influence any specific editorial group could have on missing data and time to retraction.

The Kaplan-Meier analysis for the retracted publications with known years of retraction in PubMed is shown in Figure 15.

Figure 15. Time to retraction for publications with known years of retraction in PubMed.

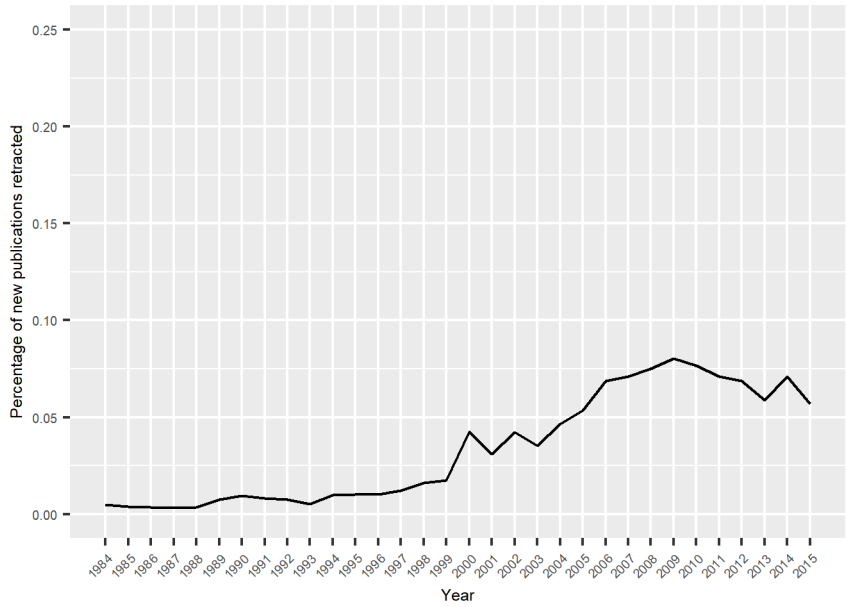


Kaplan-Meier plot for the time from publication to retraction

The time to retraction was from 0 to 29 years, with median of 2 years and IQR of 4.

Based on this analysis, we calculated the rate of retraction of new publications in PubMed annually with a cut-off of 2015, as most of the retractions likely to occur in the last two years for our data had yet to take place (Figure 16).

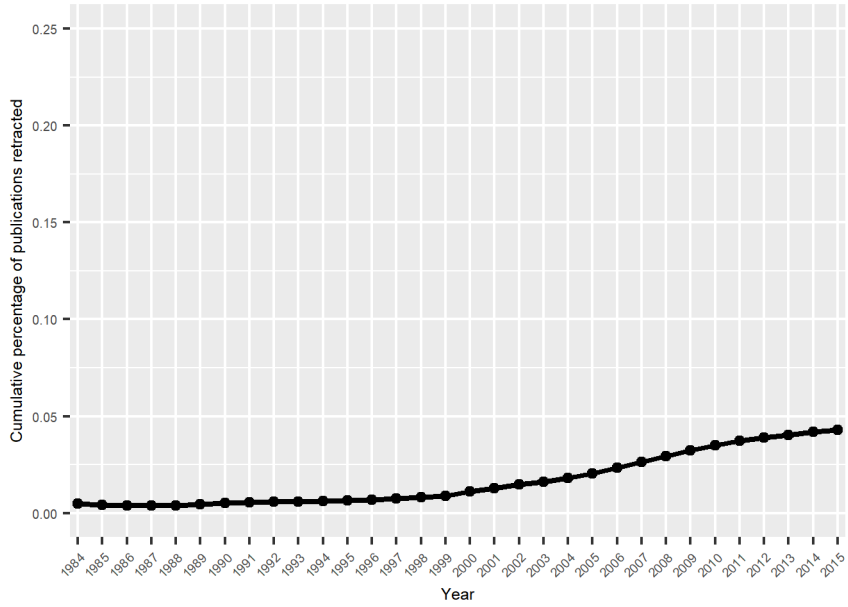
Figure 16. Annual rates of new publications in PubMed retracted by 2017, from 1984 to 2015.



Percentage of new publications each year retracted from the beginning of NLM indexing of retractions (1984) to two years before the end of our data collection (the median time till retraction). A high proportion of retractions for later years yet to occur.

Retractions remain rare events, below 0.1% of new publications from 1985 to 2015. The rate of retraction fluctuated annually, but the overall trend has been to increase, particularly since 2000. The rise in the retraction rate is also illustrated by the cumulative percentage of new publications in PubMed being retracted (Figure 17).

Figure 17. The proportion of new publications in PubMed being retracted, cumulative percentage from 1984 to 2015.



Non-compliance with selected recommended practices

Table 15 shows the results of some indicators of non-compliance in PubMed only. These indicators measure, or are proxies for, non-compliance with several NLM, COPE, and ICMJE recommendations or requirements. (31,32,34) The table also shows the proportion of retracted articles which did not have titles modified to indicate retracted status. We did not assess the extent to which journals showed retraction status at their websites.

There was a high rate of non-compliance with issuing a separate retraction notice (29.4%). When a separate retraction notice was issued, 14.5% of retraction notices retracting single publications apparently did not include the title of the retracted publication, as the titles were fewer than 6 words long (for example, “Notice of retraction”).

A minority of epub retracted publications did not have “withdrawn” or “withdrawal” as the first word in their titles (14.3%). Overall, though, most retracted publications did not have titles modified with a version of “retracted” or “withdrawn” to show this status (76.5%).

Table 15. Indicators of non-compliance with selected recommended practices in PubMed.

	Number	*Total*	*Percent*
Retracted publications without separate PubMed retraction notice*	2,587	8,814	29.4
Titles of retraction notices shorter than 6 words (retracting single publication)	805	5,570	14.5
Retracted epub publications without “Withdrawn” (or “withdrawal”) as the first word of the title*	132	926	14.3
Titles of retracted publications not modified to show retracted status*	6,741	8,814	76.5

* Note: Likely an underestimate of non-compliance as compliant records were more likely to be found in our searches.

Discussion

Article retraction is intended to end the use of seriously compromised literature by users unaware of a publication’s problematic status. For it to do so effectively, retraction needs to be unambiguous, and reflected across the scholarly community’s distributed knowledge system clearly and quickly. Our study shows this commonly fails to occur in the biomedical databases at the core of this system.

Marasović and colleagues have said that when there are serious discrepancies in major databases, this “diminishes the credibility and transparency of the research and publication system and creates confusion”. (73) In addition, it has consequences for investigations of authors’ research integrity track records. The large amount of un-

or mis-identified retracted publications also casts considerable uncertainty over the findings of studies of retractions.

Retractions are rare – less than 0.1% of publications in our data. That may make them particularly problematic. However, some of the problems contributing to their poor identification and visibility may be having an impact on common post-publication events such as errata as well. That would mean that the integrity of a substantial proportion of the scholarly record is likely to be affected.

Contributing factors to non-identified retractions

Some of the issues we identified as contributing to the non-identification of retracted status could be addressed at the NLM. NLM indexing only routinely occurs for final publications, not those that are e-publications ahead of print. Retraction often occurred when a publication is in the “ahead of print” stage and thus never reaches “in print” status. Retraction notices, too, were often only e-published.

A considerable body of literature falls into this category: as of February 2018, there were close to 200,000 “ahead of print” records in PubMed. Our experience of ongoing monitoring during this study showed that ongoing screening of new records in PubMed and PMC with limited search fields could be done in a few hours a week at the current growth rate of the literature. That could capture most retracted publications when signals of retraction appear in PubMed or PMC, “ahead of print” or not. Along with monitoring additions to the Retraction Watch database, this would enable timely indexing of retracted status in PubMed and PMC. Complete identification of retracted publications, however, depends on journal editors and publishers.

Greater policy clarity could be helpful. The NLM retraction policy specifically does not differentiate between withdrawal and retraction, and reason for removal is irrelevant. However, its policy on publications withdrawn “ahead of print” implicitly allows for a distinction. Publications withdrawn because they are out of date are assigned retraction status – unless they are published in the *CDSR*. Emerging publisher

practices, such as temporary removal, identified here, and using the same DOI for retracted and then republished articles, (73) also need to be identified and the policy implications considered.

A further contributing problem to optimal identification is time lag. (74) Identification of most of these key research integrity events could be prompt if incoming records were filtered, as we did in this study. Machine learning could also be explored.

The case of temporary removals highlights another potential contributor to identification problems. If publishers make changes to records after NLM indexing is completed, indexing may need to be reviewed. For PubMed users, not displaying previous versions of altered abstracts obscures the historical record and can lead to ambiguity. We also found that the lack of systematic harmonization of post-publication events between PubMed and PMC contributed to under-identification of retracted publications.

What NLM's role should be when journals consistently fail to adhere to NLM requirements is a critical question. Journal personnel provide the data in PubMed and PMC. They are responsible for its accuracy and timeliness. For PubMed, journal personnel are able to directly tag and link retracted publications and retraction notices. However, we identified considerable shortfall in best practices for retraction on the journal side. For example, separate notices of retraction were not submitted to PubMed for 29.4% of retracted publications.

NLM applies standardized classification to the varied nomenclature journals have always used for retracting publications. That enables users to navigate the publishing system with reduced opportunity for confusion. Editorial practices change over time, making this role both more complex, and, we believe, more valuable. For example, we observed an apparent trend toward designating "ahead of print" publications as something less than published. Schmidt speculates that using the term withdrawn instead of retracted "may well be analyzed as a potential strategic instrument of journal publishers, since withdrawn publications are much less visible, discussed, and problematized". Avoiding stigma is one of the rationales for some current

proposals to change naming conventions and practices. (75,76) Evolving practices will require periodic review of policy, processes, and search filters, and the NLM has an important role to play in safeguarding an accurate historical record and ensuring clarity for literature users.

Estimates of non-identification of retractions at journals and other databases

We did not aim to evaluate the coverage of retracted publications in all data sources we used (for example, publisher websites and CrossRef). However, we observed retractions not notified to PubMed/PMC, as well as false positives and non-identification of retraction, in both journals and databases. Several small studies since 2011 (involving 18 to 233 publications) have shown that there are problems in clear marking of retracted publications in journals and bibliographic databases.

Decullier and colleagues found that 22% of the 233 retracted articles in their sample did not show that they were retracted at the journals. (12) Wright and McDaid found that of 17 retracted articles retrievable in PubMed and EMBASE, all were tagged as retracted in PubMed, but 16 were not tagged in EMBASE (94%). Of the 15 identified in the Cochrane trials register, 40% were not tagged as retracted. (11)

Bakker and Riegelman found that of 144 articles on mental health sourced from the Retraction Watch blog post archive, 2% were not identified as retracted at the journals. For those found in bibliographic databases, 9% were not tagged in PubMed, while 16% were not in MEDLINE via Ovid, 21% were not in PsycINFO via Ovid, 29% were not in Web of Science, and 95% were not identified as retracted in EBSCO and Scopus. (14)

Marasović and colleagues studied 29 articles tagged as “corrected and republished” in MEDLINE, which replace an article that may or not also have been formally retracted. They found that about half did not identify the corrected status at the journals, CrossRef’s on-article display (CrossMark), (56) Web of Science, or Scopus. (73) We identified a total of only 1,792 publications tagged as retracted in CrossRef’s API. CrossRef is not restricted to biomedicine and includes over 80 million

publications, so the proportion of retracted publications that are not identified in CrossRef is very high.

Schmidt reported a large study examining identification of retractions. (13) Of 3,446 publications she classified as retracted in PubMed from 2008 to 2013, 25% (888 publications) were not tagged as retracted. Where she was able to locate the studies in Web of Science, 37% were not identified as retracted there.

Breakdowns in processes for following through on editorial decisions to retract

The mechanics of retracting a published article have always raised challenges. When all journals were in print and their volume was not high, subscribers and librarians could manually add notices to copies of record. However, this became prohibitive long before the internet. (14,77,78) Capturing circulated reprints and photocopies to correct them was never possible. The emergence of digital records raised new challenges, some of which were predicted. (77) Digital journals made some aspects of flagging changed status of a publication easier, if content management systems were adequately prepared for rare events such as retractions. However, some journal content management systems do not appear to be equipped for these events.

The move of scholarly communication online also enabled deletion of publications without explanation. Further threats to the integrity of the publishing record can now occur at multiple points. It appeared to us that associated record linkages were sometimes lost after a change in the publishing platform, or the transfer of a journal from one publisher to another.

Post-publication events such as retractions and errata were generally wedged into small spare spaces on pages in print journals. It also appeared to us that the digitization of legacy print journals could at times have concentrated on scanning larger articles, and neglected the need to spot small notices, create separate records for them, and then link them to the affected publications. This would explain why there is often a record of a retraction at PubMed that cannot be found at the online journal. We did not envisage this problem at the outset of our study, and so did not

collect data on it. If this has happened commonly, then a considerable body of corrections to the literature, not only retractions, could have been lost in the digitization process. Criticisms and letters, too, have suffered from under-curation and neglect, “buried in scholarly journals”, as Thomasson and Stanley put it in 1955, and the problem has endured. (64,65)

These problems highlight the critical functions provided by the NLM. Often, the only remaining record of a retraction is now PubMed/PMC. This also underscores the need for all versions of those records to be permanently and publicly displayed. Even apparently administrative changes, like adding or removing an author, can be a flag for a research integrity issue.

In 2010, journalists Marcus and Oransky launched the Retraction Watch blog, generally highlighting retractions spotted in the literature or PubMed. (65,81) In 2018, they launched a searchable database of retractions. (82) The beta version of this database was a critical source for us of retractions that had not been notified to PubMed/PMC by publishers, highlighting the vital role Retraction Watch now plays in the scholarly publication system. The largest number of retracted publications available at any single source was Retraction Watch. In 2019, Retraction Watch announced a partnership with a reference management system to annotate retractions in users’ reference libraries. (83)

Study limitations

Although we believe we have identified most publications in PubMed/PMC that have been retracted, we are sure many remain unidentified. Some may be flagged at journals but not at PubMed/PMC, but some are likely to be flagged nowhere online.

Resource constraints limited some of the strategies that may have yielded further retracted publications in PubMed, in particular searching for records with identical titles within journals, and records that have remained e-publications “ahead of print” for a prolonged period after 2013. Because some searches were not undertaken for

the whole period, our study is most complete for publications to the end of 2012, and least complete for publications from September to December 2017.

All included retractions were confirmed by at least two authors, but the initial screening of some databases was by a single author. Another limitation of our study is that we were not able to estimate trends based on time of retraction with a high level of certainty, as such a high proportion of retractions are undated, and we did not collect retraction dates from journals when they were not available in PubMed/PMC. We did not retain sources of data coupled to included publications, and so could not assess relative coverage of retracted publications in the databases and publisher websites we searched.

Our study's principal strengths were in substantially increasing the identification of retracted publications in PubMed/PMC, and diagnosing contributing factors to their currently significant under-ascertainment. Understanding and identification of other types of post-publication events would benefit from a similar intensive approach. We found that expanding the search terms typically used to look for retracted publications in PubMed/PMC resulted in valuable yields. While a minority of retractions is elusive, all but a small percentage could be captured by filtered searches that are sensitive to evolving journal practice. Making them easy for everyone to find and reducing invisible retractions to a negligible level, though, will require commitment from journals and publishers as well as the NLM and all literature services.

Acknowledgments

All authors received funding from the Intramural Research Program of the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM), National Institutes of Health (NIH). HB received support from an Australian Government Research Training Program Scholarship. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The authors are grateful to David Lipman, former director of NCBI, for his support for initiating the PubMed Commons research program, and to Alison Abritis and Ivan Oransky from Retraction Watch, for providing data on additions to their database made between September and the end of 2017. We also thank Mark Jones from Bond University for advice about time trend analysis.

The datasets for this study will be available in the Open Science Framework repository, https://osf.io/8xbqy/.

Authors' contributions

HB initiated the study and drafted the original manuscript. All authors participated in the design of this study, development of initial search strategies, screening of records, data analysis and checking, and preparation of the final manuscript. MV and DJ managed the data, and DJ ran final analyses and prepared the figures. HB and DJ finalized the search strategies. HB is guarantor for the data and contents. All authors contributed critical reading of the manuscript and approved the final version.

Disclosures

All authors worked at the NCBI when most of this work was conducted, including work on PubMed Commons, PubMed's post-publication commenting system. DJ continues to work at NCBI. The NCBI manages PubMed and PMC. No authors have a retracted publication in PubMed/PMC. HB has a retracted protocol in the *Cochrane Database of Systematic Reviews*.

References

1. Neale AV, Northrup J, Dailey R, Marks E, Abrams J. Correction and use of biomedical literature affected by scientific misconduct. Sci Eng Ethics. 2007 Mar;13(1):5–24.

2. Trikalinos NA, Evangelou E, Ioannidis JPA. Falsified papers in high-impact journals were slow to retract and indistinguishable from nonfraudulent papers. J Clin Epidemiol. 2008 May;61(5):464–70.

3. Davis PM. The persistence of error: a study of retracted articles on the Internet and in personal libraries. J Med Libr Assoc JMLA. 2012 Jul;100(3):184–9.

4. Lu SF, Jin GZ, Uzzi B, Jones B. The retraction penalty: evidence from the Web of Science. Sci Rep. 2013 Nov 6;3:3146.

5. Madlock-Brown CR, Eichmann D. The (lack of) impact of retraction on citation networks. Sci Eng Ethics. 2015 Feb;21(1):127–37.

6. Bar-Ilan J, Halevi G. Temporal characteristics of retracted articles. Scientometrics. 2018 Sep;116(3):1771–83.

7. Hamilton DG. Continued Citation of Retracted Radiation Oncology Literature-Do We Have a Problem? Int J Radiat Oncol Biol Phys. 2019 Apr 1;103(5):1036–42.

8. Mott A, Fairhurst C, Torgerson D. Assessing the impact of retraction on the citation of randomized controlled trial reports: an interrupted time-series analysis. J Health Serv Res Policy. 2019 Jan;24(1):44–51.

9. Fang FC, Casadevall A. Retracted science and the retraction index. Infect Immun. 2011 Oct;79(10):3855–9.

10. Hesselmann F, Graf V, Schmidt M, Reinhart M. The visibility of scientific misconduct: A review of the literature on retracted journal articles. Curr Sociol. 2017;65(6):814–45.

11. Wright K, McDaid C. Reporting of article retractions in bibliographic databases and online journals. J Med Libr Assoc JMLA. 2011 Apr;99(2):164–7.

12. Decullier E, Huot L, Samson G, Maisonneuve H. Visibility of retractions: a cross-sectional one-year study. BMC Res Notes. 2013 Jun 19;6:238.

13. Schmidt M. An analysis of the validity of retraction annotation in pubmed and the web of science. J Assoc Inf Sci Technol. 2018 Feb;69(2):318–28.

14. Bakker C, Riegelman A. Retracted Publications in Mental Health Literature: Discovery across Bibliographic Platforms. J Librariansh Sch Commun. 2018;6(General Issue):eP2199.

15. U.S. National Library of Medicine. MEDLINE, PubMed, and PMC (PubMed Central): How are they different? [Internet]. National Library of Medicine NIH.

2018 [cited 2019 Aug 12]. Available from: https://www.nlm.nih.gov/bsd/difference.html

16. U.S. National Library of Medicine. MEDLINE®: Description of the Database [Internet]. National Library of Medicine NIH. 2019 [cited 2019 Aug 12]. Available from: https://www.nlm.nih.gov/bsd/medline.html

17. U.S. National Library of Medicine. Technical Notes. NLM Tech Bull. 1984 Jul;(83):3–4.

18. Kotzin S, Schuyler PL. NLM's practices for handling errata and retractions. Bull Med Libr Assoc. 1989 Oct;77(4):337–42.

19. Roberts WC. Retractions. Am J Cardiol. 1987 Nov 1;60(13):1134.

20. International Committee of Medical Journal Editors. Retraction of research findings. International Committee of Medical Journal Editors. Br Med J Clin Res Ed. 1988 Feb 6;296(6619):400.

21. Huth EJ, Case K. The URM: twenty-five years old. Sci Ed. 2004 Feb;27(1):17–21.

22. U.S. National Library of Medicine. Retracted publication and retraction of publication. NLM Tech Bull. 1989 Feb;(Supplement 1):27.

23. U.S. National Library of Medicine. Coming Attractions - Publication Types. NLM Tech Bull. 1990 Oct;(256):21–4.

24. Colaianni LA. Retraction, comment, and errata policies of the US National Library of Medicine. Lancet Lond Engl. 1992 Aug 29;340(8818):536–7.

25. U.S. National Library of Medicine. Errata, Retraction, Duplicate Publication, and Comment Policy [Internet]. National Library of Medicine NIH. 1997 [cited 2019 Aug 12]. Available from: https://web.archive.org/web/19970607010700/https://www.nlm.nih.gov/pubs/factsheets/errata.html

26. U.S. National Library of Medicine. MEDLINE Data Changes - 2007. NLM Tech Bull [Internet]. 2006 Dec;(353). Available from: https://www.nlm.nih.gov/pubs/techbull/nd06/nd06_medline_data_changes2007.html

27. Gillikin D. Partial Retractions: NLM Policy Change. NLM Tech Bull. 2016 Oct;(No. 412):e3.

28. U.S. National Library of Medicine. MEDLINE/PubMed Production Improvements Underway. NLM Tech Bull. 2016 Aug;(411):e1.

29. U.S. National Library of Medicine. PubMed Data Management (PMDM) Help - Editing a Single Citation [Internet]. National Library of Medicine NIH. [cited 2019

Aug 12]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/management/help/editing-single/

30. U.S. National Library of Medicine. Policies [Internet]. PMC. [cited 2019 Sep 2]. Available from: https://www.ncbi.nlm.nih.gov/pmc/about/guidelines/

31. U.S. National Library of Medicine. Errata, Retractions, and Other Linked Citations in PubMed [Internet]. National Library of Medicine NIH. [cited 2019 Aug 12]. Available from: https://www.nlm.nih.gov/bsd/policy/errata.html

32. Wager E, Barbour V, Yentis S, Kleinert S, on behalf of COPE Council. Guidelines for Retracting Articles. Version 1. Comm Publ Ethics COPE Guidel [Internet]. 2009 Sep; Available from: https://publicationethics.org/resources/guidelines-new/guidelines-retracting-articles

33. Council of Science Editors. CSE's White Paper on Promoting Integrity in Scientific Journal Publications [Internet]. Council of Science Editors (CSE); 2018 [cited 2019 Aug 12]. Available from: https://www.councilscienceeditors.org/resource-library/editorial-policies/white-paper-on-publication-ethics/

34. International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals [Internet]. International Committee of Medical Journal Editors (ICMJE); 2018. Available from: http://www.icmje.org/icmje-recommendations.pdf

35. Moylan EC, Kowalczuk MK. Why articles are retracted: a retrospective cross-sectional study of retraction notices at BioMed Central. BMJ Open. 2016 23;6(11):e012047.

36. Qi X, Deng H, Guo X. Characteristics of retractions related to faked peer reviews: an overview. Postgrad Med J. 2017 Aug;93(1102):499–503.

37. Wager E, Williams P. Why and how do journals retract articles? An analysis of Medline retractions 1988-2008. J Med Ethics. 2011 Sep;37(9):567–70.

38. Grieneisen ML, Zhang M. A comprehensive survey of retracted articles from the scholarly literature. PloS One. 2012;7(10):e44118.

39. Vaught M, Jordan DC, Bastian H. Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central. Res Integr Peer Rev [Internet]. 2017;2. Available from: https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-017-0030-2

40. U.S. National Library of Medicine. What's New for 2018 MeSH. NLM Tech Bull. (419):e8.

41. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology

(STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007 Oct 16;4(10):e296.

42. Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, et al. So many filters, so little time: the development of a search filter appraisal checklist. J Med Libr Assoc JMLA. 2008 Oct;96(4):356–61.

43. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009 Oct;62(10):e1-34.

44. U.S. National Library of Medicine. MEDLINE®/PubMed® Data Element (Field) Descriptions [Internet]. U.S National Library of Medicine. [cited 2019 Aug 22]. Available from: https://www.nlm.nih.gov/bsd/mms/medlineelements.html

45. Udell J. Preserving Comments From PubMed Commons [Internet]. Hypothesis. 2018 [cited 2019 Aug 14]. Available from: https://web.hypothes.is/blog/archiving-pmc-comments/

46. Snodgrass GL, Pfeifer MP. The characteristics of medical retraction notices. Bull Med Libr Assoc. 1992 Oct;80(4):328–34.

47. Steen RG, Casadevall A, Fang FC. Why has the number of scientific retractions increased? PloS One. 2013;8(7):e68397.

48. Abritis AJ. An Assessment of Retractions as a Measure of Scientific Misconduct and Impact on Public Health Risks [Internet]. Tampa: University of South Florida; 2015. (Graduate Theses and Dissertations). Available from: https://scholarcommons.usf.edu/etd/5630/

49. Almeida RMVR, de Albuquerque Rocha K, Catelani F, Fontes-Pereira AJ, Vasconcelos SMR. Plagiarism Allegations Account for Most Retractions in Major Latin American/Caribbean Databases. Sci Eng Ethics. 2016 Oct;22(5):1447–56.

50. Huh S, Kim SY, Cho H-M. Characteristics of Retractions from Korean Medical Journals in the KoreaMed Database: A Bibliometric Analysis. PloS One. 2016;11(10):e0163588.

51. Jawaid SA, Jawaid M. Retractions by Pakistan Journal of Medical Sciences due to Scientific Misconduct. J Coll Physicians Surg--Pak JCPSP. 2016 Aug;26(8):718–9.

52. Yan J, MacDonald A, Baisi L-P, Evaniew N, Bhandari M, Ghert M. Retractions in orthopaedic research: A systematic review. Bone Jt Res. 2016 Jun;5(6):263–8.

53. Nogueira TE, Gonçalves AS, Leles CR, Batista AC, Costa LR. A survey of retracted articles in dentistry. BMC Res Notes. 2017 Jul 6;10(1):253.

54. Sayers E. A General Introduction to the E-utilities [Internet]. National Center for Biotechnology Information (US); 2010 [cited 2019 Aug 12]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK25497/

55. Winter D. rentrez: “Entrez” in R [Internet]. 2017. Available from: https://cran.r-project.org/web/packages/rentrez/index.html

56. Lammey R. CrossRef developments and initiatives: an update on services for the scholarly publishing community from CrossRef. Sci Ed. 2014;1(1):13–8.

57. Chamberlain S. rcrossref [Internet]. 2016. Available from: https://CRAN.R-project.org/package=rcrossref

58. RStudio Team. RStudio: integrated development environment for R. [Internet]. Boston: RStudio Inc; 2015 [cited 2019 Aug 12]. Available from: https://www.rstudio.com/

59. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing; 2016. Available from: https://www.r-project.org/

60. Wickham H. tidyverse [Internet]. RStudio Inc; 2017. Available from: https://CRAN.R-project.org/package=tidyverse

61. Kassambara A, Kosinski M, Biecek P. Survminer [Internet]. 2019. Available from: https://CRAN.R-project.org/package=survminer

62. Therneau TM, Lumley T. Survival [Internet]. 2019. Available from: https://CRAN.R-project.org/package=survival

63. Bastian H, Jordan DC, Vaught M. PubMed Commons Post-Publication Projects [Internet]. 2016 [cited 2019 Aug 12]. Available from: https://osf.io/8xbqy/

64. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016 05;5(1):210.

65. Marcus A, Oransky I. Retraction Watch: Tracking retractions as a window into the scientific process [Internet]. Retraction Watch. [cited 2019 Aug 12]. Available from: https://retractionwatch.com/

66. U.S. National Library of Medicine. MEDLINE®PubMed® XML Element Descriptions and their Attributes [Internet]. U.S National Library of Medicine. 2018 [cited 2019 Aug 26]. Available from: https://www.nlm.nih.gov/bsd/licensee/elements_descriptions.html

67. U.S. National Library of Medicine. Yearly Citation Totals from 2018 MEDLINE/PubMed Baseline: 27,837,540 Citations Found [Internet]. U.S National Library of Medicine. 2019 [cited 2019 Aug 18]. Available from: https://www.nlm.nih.gov/bsd/licensee/2018_stats/2018_Totals.html

68. U.S. National Library of Medicine. Ahead of Print Withdrawn Policy [Internet]. PubMed Overview. 2006 [cited 2019 Aug 26]. Available from: https://www.ncbi.nlm.nih.gov/entrez/query/static/aopret.html

69. National Center for Biotechnology Information (US). PubMed Help [Internet]. Bethesda: U.S. National Library of Medicine, NIH; 2005. Available from: https://www.ncbi.nlm.nih.gov/books/NBK3830/

70. Higgins JP, Green SE. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 5.1.0. Oxford: Cochrane Collaboration; 2011. Available from: http://handbook-5-1.cochrane.org/

71. U.S. National Library of Medicine. Obstetrics and Gynecology. In: NLM Catalog [Internet]. Bethesda: U.S. National Library of Medicine, NIH; Available from: https://www.ncbi.nlm.nih.gov/nlmcatalog/0401101

72. Cochrane Collaboration. Cochrane Editorial and Publishing Policy Resource [Internet]. Cochrane Collaboration; 2019 [cited 2019 Oct 15]. Available from: https://documentation.cochrane.org/display/EPPR

73. Marasović T, Utrobiĉić A, Marušić A. Transparency of retracting and replacing articles. Lancet Lond Engl. 2018 31;391(10127):1244–5.

74. Decullier E, Huot L, Maisonneuve H. What time-lag for a retraction search on PubMed? BMC Res Notes. 2014 Jun 25;7:395.

75. Barbour V, Bloom T, Lin J, Moylan EC. Amending published articles: time to rethink retractions and corrections? [version 1; peer review: 2 approved with reservations]. F1000 Res. 2017;(6):1960.

76. Fanelli D, Ioannidis JPA, Goodman S. Improving the integrity of published science: An expanded taxonomy of retractions and corrections. Eur J Clin Invest. 2018 Apr;48(4).

77. Freeman I, Spurlock S. Management of journal errata in a health sciences library. Med Ref Serv Q. 1986;5(1):41–9.

78. Hughes C. Academic medical libraries' policies and procedures for notifying library users of retracted scientific publications. Med Ref Serv Q. 1998;17(2):37–42.

79. Thomasson P, Stanley JC. Uncritical Citation of Criticized Data. Science. 1955 Apr 22;121(3147):610–1.

80. Bhopal RS, Tonks A. The role of letters in reviewing research. BMJ. 1994 Jun 18;308(6944):1582–3.

81. Marcus A, Oransky I. What studies of retractions tell us. J Microbiol Biol Educ. 2014 Dec;15(2):151–4.

82. Brainard J. Rethinking retractions. Science. 2018 Oct 26;362(6413):390–3.

83. Oransky I. Want to check for retractions in your personal library — and get alerts — for free? Now you can [Internet]. Retraction Watch. 2019 [cited 2019 Aug 12]. Available from: https://retractionwatch.com/2019/06/12/want-to-check-for-retractions-in-your-personal-library-and-get-alerts-for-free-now-you-can/

Chapter 6

Systematic reviews and the perpetuation of error in the biomedical literature: a descriptive study and pilot of research integrity search filters for PubMed

Abstract

Background

Findings of a systematic review could be unreliable if results of included studies are compromised. Research on prevalence of potentially compromised trials in systematic reviews and adequacy of ascertainment of them in PubMed is sparse.

Objectives

To develop and evaluate research integrity filters for post-publication events in PubMed, and estimate prevalence in systematic reviews and clinical trials included in them.

Methods

Fourteen research integrity filters were developed. Filtered search results for all PubMed were collected and matched against a sample of 36,462 trials included in Cochrane reviews in 2017, and 83,302 publications indexed as non-Cochrane meta-analyses in PubMed as a proxy for systematic reviews. Filters for retractions were tested against a reference set of confirmed retractions from a previous study.

Results

Retractions were rare. Errata affected 2.0% of trials and 2.1% of systematic reviews; letters to the editor affected 13.9% of trials and 8.4% of systematic reviews; and author replies affected 4.7% of trials and 2.3% of systematic reviews. In the 2,025 Cochrane reviews in this study with a median of 12 included trials, 28.0% were affected by trials potentially compromised by retraction, correction, or expressions of concern (94.8% of them errata). Most retracted trial publications in Cochrane reviews remained (60.0%), with only 53.3% reported. PubMed indexing of retracted publications identified 65% of retracted trial publications compared to 86% for the filter with the highest yield. Ascertainment for errata and other common events is not known.

Conclusions

Corrections, letters to the editors, and author replies are common for clinical trials, and may be more so for those included in systematic reviews. As systematic reviews grow, the possibility of included studies being compromised is high. PubMed indexing alone is insufficient for identifying potentially compromised clinical trials in systematic reviews, and they are under-recognised by systematic reviewers.

Hilda Bastian
Margaret J. Sampson
Jenny Doust
Melissa Vaught

Background

Systematic reviewing is intended to overcome problems caused by considering individual studies informally or relying on unsystematic selections of studies. (1) Systematic reviews of clinical trials in particular have come to occupy an influential position in health care decision making, at individual, clinical practice guideline, policy, regulatory, and funding levels. (2) These reviews can be highly reliable, but even the best capture evidence up to one point in time, while evidence can be a shifting target.

The possibility that a systematic review's findings could be outdated because of subsequently completed or published studies attracts considerable effort and research attention. (3,4) However, a systematic review could also potentially entrench an unreliable picture in people's minds, and in the literature, if any of its included studies are found after publication to be compromised. Correction of a primary evidence report does not automatically follow through to the literature that has cited it, including systematic reviews. Even where updating of systematic reviews is expected, the process tends to be forward-looking – seeking new studies for potential inclusion – rather than considering the ongoing reliability and research integrity of the studies already included. (5)

Exceptions are a 2016 Cochrane consensus paper on updating systematic reviews that discusses retractions, (6) and Cochrane's 2019 guidance on search strategies for updates of systematic reviews in the *Cochrane Handbook for Systematic Reviews of Interventions* (*Handbook*). (7) Their advice is to "search online journals or databases such as MEDLINE (if the study is indexed there) for any notifications, corrections or retractions". Cochrane's methodological standards on updating, however, are explicitly only forward-looking, (8) and until 2019, so was the *Handbook* chapter on updating. (9) Part of their guidance on managing search results for reviews has included downloading associated post-publication event fields when downloading MEDLINE citations.

Post-publication events range from those with little or no potential to affect the reliability of a systematic review's results, such as the correction of the spelling of an author's name, to those that could completely undermine them, such as retraction of a pivotal trial with fabricated data. Broadly considered, post-publication activity encompasses a wide range of commentary, as well as evaluations in subsequent literature, including systematic reviews. The categories of events addressed in this study, though, are those directly associated with a publication and can affect its status, accuracy, and/or research integrity, that is, retractions, official findings of research misconduct, editorial expressions of concern (EOCs), errata, papers republished after correction, letters to the editor, and authors' replies. EOCs are a vehicle for editors to alert readers when they have a concern about the integrity of a publication, but are not at the point of issuing an erratum or retraction notice. (10)

What should be done about post-publication events in preparing, updating, and correcting or retracting systematic reviews depends on the prevalence and impact of those events, as well as the effectiveness and efficiency of potential strategies to ameliorate their impact. There has been relatively little research on these issues, and little of it has included evaluation of relying on PubMed to identify post-publication events.

Retractions have been the most studied, (11) although – and perhaps partly because they are rare. Bastian et al concluded that fewer than 0.05% of publications in PubMed were retracted by the end of 2017, and only 69.4% of them were indexed with this status. *[Chapter 5]* Editorial EOC and articles republished after retraction or correction are more rare than retractions. (10,12) Although retracted studies are rare, two small studies found a substantial effect in around 30% to 40% of meta-analyses that included them. (13,14)

The more common post-publication activities are errata, letters to the editor, and author replies to letters to the editor. The prevalence of errata has been estimated generally between 1% and 5% of publications (although higher in some studies), with studies concluding from 5% to over 50% of these are potentially important. (15–23)

The prevalence of letters to the editor has been estimated at 12% to 52%, (24–27) with anywhere from 3% to 99% having author replies from journal to journal. (27–29) High rates of important criticisms in letters to the editors have been reported, from 30% to 90% of the letters in which content was classified. (24,26,29) Author replies are also a potential source of data and process information about studies that is not public elsewhere.

A substantial proportion of post-publication events affect methodology, data, and other aspects of content relevant for systematic reviews. Those with serious implications for clinical trials are likely to be common, at least in prominent journals. Serious implications for an included trial will not necessarily translate into serious implications for a systematic review, however, for example if the data affected by error was not included in the review or if that data carried little weight in the analysis. Evidence on prevalence, impact, and strategies to assess the implications of such events for systematic reviews is sparse, as is evidence on the effectiveness of various research integrity search methods.

This study reports on the development and piloting of research integrity filters to identify post-publication events in PubMed, and aimed to estimate the prevalence of these events in systematic reviews and clinical trials included in them. We also assessed the ascertainment of errata in PubMed's MEDLINE indexing.

Methods

This study is reported according to the STROBE statement for reporting observational studies, (30) and the UK InterTASC Information Specialists' Sub-Group (ISSG) Search Filter Appraisal Checklist. (31)

Search filter development.

In previous work with a colleague, two of the authors identified barriers to retrieval of post-publication events in PubMed, including under-ascertainment of retracted publications. [*Chapter 5]* (10) That project used search filters developed after a

literature review of studies of retractions (Appendix 4). After reviewing search strategies in the literature, an extensive search strategy was developed iteratively as retracted publications were identified from non-PubMed sources that had been missed by PubMed searching. The same process informed the development of filters for EOC. (10) The development of filter for errata and other events did not have the benefit of as extensive literature, and we did not have suitable large confirmed reference sets for evaluation for any other type of event than retractions.

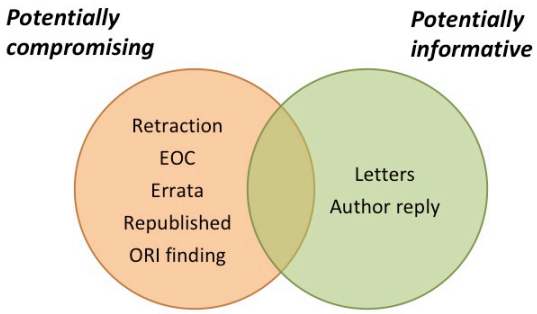
The series of research integrity search filters for this pilot study were developed by one author (HB) based on that experience and review of the literature of post-publication events detectable in PubMed (Appendix 5). There were two goals: very specific filters requiring no or minimal screening, and, where feasible, multiple versions of filters to potentially increase the retrieval of publications affected by post-publication events, following the model developed by Haynes et al for the Clinical Queries feature in PubMed. (32,33)

In addition to 13 filters for specific post-publication events, there is one to identify publications with prolonged “e-published ahead of print” status for screening. “Epub” records are added to PubMed, but are not indexed until they are added to a print issue and the record has its status changed. An “old” record with this status could either have never entered a print issue, or a new PubMed record could have been generated instead of converting the “epub” record. Not proceeding to a print issue has been shown to be a strong indicator that a publication has problems. *[Chapter 5]* Only a small proportion of “epub” records have not converted status within two years.

Each of these types of post-publication events could be compromising, informative, or neither for people preparing or maintaining systematic reviews. Some are the most extreme critical events possible for a publication, such as retraction or official findings of research misconduct such as those made by the Office of Research Integrity (ORI) for research funded by the U.S. National Institutes of Health and other health agencies. (34) However, even though each event can have catastrophic to scientifically trivial implications, some are more officially compromising to a

publication than others. We classified the event types as potentially compromising or as potentially informative or compromising. This classification is shown in Figure 18.

Figure 18. Implications of types of post-publication events for the preparation and maintenance of systematic reviews.



Estimating the prevalence of post-publication events in systematic reviews and the trials included in them.

(a) Trials and systematic reviews.

To estimate the prevalence of post-publication events, we used two datasets. (Appendix 6 includes a guide to all data files, including descriptors.)

Our trial datasets include a convenience sample of PubMed records for 36,462 studies included in 2,205 systematic reviews from the *Cochrane Database of Systematic Reviews* (*CDSR*) (called Cochrane reviews). We refer to the included studies as trials throughout, but they may not be exclusively trials. They are a subset of included trials published between 1974 and 2014 from meta-analyses with at least five trials retrieved from issue 2 of the *CDSR* in 2017 by Lamberink et al, who studied statistical power of trials across time. (35) A dataset including PubMed record numbers (PMID) was supplied by Lamberink and we obtained permission from Cochrane for our use. The trials with associated PMIDs were a small minority of trials

in Lamberink's sample of 136,212, and we do not know what proportion of the trials included in those Cochrane reviews are represented in our sample.

Although Cochrane reviews are amended and updated, the *CDSR* does not issue errata notices. In addition, although it has a commenting facility, (24) the *CDSR* does not submit letters to the editor for indexing in PubMed. For these reasons we needed a dataset of reviews not published in the *CDSR* to establish prevalence data for post-publication events among systematic reviews. We could not locate a sufficiently large dataset of non-*CDSR* systematic reviews. As a proxy, our dataset includes a download from 15 November 2019 of 83,302 non-*CDSR* PubMed records from 1990 to 2017 of publication type (36) meta-analysis, restricted to humans (search terms: meta-analysis[pt] NOT "cochrane database syst rev"[jo]).

(b) Reference set for confirmed retracted publications.

Our reference set for confirmed retracted publications is compiled from data collected by Bastian et al, *[Chapter 5]* with a total 8,814 PubMed records for retracted publications in PubMed to the end of 2017.

(c) Prevalence of post-publication events.

Filters were used in PubMed (37) on 16 November 2019, with no search limits except for the "epub ahead of print" filter (limited 2002 to 2017). Results were downloaded using the comma separated value (CSV) file option. The total number of records in PubMed that day was also determined (search term: all[sb]), and the total number of records for 2018 was determined on 20 November, including the numbers of records affected by each of the filters (except for epub ahead of print, filtering ORI citations, and filtering comments to identify letters to the editor and author replies).

Three of the searches required multiple steps: ORI findings, letters to the editor, and author replies to letters to the editor. For ORI findings, the results for the filter were downloaded as full MEDLINE records. The unique identifiers (PMIDs) for affected publications were extracted from the MEDLINE record.

The resulting files were strings of PMIDs for publications affected by post-publication events and/or post-publication events themselves. Potentially affected publications were ascertained by identifying PMID matches in the PMIDs for affected trials and non-Cochrane systematic reviews.

There is no specific MeSH indexing term for letters to the editor that discuss a publication: the publication type for letters also covers research letters in journals. If a publication is discussed in a letter to the editor, it can be linked to the publication as a comment. Comments also include editorials discussing or promoting the publication. To narrow down to post-publication letters, results of the first filter for associated comments were downloaded as a CSV file. For those that affected trials and systematic reviews in our data, PMIDs for all associated comments were extracted from the MEDLINE record. These were then filtered with the second filter for letters, and the filter for authors' replies.

Prevalence of each type of post-publication event for trials and systematic reviews was estimated conservatively, by using the data from the filter requiring the most minimal screening. That was the MeSH indexing alone, except for three cases. For both retractions and EOC, the proportion was determined by MeSH indexing alone combined and de-duplicated with reference sets.

We compiled data sets to broaden the net to include notices of retraction and EOC, to capture any potentially misclassified records. The retraction list expanded our retraction reference to a total 14,901 PubMed records for retracted publications and retraction notices. These include the 8,814 publications in the filter performance reference set, retraction notices in PubMed to the end of 2017 (n = 5,847), and duplicate PubMed records for the above (n = 108).

The dataset for EOCs includes 1,025 PubMed records related to EOC and affected publications. This combines and de-duplicates two data sources: 1) the reference set of 538 PubMed records compiled in a previous study by two of the authors of this study, Vaught et al (10) (up to 2016); and 2) a PubMed search with no search limits

on 16 November 2019 yielding 875 records (search terms: hasexpressionofconcernin OR hasexpressionofconcernfor).

To explore the impact on Cochrane reviews of the subset of trials, we used the data from the filters needing the least screening (usually MeSH indexing alone). The number of Cochrane reviews including them was calculated, as well as the number of Cochrane reviews including more than one affected trial, and the number of trials included in more than one Cochrane review.

Reviews can be retracted from the CDSR because of problems identified with the review, or for being out-of-date. The source data for trial-Cochrane review pairs included Cochrane identifiers for the reviews, but not PubMed identifiers. To identify retracted Cochrane reviews, the *CDSR* records in our retractions reference set were identified, and their Cochrane identifiers extracted. PubMed was searched for withdrawn *CDSR* reviews from 2018 to the search date (18 November 2019). These were added to those from the reference set, and de-duplicated, resulting in a dataset of 461 retracted *CDSR* reviews. Retracted Cochrane reviews, and retracted trial publications included in Cochrane reviews, were described.

Filter performance

We assessed the internal validity of the three filters for retracted publications against the reference set for confirmed retracted publications described above. We based our approach to testing filter performance on the “diagnostic test” approach to study identification. (32,38,39) The reference set was used to calculate to calculate the sensitivity of each search result with 95% confidence intervals (CI) for the three filters. (38) We defined sensitivity as the percentage of the reference set retrieved by the searches. We also calculated the percentage of missed records for the filters.

To estimate the current screening workload if applying screening to all PubMed, as well as the timeliness of MeSH indexing of post-publication events, results for all PubMed were compared with those for all PubMed in 2018 alone. We used the filters other than “epub”, and the top level filters only for comments (not the subsets of LTE

and author reply) and ORI findings (not the subset of publications affected by ORI findings).

Assessing the ascertainment of errata in PubMed's MEDLINE indexing.

We assessed the ascertainment of errata in PubMed's indexing by attempting to reproduce the results of a study published by Hauptman et al, (21) and using it as a reference set. Hauptman et al identified a corpus of 557 erratum notices by hand-searching the most-cited English language medical journals, 10 general medicine journals and 10 cardiovascular medicine journals, with at least 25 original articles published in the study period. The publications were searched for the number of eligible types of publication for 18 months (1 July 2009 to 31 December 2010), with issues up to 30 June 2012 searched for erratum notices about them. According to notation for Table 3 in the paper, there was a one-to-one relationship between erratum notices and papers.

Of the 557 erratum notices, 141 affected publications from before the study period, leaving 482 papers in the 18-month period affected by an erratum report up to the end of June 2012. The errata were also rated for severity by at least two investigators, with differences resolved with a third.

The names of the 20 journals were published, but the publications were not identified. A request to the corresponding author for the data was unsuccessful.

The search strategy to reproduce the Hauptman et al results is included in Appendix 7, together with results. The search was done on 19 September 2019 by HB, who screened each result for letter and news, and the rest by date of publication. Search results were checked by MS.

Data management and analysis

One author undertook all data collection, curation, and visualisation (HB). Data were downloaded as CSV files and managed and analysed using RStudio 1.1463 running

R 3.5.2. (9) Packages tidyverse, (20) epiR, (40) and reshape2 (21) were used for analyses and data visualisation. Summary statistics were used to describe the cohort, and sensitivity and precision were used to evaluate filter performance. Data for this project, including analytic code, will be deposited at GitHub. (41)

Results

A. Research integrity filters.

The research integrity filters include MeSH terms that identify PubMed indexing of post-publication events, as listed in Table 16. They do not apply to PubMed Central (PMC). There are no MeSH terms that directly index ORI findings of research misconduct or letters to the editor that are specifically post-publication discussions of another publication.

Table 16: NLM's MeSH indexing terms for post-publication events.

Objective	**Indexing terms**
Retracted publications	retracted publication[pt] hasretractionin
Retraction notices	retraction of publication[pt] hasretractionof
Publications with expressions of concern	hasexpressionofconcernin
Notices of expression of concern	hasexpressionofconcernfor
Publications with errata	haserratumin
Erratum notices	published erratum[pt] haserratumfor
Corrected and republished article	corrected and republished article[pt] hascorrectedrepublishedfrom hascorrectedrepublishedin
Letters to the editor *(includes research letters, not only post-publication letters)*	letter[pt]
Publications with associated letters to the editor *(includes other types of commentary)*	hascommentin

A total of 14 research integrity filters were developed. The components of the filters and how to add them as custom filters are described in Appendix 6.

The first 10 filters relating to four types of events include multiple options (Table 17). The first option is not directed specifically at publications affected by the event, including signals of that type of event as well (for example, retraction notices). The options described as publication checkers are designed to find affected publications.

Table 17. Research integrity filters with multiple options.

Name	Characteristics	Search strategy
Retractions		
Retraction 1	Requires screening	hasretractionof OR hasretractionin OR retraction of publication[pt] OR retracted publication[pt] OR retracted[ti] OR retraction[ti] OR retracted[tt] OR retraction[tt] OR www.elsevier.com/locate/withdrawalpolicy OR (Elsevier* AND policy* AND article* AND (remov* OR withdraw)) OR (withdrawn[ti] OR withdrawn[tt]) OR (withdrawn[tiab] AND (article[tiab] OR articles[tiab] OR e-publication[tiab] OR e-publications[tiab] OR epub[tiab] OR “ahead of print”[tiab] OR manuscript[tiab] OR manuscripts[tiab] OR publication[tiab] OR publications[tiab])) OR (temporary[ti] AND removal[ti]) OR (temporary[tt] AND removal[tt])
Retraction 2	A publication checker, requires screening	hasretractionin OR retracted publication[pt] OR retracted[ti] OR retracted[tt] OR www.elsevier.com/locate/withdrawalpolicy OR (Elsevier* AND policy* AND article* AND (remov* OR withdraw)) OR (withdrawn[ti] OR withdrawn[tt]) OR (temporary[ti] AND removal[ti]) OR (temporary[tt] AND removal[tt])
Retraction 3	A publication checker, MeSH indexing only	hasretractionin OR retracted publication[pt]
Expressions of concern (EOC)		
EOC 1	Requires screening	hasexpressionofconcernin OR hasexpressionofconcernfor OR (“expression of concern”[all fields] OR “notice of concern”[all fields] OR “note of concern”[all fields] or "statement of concern"[all fields]) OR ((expression*[ti] OR statement*[ti] OR note*[ti] OR notice*[ti]) AND concern*[ti])
EOC 2	A publication checker, MeSH indexing only	hasexpressionofconcernin
Errata		
Errata 1	Requires screening	published erratum[pt] OR haserratumfor OR haserratumin OR erratum*[ti] OR correction[ti] OR corrigendum*[ti] OR errata*[ti] OR corrections[ti] OR corrigenda*[ti] OR erratum*[tt] OR correction[tt] OR corrigendum*[tt] OR errata*[tt] OR corrections[tt] OR corrigenda*[tt]
Errata 2	A publication checker, requires screening	(haserratumin OR erratum*[ti] OR correction[ti] OR corrigendum*[ti] OR errata*[ti] OR corrigenda*[ti] OR erratum*[tt] OR correction[tt] OR corrigendum*[tt] OR errata*[tt] OR corrigenda*[tt])
Errata 3	A publication checker, MeSH indexing only	haserratumin
Corrected and republished		
Republished 1	Requires screening	"corrected and republished" OR hascorrectedrepublishedin OR hascorrectedrepublishedfrom OR corrected and republished article[pt]
Republished 2	A publication checker, MeSH indexing alone	hascorrectedrepublishedin OR corrected and republished article[pt]

The next three filters (ORI findings, letters to the editor, and author replies) require two-steps of filtering (Table 18). The final filter (prolonged epub status) is a non-specific filter for a group of publications at high risk of being affected by a compromising post-publication event, requiring a time limit. *[Chapter 5]*

Table 18. Research integrity filters requiring multiple steps.

Name	Characteristics	Search strategy
ORI findings of misconduct		
ORI	Requires two steps and screening	"NIH Guide Grants Contracts"[jo] *Extract and screen PMIDs from the CON field in the MEDLINE version of the PubMed record*
Letters to the editor		
Letters	Requires two steps, MeSH indexing alone	hascommentin *Extract PMIDs from CIN field in MEDLINE version of the PubMed record, then screen those PMIDs for:* letter[pt]
Author replies		
Replies	Requires two steps, MeSH indexing alone for letters to the editor	hascommentin *Extract and screen PMIDs from the MEDLINE version of the PubMed record, then screen those PMIDs for:* letter[pt] AND reply
Non-specific		
Prolonged epub status	Requires time limit and screening	Pubstatusaheadofprint time limit: 2002 up to two years ago (in this study, the end of 2017).

B. Prevalence of post-publication events for trials and systematic reviews.

The retrieval results and best prevalence estimates for types of post-publication events are shown in Table 19. There were 313 records cited in ORI findings of research misconduct and 12,293 PubMed records with prolonged epub status between 2002 and 2017. However, none of either of these sets of records included any of the trials or non-Cochrane systematic reviews in our study.

Three records for trials from three Cochrane reviews were no longer in PubMed, one because it had been a duplicate record. Both of the other publications remain online at the journals. One is an author reply to a letter to the editor providing extra data about a trial. (42) We did not deduct these three from our calculations.

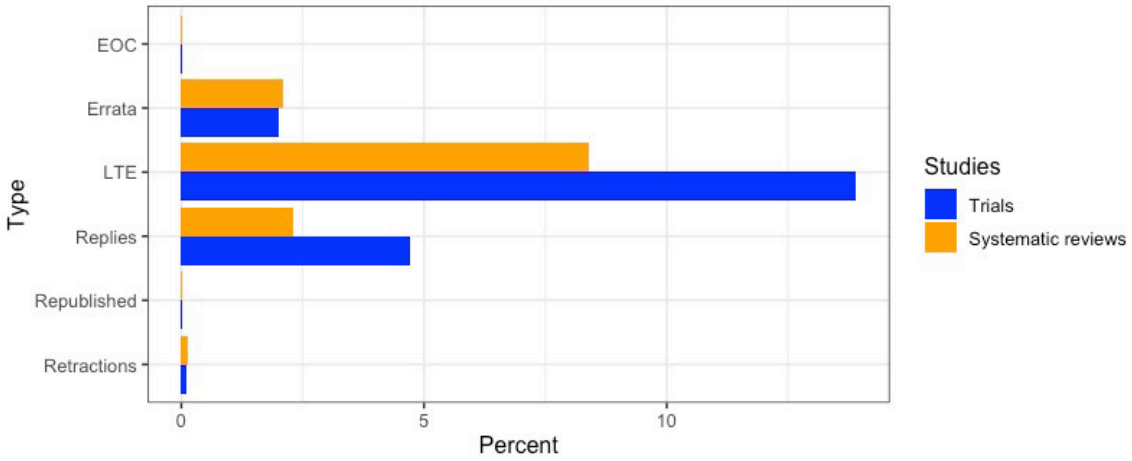
Table 19. Retrievals and retrieval rates of post-publication events[1] by filter for trials and systematic reviews in PubMed, with best prevalence estimate.

Studies Trials (n = 36,462) Systematic reviews (n = 83,302)	**Filter 1**		**Filter 2**		**Filter 3**		**Best prevalence estimate[2]**	
	No.	*%*	*No.*	*%*	*No.*	*%*	*No.*	*%*
Retracted publications								
Trials	41	0.11	39	0.11	35	0.10	36	0.10
Systematic reviews	128	0.20	110	0.13	105	0.13	107	0.13
Publications with expression of concern (EOC)								
Trials	3	<0.01	3	<0.01	*n.a.*	*n.a.*	3	<0.01
Systematic reviews	5	<0.01	2	<0.01	*n.a.*	*n.a.*	2	<0.01
Publications with errata								
Trials	785	2.1	785	2.1	749	2.0	749	2.0
Systematic reviews	1,821	2.2	1,817	2.2	1,718	2.1	1,718	2.1
Corrected and republished publications								
Trials	3	<0.01	3	<0.01	*n.a.*	*n.a.*	3	<0.01
Systematic reviews	17	0.02	17	0.02	*n.a.*	*n.a.*	17	0.02
Publications with letters to the editor								
Trials	5,601	13.9	*n.a.*	*n.a.*	*n.a.*	*n.a.*	5,601	13.9
Systematic reviews	6,991	8.4	*n.a.*	*n.a.*	*n.a.*	*n.a.*	6,991	8.4
Publications with author replies to letters to the editor								
Trials	1,706	4.7	*n.a.*	*n.a.*	*n.a.*	*n.a.*	1,706	4.7
Systematic reviews	1,926	2.3	*n.a.*	*n.a.*	*n.a.*	*n.a.*	1,926	2.3

[1] No records were retrieved with filters for ORI findings or prolonged epub status.
[2] Best prevalence estimate for retractions: based on the combination of the retraction reference set and MeSH indexing term for a retracted publication. Best prevalence estimate for EOC: based on the combination of the EOC reference set and MeSH indexing term for a publication with an EOC. For other types, the estimate is that using MeSH indexing terms for affected publications alone.

Prevalence was similar for trials and systematic reviews, except for letters and author replies, which were more common for trials. Figure 10 illustrates the best prevalence estimates from Table 19.

Figure 19. Best prevalence estimates of types of post-publication events for trials and systematic reviews.



C. Relationship between trials affected by post-publication events and the Cochrane reviews that included them.

There were 3,073 trials affected by post-publication events that were included in more than one Cochrane review (8.4%), with a range from 1 to 31 (IQR 1). The trials in our sample are a minority of those included in meta-analyses with five or more trials. The 2,025 reviews in which they are included are, in turn, a minority of Cochrane reviews: 25.7% of the 7,874 at the end of 2018. (43) This non-random sample has a disproportionate number of included trials. The median number of trials our sample included per Cochrane review was 12 (Table 20).

A total of 567 of the 2,025 reviews were affected by a potentially compromised included study (28.0%). Table 20 shows that the number of Cochrane reviews affected by potentially compromising post-publication events had larger numbers of included trials.

Table 20. Proportion of trials per Cochrane review: total sample and those affected by a potentially compromised included study.

Individual CRGs	**All Cochrane reviews** *(n = 2,025)*	**Subset with potentially compromised trials** *(n = 567)*
Number of trials: median	12	23
Number of trials: range	1 – 493 (IQR 17)	2 – 493 (IQR 28)
Number and proportion of reviews with more than 100 included trials	29 (1.4%)	26 (4.6%)

The prevalence of potentially compromised trials is broken down in Table 21. All but 5% of the affected trials had errata. Nine of the Cochrane reviews were themselves retracted (27.3% of the 33 retracted reviews in the whole sample of 2,025). We do not know whether these reviews were retracted because problems identified in the review, whether they were all withdrawn because they were outdated, or whether post-publication events played a part in either case.

Table 21. Potentially compromised included trials in Cochrane reviews.

Post-publication event	**Trials affected** *(n = 790[1])*		**Reviews affected** *(n = 2,025[2])*		**Reviews retracted**
	No.	*%*	*No.*	*%*	*No.*
Retracted	36	4.6	15	0.7	1
EOC	3	0.4	3	0.1	0
Errata	749	94.8	547	27.0	8
Republished	3	0.4	4	0.2	0
Total	***791***[3]	***100***	***569***[4]	***n.a.***	***9***

[1] The proportion of affected trials, not the proportion of all trials.
[2] Number of Cochrane reviews affected, and the proportion of all Cochrane reviews.
[3] One trial with an editorial expression of concern also had errata.
[4] One Cochrane review was affected by the trial with more than one event, and one republished trial affected two Cochrane reviews.

We analysed the 15 Cochrane reviews that included the 36 retracted trial publications (Table 22). There was a median of one retracted trial publication in each of the 15 reviews (range 1–19; IQR 1). Two of the retracted trial publications were each included in two reviews. The one Cochrane review with retracted trial publications that was itself retracted was the one with 19 included in this sample, all related to one trialist.

Table 22. Cochrane reviews including retracted trial publications in meta-analyses (n = 15).

Number of retracted trial publications	Retracted trial publication(s) still included		Retraction reported		Retracted before inclusion		Retracted before current version	
	No.	*%*	*No.*	*%*	*No.*	*%*	*No.*	*%*
1 only (n = 10)	8	80.0	4	40.0	5	50.0	8	80.0
> 1 (n = 5)	1	20.0	4	80.0	0	0	4	80.0
Total (n = 15)	***9***	***60.0***	***8***	***53.3***	***5***	***33.3***	***12***	***80.0***

In five of the 15 reviews, there was no indication that the reviewers were aware of the retractions. In one of those five, the publication had been retracted in 2004, and the review first published in 2007. In three of the five, the publications were retracted after the current version of the review was published. The time since those three publications had been retracted was two, six, and eight years. In the last of the five reviews not reporting the retraction, the publication's retraction notice was not linked to it at PubMed or at the journal: it too had been retracted (in 2003) before the review was first published (in 2015). (44)

For the other three reviews that did not report the retractions, Cochrane authors' replies to comments by readers concerned at the inclusion of retracted studies revealed that they were aware of them. For one review, the authors replied that the errors causing the retraction of a publication only applied to one of the study's

reports, and did not compromise the results used in the review. For another, the authors replied that removing it would make no difference to conclusions: no amendment was made to the review. For the third, a retraction for lack of ethics committee approval was not considered grounds for removing the trial from the review.

D. Filter performance and MEDLINE-indexing ascertainment of potentially compromising post-publication events.

Filter performance

We assessed the internal validity of the five filters for retracted publications and those with expressions of concern against our reference sets of confirmed affected publications. The reference sets included 8,814 retracted publications in PubMed to the end of 2017, and 300 publications with EOC to the end of 2016.

Retrieved versus missing records from the filter searches are illustrated in Figure 20. Use of the MeSH indexing terms only (Retraction 3 and EOC 2) was relatively poor, and the additional filter for EOCs was not an improvement. However, both other retraction filters substantially improved ascertainment of retracted publications (Retraction 1 and Retraction 2). Sensitivity was 65% (95% CI: 0.64 – 0.66) for MeSH indexing and 86% using the filter with the highest yield (95% CI: 0.86 – 0.87) (Retraction 1) (Table 23).

Figure 20. Performance of three retraction filters in PubMed (records retrieved and missed).

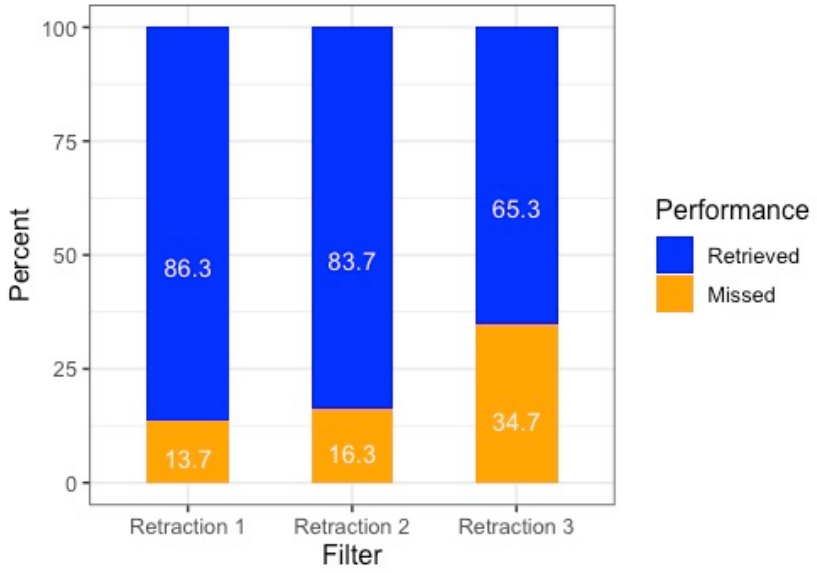


Table 23. Retraction filter performance: retrieved and missed records, and filter sensitivity.

Filter	**Records retrieved**	**Affected records retrieved**	**Affected records missed**	**Sensitivity[1]** *(95% CI)*
Retraction1	22,063	7,604	1,210	0.86 (0.86–0.87)
Retraction2	11,038	7,374	1,440	0.84 (0.83–0.84)
Retraction3	7,124	5,756	3,058	0.65 (0.64–0.66)

[1] Sensitivity is the percentage of the reference set records retrieved (retractions: n = 8,814).

All PubMed

There were 30,333,299 records in PubMed, and 1,335,691 with 2018 publication or epub ahead of print dates. The number and proportion of hits for filters in all of

PubMed, and in 2018 only, are shown in Table 24. Based on the filters with the largest yield of potentially compromising post-publication events (marked grey in the table), the number of records the NLM or other literature service providers would need to screen for the critical events would be less than 620 records a week.

Table 24. Retrieval of records by filter in all of PubMed, and in 2018 only.

Filter	**All PubMed** *(n = 30,333,299)*		**In 2018 only** *(n = 1,335,691)*	
	No.	%	No.	%
Retraction 1	22,063	0.073	1,446	0.108
Retraction 2	11,038	0.036	592	0.044
Retraction 3 *	7,124	0.023	264	0.020
Expression of concern 1	1,551	0.005	100	0.007
Expression of concern 2 *	487	0.002	20	0.001
Errata 1	316,518	1.043	30,623	2.293
Errata 2	303,772	1.001	28,697	2.148
Errata 3 *	202,636	0.668	13,354	1.000
Corrected republished 1	2,929	0.010	42	0.003
Corrected republished 2 *	2,913	0.010	40	0.003
All commented-on *	722,191	2.381	42,800	3.204
ORI *	313	0.001	0	0

* MEDLINE indexing only. Grey colouring indicates largest-yield filters for potentially compromising post-publication types.

The results in Table 24 give some indication of the extent of under-ascertainment of post-publication events from reliance on MEDLINE indexing alone. Our reference set included 7,828 publications confirmed as retracted to the end of 2017, compared with 7,124 to November 2019 identified by MEDLINE-indexed retracted publications (Table 24). ORI findings have not been completely indexed. (45) Our assessment of MEDLINE ascertainment of errata follows.

Ascertainment of errata in PubMed's MEDLINE indexing.

Hauptman et al hand-searched the 20 most highly cited general medical and cardiovascular medicine journals to ascertain the prevalence of errata (21). We were

able to identify 451 publications in PubMed that were possibly eligible in this study, and were indexed as errata in the time period of the Hauptman (Appendix 7). If all these studies were among the 482 found by Hauptman et al, the rate of ascertainment of errata in PubMed indexing in 2019 for prominent medical journals in 2012 was 93.6%.

Discussion

Retractions are rare, but errata, letters to the editor, and author replies are common for clinical trials and systematic reviews. In this study, the prevalence of post-publication events was similar for both, except for letters and replies. They occurred more often for trials (13.9% versus 8.4% for letters). That might be related to the amount of attention clinical trials attract.

We were able to analyse the relationship between a large group of trial publications with data in meta-analyses, and the Cochrane reviews that included them. Although they were a minority of the reviews' included trials, the reviews accounted for about a quarter of Cochrane reviews. They had a disproportionately high number of included trials, which increased the risk of one or more being affected. We found that 28.0% of these reviews included trial publications that were potentially compromised by errata/correction, retraction, or expression of concern.

Reviews with fewer included studies may be less likely to be affected by errors in them, but these could have a bigger impact in a smaller study pool. We did not assess the potential for impact of these events on the reviews' results. Many would have no implications for the reviews, but the consequences for some could be critical. (13) Fanelli and Moher found when studies were retracted due to data, methods, or results, effect size was over-estimated by an average of 30% (median of 13%). (14)

Other types of events can have similar repercussions. Farrah and Rabb concluded that 16% of errata in included drug trials could affect a systematic review's results.

(23) Royle and Waugh found that 10% of errata would affect interpretation of trials, and 5% could change a systematic review's conclusion. (15) Studies of errata generally have found rates of errors with serious implications or involving data ranging from 6.3% to 54.2%. (17–19,22) Important proportions of letters have also been found to have serious implications for both trials and systematic reviews. (24,26,29)

Analysis of post-publication events could affect systematic reviews in different ways. For example, we encountered two Cochrane reviews in this study where authors' replies to letters were the source for data included in meta-analyses, (46,47) but we did not assess the prevalence of this. We identified an overall prevalence for author replies to letters about trials of 4.7%, based on a filter process that relied on MeSH indexing for comments and letters.

Our analysis of the 15 Cochrane reviews including data from retracted trial publications showed that only half were reported in the review. Five of the 15 author teams were apparently unaware of the retraction(s), including two cases where the retractions had occurred years before the review was first published. There were conflicting practices on matters such as whether a trial affected by research misconduct on ethical grounds could continue to be included in a review.

In one of these cases of unreported retracted trial publications, the retraction was not linked to the trial at the journal or at PubMed. Thus, neither of Cochrane's recommended strategies for identifying retractions and errata would have found it. (7) Our study confirms that relying on MeSH indexing alone will lead to under-ascertainment of important post-publication events, especially recent ones. We established that screening using our two research integrity filters with terms additional to MeSH indexing could increase the ascertainment of retracted publications. Even higher ascertainment levels could be reached by checking journals as well, and using the Retraction Watch database either directly, (48) or via the free and open source reference management system, Zotero. (49)

Based on our evaluation of a reference set from prominent medical journals in 2012 in the Hauptman et al study, (21) there is also under-ascertainment of errata using MeSH terms only as a filter. We found that 93.6% of these errata were indexed by 2019. However, Farrah et al found that only 69% of errata in a reference set for clinical trials of drugs was MeSH indexed, with 73% in EMBASE and Scopus, and 27% ascertainment in Cochrane's CENTRAL database of trials. Our work suggests that our filters with terms additional to MeSH indexing could also increase timely ascertainment of errata.

Since 2016, journal publishers have had the ability to tag and link their PubMed records as they submit them, and to edit them later. (50) They could resolve a large part of this problem with linking, and with more uniform use of naming conventions for post-publication events in titles and abstracts. Our 2018 data shows they have yet to do so. This raises issues for NLM, and strengthens the case for timely screening of all incoming records to PubMed. We estimate the number of records to screen for the critical events is currently less than 620 records a week. That would take a relatively small investment to ensure the integrity of PubMed and save considerable time across the community.

The workload for systematic reviewers could also be reduced if Cochrane's CENTRAL register of trials was screened, and if Cochrane undertook a collaboration with Retraction Watch similar to that at Zotero. (49) There may also be a case for coordinated coding of reviewed events. In our study, 8.4% of the potentially compromised trials were included in more than one Cochrane review. If author replies prove to be a rich source of data on trials, screening centrally for these could also prove worthwhile.

Identifying published post-publication events addresses a major part of the research integrity checking problem, but not all of it. That which is not published or co-located with publications is also a problem. As systematic reviews expand further into using correctible sources such as reports in clinical trial registries, (7) keeping up with those records pose an additional challenge.

Findings of research misconduct are critical events that require a clear profile in PubMed and at journals, and we have demonstrated how to include those from ORI in research integrity checks. Only a minority of these publications is retracted. (34) Seife has reported on the degree of invisibility in publications for trials where the U.S. Food and Drug Administration (FDA) reported important conduct violations. (51) Public research misconduct findings at other national agencies could and should be assembled into a corpus that can be incorporated into PubMed and linked to publications. Some academic institutions have published findings of research integrity investigations, as Tilburg University did in the Stapel case, (52) but this remains unusual.

A more widespread problem is journal identification of, and willingness to issue errata and publish letters on, the research they publish. (53–56) Some journals do not publish letters to the editor at all, or do so for only a limited time after the publication of a study. (57,58) This underscores the potential importance of additional sources of red flags about publications, such as PubPeer. (59)

Our study of published post-publication events had multiple limitations, several of which have already been mentioned. We did not address PubMed Central (PMC), even though it and PubMed are not totally in sync. *[Chapter 5]* We also could not show how often major errors or retractions result in correction or retraction of systematic reviews. We also did not collect the number of events per affected study.

The proportion of errata notices is similar between trials and systematic reviews. We do not know to what extent the errata at systematic reviews refer to their own errors, (60) correction of errors that flow from identification of error in included studies, or identification of information resulting in re-consideration of previously excluded studies.

However, we have addressed problems and strategies that could be tackled by individuals, journals, and bibliographic services. In 1999, Chalmers and Altman stressed the potential for online publication to perform the critical function of “threading” disparate sources of data on clinical trials, (57) but in 2011, Altman et al

showed this had not yet happened. (61) It still has not. In the meantime, systematic reviewers and others have a roadmap for more thorough research integrity checks of publications in PubMed.

Funding

There was no specific funding of this study. HB received support from an Australian Government Research Training Program Scholarship.

Acknowledgements

We are grateful to Diana Jordan, who made a major contribution to the work leading up to this project. *[Chapter 5]* We are also grateful to Herm Lamberink for providing the list of PubMed identifiers for trials included in meta-analyses of Cochrane reviews, (35) and to Cochrane for permission to use and release this data.

Contributions

HB initiated the study, and sourced, collected, curated, and analysed the data. HB drafted the manuscript, and MS contributed to methods and manuscript development. HB and MV both contributed to the development of methods and concepts that informed this study. All authors contributed to interpretation of the data and critical revision of the manuscript.

Disclosures

HB and MV previously worked at the publishers of PubMed, the National Center for Biotechnology Information (NCBI) in the NLM, on projects related to post-publication events in PubMed. DJ currently works at NCBI.

References

1. Moher D, Stewart L, Shekelle P. All in the Family: systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. Syst Rev. 2015 Dec 22;4(1):183.

2. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLOS Med. 2010 Sep 21;7(9):e1000326.

3. Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. When and how to update systematic reviews. Cochrane Database Syst Rev. 2008 Jan 23;(1):MR000023.

4. Garritty C, Tsertsvadze A, Tricco AC, Sampson M, Moher D. Updating systematic reviews: an international survey. PloS One. 2010 Apr 1;5(4):e9914.

5. Tsertsvadze A, Maglione M, Chou R, Garritty C, Coleman C, Lux L, et al. Updating Comparative Effectiveness Reviews: Current Efforts in AHRQ's Effective Health Care Program. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2011 [cited 2019 Sep 16]. (AHRQ Methods for Effective Health Care). Available from: http://www.ncbi.nlm.nih.gov/books/NBK66066/

6. Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ. 2016 20;354:i3507.

7. Higgins JPT, Thomas J. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 6.0. London: Cochrane Collaboration; 2019 [cited 2019 Oct 2]. Available from: https://training.cochrane.org/handbook/current

8. Higgins JPT, Lasserson T, Chandler J, Tovey D, Churchill R. Methodological Expectations of Cochrane Intervention Reviews (MECIR) [Internet]. London: Cochrane Collaboration; 2019. Available from: https://community.cochrane.org/mecir-manual

9. Higgins JP, Green SE. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 5.1.0. Oxford: Cochrane Collaboration; 2011. Available from: http://handbook-5-1.cochrane.org/

10. Vaught M, Jordan DC, Bastian H. Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central. Res Integr Peer Rev [Internet]. 2017;2. Available from: https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-017-0030-2

11. Hesselmann F, Graf V, Schmidt M, Reinhart M. The visibility of scientific misconduct: A review of the literature on retracted journal articles. Curr Sociol. 2017;65(6):814–45.

12. Marasović T, Utrobiĉić A, Marušić A. Transparency of retracting and replacing articles. Lancet Lond Engl. 2018 31;391(10127):1244–5.

13. Garmendia CA, Nassar Gorra L, Rodriguez AL, Trepka MJ, Veledar E, Madhivanan P. Evaluation of the Inclusion of Studies Identified by the FDA as Having Falsified Data in the Results of Meta-analyses: The Example of the Apixaban Trials. JAMA Intern Med. 2019 Mar 4;

14. Fanelli D, Moher D. What difference do retractions make? An estimate of the epistemic impact of retractions on recent meta-analyses [Internet]. bioRxiv; 2019. Available from: https://www.biorxiv.org/content/10.1101/734137v1

15. Royle P, Waugh N. Should systematic reviews include searches for published errata? Health Inf Libr J. 2004 Mar;21(1):14–20.

16. Trikalinos TA. Does it mean anything if your own name is wrong in your published paper? FASEB J Off Publ Fed Am Soc Exp Biol. 2009 Aug;23(8):2345–8.

17. Molckovsky A, Vickers MM, Tang PA. Characterization of published errors in high-impact oncology journals. Curr Oncol Tor Ont. 2011 Jan;18(1):26–32.

18. Bhatt VR, Aryal MR, Panta S, Mosalpuria K, Armitage JO. A retrospective analysis of reported errata in five leading medical journals in 2012. J Community Hosp Intern Med Perspect. 2014;4(5):25738.

19. Castillo M, Northam M, Halm KE. Postpublication errors in imaging-related journals. AJNR Am J Neuroradiol. 2012 Sep;33(8):1447–8.

20. Grcar J. Comments and Corrigenda in Scientific Literature. Am Sci. 2013 Feb;101(1):16.

21. Hauptman PJ, Armbrecht ES, Chibnall JT, Guild C, Timm JP, Rich MW. Errata in medical publications. Am J Med. 2014 Aug;127(8):779-785.e1.

22. Strothmann M. No Bad Publicity: The Bibliometric Advantage of Corrected Articles in PLOS ONE. Sci Technol Libr. 2018 Oct 2;37(4):394–408.

23. Farrah K, Rabb D. Errata for trial publications are not uncommon, are frequently not trivial, and can be challenging to access: a retrospective review. J Med Libr Assoc JMLA. 2019;107(2):187–93.

24. Ober M, Shohara R, Rennie D, Yank V, Bero LA. The Criticism Management System for the Cochrane Library. Eval Health Prof. 2002 Mar;25(1):98–115.

25. Baethge C, Seger G. Our readers' voice. Letters to the editor are an important component of the discussion of scientific articles, in Deutsches Arzteblatt as in

other journals. our correspondence pages reflect a diversity of opinion thanks to the love of debate among our readers--and thanks to a few rules. Dtsch Arzteblatt Int. 2009 Mar;106(12):207–9.

26. Kastner M, Menon A, Straus SE, Laupacis A. What do letters to the editor publish about randomized controlled trials? A cross-sectional study. - PubMed - NCBI. BMC Res Notes. 2013 Oct;6:414.

27. Wiebe DJ, Flynn K, Branas CC. Letters to the editor in response to studies of guns in the home and homicide and suicide. Inj Epidemiol. 2017 Dec;4(1):3.

28. von Elm E, Wandel S, Jüni P. The role of correspondence sections in post-publication peer review: A bibliometric study of general and internal medicine journals. Scientometrics. 2009 Dec;(81):747.

29. Gøtzsche PC, Delamothe T, Godlee F, Lundh A. Adequacy of authors' replies to criticism raised in electronic letters to the editor: cohort study. BMJ. 2010 Aug 10;341:c3926.

30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007 Oct 16;4(10):e296.

31. Glanville J, Bayliss S, Booth A, Dundar Y, Fernandes H, Fleeman ND, et al. So many filters, so little time: the development of a search filter appraisal checklist. J Med Libr Assoc JMLA. 2008 Oct;96(4):356–61.

32. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc JAMIA. 1994 Dec;1(6):447–58.

33. Haynes RB, Wilczynski N. Finding the gold in MEDLINE: clinical queries. ACP J Club. 2005 Feb;142(1):A8-9.

34. Abritis AJ. An Assessment of Retractions as a Measure of Scientific Misconduct and Impact on Public Health Risks [Internet]. Tampa: University of South Florida; 2015. (Graduate Theses and Dissertations). Available from: https://scholarcommons.usf.edu/etd/5630/

35. Lamberink HJ, Otte WM, Sinke MRT, Lakens D, Glasziou PP, Tijdink JK, et al. Statistical power of clinical trials increased while effect size remained stable: an empirical analysis of 136,212 clinical trials between 1975 and 2014. J Clin Epidemiol. 2018;102:123–8.

36. U.S. National Library of Medicine. Publication Characteristics (Publication Types) with Scope Notes [Internet]. NLM Medical Subject Headings. 2018 [cited 2019 Sep 15]. Available from: https://www.nlm.nih.gov/mesh/pubtypes.html

37. U.S. National Library of Medicine. MEDLINE, PubMed, and PMC (PubMed Central): How are they different? [Internet]. National Library of Medicine NIH.

2018 [cited 2019 Aug 12]. Available from: https://www.nlm.nih.gov/bsd/difference.html

38. Lefebvre C, Glanville J, Beale S, Boachie C, Duffy S, Fraser C, et al. Assessing the performance of methodological search filters to improve the efficiency of evidence information retrieval: five literature reviews and a qualitative study. Health Technol Assess Winch Engl. 2017;21(69):1–148.

39. Leeflang MMG, Scholten RJPM, Rutjes AWS, Reitsma JB, Bossuyt PMM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. J Clin Epidemiol. 2006 Mar;59(3):234–40.

40. Stevenson M. epiR: Tools for the Analysis of Epidemiological Data [Internet]. Melbourne; 2019. Available from: https://CRAN.R-project.org/package=epiR

41. Bastian H. Perpetuation of error. GitHub [Internet]. 2019 Nov; Available from: https://github.com/hildabast/perpetuation-of-error

42. Araco F, Gravante G, Sorge R, Piccione E. TVT-O vs. TVT: reply to comments by Burton et al. Int Urogynecology J. 2009 Mar 1;20(3):371–2.

43. Cochrane Collaboration. Dashboard [Internet]. Cochrane Community. [cited 2019 Oct 30]. Available from: https://community.cochrane.org/organizational-info/resources/strategy-2020/dashboard

44. Voluntary Retraction of: Surgical outcome in mesial temporal sclerosis correlates with prion protein gene variant. Neurology [Internet]. 2007 Jul 24;69(4). Available from: https://n.neurology.org/content/69/4/405

45. National Institutes of Health (US), United States Department of Health Education and Welfare, United States Department of Health and Human Services. NIH guide for grants and contracts [Internet]. NLM Catalog. [cited 2019 Nov 20]. Available from: https://www.ncbi.nlm.nih.gov/nlmcatalog/9004625

46. Ford AA, Rogerson L, Cody JD, Aluko P, Ogah JA. Mid-urethral sling operations for stress urinary incontinence in women. Cochrane Database Syst Rev. 2017 31;7:CD006375.

47. Lewis SR, Pritchard MW, Evans DJ, Butler AR, Alderson P, Smith AF, et al. Colloids versus crystalloids for fluid resuscitation in critically ill people. Cochrane Database Syst Rev. 2018 03;8:CD000567.

48. Brainard J. Rethinking retractions. Science. 2018 Oct 26;362(6413):390–3.

49. Oransky I. Want to check for retractions in your personal library — and get alerts — for free? Now you can [Internet]. Retraction Watch. 2019 [cited 2019 Aug 12]. Available from: https://retractionwatch.com/2019/06/12/want-to-check-for-retractions-in-your-personal-library-and-get-alerts-for-free-now-you-can/

50. U.S. National Library of Medicine. MEDLINE/PubMed Production Improvements Underway. NLM Tech Bull. 2016 Aug;(411):e1.

51. Seife C. Research Misconduct Identified by the US Food and Drug Administration: Out of Sight, Out of Mind, Out of the Peer-Reviewed Literature. JAMA Intern Med. 2015 Apr 1;175(4):567–77.

52. Tilburg University. Commissie Levelt [Internet]. Tilburg University. [cited 2019 Nov 25]. Available from: https://www.tilburguniversity.edu/nl/over/gedrag-integriteit/commissie-levelt

53. Allison DB, Brown AW, George BJ, Kaiser KA. Reproducibility: A tragedy of errors. Nat News. 2016 Feb 4;530(7588):27.

54. Shashok K, Matarese V. Post-publication peer review in biomedical journals: overcoming obstacles and disincentives to knowledge sharing. RT J Res Policy Eval [Internet]. 2018 Aug 29 [cited 2019 Nov 25];6(1). Available from: https://riviste.unimi.it/index.php/roars/article/view/10125

55. Marcus A, Oransky I. The data thugs. Science. 2018 Feb 16;359(6377):730–2.

56. Goldacre B, Drysdale H, Marston C, Mahtani KR, Dale A, Milosevic I, et al. COMPare: Qualitative analysis of researchers' responses to critical correspondence on a cohort of 58 misreported trials. Trials. 2019 Feb 14;20(1):124.

57. Chalmers I, Altman DG. How can medical journals help prevent poor medical research? Some opportunities presented by electronic publishing. Lancet Lond Engl. 1999 Feb 6;353(9151):490–3.

58. Altman DG. Unjustified restrictions on letters to the editor. PLoS Med. 2005 May;2(5):e126; discussion e152.

59. Wager E, Veitch E. The Role of PubPeer Comments in Alerting Editors to Serious Problems With Clinical Research Publications. In Chicago: American Medical Association; 2017 [cited 2019 Oct 28]. p. prc17-0144. Available from: https://peerreviewcongress.org/prc17-0144

60. Ford AC, Guyatt GH, Talley NJ, Moayyedi P. Errors in the conduct of systematic reviews of pharmacological interventions for irritable bowel syndrome. Am J Gastroenterol. 2010 Feb;105(2):280–8.

61. Altman DG, Hrynaszkiewicz I, Furberg C, Grimshaw JM, Rothwell PM. Five years of Trials. Trials. 2011 Nov 23;12:248.

Chapter 7

Discussion

Summary

The increase in new evidence, coupled with the problems identified in trials already reviewed, is outpacing the methods, infrastructure, and collaborations needed to enable systematic reviewers to keep up with shifting evidence. Yet research gaps remain, and there are also problems of critical unidentified trials and unidentified or unreported errors. Our current systems are inadequate for the challenges of correcting the record, but valuable gains are feasible. A more serious approach to error correction might also help with the cultural change required to limit the harm done by compromised research reports.

Hilda Bastian

Keeping up with evidence that is growing quickly is challenging. Eriksen has described how our solutions and filters for dealing with information in a time of surplus inevitably break down rapidly, and we need to continually find new solutions. (1) Systematic reviewing is one solution, providing filtering and synthesis that is essential to interpreting bodies of evidence in biomedicine, especially as they grow and shift direction. But the practice is itself in constant need of new solutions as the corpus of clinical trials grows, and our knowledge about the reliability of aspects of review methodology expands as well. This dissertation has focused on several aspects of systematic reviewing that can affect the reliability of conclusions as the evidence shifts in a variety ways.

Continuing growth of evidence

In the time since this doctoral work began, Cochrane's CENTRAL register of trials added over 800,000 records for new or completed trials, published and not. That quantity far exceeded our capacity to digest it, even supported by current systematic review practices. The resulting challenges described at the outset of this project in Chapter 2 remain, (2) and they are becoming steadily more acute as the numbers of clinical trials and systematic reviews increase. The updated data from a variety of sources in Chapter 2 indicate that the number of new studies of both types appearing daily more than quadrupled between 2007 and 2017.

An analysis by Ioannidis in 2016 concluded that the rate of increase of systematic reviews was much higher than that estimated here. (3) However, his analysis relied on using only PubMed's Clinical Queries filter to estimate the prevalence of systematic reviews, (4) without accounting for the non-systematic reviews inflating the results. That method also incorporates systematic reviews of a broad range of study types that may only rarely have an impact on healthcare decision-making, such as genetic and animal studies.

Ioannidis also argued that the majority of systematic reviews are redundant, of poor quality, or otherwise unfit for healthcare decision-making. His consideration of the prevalence of redundant reviews did not account for systematic reviews being

ephemeral. Reviews on the same question may only be redundant when they are conducted at the same time with the same scope, and are in excess of those that can be helpful replications or challenges to review authors' methods, results, or interpretations. Studies that could estimate a rate of redundancy have not assessed large enough samples of systematic reviews likely to be representative. (5,6) Further research on the prevalence of systematic reviews would benefit from a robust estimate of what proportion represents true redundancy.

Clinical trials may often fail to address questions of vital importance to patients, clinicians, and other decision-makers. (7) Identifying critical research gaps is a fundamental role of systematic reviews, even though, unfortunately, using systematic reviews to determine clinical trial priorities may still be uncommon. (8,9) Systematic reviews can deepen the fragmented picture from clinical trials when they contain substantial gaps, missing trials or not including important treatment comparisons from within trials. (10)

In 1987, Mulrow analysed the shortcomings of non-systematic reviews of the literature. (11) Since then, as we show in Figure 4, the growth of non-systematic reviews has continued to vastly outstrip that of systematic reviews. In recent years, there has also been a proliferation of variants of systematic reviews, such as rapid reviews, network meta-analyses, living systematic reviews, and umbrella reviews or overviews. (Chapter 1) Further research on the prevalence of systematic reviews should consider the contribution and respective growth of major types of reviews, especially rapid reviews. The impact of systematic reviews on non-systematic reviews could also be explored. These presumably now cite and rely on systematic reviews to some extent, and their conclusions may be converging.

Systematic reviews and new evidence

A major strategy for keeping up with the evidence from clinical trials was the development of the international network and infrastructure of the Cochrane Collaboration. It began at a time when systematic reviewing was uncommon, and these studies required advocacy to gain support for undertaking and using them. Between 1996 and 2003, about one in five systematic reviews in the Database of Abstracts of Reviews of Effects (DARE) were Cochrane reviews. (12) However, while the growth of systematic reviewing overall has been dramatic, chapter 3 shows that the number of new Cochrane reviews has plateaued, and Cochrane reviews may now be more likely to out of date than not. (13) The most current systematic review on a question is probably not a Cochrane one. The situation for other organisations with systematic review portfolios is likely to be similar. A 2010 survey by Garritty and co-authors found that over half of the 103 organisations that responded believed the majority of their systematic reviews were out of date. (14)

Although they are now a small minority of systematic reviews, Cochrane reviews constitute a large corpus of studies that had the aspiration of keeping up with clinical trials. The original conception of Cochrane reviews had been the model now called living systematic reviews, albeit without the help of trials in digital formats and technologies such as machine-assisted searching. (15) The reality of attempting to sustain this goal, however, led to a succession of downgraded expectations for updating Cochrane reviews, first to at least annually, then to biennially, and now to not necessarily updated at all (Chapter 3). It is too early in this development to know to what extent this new generation of living systematic reviews will be adopted, and whether they will prove to be reliable and sustainable.

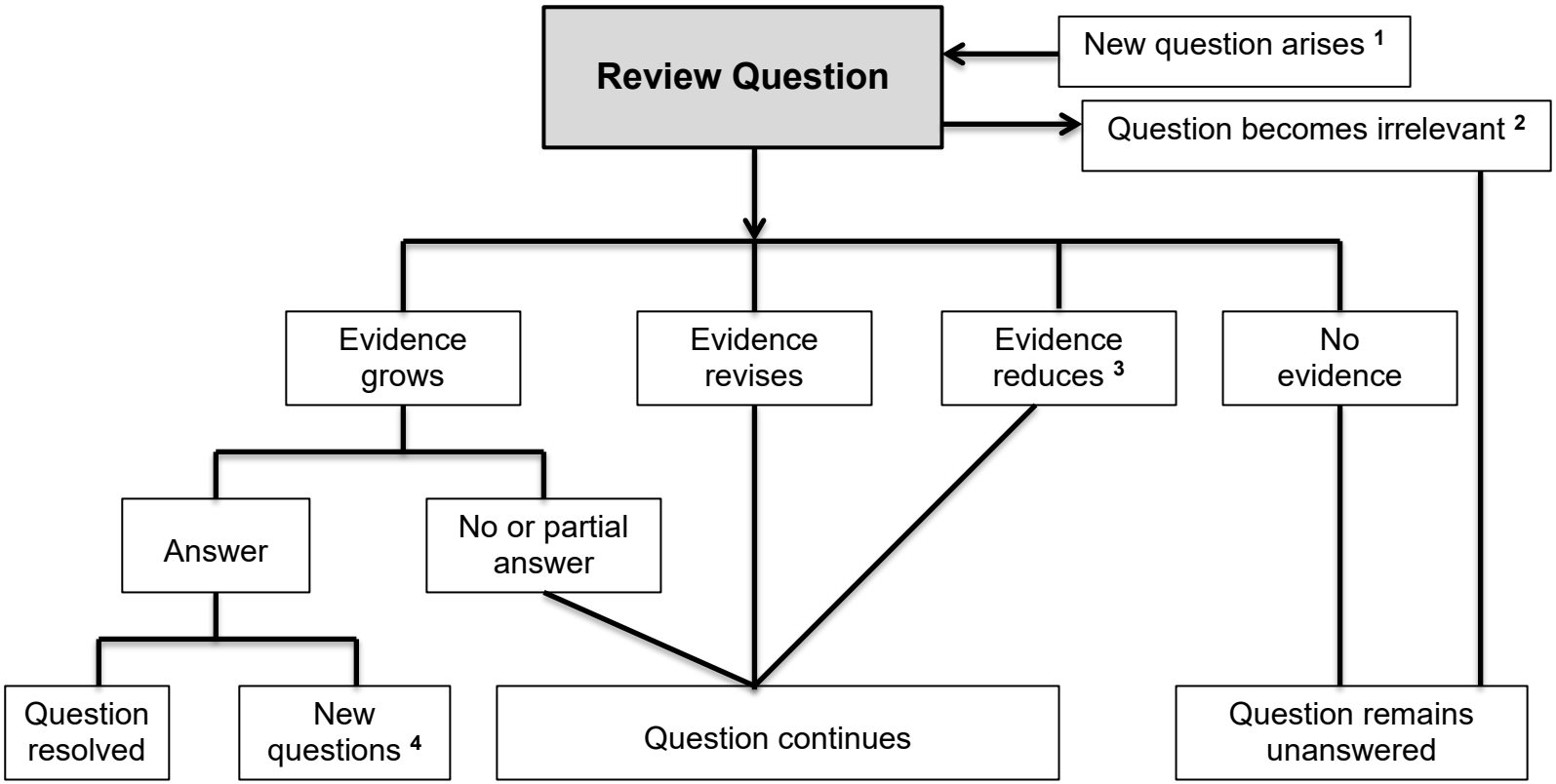
We found that over time, relatively few Cochrane reviews were regularly updated, and the intervals between updates lengthened over time. (13) Although the median number of included studies increased in our cohort, a major change in conclusion as a result of new evidence may have been uncommon. Our results are consistent with Cochrane groups having moved to concentrating updating effort on questions where new trials were more likely.

Concentrating updating effort in priority areas was one of the strategies identified in chapter 2 as critical for coping with the rapid growth of evidence. (2) It is not clear, however, how effective prioritisation of reviews for updating has been. One option for exploring this would be assessing subsequent non-Cochrane systematic reviews on the questions for which Cochrane reviews were not updated to determine how often new studies had shifted the body of evidence.

A striking feature of the models developed for streamlining the work of updating systematic reviews has been their individualistic and technical nature, centring on efficiency of search strategies for updates, for example. Yet, as the Garritty survey showed, a large number of organisations internationally are monitoring evidence for systematic review updates. (14) Update searches and assessments on whether there is any evidence that justifies updating a systematic review represent key research that should be publicly reported. The practice of not reporting this work unless a review proceeds to update is contributing to substantial research waste, as the same futile searches will be repeated many times globally.

We also found that the dates of searches were not reported in 7% of updated Cochrane reviews; (13) Beller and colleagues found these dates were not reported in 10% of a sample of systematic reviews identified via PubMed. (16) Developing consensus on update reporting standards through the PRISMA reporting standards for systematic reviews might encourage major groups of systematic reviewers to resolve these problems.

Assessing the impact of new studies has been the principle form of evidence shift focused on for the practice and study of systematic reviewing. This has been based on a rather linear approach to evidence: as though evidence moves only forward until a question is answered. However, the analyses in Chapters 3 to 6 show that evidence often shifts in other ways that can have just as significant an impact on a systematic review as new evidence. Figure 21 sketches out the contours of shifting evidence in relation to research questions addressed by systematic reviews.



[1] E.g. new adverse effect or subgroup question; [2] E.g. intervention superseded or withdrawn;
[3] E.g. retracted or inclusion decision changes; [4] E.g. optimum dose or intervention combination.

Figure 21. Schema of review questions and potential shifts in evidence.

As evidence grows, converges, or shifts in direction, reviews can be refocused, merged, or split to better grapple with the related questions. Sometimes, the review's question becomes irrelevant, as for example when a drug is withdrawn from market, or an intervention falls out of use after the adoption of a new practice.

Even when enough evidence accumulates to answer a question about the effectiveness of an intervention, the need for a systematic review can continue, as the success or failure of an interventions generates new questions: Does the intervention help some people more than others? What is the optimum use of the intervention (dose, length, etc)? Can health outcomes be optimised if this intervention is combined with others? Shifting evidence and questions may be the rule for systematic reviews.

In addressing the various scenarios shown in Figure 21, this work pointed to a critical methodological gap in determining when a question has been adequately answered. The inconsistencies we found in the application of the "stable" category to Cochrane reviews highlights the need for reliable analytic methods to support decision-making about the conclusiveness of evidence. (17)

The potential for centralised infrastructure

In the early stages of this work, we argued that any proposals for adding to the workload of systematic reviewing needed to consider whether slowing down the process and reducing was justified. (2) In proposing research integrity checks for the content of systematic reviews, I attempted to address this issue by developing tools to make it easier for systematic reviewers to meet this demand. However, another critical strategy is for improved centralised infrastructure that can mitigate the workload for individuals.

I have focused particularly on parts of the information management ecosystem in which I have been heavily involved over the years, the information infrastructure at PubMed and Cochrane. Both of these resources could add investment to enable

timely identification and notification/indexing of key research integrity issues in studies, reducing the time required by individual systematic reviewers. In particular, the U.S. National Library of Medicine, the producer of PubMed, could resume the work undertaken by our team, filtering and indexing all records – including "epublished ahead of print" records – for all key post-publication events relevant to research integrity. Records should also be made directly searchable, indexing synchronised between PubMed and PMC, and a feed of data on research integrity easily usable by individuals and services such as Cochrane's trial registry and other databases and reference management systems, as well as enabling checking of lists of publications. Some problems are best fixed globally, taking advantage of important economies of scale. Detailed audit of other databases and platforms similar to that undertaken here for PubMed and PMC could usefully identify barriers and problems related to the visibility and ascertainment of post-publication events.

Another theme throughout this work has been the need for meta-research to identify and solve methodological problems. That issue, too, has a critical infrastructure component. The ending of the Cochrane Methodology Register (18) and its successor, the Scientific Resource Center's Methods Library, (19) because of the relatively modest resource investment they require is emblematic of several problems. One of those is that well-resourced groups in the evidence synthesis arena are under-estimating how vital it is that we have science on how best to do science. The second is that we are not learning the lessons of what happens if research keeps piling up without an adequately functioning system for us to be able to find, use, and synthesise it. We are repeating this pattern with meta-research.

Reducing the perpetuation of error by systematic reviews

The final major theme that emerged during this work was unexpected: the need for better systems for correcting the scientific record, and effective research integrity filters. This is an important area of development for systematic reviewing, but also for almost anything that relies on cited research or lists of published studies, including grant applications. Although the work in chapters 5 and 6 make a contribution here, we still know very little about many critical issues, notably errata and the data content

of authors' replies to letters to the editor, and these require further research. A better understanding of their prevalence and implications for systematic reviews is essential to know whether the effort in finding and evaluating these potential sources of information are worthwhile.

The work here has focused on publications about trials. Systematic reviews may increasingly use data from other sources, as well, such as trial register entries, clinical study reports, and author replies to requests for data. Some systematic reviews use incorporate conference abstracts. While trial register entries can be updated, other forms of trial report may not be corrected. The impact of error in non-journal forms of trial reports does not appear to have been assessed. In addition, more research is needed about de novo errors in systematic reviews. (20,21) Considerable effort goes into assessing methodological quality of systematic reviews. However, as systematic reviews build on previous reviews, we need to know if there are other ways we can detect if they are unreliable.

Using compromised research does not inevitably cause harm and waste, but it can. The additional effort required to assess the ongoing validity of data included in systematic reviews may be substantial, and the impact on the efficiency of systematic reviews needs to be assessed. The situation is analogous to the issue of updating: adding new studies may only rarely have a major impact on a systematic review's results, yet it is not always easy to predict when the effort of updating will justify the effort.

Systematic reviews can be unreliable because they were poor quality from the outset. They can begin with error, such as missing critical studies, but they can accrue it, and perpetuate it too. As chapter 6 showed, shifts in the evidence a review has already synthesised has the potential to affect reliability, and that might happen more often than we realise. This could perpetuate error because so many decision makers and information producers rely on systematic reviews. Compromised information is carried forward through clinical practice guidelines and health information for the public. Other systematic reviews may pick up where a previous systematic review left off, taking whatever compromised results it may have at face

value. A compromised systematic review could inform further research in erroneous ways as well. A clinical trial might be clearly corrected or retracted, but it may not be noticed because of reliance on uncorrected systematic reviews.

Chapter 6 demonstrates that the methods advised by the *Cochrane Handbook* (22) for dealing with post-publication events affecting included studies are inadequate, although they stand out in the systematic review community for having a recommendation on a systematic approach at all. The under-appreciation and low profile of these issues is also highlighted by the striking differences in how trials retracted for ethical issues are being handled by systematic reviewers. (23)

The invisibility of published error correction is part of a larger problem of a view of research that sees dissemination of reports as the end of a process, instead of the beginning of a new one. Scientific communities need to value what Rennie (24) called in the 1990s the “aftercare” of research, as well as having a stronger post-publication peer review culture. (25) Research funders need to value it, too, and support the time it takes, for trialists with a responsibility to pursue correction of the record through to systematic reviewers and other researchers. Journals need to value it as well, (26,27) or at least not pull us backwards, as it seems may be happening with practices around e-publication ahead of print. The prevalence of errata does not indicate the extent of error, which remains largely unaddressed. (28) Thinking and talking about research is part of the science community’s vibrant culture. Capturing that post-publication intellectual effort more rigorously is essential for better science. (25)

References

1. Eriksen TH. Tyranny of the moment: fast and slow time in the information age. London: Pluto Press; 2001.

2. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? PLOS Med. 2010 Sep 21;7(9):e1000326.

3. Ioannidis JPA. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. Milbank Q. 2016;94(3):485–514.

4. Search Strategy Used to Create the PubMed Systematic Reviews Filter [Internet]. [cited 2020 Apr 22]. Available from: https://www.nlm.nih.gov/bsd/pubmed_subsets/sysreviews_strategy.html

5. Siontis KC, Hernandez-Boussard T, Ioannidis JPA. Overlapping meta-analyses on the same topic: survey of published studies. BMJ. 2013 Jul 19;347:f4501.

6. Khamis AM, Moheb ME, Nicolas J, Iskandarani G, Refaat MM, Akl EA. Several reasons explained the variation in the results of 22 meta-analyses addressing the same question. J Clin Epidemiol. 2019 Sep 1;113:147–58.

7. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet Lond Engl. 2009 Jul 4;374(9683):86–9.

8. Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. J R Soc Med. 2007 Apr;100(4):187–90.

9. Robinson KA, Goodman SN. A systematic examination of the citation of prior research in reports of randomized, controlled trials. Ann Intern Med. 2011 Jan 4;154(1):50–5.

10. Créquit P, Trinquart L, Yavchitz A, Ravaud P. Wasted research when systematic reviews fail to provide a complete and up-to-date evidence synthesis: the example of lung cancer. BMC Med. 2016 Jan 20;(14):8.

11. Mulrow CD. The medical review article: state of the science. Ann Intern Med. 1987 Mar;106(3):485–8.

12. Bastian H, Glasziou P. Four systematic reviews per day: the epidemiology of research synthesis. In Melbourne: Cochrane Collaboration; 2005 [cited 2020 Apr 25]. Available from: /2005-melbourne/four-systematic-reviews-day-epidemiology-research-synthesis

13. Bastian H, Doust J, Clarke M, Glasziou P. The epidemiology of systematic review updates: a longitudinal study of updating of Cochrane reviews, 2003 to 2018. | medRxiv. medRxiv [Internet]. 2019 Dec 11 [cited 2020 Apr 24]; Available from: https://www.medrxiv.org/content/10.1101/19014134v1.full

14. Garritty C, Tsertsvadze A, Tricco AC, Sampson M, Moher D. Updating systematic reviews: an international survey. PloS One. 2010 Apr 1;5(4):e9914.

15. Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, et al. Living systematic review: 1. Introduction—the why, what, when, and how. J Clin Epidemiol. 2017 Nov 1;91:23–30.

16. Beller EM, Chen JK-H, Wang UL-H, Glasziou PP. Are systematic reviews up-to-date at the time of publication? Syst Rev. 2013 May 28;2(1):36.

17. Bastian H, Hemkens LG. Enough evidence and other endings: a descriptive study of stable Cochrane systematic reviews in 2019. | medRxiv. medRxiv [Internet]. 2019 Dec 9 [cited 2020 Apr 28]; Available from: https://www.medrxiv.org/content/10.1101/19013912v1

18. Cochrane Collaboration. Cochrane Methodology Register [Internet]. 2019 [cited 2019 Oct 29]. Available from: https://cmr.cochrane.org/

19. Scientific Resource Center. SRC Methods Library [Internet]. [cited 2019 Nov 28]. Available from: https://www.zotero.org/groups/355490/src_methods_library

20. Ford AC, Guyatt GH, Talley NJ, Moayyedi P. Errors in the conduct of systematic reviews of pharmacological interventions for irritable bowel syndrome. Am J Gastroenterol. 2010 Feb;105(2):280–8.

21. Saldanha I. Evaluating Data Abstraction Assistant, a novel software application for data abstraction during systematic reviews: protocol for a randomized controlled trial. [cited 2019 Nov 25]; Available from: https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-016-0373-7

22. Higgins JPT, Thomas J. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 6.0. London: Cochrane Collaboration; 2019 [cited 2019 Oct 2]. Available from: https://training.cochrane.org/handbook/current

23. Elia N, Elm E von, Chatagner A, Pöpping DM, Tramèr MR. How do authors of systematic reviews deal with research malpractice and misconduct in original studies? A cross-sectional analysis of systematic reviews and survey of their authors. BMJ Open [Internet]. 2016 Mar 1 [cited 2019 Nov 25];6(3). Available from: https://bmjopen.bmj.com/content/6/3/e010442

24. Rennie D. Freedom and Responsibility in Medical Publication: Setting the Balance Right. JAMA. 1998 Jul 15;280(3):300–2.

25. Bastian H. A Stronger Post-Publication Culture Is Needed for Better Science. PLOS Med. 2014 Dec 30;11(12):e1001772.

26. Altman DG. Poor-quality medical research: what can journals do? JAMA. 2002 Jun 5;287(21):2765–7.

27. Altman DG. Unjustified restrictions on letters to the editor. PLoS Med. 2005 May;2(5):e126; discussion e152.

28. Costantino G, Casazza G, Cernuschi G, Solbiati M, Birocchi S, Ceriani E, et al. Errors in medical literature: not a question of impact. Intern Emerg Med. 2013 Mar 1;8(2):157–60.

Appendix 1:

Methods for 2019 update on the growth of trials and systematic reviews in health.

Non-systematic reviews

We used the review publication type in PubMed: review[pt]

Case reports

We used the case reports publication type in PubMed: case reports[pt]
The definition used for this subject/publication type were narrowed in 1967 and then expanded in 1976, (1) leading to an anomalous dip in the apparent number of case reports.

Systematic reviews

1. The annual count of Cochrane reviews to the end of 2008 was provided by Claire Allen from the Cochrane Collaboration secretariat. The annual count of Cochrane reviews from 2014 to 2017 came from the Cochrane website. (2) We could identify no official reports of numbers of reviews between 2009 and 2013, and dates of first publication could not be searched in the *CDSR*. For those years, PubMed records for the *CDSR* were retrieved, and R was used to find records with no ".pub" or with ".pub2" in the DOI: protocols were excluded. This method was also used for 2014 to 2017 to measure results against the official Cochrane numbers for those years: it resulted in an over-estimation of the number of Cochrane reviews (from 420 versus 406 for 2017, to 513 to 417 in 2016).

2. In the 2010 paper, the count on INAHTA (International Network of Agencies for Health Technology Assessment) came from the INAHTA database at the Centre for Reviews and Dissemination. However, this was no longer updated and was not used for the 2019 update. Epistemonikos was added. (3) Search

results were limited by “systematic review” type. As there was a discrepancy between hits reported and numbers on download, all results were downloaded and the total downloaded records used. As this exceeded download limits, the first search on year of publication was divided into small enough chunks to download by segmenting according to date added to Epistemonikos.

3. We used two counts for MEDLINE. First we used the meta-analysis [publication type] limited to Humans. But this is known to miss many systematic reviews. So we also used one of the Montori filters (4): Medline[tiab] OR (systematic[tiab] AND review[tiab]) OR meta-analysis[ptyp], again limited to Humans. This filter has been assessed to have a sensitivity of 71% and precision of 57%. Given the 57% precision, we then multiplied the results by 0.57. Given the sensitivity of 71% this will underestimate the total number of systematic reviews. As we were interested in therapeutic reviews which are around 70% of all reviews (5), we kept this estimate as a reasonable one for therapeutic reviews.

Trials

For trials we used three sources.

1. We used all trials in the Cochrane Central Register of Controlled Trials.

2. For MEDLINE we used Pubmed and 3 publication types: ("Controlled Clinical Trial " OR "Clinical Trial, Phase III " OR "Randomized Controlled Trial") limited to humans

3. For MEDLINE, we also used the "narrow" version of the Clinical Queries filter for trials on PubMed.

To estimate the proportion of unpublished in trial registries, we searched for all interventional studies on clinicaltrials.gov (254,135 hits), then restricted the search to completed studies (139,811 hits) (55.0%).

The original searches were done on 18 April 2009. The update searches were done in the last week of October 2019. The search of clinicaltrials.gov was done on 18 November 2019.

References

1. Charen T. 1976 indexing orientation. NLM Tech Bull. 1975 Oct;(78):5.

2. Cochrane Collaboration. Dashboard [Internet]. Cochrane Community. [cited 2019 Oct 30]. Available from: https://community.cochrane.org/organizational-info/resources/strategy-2020/dashboard

3. Rada G, Pérez D, Capurro D. Epistemonikos: a free, relational, collaborative, multilingual database of health evidence. Stud Health Technol Inform. 2013;192:486–90.

4. Boluyt N, Tjosvold L, Lefebvre C, Klassen TP, Offringa M. Usefulness of Systematic Review Search Strategies in Finding Child Health Systematic Reviews in MEDLINE. Arch Pediatr Adolesc Med. 2008 Feb 1;162(2):111–6.

5. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and Reporting Characteristics of Systematic Reviews. PLOS Med. 2007 Mar 27;4(3):e78.

Appendix 2.

Placing the results of the longitudinal study into context with other studies of Cochrane review updating.

One author (HB) searched for studies of Cochrane review updating in the Cochrane Methodology Register (which ended in May 2012), (1) as well as the author's collection of studies on updating systematic reviews. Reference lists of relevant papers were searched for further studies. There was no restriction on language or type of document, as long as it included data on updating, changes in conclusion, or analysis of reporting quality of updates in a cohort of Cochrane reviews.

Data were collected on the number of included Cochrane reviews, the time coverage of the cohort, and whether there was a topic restriction for the reviews. Each was reviewed for data on measures of currency of reviews, frequency of updating, frequency of changes of conclusion, and quality of reporting on updates, and relevant findings were summarised in the overview that follows.

There were 13 studies with relevant data. There were another two studies that measured different outcomes, (2,3) and one study (4) referred to two studies in conference abstracts to which we did not have access.

Overview of studies of updating in Cochrane reviews.

95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	Study	Findings
36 new 2/95																								1998 (5)	50% updated within 2 years
1268 from all in 1/02																								2002 (6)	95% amended, not clear if updated
1571 from all in 1/03																								2003 (7)	46.5% date of search within last 2 years
147 from all in 2/03																								2004 (8)	Total updated: 48%
377 from all in 2/98																								2005 (4)	70% updated, 9% changed conclusion
86 reviews & protocols from all in 4/06																								2006 (9)	60% updated within 2 years; 26% retracted
2607 from all in 1/06																								2006 (10)	43% updated in last 2 years
53 updated in 4/02																								2006 (11)	Mean time to update 2.7 years
					301 with protocols in 2/00-1/01																			2008 (12)	A third updated; 0.7% every 2 years
313 from all in 3/07																								2010 (13)	32.7% in 2 yrs, median 3.3 yrs to 1st; 2% major conclusion change
					623 new in 2000-2005																			2011 (14)	45% in last 2 yrs; median to update 1.8 yrs
5418 all to end of 2012																								2013 (15)	~20% updated every 2 years
															682 new 2010, updated by 2017									2018 (16)	43% update rate; 3.9% of subset changed conclusion
177 updated in 2003																								2019 *Cohorts in this paper*	Median 3 yrs update, 7 since last; 2.8% major conclusion change
1532 from all in 4/02																									11.3% updated

Green = no subject restriction, Orange = subset of Cochrane reviews in a subject area.

Until 2012, Cochrane reviews were published in 4 issues per year. Issues are referred to by issue/year: e.g. 2/03 is Issue 2 of 2003. First 23 columns are each year of publication of *CDSR*, from issue 1 in 1995 to end of December 2018.

Length of the bar is the time from first included Cochrane reviews. If a study reported it was from the *CDSR*'s inception, the bar begins in 1995, whether or not reviews from 1995 were in the cohort.

References

1. Cochrane Collaboration. Cochrane Methodology Register [Internet]. 2019 [cited 2019 Oct 29]. Available from: https://cmr.cochrane.org/

2. Higgins JP. How Should We Interpret Updated Meta-analyses? In Rome: Cochrane Collaboration; 1999 [cited 2019 Nov 13]. Available from: /1999-rome/how-should-we-interpret-updated-meta-analyses

3. Eshun-Wilson I, Jaffer S, Smith R, Johnson S, Hine P, Mateo A, et al. Maintaining relevance in HIV systematic reviews: an evaluation of Cochrane reviews. Syst Rev. 2019 07;8(1):46.

4. French SD, McDonald S, McKenzie JE, Green SE. Investing in updating: how do conclusions change when Cochrane systematic reviews are updated? BMC Med Res Methodol. 2005 Oct 14;5:33.

5. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al. Methodology and Reports of Systematic Reviews and Meta-analyses: A Comparison of Cochrane Reviews With Articles Published in Paper-Based Journals. JAMA. 1998 Jul 15;280(3):278–80.

6. Koch G. Are Cochrane Reviews Updated Regularly or Not? In Stavanger: Cochrane Collaboration; 2002. Available from: https://cmr.cochrane.org/?CRGReportID=4572

7. Koch G. Less than half of the Cochrane Reviews are current. In Barcelona: Cochrane Collaboration; 2003. Available from: https://cmr.cochrane.org/?CRGReportID=5113

8. González de Dios J. [Bibliometric analysis of systematic reviews in the Neonatal Cochrane Collaboration. Its role in evidence-based decision making in neonatology]. An Pediatr Barc Spain 2003. 2004 May;60(5):417–27.

9. Henderson S, Hampson L, Atherton D, Neilson J. Ten years of updating Cochrane reviews: the experiences of the Pregnancy and Childbirth Group. In Dublin: Cochrane Collaboration; 2006 [cited 2019 Jul 29]. Available from: https://cmr.cochrane.org/?CRGReportID=9981

10. Koch G. No improvement - still less than half of the Cochrane reviews are up to date. In Dublin: Cochrane Collaboration; 2006. Available from: https://cmr.cochrane.org/?CRGReportID=10066

11. Shea B, Boers M, Grimshaw JM, Hamel C, Bouter LM. Does updating improve the methodological and reporting quality of systematic reviews? BMC Med Res Methodol. 2006 Jun 13;6:27.

12. Tricco AC, Brehaut J, Chen MH, Moher D. Following 411 Cochrane protocols to completion: a retrospective cohort study. PloS One. 2008;3(11):e3684.

13. Jaidee W, Moher D, Laopaiboon M. Time to update and quantitative changes in the results of cochrane pregnancy and childbirth reviews. PloS One. 2010 Jul 13;5(7):e11553.

14. Ngamjarus C, Jaidee W, Laopaiboon M. Progress in time to update of Cochrane Systematic Reviews. In: Cochrane Database of Systematic Reviews [Internet]. Cochrane Collaboration; 2011. p. 159–60. Available from: https://cmr.cochrane.org/?CRGReportID=16807

15. Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. BMJ. 2013 Dec 13;347:f7191.

16. Bashir R, Surian D, Dunn AG. Time-to-update of systematic reviews relative to the availability of new evidence. Syst Rev. 2018 17;7(1):195.

Appendix 3.

Search strategies for retractions and retrievals.

A. Search strategies for databases.

1. PubMed.

hasretractionof
hasretractionin
retraction of publication[pt]
retracted publication[pt]
(retracted[ti] OR retraction*[ti] OR retracted[tt] OR retraction*[tt])
(erratum*[ti] OR correction[ti] OR corrigendum*[ti] OR errata*[ti] OR corrections[ti] OR corrigenda*[ti] OR erratum*[tt] OR correction[tt] OR corrigendum*[tt] OR errata*[tt] OR corrections[tt] OR corrigenda*[tt] OR published erratum[pt] or haserratumfor) AND (retract* OR withdraw*) AND (article[tiab] OR articles[tiab] OR e-publication[tiab] OR e-publications[tiab] OR epub*[tiab] OR "ahead of print"[tiab] OR manuscript[tiab] OR manuscripts[tiab] OR publication[tiab] OR publications[tiab])
http://www.elsevier.com/locate/withdrawalpolicy OR (Elsevier* AND policy* AND article* AND (remov* OR withdraw*))
(withdrawn[ti] OR withdrawn[tt])
(withdrawn[tiab] AND (article[tiab] OR articles[tiab] OR e-publication[tiab] OR e-publications[tiab] OR epub*[tiab] OR "ahead of print"[tiab] OR manuscript[tiab] OR manuscripts[tiab] OR publication[tiab] OR publications[tiab]))
(temporary*[ti] AND removal*[ti]) OR (temporary*[tt] AND removal*[tt])
retrait[tiab] AND article[tiab]
retrait[ti] OR retrait[tt]
(retirado[tiab] OR retraído[tiab]) AND (artículo[tiab] OR artigo[tiab])
retirado[ti] OR retraído[ti] OR retirado[tt] OR retraído[tt]
pubstatusaheadofprint AND ("0001/01/01"[PDAT] : "2012/12/31"[PDAT])

2. PubMed Central (PMC).

retraction[filter]
is retracted[filter]
withdrawn[ti]
retracted[ti] OR retraction*[ti]
(erratum*[ti] OR correction[ti] OR corrigendum*[ti] OR errata*[ti] OR corrections[ti] OR corrigenda*[ti] OR correction[filter]) AND (article OR articles) AND (retract* OR withdraw OR withdrawn OR withdraws OR withdrawal OR withdrawing)
http://www.elsevier.com/locate/withdrawalpolicy OR "Elsevier Policy"

3. Web of Science.

(PY=1980-2017) AND (DT= Retracted Publication)
(PY=1980-2017) AND (DT=Retraction)

4. Retraction Watch.

All records in http://retractiondatabase.org/RetractionSearch.aspx to 14 August 2017.

Additional subsequent records provided by Retraction Watch on 3 January 2018.

5. CrossRef.

Retractions per API.

6. PubMed Commons.

Active comments: ‘withdrawn’ or ‘retract’

B. Reference searching

Fanelli D. Why growing retractions are (mostly) a good sign. PLoS medicine. 2013;10(12):e1001563. doi: 10.1371/journal.pmed.1001563. PubMed PMID: 24311988; PubMed Central PMCID: PMC3848921.
Fanelli D, Costas R, Lariviere V. Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. PloS one. 2015;10(6):e0127556. doi: 10.1371/journal.pone.0127556. PubMed PMID: 26083381; PubMed Central PMCID: PMC4471332.
Fang FC, Casadevall A. Retracted science and the retraction index. Infection and immunity. 2011;79(10):3855-9. doi: 10.1128/iai.05661-11. PubMed PMID: 21825063; PubMed Central PMCID: PMC3187237.
Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(42):17028-33. doi: 10.1073/pnas.1212247109. PubMed PMID: 23027971; PubMed Central PMCID: PMC3479492.
Foo JY. A retrospective analysis of the trend of retracted publications in the field of biomedical and life sciences. Science and engineering ethics. 2011;17(3):459-68. doi: 10.1007/s11948-010-9212-8. PubMed PMID: 20517712.
Furman JL, Jensen K, Murray F. Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. Research Policy. 2012;41(2):276-90. doi: 10.1016/j.respol.2011.11.001.
Gabehart ME. An analysis of citations to retracted articles in the scientific literature. [Master's Paper]. Chapel Hill, NC: University of North Carolina at Chapel Hill; 2005.
Gasparyan AY, Ayvazyan L, Akazhanov NA, Kitas GD. Self-correction in biomedical publications and the scientific impact. Croatian medical journal. 2014;55(1):61-72. doi: 10.3325/cmj.2014.55.61. PubMed PMID: 24577829; PubMed Central PMCID: PMC3944419.
Grieneisen ML, Zhang M. A comprehensive survey of retracted articles from the scholarly literature. PloS one. 2012;7(10):e44118. doi: 10.1371/journal.pone.0044118. PubMed PMID: 23115617; PubMed Central PMCID: PMC3480361.
He T. Retraction of global scientific publications from 2001 to 2010. Scientometrics. 2013;96(2):555-61. doi: 10.1007/s11192-012-0906-3.
Jalalzadeh S, Abbasi F, Pashazadeh F. 212: SCIENTOMETRICS STUDY OF PUBLISHED RETRACTIONS IN PUBMED. BMJ open. 2017;7(Suppl 1). doi: 10.1136/bmjopen-2016-015415.212.
Jin GZ, Jones B, Lu SF, Uzzi B. The Reverse Matthew Effect: Catastrophe and Consequence in Scientific Teams. National Bureau of Economic Research Working Paper Series. 2013;No. 19489. doi: 10.3386/w19489.
Karabag SF, Berggren C. Misconduct, Marginality and Editorial Practices in Management, Business and Economics Journals. PloS one. 2016;11(7):e0159492. doi: 10.1371/journal.pone.0159492. PubMed PMID: 27454761; PubMed Central PMCID: PMC4959770.
Lu SF, Jin GZ, Uzzi B, Jones B. The retraction penalty: evidence from the Web of Science. Scientific reports. 2013;3:3146. doi: 10.1038/srep03146. PubMed PMID: 24192909; PubMed Central PMCID: PMC3818648.
Madlock-Brown CR, Eichmann D. The (lack of) impact of retraction on citation networks. Science and engineering ethics. 2015;21(1):127-37. doi: 10.1007/s11948-014-9532-1. PubMed PMID: 24668038.
Mongeon P, Larivière V. The Collective Consequences of Scientific Fraud: An Analysis of Biomedical Research. In: Gorraiz J, Schiebel E, Gumpenberger C, Hörlesberger M, Moed H, editors. 14th International Society of Scientometrics and Informetrics

Conference; 15-19 July; Vienna, Austria2013. p. 1897-9.

Mongeon P, Larivière V. Costly collaborations: The impact of scientific fraud on co-authors' careers. Journal of the Association for Information Science and Technology. 2016;67(3):535-42. doi: 10.1002/asi.23421.

Nath SB, Marcus SC, Druss BG. Retractions in the research literature: misconduct or mistakes? The Medical journal of Australia. 2006;185(3):152-4. PubMed PMID: 16893357.

Neale AV, Northrup J, Dailey R, Marks E, Abrams J. Correction and use of biomedical literature affected by scientific misconduct. Science and engineering ethics. 2007;13(1):5-24. doi: 10.1007/s11948-006-0003-1. PubMed PMID: 17703606; PubMed Central PMCID: PMC2796981.

Parrish DM. Scientific misconduct and correcting the scientific literature. Academic medicine : journal of the Association of American Medical Colleges. 1999;74(3):221-30. PubMed PMID: 10099640.

Rai R, Sabharwal S. Retracted Publications in Orthopaedics: Prevalence, Characteristics, and Trends. The Journal of bone and joint surgery American volume. 2017;99(9):e44. doi: 10.2106/jbjs.16.01116. PubMed PMID: 28463926.

Redman BK, Yarandi HN, Merz JF. Empirical developments in retraction. Journal of medical ethics. 2008;34(11):807-9. doi: 10.1136/jme.2007.023069. PubMed PMID: 18974415.

Resnik DB, Dinse GE. Scientific retractions and corrections related to misconduct findings. Journal of medical ethics. 2013;39(1):46-50. doi: 10.1136/medethics-2012-100766. PubMed PMID: 22942373; PubMed Central PMCID: PMC3525741.

Riegelman A, Bakker C. The Complexities of Retractions: Data Integrity, Discovery, and Tools. Research Data Access and Preservation Summit; 19-20 Apr; Seattle, WA. 2017.

Rosenkrantz AB. Retracted Publications Within Radiology Journals. AJR American journal of roentgenology. 2016;206(2):231-5. doi: 10.2214/ajr.15.15163. PubMed PMID: 26797347.

Samp JC, Schumock GT, Pickard AS. Retracted publications in the drug literature. Pharmacotherapy. 2012;32(7):586-95. doi: 10.1002/j.1875-9114.2012.01100.x. PubMed PMID: 22581659.

Steen RG. Retractions in the medical literature: how many patients are put at risk by flawed research? Journal of medical ethics. 2011;37(11):688-92. doi: 10.1136/jme.2011.043133. PubMed PMID: 21586404.

Steen RG. Retractions in the scientific literature: is the incidence of research fraud increasing? Journal of medical ethics. 2011;37(4):249-53. doi: 10.1136/jme.2010.040923. PubMed PMID: 21186208.

Steen RG. Retractions in the scientific literature: do authors deliberately commit research fraud? Journal of medical ethics. 2011;37(2):113-7. doi: 10.1136/jme.2010.038125. PubMed PMID: 21081306.

Steen RG. Retractions in the medical literature: how can patients be protected from risk? Journal of medical ethics. 2012;38(4):228-32. doi: 10.1136/medethics-2011-100184. PubMed PMID: 22138727.

Stern AM, Casadevall A, Steen RG, Fang FC. Financial costs and personal consequences of research misconduct resulting in retracted publications. eLife. 2014;3:e02956. doi: 10.7554/eLife.02956. PubMed PMID: 25124673; PubMed Central PMCID: PMC4132287.

Stretton S, Bramich NJ, Keys JR, Monk JA, Ely JA, Haley C, et al. Publication misconduct and plagiarism retractions: a systematic, retrospective study. Current medical research and opinion. 2012;28(10):1575-83. doi: 10.1185/03007995.2012.728131. PubMed PMID: 22978774.

Tchao R. A Need to Archive Correct Biomedical Scientific Data and to Prevent

Continued Citation of Retracted Scientific Publications. International Journal of Humanities and Arts Computing. 2014;8(Suppl):29-37. doi: 10.3366/ijhac.2014.0097.

Trikalinos NA, Evangelou E, Ioannidis JP. Falsified papers in high-impact journals were slow to retract and indistinguishable from nonfraudulent papers. Journal of clinical epidemiology. 2008;61(5):464-70. doi: 10.1016/j.jclinepi.2007.11.019. PubMed PMID: 18394539; PubMed Central PMCID: PMC4699806.

Van Leeuwen T, Luwel M. An in-depth analysis of papers retracted in the Web of Science. In: Noyons E, editor. Proceedings of the 19th International Conference on Science and Technology Indicators; Leiden: Universiteit Leiden - CWTS; 2014. p. 337-44.

Wager E, Williams P. Why and how do journals retract articles? An analysis of Medline retractions 1988-2008. Journal of medical ethics. 2011;37(9):567-70. doi: 10.1136/jme.2010.040964. PubMed PMID: 21486985.

Wang J, Ku JC, Alotaibi NM, Rutka JT. Retraction of Neurosurgical Publications: A Systematic Review. World neurosurgery. 2017;103:809-14.e1. doi: 10.1016/j.wneu.2017.04.014. PubMed PMID: 28412480.

Woolley KL, Lew RA, Stretton S, Ely JA, Bramich NJ, Keys JR, et al. Lack of involvement of medical writers and the pharmaceutical industry in publications retracted for misconduct: a systematic, controlled, retrospective study. Current medical research and opinion. 2011;27(6):1175-82. doi: 10.1185/03007995.2011.573546. PubMed PMID: 21473670.

C. Publisher websites.

Bentham Science	eurekaselect.com
Cell Press	cell.com
SagePub	sagepub.com
ScienceDirect	sciencedirect.com
SpringerLink	link.springer.com
Taylor & Francis	tandfonline.com
Wiley Online Library	onlinelibrary.wiley.com

D. IEEE.

http://ieeexplore.ieee.org/Xplore/home.jsp

((Notice AND (withdraw* OR retract*)) OR withdrawn)

Limits: "subscribed content" & "Metadata only"

E. Other.

Tumor Biology	https://link.springer.com/article/10.1007/s13277-017-5487-6
Springer and BioMed Central	http://retractionwatch.com/wp-content/uploads/2016/10/BMC-article-list.pdf
ACOG (“College publications”)	http://journals.lww.com/greenjournal/pages/default.aspx

F. Retrieval totals.

PubMed and PMC	25,284
Web of Science	4638
CrossRef	1805
IEEE	10374
Publishers	3452
Retraction Watch	15,266
Reference searching	515
PubMed Commons	2
Other Sources	252
Totals	**61588**

Appendix 4.

Studies of retractions in the biomedical literature.

A. Search for studies for filter development.

Study criteria

- Included if data from a search for retractions/retraction notices reported.
- Excluded if no data from the 1990s or later.

Although the search strategy was designed for the biomedical literature, studies of other literature were not excluded.

Search dates

Initial search finalized on 8 September 2015. Further searches were undertaken in August 2017 and August 2019. As only studies found up to August 2017 were included in data collection, the 2017 and 2019 results are included in separate tables below. Searches and study selection done by Hilda Bastian.

Search strategy

1. Snowball search of references beginning with initial seed set of articles.
2. Included studies in 2016 review (1).
3. Citations of seminal study (2) in Google Scholar in August 2017.
4. PubMed search for studies from 2016 and 2017 on 21 August 2017 using: "retractions" OR "retracted publications" OR "retracted articles".
5. Additional studies identified during searches for retractions.
6. Update of PubMed search (#4 above) from 1 August 2017 on 5 August 2019.

7. Update of seminal study citations (#3 above) from 2017 to 10 August 2019.
8. PubMed search on 11 August 2019 using: retracted[ti] AND (publications[ti] OR articles[ti])

B. Studies and abstracted data.

Table 1. Included studies – original set used for reference searching and filter development.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
Abritis 2015 (3)	Scope: Office of Research Integrity (ORI) findings of research misconduct, articles related. Search: Not stated.	No.
Almeida 2016a (4)	Scope: SciELO and LILACS. Search: December 2014.	No.
Almeida 2016b (5)	Scope: Web of Knowledge. Search: November 2014.	Yes.
Amos 2014 (6)	Scope: PubMed. Strategy: PT Retracted publication Date of search: 27 Jan 2013.	Yes.
Azoulay 2012 (7)	Scope: PubMed. Strategy: Not stated. Date of search: Not stated.	Yes.
Balhara 2014 (8)	Scope: PubMed, MeSH heading (mental disorders only). Date of search: 15 September 2012.	No.
Balhara 2015 (9)	Scope: PubMed, MeSH heading (mental	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	disorders only). Date of search: Not stated.	
Bilbrey 2014 (10)	Scope: 15 journals, PubMed & Web of Knowledge. Strategy: Web of Knowledge – retraction in topic, PubMed corrected/retracted search limit Date: 2012.	Yes.
Bozzo 2017 (11)	Scope: MEDLINE, EMBASE, Cochrane Library (cancer only). Search: 23 August 2015.	No.
Budd 1998 (12) Additional reports (13,14)	Scope: MEDLINE (unspecified). Strategy: Publication Type “retraction of publication”. Citation impact study. Search: August 1997, restricted to end of 1996.	Yes.
Cicero 2014 (15)	Scope: Web of Science. Search: 7 January 2014.	Yes.
Claxton 2005 (16)	Scope: PubMed, secondly, retractions for 10 years in 11 journals. Strategy and date not reported.	No.
Cokol 2008 (17)	Scope: MEDLINE (unspecified). Strategy: Apparently as NLM-recorded. Search: 21 Oct 2007.	Yes.
Colaianni 1992 (18)	Scope: MEDLINE (unspecified). Strategy: NLM retraction notices. Search: date not reported.	No.
Damineni 2015 (19)	Scope: PubMed. Strategy: “Keyword search” – Retraction of articles	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	Retraction notice Withdrawal of article Search: January 2014.	
Davis 2012 (20)	Scope: PubMed. Strategy: “Retracted publication” in PT field Search: Between July & September 2011.	Yes.
Decullier 2013 (21)	Scope: MEDLINE (unspecified). Strategy: Publication type “retraction of publication” Search: August 2011, limited to 2008.	Yes.
Decullier 2014 (22)	As for Decullier 2013. Updates to May 2012.	Yes.
Elia 2014 (23)	Scope: 88 articles retracted due to ethical concerns (Boldt case). Search: January 2013.	No.
Fanelli 2013 (24) Additional report (25)	Scope: Web of Science. Strategy: Corrections, & corrections with the term “retraction” in their title. Search: date not reported.	Yes.
Fang 2011 (26)	Scope: PubMed search for 17 journals, 2001 to 2010. Details of strategy and search date not reported.	Yes.
Fang 2012 (27) Additional reports (28,29)	Scope: PubMed. Date of search: May 2012.	Yes.
Foo 2010 (30)	Scope: PubMed. Strategy: “retracted publication” in PT Search: 23 July 2009.	Yes.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
Furman 2012 (31)	Scope: PubMed and Web of Science. Search strategy and date not reported.	Yes.
Gasparyan 2014 (32)	Scope: PubMed. Search: January 2014.	Yes.
Grieneisen 2012 (2)	Scope: PubMed & 41 other databases (to a point of diminishing returns) Retracted publication PT or retraction/s in title, or withdrawn/withdrawal in title Search: 22 August 2011.	Yes.
He 2013 (33)	Scope: Web of Science 2001 - 2010. Search: February 2012.	Yes.
Huh 2016 (34)	Scope: KoreaMed database. Search: January 1990 to January 2016.	No.
Jawaid 2016 (35)	Scope: Pakistan Journal of Medical Sciences.	No.
Jin 2013 (36)	Scope: an effect of retraction study. Web of Science, authors with a single retraction between 1993 and 2009.	Yes.
Karabag 2016 (37)	Scope: Management, business & economic journals 2005-2015. Not biomed. Strategy: keyword searches in 7 publishers' databases. Date not reported.	Yes.
Lu 2013 (38)	Scope: Web of Science, 2000 to 2011, backend data. Impact of publications and citation study.	Yes.
Madlock-Brown 2015 (39)	Scope: MEDLINE – not specified. Details of search strategy not reported. Apparently based on NLM tagging. Search: 2012 (not further specified).	Yes.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
Mongeon 2015 (40) Additional report (41)	Scope: PubMed 1996-2006. Impact on authors of retractions study. Strategy: “retracted publication” PT and “retraction of publication” PT Date not reported.	Yes.
Moylan 2016 (42)	Scope: BioMed Central backend. Search: January 2000 to December 2015.	No.
Nath 2006 (43)	Scope: MEDLINE – not specified. “retracted publication” PT and “retraction of publication” PT Search: date not reported.	Yes.
Neale 2007 (44)	Scope: ORI findings 1991 – 2001. Checked for retractions in PubMed and Web of Science. (Last check May 2005).	No.
Nogueira 2017 (45)	Scope: SCImago, Retraction Watch, PubMed. Strategy: PubMed included publication type retraction, and “withdrawn [title] OR retraction [title] OR retracted [title] AND dentistry”). (Dentistry only.) Search: March 2016.	No.
Parrish 1999 (46)	Scope: ORI findings, 25 cases (1983 to 1997).	No.
Qi 2017 (47)	Scope: Retraction Watch. Strategy: “Fake peer review” OR “faked peer review” in PubMed and Google Scholar. Search: November 2015.	No.
Rai 2017 (48)	Scope: PubMed/MEDLINE, CINAHL, Google Scholar, Scopus. (Orthopedics only.) Search: 1984 to 4 June 2016.	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
Redman 2008 (49)	Scope: PubMed, articles published 1995-2004. Strategy: PT “retraction in” Search: 9 June 2005	Yes.
Resnik 2013 (50)	Scope: ORI findings, 1992-2011, PubMed notifications. Time of searching PubMed not reported.	No.
Rosenkrantz 2016 (51)	Scope: PubMed, Web of Science. (Radiology journals only.) Strategy: PubMed - retraction PT, “retraction notice” at start of title Search: May 2015.	Yes.
Samp 2012 (52)	Scope: PubMed, 2000-2011. Strategy: PT retraction notice and retracted publication Search: 11 August 2011.	Yes.
Snodgrass 1992 (53)	Scope: PubMed – plus 7 extras added, provenance unclear. Strategy: “retraction of publication” PT Search: August 1991.	Yes.
Steen 2011 (54) Additional reports (55–57)	Scope: PubMed Strategy: Limits “items with abstracts, retracted publication, English”, retracted in 2000 to 2010 Search: 22 January 2010.	Yes.
Steen 2013 (58)	Scope: PubMed Strategy: Limits retracted publication, English language Search: 3 May 2012.	Yes.
Trikalinos 2008 (59)	Scope: 21 top-cited journals. Strategy: Limit retracted publication in PubMed, also searched Web of Science.	Yes.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	Search: date not reported.	
Wang 2017 (60)	Scope: MEDLINE, EMBASE, neurosurgery journals. (Neurosurgery only.) Search: January 2017.	No.
Wager 2011 (61)	Scope: MEDLINE not specified Retractions in English, Strategy not specified. Search: Date not reported.	Yes.
Woolley 2011 (62) Additional report (63)	Scope: PubMed Strategy: Limit retracted publications, English language, human research. Search: 18 February 2008.	Yes
Yan 2016 (64)	Scope: MEDLINE, EMBASE, Cochrane Library. (Orthopedic only.) Search: September 2015.	No.

Table 2. Included studies – update search, used for ongoing filter development.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
Al-Ghareeb 2018 (65)	Scope: MEDLINE via Ovid, Retraction Watch website. Strategy: Limited to nursing research: 1. (retract* OR remove* OR recall* OR withdraw* OR 'retract* public*' ~ 10) [mp = title, abstract, original title, name of substance word, subject heading word, keyword heading word; protocol supplementary concept word; rare disease supplementary concept word; unique identifier; synonyms].	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	2. Retraction of publication (MeSH) 3. 1 OR 2 4. Limit 3 to filters (retracted publication OR retraction of publication) Date: July 2017.	
Ajiferuke 2018 (66)	Scope: Web of Science. Strategy: Library and information science journals, includes citation study Date not reported.	No.
Al-Hidabi 2018 (67)	Scope: Web of Science. Strategy: Computer science only, January 2007 to July 2017 Date not reported.	No.
Alrawadieh 2019 (68)	Scope: ScienceDirect, Google Scholar, Retraction Watch website. Strategy: Tourism and hospitality journals only Date not reported.	No.
Ayodele 2018 (69)	Scope: CrossRef, Google Scholar. Strategy: Management research only (2005-2016) Date not reported.	No.
Bakker 2018 (70)	Scope: Retraction Watch website then discoverability at MEDLINE via Ovid, PsycInfo via Ovid, EBSCO databases, Scopus, Web of Science, PubMed. Strategy: Mental health only Date: 27 June 2016 to 8 July 2016.	No.
Bar-Ilan 2018 (71)	Scope: ScienceDirect (Elsevier). Strategy: Citation and Mendeley readership study. Date: October 2014.	No.
Bolboacā 2019 (72)	Scope: PubMed.	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	Strategy: Retracted publication and retraction subsets, radiology-imaging diagnostic methods only, impact on citations study using Scopus. Date: June 2017.	
Brainard 2018 (73)	Scope: Retraction Watch database. Date: 30 August 2018.	No.
Campos-Varela 2018 (74)	Scope: PubMed. Strategy: “indexed as ‘retracted publication’,” from 2013 to 2016. Date: 30 April 2017.	No.
Cassão 2018 (75)	Scope: 100 surgery journals. Strategy and date not reported.	No.
Chambers 2019 (76)	Scope: PubMed. Strategy: Indexed as retracted or withdrawn, plus search for “retracted”, “retraction” and “withdrawn”. Obstetrics and gynaecology only. Date: June 2018.	No.
Chauvin 2019 (77)	Scope: MEDLINE unspecified, Web of Science, Cochrane Central Register of Controlled Trials, Retraction Watch website. Strategy: Retracted publication and retraction of publication PTs. Emergency medicine only. Date: 5 July 2016.	No.
Coudert 2019 (78)	Scope: Scopus. Strategy: Chemistry, material science and chemical engineering only, 2017-2018 Date not reported.	No.
Dal-Ré 2019a (79) Additional report	Scope: Retraction Watch database Strategy: All entries for misconduct or	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
(Spanish subset) (80)	investigation, target genetics only, but multiple disciplines compared. Date: 14-16 January 2019	
Dal-Ré 2019b (81)	Scope: Retraction Watch database Strategy: Retractions and “Medicine-pharmacology”, “clinical study”, “research article”, “meta-analysis”. Date: 17-20 May 2019.	No.
Decullier 2018 (82)	Scope: MEDLINE (unspecified). Strategy: Publication type “retraction of publication” Date: 1 February 2017	No.
Drimer-Batca 2019 (83)	Scope: PubMed. Strategy: Reviewing PubMed status of 200 papers with ORI finding of misconduct Date: November 2017.	No.
Drury 2009 (84)	Scope: MEDLINE (not specified). Strategy: MeSH “retracted publication”, 1990-2008, citation study for journals in the Cardiothoracic Surgery Network Date not reported.	No.
Faggion 2018 (85)	Scope: PubMed, PubMed Central, Web of Science, Google Scholar, Retraction Watch (website). Strategy: PubMed retraction and retracted publication filters Date: 2 July 2018.	No.
Gray 2018 (86)	Scope: MEDLINE Strategy: Not specified. Nursing only, trials to follow up if they were in systematic reviews. Date not reported.	No.
Gray 2019 (87)	Scope: PubMed. Uses retractions identified in (65)	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	Strategy: Study of availability of retraction notices at PubMed and/or journal, nursing only Date not reported.	
Hamilton 2019 (88)	Scope: Google Scholar, Web of Science, Scopus. Strategy: Citation impact, radiation oncology only. Date: June 2017.	No.
Horbach 2019 (89)	Scope: Retraction Watch database. Strategy: Analysis of retraction rates at journals according to reported peer review procedures. Date: 11 December 2017.	No.
Hosseini 2018 (90)	Scope: Web of Science. Strategy: Limited countries, “honest error” as reason, 2010-2015. Date not reported.	No.
Hwang 2018 (91)	Scope: PubMed, Scopus. Strategy: PubMed “(retracted OR withdrawn) AND (article OR publication OR paper)”, plastic surgery journals only Date not reported.	No.
Jan 2018 (92)	Scope: Retraction Watch website. Strategy: Citation study, top 7 from highly cited retractions list Date not reported.	No.
King 2018 (93)	Scope: PubMed. Strategy: Retracted or retraction, filtered English language (1991-2015), surgery only. Date not reported.	No.
Kuroki 2018 (94)	Scope: PubMed, Web of Science. Strategy: Used supplementary file from	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	Steen 2013 (58). Study of rate of repeat retractions for authors. Date: Web of Science 20 January 2016.	
Lei 2018 (95)	Scope: Web of Science. Strategy: Chinese researchers only. Date: 5 February 2017.	No.
Li 2018 (96)	Scope: PubMed, Retraction Watch website. Strategy: retract*, limited to retracted publication PT. Limited to human research and English language filters. Date: February 2016.	No.
Mena 2019 (97)	Scope: PubMed, EMBASE, Retraction Watch database. Strategy: PubMed retraction and retracted PTs, urology only. Date: 11 May 2018.	No.
Mistry 2019 (98)	Scope: Retraction Watch database. Strategy: Authors with multiple retractions only, impact on publication rates using Scopus. Date: November 2018.	No.
Mott 2019 (99)	Scope: PubMed, CINAHL, Google, Retraction Watch database. Strategy: “terms specific to each database”, randomized clinical trials only, citation impact study. Date: August 2017.	No.
Pierson 2018 (100)	Scope: PubMed. Strategy: Retracted publication PT, nursing Date: 15 January 2018.	No.
Ribeiro 2018 (101)	Scope: Retraction Watch database. Strategy: Retractions from 2013 to 2015.	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	Date: August 2017.	
Rubbo 2019 (102)	Scope: Web of Science. Strategy: Engineering only Date: April to May 2016.	No.
Schmidt 2018 (103)	Scope: PubMed, Web of Science. Strategy: Words (unspecified) in title indicating retraction but not a retracted PT; and ((“withdrawn"[Title] OR “withdrawal"[Title]) NOT “Retracted publication"[Publication Type]) NOT “Retraction of publication"[Publication Type]; (“withdrawn"[Title] OR “withdrawal"[Title]) AND “Retraction of publication"[Publication Type]; (“withdrawn"[Title] OR “withdrawal"[Title]) AND “Retracted publication"[Publication Type] Date not reported.	No.
Shamim 2018 (104)	Scope: Analysis of Indian articles from (45) Strategy: n/a Date: n/a	No.
Shema 2019 (105)	Scope: PubMed. Strategy: Unspecified, 2012-2017, Altmetric study Date: 2 August 2017.	No.
Shuai 2017 (106)	Scope: Web of Science. Strategy: Citation impact study. Date not reported.	No.
Singh 2014 (107)	Scope: PubMed and MEDLINE (not specified). Strategy: Keywords retraction of articles, retraction notice, and withdrawal of article, 2004-2013 Date: April 2014.	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
Stavale 2019 (108)	Scope: PubMed, Web of Science, Biblioteca Virtual em Saúde, Google Scholar, Retraction Watch database. Strategy: From 2004, articles from Brazilian institutions only. In PubMed: ("Retraction of publication"[All Fields] OR "retraction of publication as topic"[All Fields] OR "retracted publication"[All Fields]) NOT (retraction[All Fields] AND ("dentistry"[MeSH Terms] OR "dentistry"[All Fields])) Date: Not reported.	No.
Stricker 2019 (109)	Scope: PsycInfo. Strategy: Psychology only Date: January 2018.	No.
Tang 2017 (110)	Scope: Web of Science. Strategy: 1987 to 2013 Date: January 2014.	No.
Tripathi 2018 (111)	Scope: Scopus. Strategy: 2000-2017 Date: January 2018.	No.
Van Leeuwen 2014 (112)	Scope: Web of Science. Strategy: Tagged retracted. Search: January 2014.	No.
Wang 2018 (113)	Scope: PubMed. Strategy: “Retracted publication” PT, Directory of Open Access Journals only Date: 5 October 2017.	No.
Wasiak 2018 (114)	Scope: Ovid MEDLINE, PubMed, Ovid EMBASE, Cochrane Library. Strategy: "retraction note," "retracted note," "withdrawn" Date: May 2017.	No.
Wray 2018 (115)	Science journal 1983-1997 – method and	No.

ID	Method for finding retractions (PubMed strategies)	Included in (1)
	date not reported.	
Xu 2018 (116)	Scope: Web of Science. Strategy: Cell biology, business, finance, management only. Authorship of retraction notices. Date: March 2017.	No.
Yanti Idaya Aspura 2018 (117)	Scope: Web of Science, Scopus. Strategy: Malaysian articles only. Date: 30 July 2017.	No.

References

1. Hesselmann F, Graf V, Schmidt M, Reinhart M. The visibility of scientific misconduct: A review of the literature on retracted journal articles. Curr Sociol. 2017;65(6):814–45.

2. Grieneisen ML, Zhang M. A comprehensive survey of retracted articles from the scholarly literature. PloS One. 2012;7(10):e44118.

3. Abritis AJ. An Assessment of Retractions as a Measure of Scientific Misconduct and Impact on Public Health Risks [Internet]. Tampa: University of South Florida; 2015. (Graduate Theses and Dissertations). Available from: https://scholarcommons.usf.edu/etd/5630/

4. Almeida RMVR, de Albuquerque Rocha K, Catelani F, Fontes-Pereira AJ, Vasconcelos SMR. Plagiarism Allegations Account for Most Retractions in Major Latin American/Caribbean Databases. Sci Eng Ethics. 2016 Oct;22(5):1447–56.

5. Almeida RMVR, Catelani F, Fontes-Pereira AJ, Gave N de S. Retractions in general and internal medicine in a high-profile scientific indexing database. Sao Paulo Med J Rev Paul Med. 2016 Feb;134(1):74–8.

6. Amos KA. The ethics of scholarly publishing: exploring differences in plagiarism and duplicate publication across nations. J Med Libr Assoc JMLA. 2014 Apr;102(2):87–91.

7. Azoulay P, Furman JL, Krieger JL, Murray FE. Retractions [Internet]. Cambridge, Massachusetts: National Bureau of Economic Research; 2012. (NBER Working Papers). Available from: https://www.nber.org/papers/w18499

8. Balhara YPS, Mishra A. Compliance of retraction notices for retracted articles on mental disorders with COPE guidelines on retraction. Curr Sci. 2014 Sep 10;107(5):757–60.

9. Balhara YPS, Mishra A. A study exploring attributes and nature of the retracted literature on mental disorders. Indian J Med Ethics. 2015 Mar;12(1):30–7.

10. Bilbrey E, O'Dell N, Creamer J. A Novel Rubric for Rating the Quality of Retraction Notices. Publications. 2014;2:14–26.

11. Bozzo A, Bali K, Evaniew N, Ghert M. Retractions in cancer research: a systematic survey. Res Integr Peer Rev. 2017;2:5.

12. Budd JM, Sievert M, Schultz TR. Phenomena of retraction: reasons for retraction and citations to the publications. JAMA. 1998 Jul 15;280(3):296–7.

13. Budd JM, Sievert M, Schultz TR, Scoville C. Effects of article retraction on citation and practice in medicine. Bull Med Libr Assoc. 1999 Oct;87(4):437–43.

14. Budd JM, Coble ZC, Anderson KM. Retracted Publications in Biomedicine: Cause for Concern. In Philadelphia: Association of College and Research Libraries, American Library Association; 2011. p. 390–5. Available from: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/national/2011/papers/retracted_publicatio.pdf

15. Cicero T, Malgarini M, Benedetto S. Research quality, characteristics of publications and sociodemographic features of Universities and Researchers: evidence from the Italian VQR 2004-2010 evaluation exercise. In Leiden: Leiden University; 2014. p. 106–13. Available from: https://www.researchgate.net/publication/265396216_Research_quality_characteristics_of_publications_and_socio-demographic_features_of_Universities_and_Researchers_evidence_from_the_Italian_VQR_2004-2010_evaluation_exercise#page=349

16. Claxton LD. Scientific authorship. Part 1. A window into scientific fraud? Mutat Res. 2005 Jan;589(1):17–30.

17. Cokol M, Ozbay F, Rodriguez-Esteban R. Retraction rates are on the rise. EMBO Rep. 2008 Jan;9(1):2.

18. Colaianni LA. Retraction, comment, and errata policies of the US National Library of Medicine. Lancet Lond Engl. 1992 Aug 29;340(8818):536–7.

19. Damineni RS, Sardiwal KK, Waghle SR, Dakshyani MB. A comprehensive comparative analysis of articles retracted in 2012 and 2013 from the scholarly literature. J Int Soc Prev Community Dent. 2015 Feb;5(1):19–23.

20. Davis PM. The persistence of error: a study of retracted articles on the Internet and in personal libraries. J Med Libr Assoc JMLA. 2012 Jul;100(3):184–9.

21. Decullier E, Huot L, Samson G, Maisonneuve H. Visibility of retractions: a cross-sectional one-year study. BMC Res Notes. 2013 Jun 19;6:238.

22. Decullier E, Huot L, Maisonneuve H. What time-lag for a retraction search on PubMed? BMC Res Notes. 2014 Jun 25;7:395.

23. Elia N, Wager E, Tramèr MR. Fate of articles that warranted retraction due to ethical concerns: a descriptive cross-sectional study. PloS One. 2014;9(1):e85846.

24. Fanelli D. Why growing retractions are (mostly) a good sign. PLoS Med. 2013 Dec;10(12):e1001563.

25. Fanelli D, Costas R, Larivière V. Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. PloS One. 2015;10(6):e0127556.

26. Fang FC, Casadevall A. Retracted science and the retraction index. Infect Immun. 2011 Oct;79(10):3855–9.

27. Fang FC, Steen RG, Casadevall A. Misconduct accounts for the majority of retracted scientific publications. Proc Natl Acad Sci U S A. 2012 Oct 16;109(42):17028–33.

28. Casadevall A, Steen RG, Fang FC. Sources of error in the retracted scientific literature. FASEB J Off Publ Fed Am Soc Exp Biol. 2014 Sep;28(9):3847–55.

29. Stern AM, Casadevall A, Steen RG, Fang FC. Financial costs and personal consequences of research misconduct resulting in retracted publications. eLife. 2014 Aug 14;3:e02956.

30. Foo JYA. A retrospective analysis of the trend of retracted publications in the field of biomedical and life sciences. Sci Eng Ethics. 2011 Sep;17(3):459–68.

31. Furman JL, Jensen K, Murray F. Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. Res Policy. 2012 Mar 1;41(2):276–90.

32. Gasparyan AY, Ayvazyan L, Akazhanov NA, Kitas GD. Self-correction in biomedical publications and the scientific impact. Croat Med J. 2014 Feb;55(1):61–72.

33. He T. Retraction of global scientific publications from 2001 to 2010. Scientometrics. 2013 Aug 1;96(2):555–61.

34. Huh S, Kim SY, Cho H-M. Characteristics of Retractions from Korean Medical Journals in the KoreaMed Database: A Bibliometric Analysis. PloS One. 2016;11(10):e0163588.

35. Jawaid SA, Jawaid M. Retractions by Pakistan Journal of Medical Sciences due to Scientific Misconduct. J Coll Physicians Surg--Pak JCPSP. 2016 Aug;26(8):718–9.

36. Jin GZ, Jones B, Lu SF, Uzzi B. The reverse Matthew Effect: catastrophe and consequence in scientific teams [Internet]. Cambridge, Massachusetts: National Bureau of Economic Research; 2013. (NBER Working Papers). Available from: https://www.nber.org/papers/w19489

37. Karabag SF, Berggren C. Misconduct, Marginality and Editorial Practices in Management, Business and Economics Journals. PloS One. 2016;11(7):e0159492.

38. Lu SF, Jin GZ, Uzzi B, Jones B. The retraction penalty: evidence from the Web of Science. Sci Rep. 2013 Nov 6;3:3146.

39. Madlock-Brown CR, Eichmann D. The (lack of) impact of retraction on citation networks. Sci Eng Ethics. 2015 Feb;21(1):127–37.

40. Mongeon P, Larivière V. Costly collaborations: The impact of scientific fraud on co-authors' careers. J Assoc Inf Sci Technol. 2016;67(3):535–42.

41. Mongeon P, Larivière V. The collective consequences of scientific fraud: an analysis of biomedical research. In Vienna: International Society of Scientometrics and Informetrics; 2013. p. 1897–9. Available from: http://www.issi2013.org/proceedings.html

42. Moylan EC, Kowalczuk MK. Why articles are retracted: a retrospective cross-sectional study of retraction notices at BioMed Central. BMJ Open. 2016 23;6(11):e012047.

43. Nath SB, Marcus SC, Druss BG. Retractions in the research literature: misconduct or mistakes? Med J Aust. 2006 Aug 7;185(3):152–4.

44. Neale AV, Northrup J, Dailey R, Marks E, Abrams J. Correction and use of biomedical literature affected by scientific misconduct. Sci Eng Ethics. 2007 Mar;13(1):5–24.

45. Nogueira TE, Gonçalves AS, Leles CR, Batista AC, Costa LR. A survey of retracted articles in dentistry. BMC Res Notes. 2017 Jul 6;10(1):253.

46. Parrish DM. Scientific misconduct and correcting the scientific literature. Acad Med J Assoc Am Med Coll. 1999 Mar;74(3):221–30.

47. Qi X, Deng H, Guo X. Characteristics of retractions related to faked peer reviews: an overview. Postgrad Med J. 2017 Aug;93(1102):499–503.

48. Rai R, Sabharwal S. Retracted Publications in Orthopaedics: Prevalence, Characteristics, and Trends. J Bone Joint Surg Am. 2017 May 3;99(9):e44.

49. Redman BK, Yarandi HN, Merz JF. Empirical developments in retraction. J Med Ethics. 2008 Nov;34(11):807–9.

50. Resnik DB, Dinse GE. Scientific retractions and corrections related to misconduct findings. J Med Ethics. 2013 Jan;39(1):46–50.

51. Rosenkrantz AB. Retracted Publications Within Radiology Journals. AJR Am J Roentgenol. 2016 Feb;206(2):231–5.

52. Samp JC, Schumock GT, Pickard AS. Retracted publications in the drug literature. Pharmacotherapy. 2012 Jul;32(7):586–95.

53. Snodgrass GL, Pfeifer MP. The characteristics of medical retraction notices. Bull Med Libr Assoc. 1992 Oct;80(4):328–34.

54. Steen RG. Retractions in the scientific literature: do authors deliberately commit research fraud? J Med Ethics. 2011 Feb;37(2):113–7.

55. Steen RG. Retractions in the medical literature: how many patients are put at risk by flawed research? J Med Ethics. 2011 Nov;37(11):688–92.

56. Steen RG. Retractions in the scientific literature: is the incidence of research fraud increasing? J Med Ethics. 2011 Apr;37(4):249–53.

57. Steen RG. Retractions in the medical literature: how can patients be protected from risk? J Med Ethics. 2012 Apr;38(4):228–32.

58. Steen RG, Casadevall A, Fang FC. Why has the number of scientific retractions increased? PloS One. 2013;8(7):e68397.

59. Trikalinos NA, Evangelou E, Ioannidis JPA. Falsified papers in high-impact journals were slow to retract and indistinguishable from nonfraudulent papers. J Clin Epidemiol. 2008 May;61(5):464–70.

60. Wang J, Ku JC, Alotaibi NM, Rutka JT. Retraction of Neurosurgical Publications: A Systematic Review. World Neurosurg. 2017 Jul;103:809-814.e1.

61. Wager E, Williams P. Why and how do journals retract articles? An analysis of Medline retractions 1988-2008. J Med Ethics. 2011 Sep;37(9):567–70.

62. Woolley KL, Lew RA, Stretton S, Ely JA, Bramich NJ, Keys JR, et al. Lack of involvement of medical writers and the pharmaceutical industry in publications retracted for misconduct: a systematic, controlled, retrospective study. Curr Med Res Opin. 2011 Jun;27(6):1175–82.

63. Stretton S, Bramich NJ, Keys JR, Monk JA, Ely JA, Haley C, et al. Publication misconduct and plagiarism retractions: a systematic, retrospective study. Curr Med Res Opin. 2012 Oct;28(10):1575–83.

64. Yan J, MacDonald A, Baisi L-P, Evaniew N, Bhandari M, Ghert M. Retractions in orthopaedic research: A systematic review. Bone Jt Res. 2016 Jun;5(6):263–8.

65. Al-Ghareeb A, Hillel S, McKenna L, Cleary M, Visentin D, Jones M, et al. Retraction of publications in nursing and midwifery research: A systematic review. Int J Nurs Stud. 2018 May;81:8–13.

66. Ajiferuke I, Adekannbi JO. Correction and retraction practices in library and information science journals. J Librariansh Inf Sci [Internet]. 2018; Available from: https://journals.sagepub.com/doi/abs/10.1177/0961000618785408

67. Al-Hidabi MDA, Teh PL. Multiple Publications: The Main Reason for the Retraction of Papers in Computer Science. In: Future of Information and Communication Conference [Internet]. Singapore: Science and Information Conferences; 2018. p. 511–26. Available from: https://link.springer.com/chapter/10.1007/978-3-030-03402-3_35

68. Alrawadieh Z, Zareer A. Exploring Retraction in Tourism and Hospitality Journals. Eur J Tour Res. 2019;22:20–30.

69. Ayodele FO, Yao L, Haron H. Promoting Ethics and Integrity in Management Academic Research: Retraction Initiative. Sci Eng Ethics. 2019;25(2):357–82.

70. Bakker C, Riegelman A. Retracted Publications in Mental Health Literature: Discovery across Bibliographic Platforms. J Librariansh Sch Commun. 2018;6(General Issue):eP2199.

71. Bar-Ilan J, Halevi G. Temporal characteristics of retracted articles. Scientometrics. 2018 Sep;116(3):1771–83.

72. Bolboacă SD, Buhai D-V, Aluaș M, Bulboacă AE. Post retraction citations among manuscripts reporting a radiology-imaging diagnostic method. PloS One. 2019;14(6):e0217918.

73. Brainard J. Rethinking retractions. Science. 2018 Oct 26;362(6413):390–3.

74. Campos-Varela I, Ruano-Raviña A. Misconduct as the main cause for retraction. A descriptive study of retracted publications and their authors. Gac Sanit. 2019 Aug;33(4):356–60.

75. Cassão BD, Herbella FAM, Schlottmann F, Patti MG. Retracted articles in surgery journals. What are surgeons doing wrong? Surgery. 2018;163(6):1201–6.

76. Chambers LM, Michener CM, Falcone T. Plagiarism and data falsification are the most common reasons for retracted publications in obstetrics and gynaecology. BJOG Int J Obstet Gynaecol. 2019 Aug;126(9):1134–40.

77. Chauvin A, De Villelongue C, Pateron D, Yordanov Y. A systematic review of retracted publications in emergency medicine. Eur J Emerg Med Off J Eur Soc Emerg Med. 2019 Feb;26(1):19–23.

78. Coudert F-X. Correcting the Scientific Record: Retraction Practices in Chemistry and Materials Science. Chem Mater. 2019;31(10):3593–8.

79. Dal-Ré R, Ayuso C. Reasons for and time to retraction of genetics articles published between 1970 and 2018. J Med Genet. 2019 Jul 12;

80. Dal-Ré R. Analysis of biomedical Spanish articles retracted between 1970 and 2018. Med Clin (Barc). 2019 Jun 22;

81. Dal-Ré R. Analysis of retracted articles on medicines administered to humans. Br J Clin Pharmacol. 2019 Jun 24;

82. Decullier E, Maisonneuve H. Correcting the literature: Improvement trends seen in contents of retraction notices. BMC Res Notes. 2018 Jul 17;11(1):490.

83. Drimer-Batca D, Iaccarino JM, Fine A. Status of retraction notices for biomedical publications associated with research misconduct. Res Ethics. 2019;15(2):1–15.

84. Drury NE, Karamanou DM. Citation of retracted articles: a call for vigilance. Ann Thorac Surg. 2009 Feb;87(2):670.

85. Faggion CM, Ware RS, Bakas N, Wasiak J. An analysis of retractions of dental publications. J Dent. 2018 Dec;79:19–23.

86. Gray R, Al-Ghareeb A, Davis J, McKenna L, Amichai Hillel S. Inclusion of nursing trials in systematic reviews after they have been retracted: Does it happen and what should we do? Int J Nurs Stud. 2018;79:154.

87. Gray R, Al-Ghareeb A, McKenna L. Why articles continue to be cited after they have been retracted: An audit of retraction notices. Int J Nurs Stud. 2019;90:11–2.

88. Hamilton DG. Continued Citation of Retracted Radiation Oncology Literature-Do We Have a Problem? Int J Radiat Oncol Biol Phys. 2019 Apr 1;103(5):1036–42.

89. Horbach SPJM, Halffman W. The ability of different peer review procedures to flag problematic publications. Scientometrics. 2019;118(1):339–73.

90. Hosseini M, Hilhorst M, de Beaufort I, Fanelli D. Doing the Right Thing: A Qualitative Investigation of Retractions Due to Unintentional Error. Sci Eng Ethics. 2018;24(1):189–206.

91. Hwang K, Wu X. Retracted or Withdrawn Publications in Journals Relating to Plastic Surgery. J Craniofac Surg. 2018 Jul;29(5):1114–6.

92. Jan R, Bano S, ur Rehman SI, Mehraj M. Context Analysis of Top Seven Retracted Articles: Should Retraction Watch Revisit the List? Libr Philos Pract [Internet]. 2018 Dec; Available from: https://digitalcommons.unl.edu/libphilprac/2016/

93. King EG, Oransky I, Sachs TE, Farber A, Flynn DB, Abritis A, et al. Analysis of retracted articles in the surgical literature. Am J Surg. 2018 Nov;216(5):851–5.

94. Kuroki T, Ukawa A. Repeating probability of authors with retracted scientific publications. Account Res. 2018;25(4):212–9.

95. Lei L, Zhang Y. Lack of Improvement in Scientific Integrity: An Analysis of WoS Retractions by Chinese Researchers (1997-2016). Sci Eng Ethics. 2018;24(5):1409–20.

96. Li G, Kamel M, Jin Y, Xu MK, Mbuagbaw L, Samaan Z, et al. Exploring the characteristics, global distribution and reasons for retraction of published articles involving human research participants: a literature survey. J Multidiscip Healthc. 2018;11:39–47.

97. Mena JD, Ndoye M, Cohen AJ, Kamal P, Breyer BN. The landscape of urological retractions: the prevalence of reported research misconduct. BJU Int. 2019 Jul;124(1):174–9.

98. Mistry V, Grey A, Bolland MJ. Publication rates after the first retraction for biomedical researchers with multiple retracted publications. Account Res. 2019 Jul;26(5):277–87.

99. Mott A, Fairhurst C, Torgerson D. Assessing the impact of retraction on the citation of randomized controlled trial reports: an interrupted time-series analysis. J Health Serv Res Policy. 2019 Jan;24(1):44–51.

100. Pierson CA. Retractions in nursing literature: Responsibilities of nurse authors, reviewers, and editors. J Am Assoc Nurse Pract. 2018 Mar;30(3):115–6.

101. Ribeiro MD, Vasconcelos SMR. Retractions covered by Retraction Watch in the 2013–2015 period: prevalence for the most productive countries. Scientometrics. 2018 Feb;114(2):719–34.

102. Rubbo P, Helmann CL, dos Santos CB, Pilatti LA. Retractions in the Engineering Field: A Study on the Web of Science Database. Ethics Behav. 2019;29(2):141–55.

103. Schmidt M. An analysis of the validity of retraction annotation in pubmed and the web of science. J Assoc Inf Sci Technol. 2018 Feb;69(2):318–28.

104. Shamim T. Data regarding articles retracted from PubMed indexed dental journals from India. Data Brief. 2018 Jun;18:1069–72.

105. Shema H, Hahn O, Mazarakis A, Peters I. Retractions from altmetric and bibliometric perspectives. Inf - Wiss Prax [Internet]. 2019;70(2–3). Available from: https://www.degruyter.com/abstract/j/iwp.2019.70.issue-2-3/iwp-2019-2006/iwp-2019-2006.xml

106. Shuai X, Rollins J, Moulinier I, Custis T, Edmunds M, Schilder F. A Multidimensional Investigation of the Effects of Publication Retraction on Scholarly Impact. J Assoc Inf Sci Technol. 2017 Sep;68(9):2225–36.

107. Singh HP, Mahendra A, Yadav B, Singh H, Arora N, Arora M. A comprehensive analysis of articles retracted between 2004 and 2013 from biomedical literature - a call for reforms. J Tradit Complement Med. 2014 Jul;4(3):136–9.

108. Stavale R, Ferreira GI, Galvão JAM, Zicker F, Novaes MRCG, Oliveira CM de, et al. Research misconduct in health and life sciences research: A systematic review of retracted literature from Brazilian institutions. PloS One. 2019;14(4):e0214272.

109. Stricker J, Günther A. Scientific misconduct in psychology: A systematic review of prevalence estimates and new empirical data. Z Für Psychol. 2019;227(1):53–63.

110. Tang L, Hu G, Yang S, Yang Y, Cao C. Retraction: The “Other Face” of Research Collaboration [Internet]. 2017 [cited 2019 Aug 11]. Available from: https://works.bepress.com/li_tang/39/

111. Tripathi M, Dwivedi G, Sonkar SK, Kumar S. Analysing Retraction Notices of Scholarly Journals: A Study. DESIDOC J Libr Inf Technol. 2018;38(5):305–11.

112. van Leeuwen TN, Luwel M. An in-depth analysis of papers retracted in the Web of Science. In: Proceedings of the science and technology indicators conference 2014 Leiden [Internet]. Leiden: Universiteit Leiden; p. 337–43. Available from: https://www.researchgate.net/profile/Tindaro_Cicero/publication/265396216_Research_quality_characteristics_of_publications_and_socio-demographic_features_of_Universities_and_Researchers_evidence_from_the_Italian_VQR_2004-2010_evaluation_exercise/links/540d89180cf2df04e754b658.pdf#page=349

113. Wang T, Xing Q-R, Wang H, Chen W. Retracted Publications in the Biomedical Literature from Open Access Journals. Sci Eng Ethics. 2019 Jun;25(3):855–68.

114. Wasiak J, Hamilton DG, Foroudi F, Faggion CM. Surveying Retracted Studies and Notices Within the Field of Radiation Oncology. Int J Radiat Oncol Biol Phys. 2018 01;102(3):660–5.

115. Wray KB, Andersen LE. Retractions in Science. Scientometrics. 2018 Dec;117(3):2009–19.

116. Xu S (Brian), Hu G. Retraction Notices: Who Authored Them? Publications. 2018;6(1):2.

117. Yanti Idaya Aspura MK, Noorhidawati A, Abrizah A. An analysis of Malaysian retracted papers: Misconduct or mistakes? Scientometrics. 2018 Jun;115(3):1315–28.

118. Yasri S, Wiwanitkit V. Retracted or Withdrawn Publications in Plastic Surgery Journals. J Craniofac Surg. 2018;29(8):2368.

Appendix 5.

Studies on the prevalence, content, and impact of post-publication events detectable in PubMed.

1. Prevalence studies

(a) Retractions

Retraction of publications is the post-publication event that has been researched the most, (1) *[Appendix 4]* although – and perhaps partly because – they are rare. According to a study by Bastian et al to the end of 2017, less than 0.1% of new publications in PubMed have been retracted annually, only 64% of which were indexed with this status. *[Chapter 5]* This study found a very high rate of unindexed retracted publications among records which were still "e-published ahead of print" status without ever being formally added to an issue of a journal. We found one study reporting on the prevalence of retracted clinical trials, but none on trials included in systematic reviews. Mott et al identified 383 retracted randomized trials in PubMed up to February 2017, 95% of which were indexed with retracted status. (2) We found no studies on retractions of systematic reviews generally, but Ma et al reported withdrawal of 196 Cochrane systematic reviews to the end of 2012, representing 2.7% of reviews. (3) Cochrane retracted reviews include a mixture of reasons, from serious error to being out of date, or superseded by another review(s). Some Cochrane reviews have been removed from the journal with no explanation. *[Chapter 3]*

(b) Misconduct

Although it is not possible to assess the findings of research misconduct by research institutions and agencies worldwide, publications affected by them are likely to be relatively rare. Abritis found that of the 167 researchers with findings of misconduct by the U.S. Office of Research Integrity between 1993 and 2013, half had no publications associated with the misconduct. (4)

(c) Expressions of concern

Editorial expressions of concern are also rare, but increasing: only 320 publications in PubMed were known to have had expressions of concern at the end of 2016. (5) At that time, this status was not indexed in PubMed. A publication type for expressions of concern was introduced and used retrospectively in 2018. (6)

(d) Errata

We found one study of the rate of errata for included studies in systematic reviews. Farrah and Rabb searched PubMed and journal websites for errata affecting 127 of 669 included articles (19%) in 40 systematic reviews of drugs by the Canadian Agency for Drugs and Technology in Health (CADTH). (7) There was a median of 1.5 errors per erratum. Databases varied greatly in how often errata were included, or original records modified, from 27% in Cochrane's CENTRAL database, to 69% in MEDLINE and 73% in Embase and Scopus.

Estimates of the prevalence of errata in the biomedical literature have varied across studies and journals. Royle and Waugh concluded that the errata rate for randomized trials in MEDLINE was 1.2% from 1995 to 2001, compared with 0.6% of other publications. (8) Most of the errata were in 4 journals, with a published errata rate over 8% in *The Lancet*, *JAMA*, and *NEJM*, and over 5% in *BMJ*. They did not evaluate the adequacy of identification of errata in MEDLINE/PubMed.

Molckovsky et al reviewed errata in two prominent oncology journals from 2004 to 2007, finding an errata rate of 4% of research articles, and 5% for randomized trials. (9) They also reported that oncology journals do not always link errata to the affected publication online. Trikalinos used PubMed to estimate the errata rate in five prominent medical journals between 1996 and 2008, reporting a 2.5% errata rate, compared to 0.6% for a group of 100 journals not chosen for their prominence. (10) Hauptman et al analysed errata for 557 papers identified by manually searching in 20 general medical and cardiovascular journals across 18 months in 2009/2010. (11)

They found the errata rate was correlated with journal impact factor, and the overall publication rate of errata was 4.2% for original research and reviews (with a range from 0 to 18.8%). Castillo et al analyzed errata in five prominent clinical imaging journals for five years to June 2011, and found a 1.8% errata rate. (12) Strothmann analysed 564 papers with errata at *PLOS One* for just over four years (to November 2011), an errata rate of 2.1%. (13)

Grcar (14) and Fanelli et al (15) used Scopus and Web of Science respectively to estimate rates of correction, but neither evaluated the adequacy of ascertainment of errata in these databases. Grcar reported 1.5% to 2.0% rate of errata for health, social sciences, and liberal arts from 1990 to 2010, and 3.0% for multidisciplinary journals. Fanelli et al concluded that the rate of errata was not increasing, but did not report the rate. Erfanmanesh and Teixeira da Silva also (apparently) used Web of Science, reporting on rate of errata for 16 open access mega-journals between 2012 and 2018, with a range from 0 to 3.2%. (16)

(e) Correction and republication

Correction with republication is rare. We found only one study with prevalence data for this type of post-publication action. Marasović et al investigated 35 corrected and republished articles in PubMed between 2015 and 2016, of which 29 were indexed with this status (83%). (17)

(f) Letters to the editor

Letters to the editor are the most common of the post-publication events we are concerned with in this study. We found 7 analyses of them after 2000. Baethge et al found 13% as many letters as original articles in PubMed in 2007, but without assessing how many of the letters were in response to publications. (18) They reported that in the year ending October 2007, *Deutsches Ärzteblatt* published correspondence relating to 49% of the original articles and reviews they published.

Kastner et al used PubMed to assess letters to the editor on randomized trials in five leading medical journals in 2007, finding 52% of 334 trials had letters published. (19) Ober et al reported on the usage of the criticisms management system for Cochrane systematic reviews in 2002, with 171 criticisms submitted per 1,388 reviews (12.3%). (20) Wiebe et al looked for letters to the editor about 16 studies in a systematic review of a controversial topic (the epidemiology murder and suicide in households with firearms). (21) There were letters to the editor for 8 of the 16 studies (50%) with author replies for 5 of them: a total of 24 letters and 6 author replies. For 7 letters there was no scientific content, political and ideological content was common, with some letters including character critiques of the authors.

(g) Author replies to letters to the editor

The rate of author replies also varies substantially by journal. Von Elm et al used PubMed to study letters and authors' replies in 8 leading general and internal medicine journals that published letters in 2002 and in 2007, and found that 63.4% received author replies in 2007, ranging from 3.0% to 99.1% between journals. (22) Gøtzsche et al found a 45% rate of author reply to rapid responses in the BMJ with substantive criticisms in the two years up to September 2007, (23) and Baethge reported an author reply rate of 51% in *Deutsches Ärzteblatt*. (18) Horton reported on letters published in response to three trials in *The Lancet*, with less than half receiving author replies. (24) Goldacre et al reported that they received 20 author replies to 23 letters: the 23 were the only letters accepted out of 58 submitted about misreported randomized trials. (25)

2. Content and impact studies

(a) Retractions

There have been two studies that we know of specifically related to retractions and systematic reviews. Garmendia et al studied the inclusion of falsified data from a single trial in 22 meta-analyses, concluding that the median weight of the trial was 37.3% (range, 7%–100%), and removal of the trial would potentially change the

results for 46% of the studies. (26) Fanelli and Moher re-analysed meta-analyses in 31 systematic reviews with retracted included studies, finding that the inclusion of studies retracted due to data, methods, or results had resulted in an average over-estimation across effect size (b=31.239±15.94, median of 13%). (27)

(b) Errata

Farrah and Rabb concluded that 16% of errata impacted analysis or interpretation of the study's primary outcome, and so could affect a review's results. (7) Royle and Waugh classified 5% of errata in trials as likely to have an effect on a meta-analysis, 10% as having significant errors that would affect interpretation of the trial, and 5% potentially changing a systematic review's conclusion. (8) In addition, they concluded that errata "can reduce confusion and save reviewers' time".

Molckovsky et al judged 14% of reported errata to be serious. (9) Hauptman et al concluded that 24.2% of errata "contained at least one major error that materially altered data interpretation". (11) Castillo et al rated 6.3% of errata in clinical imaging journals as major. (12) Strothmann reported that 26.5% of the errata in *PLOS One* related to figures or tables. (13) Bhatt et al manually searched the leading five medical journals, concluding that 54.2% of the 3,200 errata in 2012 related to errors of fact or data. (28) Trikalinos analyzed errata correcting author names, reporting that these were rare: only 83 out of 2,455 errata in five prominent medical journals. (10) These would not be classified as serious in the other studies, but Trikalinos raised the question of whether sloppiness in proof-reading authors' names could be an indicator that other errors in the paper are more likely to have escaped notice.

(c) Letters to the editor

Ober et al judged 394 out of 661 online letters (60%) about Cochrane systematic reviews to be major, and found that 11% of them of Cochrane systematic reviews resulted in a change to the review. (20) Kastner et al concluded that 90% of letters about randomized trials addressed issues defined as "study design, outcomes, population, intervention, and analysis". (19) Gøtzsche et al judged that of papers with

online letters, 30% made substantive criticisms. (23) Two editors assessed those online letters, with one judging rating 10% of the criticisms as possibly invalidating the paper, and the other judging that rate to be 15%. Baethge et al assessed 71 out of 97 letters to *Deutsches Ärzteblatt* as critical of the paper they discussed. (18) Horton listed multiple serious problems identified by letter writers, and noted that "important weaknesses in these trials" highlighted by published letters to the editor "were ignored in subsequently published practice guidelines". (24) Goldacre et al reported that in response to their 23 letters about misreported trials, only 1 trial issued a correction, although 11 replies contained at least some admission of error.

References

1. Hesselmann F, Graf V, Schmidt M, Reinhart M. The visibility of scientific misconduct: A review of the literature on retracted journal articles. Curr Sociol. 2017;65(6):814–45.

2. Mott A, Fairhurst C, Torgerson D. Assessing the impact of retraction on the citation of randomized controlled trial reports: an interrupted time-series analysis. J Health Serv Res Policy. 2019 Jan;24(1):44–51.

3. Ma J, Sun X, Zhao Y, Tian J. Why 196 Cochrane Systematic Reviews were withdrawn in Cochrane Database of systematic review? In: Abstracts of the 21st Cochrane Colloquium, Quebec City, Canada Cochrane Database of Systematic Reviews [Internet]. Quebec: Cochrane Collaboration; 2013. p. 56–7. Available from: https://abstracts.cochrane.org/2013-qu%C3%A9bec-city/why-196-cochrane-systematic-reviews-were-withdrawn-cochrane-database-systematic

4. Abritis AJ. An Assessment of Retractions as a Measure of Scientific Misconduct and Impact on Public Health Risks [Internet]. Tampa: University of South Florida; 2015. (Graduate Theses and Dissertations). Available from: https://scholarcommons.usf.edu/etd/5630/

5. Vaught M, Jordan DC, Bastian H. Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central. Res Integr Peer Rev [Internet]. 2017;2. Available from: https://researchintegrityjournal.biomedcentral.com/articles/10.1186/s41073-017-0030-2

6. Tybaert S. MEDLINE Data Changes—2018. NLM Tech Bull. 2017 Dec;(419):e5a.

7. Farrah K, Rabb D. Errata for trial publications are not uncommon, are frequently not trivial, and can be challenging to access: a retrospective review. J Med Libr Assoc JMLA. 2019;107(2):187–93.

8. Royle P, Waugh N. Should systematic reviews include searches for published errata? Health Inf Libr J. 2004 Mar;21(1):14–20.

9. Molckovsky A, Vickers MM, Tang PA. Characterization of published errors in high-impact oncology journals. Curr Oncol Tor Ont. 2011 Jan;18(1):26–32.

10. Trikalinos TA. Does it mean anything if your own name is wrong in your published paper? FASEB J Off Publ Fed Am Soc Exp Biol. 2009 Aug;23(8):2345–8.

11. Hauptman PJ, Armbrecht ES, Chibnall JT, Guild C, Timm JP, Rich MW. Errata in medical publications. Am J Med. 2014 Aug;127(8):779-785.e1.

12. Castillo M, Northam M, Halm KE. Postpublication errors in imaging-related journals. AJNR Am J Neuroradiol. 2012 Sep;33(8):1447–8.

13. Strothmann M. No Bad Publicity: The Bibliometric Advantage of Corrected Articles in PLOS ONE. Sci Technol Libr. 2018 Oct 2;37(4):394–408.

14. Grcar J. Comments and Corrigenda in Scientific Literature. Am Sci. 2013 Feb;101(1):16.

15. Fanelli D. Why growing retractions are (mostly) a good sign. PLoS Med. 2013 Dec;10(12):e1001563.

16. Erfanmanesh M, Teixeira da Silva JA. Is the soundness-only quality control policy of open access mega journals linked to a higher rate of published errors? Scientometrics. 2019 Aug;120(2):917–23.

17. Marasović T, Utrobičić A, Marušić A. Transparency of retracting and replacing articles. Lancet Lond Engl. 2018 31;391(10127):1244–5.

18. Baethge C, Seger G. Our readers' voice. Letters to the editor are an important component of the discussion of scientific articles, in Deutsches Arzteblatt as in other journals. our correspondence pages reflect a diversity of opinion thanks to the love of debate among our readers--and thanks to a few rules. Dtsch Arzteblatt Int. 2009 Mar;106(12):207–9.

19. Kastner M, Menon A, Straus SE, Laupacis A. What do letters to the editor publish about randomized controlled trials? A cross-sectional study. - PubMed - NCBI. BMC Res Notes. 2013 Oct;6:414.

20. Ober M, Shohara R, Rennie D, Yank V, Bero LA. The Criticism Management System for the Cochrane Library. Eval Health Prof. 2002 Mar;25(1):98–115.

21. Wiebe DJ, Flynn K, Branas CC. Letters to the editor in response to studies of guns in the home and homicide and suicide. Inj Epidemiol. 2017 Dec;4(1):3.

22. von Elm E, Wandel S, Jüni P. The role of correspondence sections in post-publication peer review: A bibliometric study of general and internal medicine journals. Scientometrics. 2009 Dec;(81):747.

23. Gøtzsche PC, Delamothe T, Godlee F, Lundh A. Adequacy of authors' replies to criticism raised in electronic letters to the editor: cohort study. BMJ. 2010 Aug 10;341:c3926.

24. Horton R. Postpublication criticism and the shaping of clinical knowledge. JAMA. 2002 Jun 5;287(21):2843–7.

25. Goldacre B, Drysdale H, Marston C, Mahtani KR, Dale A, Milosevic I, et al. COMPare: Qualitative analysis of researchers' responses to critical correspondence on a cohort of 58 misreported trials. Trials. 2019 Feb 14;20(1):124.

26. Garmendia CA, Nassar Gorra L, Rodriguez AL, Trepka MJ, Veledar E, Madhivanan P. Evaluation of the Inclusion of Studies Identified by the FDA as Having Falsified Data in the Results of Meta-analyses: The Example of the Apixaban Trials. JAMA Intern Med. 2019 Mar 4;

27. Fanelli D, Moher D. What difference do retractions make? An estimate of the epistemic impact of retractions on recent meta-analyses [Internet]. bioRxiv; 2019. Available from: https://www.biorxiv.org/content/10.1101/734137v1

28. Bhatt VR, Aryal MR, Panta S, Mosalpuria K, Armitage JO. A retrospective analysis of reported errata in five leading medical journals in 2012. J Community Hosp Intern Med Perspect. 2014;4(5):25738.

Appendix 6.

Key to research integrity searches, and instructions for adding custom filter to PubMed.

A. Instructions for custom filters in PubMed.

You can turn any search term or search string you can use in PubMed into a custom filter. When activated, your custom filter(s) will run with every PubMed search you do, giving you the option of checking the subset of any filtered results with one click. In the new PubMed (from November 2019), custom filtered results show at the top left hand side, just next to the first PubMed record in your complete search. In “legacy” PubMed, the custom filters appear at the top on the right.

To set up and activate custom filters:

1. You need a My NCBI account: https://www.ncbi.nlm.nih.gov/sites/myncbi/ . The filters will only work while you are signed into your account.
2. In My NCBI, scroll down to the “Filters” box, and click on “Manage filters”. (Alternative: click on “NCBI Site References” to the top right, and look for “Filters and icons” to arrive at the same place.)
3. Click on the blue “Create custom filter” button.
4. The search filter you want to add goes into the large box “Query terms”. You need to give it a name – like “Retractions” – that will appear next to your PubMed searches, hyperlinked to the filtered search results. That name goes into the small box “Save filter as”. (If you don’t specify something, that box will auto-populate. You can edit it at any time.)
5. There is a “Test This Query” button: that will tell you the number of hits in the whole of PubMed from your filter.
6. Save the filter. You should now see it on “Your PubMed filter list”. Click “Active” to activate the filter.

7. To edit the words that appear next to your search, return to the My NCBI Filters page (number 2 above), and click on the gear symbol. To de-activate the filter, click on the “Active” box again. To delete it completely, click “delete”.

There is more about custom filters in NCBI’s guide. (1)

B. Key to components of the searches.

Ellipsis (…) indicates that a filter has been broken into a chunk for explanation.

The option that contains all the components for that category of post-publication event is shown.

Explanations for repeating components are not repeated for chunks within a type.

Name	Search strategy	Explanation
Retractions		
Retraction 1	hasretractionof OR hasretractionin OR …	Indicates relationship to another record that has been linked (or that one record serves both purposes). “of” indicates it is a retraction notice, “in” links from a retracted publication to a retraction notice.
	retraction of publication[pt] OR retracted publication[pt] …	[pt] = publication type (NLM Medical Subject Heading (MeSH) indexed) (2). “Retraction of publication” is a retraction notice; “Retracted publication” is a publication.
	retracted[ti] OR retraction[ti] OR retracted[tt] OR retraction[tt] OR …	[ti] term appears in record title [tt] terms appears in a transliterated title of a record (eg the record was not in

		English) (3)
	www.elsevier.com/locate/withdrawalpolicy OR (Elsevier* AND policy* AND article* AND (remov* OR withdraw)) OR (withdrawn[ti] OR withdrawn[tt]) OR (withdrawn[tiab] AND ….	Withdrawn reviews, including those under Elsevier policy [ti] and [tt] (see above) [tiab] term appears in the title or abstract of a record. Parentheses group terms to narrow down results specifically. The asterisk (*) is a wild card in PubMed: on a full word it prevents PubMed expanding to synonyms or a full MeSH set for the word; on a truncated word, it searches for options (eg remov* will retrieve remove, removed, removal, etc).
	(article[tiab] OR articles[tiab] OR e-publication[tiab] OR e-publications[tiab] OR epub[tiab] OR "ahead of print"[tiab] OR manuscript[tiab] OR manuscripts[tiab] OR publication[tiab] OR publications[tiab])) OR …	E-publications ahead of print via text terms. Quotation marks ("") force PubMed to search for that phrase (if the phrase is recognised by PubMed).
	(temporary[ti] AND removal[ti]) OR (temporary[tt] AND removal[tt])	Temporary removal of records by publishers.
Expressions of concern (EOC)		
EOC 1	hasexpressionofconcernin OR hasexpressionofconcernfor OR …	Indicates relationship to another record that has been linked (or that one record serves both purposes). "in" is on a publication affected by an EOC, linking to the EOC. "for" links an EOC to the affected publication.
	("expression of concern"[all fields] OR "notice of concern"[all fields] OR "note of concern"[all fields] or "statement of concern"[all fields]) OR …	[allfields] means every part of the MEDLINE record will be searched. Quotation marks ("") force PubMed to search for that phrase (if the phrase is

		recognised by PubMed).
Errata		
Errata 1	published erratum[pt] OR …	[pt] = publication type (NLM Medical Subject Heading (MeSH) indexed) (2). Published erratum indicates the record is an erratum notice.
	haserratumfor OR haserratumin OR …	Indicates relationship to another record that has been linked (or that one record serves both purposes). “for” indicates it is an erratum notice, “in” links from an affected publication to an erratum notice.
	erratum*[ti] OR correction[ti] OR corrigendum*[ti] OR errata*[ti] OR corrections[ti] OR corrigenda*[ti] OR erratum*[tt] OR correction[tt] OR corrigendum*[tt] OR errata*[tt] OR corrections[tt] OR corrigenda*[tt]	[ti] term appears in record title [tt] terms appears in a transliterated title of a record (eg the record was not in English) (3). All the words in this chunk will be searched for in titles (therefore picking up publications where surgical correction is the subject for example, as well as errata). The asterisk (*) is a wild card in PubMed: on a full word it prevents PubMed expanding to synonyms or a full MeSH set for the word. (The wild card can also truncate a word, so that variants are searched.)
Corrected and republished		
Republished 1	"corrected and republished" OR …	Quotation marks (“”) force PubMed to search for that phrase (if the phrase is recognised by PubMed). This chunk does not specify where the phrase could appear.

	hascorrectedrepublishedin OR hascorrectedrepublishedfrom OR …	Indicates relationship to another record that has been linked (or that one record serves both purposes). “in” indicates the original publication that has been corrected and republished in the record to which it links, “from” is the republished article, linking back to the original.
	corrected and republished article[pt]	[pt] = publication type (NLM Medical Subject Heading (MeSH) indexed) (2). This is the publication type for a publication that has been republished.
ORI findings of misconduct		
ORI	"NIH Guide Grants Contracts"[jo]	[jo] = journal (can also be jour or journal). This is a journal that catalogues ORI findings. Select or save the MEDLINE display option for the record(s) in “legacy” PubMed. The CON field (abbreviation for comment on) has a citation and PMID (PubMed ID) for cited publications. Screen these to see the subject of the finding. Everything cited is not necessarily the subject of the finding.
Letters to the editor (LTE)		
LTE	hascommentin …	Indicates relationship to another record that has been linked. This is a publication that has some form of commentary, not necessarily a letter to the editor about a publication. Select or save the MEDLINE display option for the record(s) in

		"legacy" PubMed. The CIN field (abbreviation for comment in) has a citation and PMID for the comment. Search for the PMID(s) and limit with the next filter…
	(Second step) letter[pt]	[pt] = publication type (NLM Medical Subject Heading (MeSH) indexed) (2). This is the publication type for a letter. As it has already been restricted to a comment, it is very likely to be a letter to the editor about the publication(s) from the first step.
Author replies		
Replies	hascommentin …	First step as above for LTE.
	letter[pt] AND reply	As above for LTE, but with the addition of "reply" without any specification of where it could appear in the record or in what form.
Non-specific		
Prolonged epub status	pubstatusaheadofprint	Retrieves records that have the status epublished ahead of a print issue. These records are added to PubMed, but are not indexed until they are added to a print issue and the record has its status changed. An "old" record with this status could either never have entered a print issue, or a new PubMed record was generated instead of converting the "epub" record. Used time limited for 2002 until two years ago.

Footnote: PMC indexes and link retracted publications and retraction notices, searchable by:

retraction[filter]
is retracted[filter]

References

1. Information NC for B, Pike USNL of M 8600 R, MD B, Usa 20894. Working with Filters [Internet]. National Center for Biotechnology Information (US); 2018 [cited 2019 Sep 3]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK53591/

2. U.S. National Library of Medicine. Publication Characteristics (Publication Types) with Scope Notes [Internet]. NLM Medical Subject Headings. 2018 [cited 2019 Sep 15]. Available from: https://www.nlm.nih.gov/mesh/pubtypes.html

3. Canese K. Transliterated/Vernacular Title Search Added to PubMed®. NLM Tech Bull. 2006 Feb;(348):e4.

Appendix 7.

Assessment of NLM coverage of errata reference set.

A. Search strategy

date of search	19-Sep-19
search string	haserratumin AND ("N Engl J Med"[jour] OR "Lancet"[jour] OR "BMJ"[jour] OR "Circulation"[jour] OR "JAMA"[jour] OR "Ann Intern Med"[jour] OR "Arch Intern Med"[jour] OR "J Am Coll Cardiol"[jour] OR "Am Heart J"[jour] OR "Heart Rhythm"[jour] OR "Am J Prev Med"[jour] OR "Eur Heart J"[jour] OR "Am J Med"[jour] OR "Eur J Heart Fail"[jour] OR "Heart"[jour] OR "Am J Cardiol"[jour] OR "Ann Thorac Surg"[jour] OR "Prev Med"[jour] OR "J Intern Med"[jour] OR "J Cardiovasc Electrophysiol"[jour])
time limit	2009/07/01 to 2010/12/31
hits	521
search string for publication type letter	haserratumin AND letter[pt] AND ("N Engl J Med"[jour] OR "Lancet"[jour] OR "BMJ"[jour] OR "Circulation"[jour] OR "JAMA"[jour] OR "Ann Intern Med"[jour] OR "Arch Intern Med"[jour] OR "J Am Coll Cardiol"[jour] OR "Am Heart J"[jour] OR "Heart Rhythm"[jour] OR "Am J Prev Med"[jour] OR "Eur Heart J"[jour] OR "Am J Med"[jour] OR "Eur J Heart Fail"[jour] OR "Heart"[jour] OR "Am J Cardiol"[jour] OR "Ann Thorac Surg"[jour] OR "Prev Med"[jour] OR "J Intern Med"[jour] OR "J Cardiovasc Electrophysiol"[jour])
search string for news	haserratumin AND news[pt] AND ("N Engl J Med"[jour] OR "Lancet"[jour] OR "BMJ"[jour] OR "Circulation"[jour] OR "JAMA"[jour] OR "Ann Intern Med"[jour] OR "Arch Intern Med"[jour] OR "J Am Coll Cardiol"[jour] OR "Am Heart J"[jour] OR "Heart Rhythm"[jour] OR "Am J Prev Med"[jour] OR "Eur Heart J"[jour] OR "Am J Med"[jour] OR "Eur J Heart Fail"[jour] OR "Heart"[jour] OR "Am J Cardiol"[jour] OR "Ann Thorac Surg"[jour] OR "Prev Med"[jour] OR "J Intern Med"[jour] OR "J Cardiovasc Electrophysiol"[jour])
search string for editorials	haserratumin AND hascommenton AND ("N Engl J Med"[jour] OR "Lancet"[jour] OR "BMJ"[jour] OR "Circulation"[jour] OR "JAMA"[jour] OR "Ann Intern Med"[jour] OR "Arch Intern Med"[jour] OR "J Am Coll Cardiol"[jour] OR "Am Heart J"[jour] OR "Heart Rhythm"[jour] OR "Am J Prev Med"[jour] OR "Eur Heart J"[jour] OR "Am J Med"[jour] OR "Eur J Heart Fail"[jour] OR "Heart"[jour] OR "Am J Cardiol"[jour] OR "Ann Thorac Surg"[jour] OR "Prev Med"[jour] OR "J Intern Med"[jour] OR "J Cardiovasc Electrophysiol"[jour])
excluded	70
included	451
retracted among included	0

META-DATA

Journals:

name	Journal name as appeared in Hauptman et al
pubmed_search	Journal identifier used in PubMed search

Excluded:

reason	Reason for exclusion
reason - Publication date	Excluded because publication date outside search period
reason - Letter to editor - non-research	Excluded because non-research letter to the editor
reason - News	Excluded because news report
reason - Late erratum	Excluded because erratum date after June 2012
citation	Citation details from PubMed download
Pmid	PubMed ID
date_author	PubMed create date and first author

Included:

citation	Citation details from PubMed download
Pmid	PubMed ID
date_author	PubMed create date and first author
research_letter	Identified by publication type letter search; assessed as research letter

B. Key to journals

name	pubmed_search
New England Journal of Medicine	"N Engl J Med"[jour]
Lancet	"Lancet"[jour]
British Medical Journal	"BMJ"[jour]
Circulation	"Circulation"[jour]
JAMA	"JAMA"[jour]
Annals of Internal Medicine	"Ann Intern Med"[jour]
Archives of Internal Medicine	"Arch Intern Med"[jour]
Journal of the American College of Cardiology	"J Am Coll Cardiol"[jour]
American Heart Journal	"Am Heart J"[jour]
Heart Rhythm	"Heart Rhythm"[jour]
American Journal of Preventive Medicine	"Am J Prev Med"[jour]
European Heart Journal	"Eur Heart J"[jour]
American Journal of Medicine	"Am J Med"[jour]

European Journal of Heart Failure	"Eur J Heart Fail"[jour]
Heart	"Heart"[jour]
American Journal of Cardiology	"Am J Cardiol"[jour]
Annals of Thoracic Surgery	"Ann Thorac Surg"[jour]
Preventive Medicine	"Prev Med"[jour]
Journal of Internal Medicine	"J Intern Med"[jour]
Journal of Cardiovascular Electrophysiology	"J Cardiovasc Electrophysiol"[jour])

C. Excluded PubMed records (n = 70).

reason	citation	pmid
Publication date	Heart. 2011 Feb;97(4):287-94. doi: 10.1136/hrt.2010.205815. Epub 2010 Dec 30. Erratum in: Heart. 2011 Mar;97(6):516. Xie, Jun-Mei [corrected to Xie, Jun-Min].	21193686
Publication date	Eur Heart J. 2011 Sep;32(17):2168-78. doi: 10.1093/eurheartj/ehq467. Epub 2010 Dec 22. Erratum in: Eur Heart J. 2013 Feb;34(8):604.	21183501
Publication date	Lancet. 2011 Jan 1;377(9759):42-51. doi: 10.1016/S0140-6736(10)62175-7. Epub 2010 Dec 20. Erratum in: Lancet. 2011 Apr 2;377(9772):1154.	21176949
Publication date	Circulation. 2011 Jan 4;123(1):104-23. doi: 10.1161/CIR.0b013e3181fa3cf4. Epub 2010 Dec 20. No abstract available. Erratum in: Circulation. 2011 Aug 2;124(5):e173.	21173346
Publication date	Circulation. 2011 Feb 1;123(4):e18-e209. doi: 10.1161/CIR.0b013e3182009701. Epub 2010 Dec 15. No abstract available. Erratum in: Circulation. 2011 Feb 15;123(6):e240. Circulation. 2011 Oct 18;124(16):e426.	21160056
Publication date	J Intern Med. 2011 Jan;269(1):88-93. doi: 10.1111/j.1365-2796.2010.02273.x. Epub 2010 Sep 24. Erratum in: J Intern Med. 2011 Mar;269(3):363.	21158981
Publication date	Heart. 2011 Jan;97(2):91-3. doi: 10.1136/hrt.2010.205096. Epub 2010 Dec 13. No abstract available. Erratum in: Heart. 2011 Jul;97(13):1112.	21149862
Publication date	Heart Rhythm. 2011 Apr;8(4):493-9. doi: 10.1016/j.hrthm.2010.11.046. Epub 2010 Nov 30. Erratum in: Heart Rhythm. 2011 Jun;8(6):953.	21129504
Publication date	Eur Heart J. 2011 Feb;32(4):437-42. doi: 10.1093/eurheartj/ehq438. Epub 2010 Nov 30. Erratum in: Eur Heart J. 2013 Oct;34(40):3160. David Batty, G [corrected to Batty, G David].	21118851
Publication date	Lancet. 2011 Mar 26;377(9771):1113-21. doi: 10.1016/S0140-6736(10)61961-7. Epub 2010 Nov 10. Erratum in: Lancet. 2011 Mar 26;377(9771):1076.	21074256
Publication date	Heart Rhythm. 2011 Mar;8(3):377-82. doi: 10.1016/j.hrthm.2010.11.013. Epub 2010 Nov 9. Erratum in: Heart Rhythm. 2011 Apr;8(4):640.	21070886
Publication date	Heart Rhythm. 2011 Feb;8(2):221-7. doi: 10.1016/j.hrthm.2010.10.010. Epub 2010 Oct 14. Erratum in: Heart Rhythm. 2011 Jun;8(6):953. Papiashvilli, Giorgi [corrected to Papiashvili, Giorgi].	20951231
Publication	Heart Rhythm. 2011 Feb;8(2):199-206. doi:	20950713

date	10.1016/j.hrthm.2010.10.015. Epub 2010 Oct 13. Erratum in: Heart Rhythm. 2011 Apr;8(4):640.	
Publication date	J Intern Med. 2011 Mar;269(3):289-98. doi: 10.1111/j.1365-2796.2010.02280.x. Epub 2010 Sep 10. Erratum in: J Intern Med. 2011 Jul;270(1):85-6.	20831629
Publication date	Lancet. 2011 Jan 29;377(9763):393-402. doi: 10.1016/S0140-6736(10)60959-2. Epub 2010 Aug 23. Erratum in: Lancet. 2011 Jan 29;377(9763):382.	20739054
Letter to editor - non-research	Am Heart J. 2009 Jul;158(1):e15. doi: 10.1016/j.ahj.2009.05.009. No abstract available. Erratum in: Am Heart J. 2009 Oct;158(4):666.	19540379
Letter to editor - non-research	JAMA. 2009 Jul 8;302(2):142; author reply 145-6. doi: 10.1001/jama.2009.908. No abstract available. Erratum in: JAMA. 2009 Aug 19;302(7):739.	19584334
Letter to editor - non-research	N Engl J Med. 2009 Jul 9;361(2):206-7; author reply 207-8. No abstract available. Erratum in: N Engl J Med. 2013 Jan 31;368(5):490.	19593855
Letter to editor - non-research	Arch Intern Med. 2009 Jul 13;169(13):1244. doi: 10.1001/archinternmed.2009.182. No abstract available. Erratum in: Arch Intern Med. 2011 Sep 26;171(17):1540. Costa, Sandra [corrected Fuchs, Sandra C].	19597077
Letter to editor - non-research	JAMA. 2009 Aug 5;302(5):487-8; author reply 488. doi: 10.1001/jama.2009.1080. No abstract available. Erratum in: JAMA. 2009 Sep 16;302(11):1170.	19654380
Letter to editor - non-research	BMJ. 2009 Sep 22;339:b3856. doi: 10.1136/bmj.b3856. No abstract available. Erratum in: BMJ. 2009;339. doi: 10.1136/bmj.b4044. Vahdani, Kave [corrected to Vahdani, Kaveh].	19773311
Letter to editor - non-research	Heart. 2009 Nov;95(21):1802; author reply 1802. doi: 10.1136/hrt.2009.179283. No abstract available. Erratum in: Heart. 2009 Dec;95(23):1953.	19828823
Letter to editor - non-research	JAMA. 2009 Nov 4;302(17):1862; author reply 1862-3. doi: 10.1001/jama.2009.1580. No abstract available. Erratum in: JAMA. 2010 Feb 24;303(8):734.	19887663
Letter to editor - non-research	BMJ. 2009 Nov 24;339:b4972. doi: 10.1136/bmj.b4972. No abstract available. Erratum in: BMJ. 2009;339:b5312.	19934198
Letter to editor - non-research	Lancet. 2009 Dec 12;374(9706):1966. doi: 10.1016/S0140-6736(09)62118-8. No abstract available. Erratum in: Lancet. 2010 Jan 23;375(9711):282.	20006129
Letter to editor - non-research	BMJ. 2009 Dec 10;339:b5340. No abstract available. Erratum in: BMJ. 2010;340:c369.	20007234
Letter to editor - non-research	Ann Intern Med. 2009 Dec 15;151(12):895-7. doi: 10.7326/0003-4819-151-12-200912150-00020. No abstract available. Erratum in: Ann Intern Med. 2010 Feb 2;152(3):200.	20008769
Letter to editor - non-	JAMA. 2009 Dec 16;302(23):2550; author reply 2550-1. doi: 10.1001/jama.2009.1826. No abstract available. Erratum in: JAMA. 2010 Mar 17;303(11):1036.	20009050

research		
Letter to editor - non-research	J Am Coll Cardiol. 2010 Feb 2;55(5):502-3. doi: 10.1016/j.jacc.2009.10.025. No abstract available. Erratum in: J Am Coll Cardiol. 2010 Mar 30;55(13):1401.	20117470
Letter to editor - non-research	BMJ. 2010 Feb 16;340:c591. doi: 10.1136/bmj.c591. No abstract available. Erratum in: BMJ. 2010;340:c1135.	20160317
Letter to editor - non-research	Lancet. 2010 Feb 27;375(9716):722-3; author reply 723. doi: 10.1016/S0140-6736(10)60299-1. No abstract available. Erratum in: Lancet. 2011 Nov 19;378(9805):1778.	20189018
Letter to editor - non-research	Lancet. 2010 Apr 10;375(9722):1250-1; author reply 1252-3. doi: 10.1016/S0140-6736(10)60534-X. No abstract available. Erratum in: Lancet. 2010 Jul 3;376(9734):22.	20382317
Letter to editor - non-research	N Engl J Med. 2010 May 6;362(18):1738-9; author reply 1739-40. No abstract available. Erratum in: N Engl J Med. 2010 Jul 15;363(3):298.	20449878
Letter to editor - non-research	N Engl J Med. 2010 Jul 22;363(4):394-5. doi: 10.1056/NEJMc1005898. No abstract available. Erratum in: N Engl J Med. 2010 Nov 25;363(22):2178.	20660410
Letter to editor - non-research	BMJ. 2010 Sep 28;341:c5303. doi: 10.1136/bmj.c5303. No abstract available. Erratum in: BMJ. 2010;341. doi: 10.1136/bmj.c5495.	20880916
Letter to editor - non-research	Lancet. 2010 Oct 16;376(9749):1301. doi: 10.1016/S0140-6736(10)61916-2. No abstract available. Erratum in: Lancet. 2011 Mar 12;377(9769):904.	20951890
News	BMJ. 2009 Jul 21;339:b2936. doi: 10.1136/bmj.b2936. No abstract available. Erratum in: BMJ. 2009;339:b3085.	19622555
News	Lancet. 2009 Aug 8;374(9688):442. No abstract available. Erratum in: Lancet. 2009 Oct 3;374(9696):1148.	19681181
News	Lancet. 2009 Sep 12;374(9693):868. No abstract available. Erratum in: Lancet. 2009 Nov 7;374(9701):1596.	19757511
News	Lancet. 2009 Sep 19;374(9694):963. No abstract available. Erratum in: Lancet. 2009 Nov 7;374(9701):1596.	19774704
News	Lancet. 2009 Sep 26;374(9695):1053-4. No abstract available. Erratum in: Lancet. 2009 Nov 21-2009 Nov 27;374(9703):1744.	19790287
News	Lancet. 2010 Jan 16;375(9710):187-8. doi: 10.1016/S0140-6736(10)60076-1. No abstract available. Erratum in: Lancet. 2010 Feb 20;375(9715):640.	20109905
News	JAMA. 2010 Apr 28;303(16):1586. doi: 10.1001/jama.2010.455. No abstract available. Erratum in: JAMA. 2010 May 26;303(20):2034.	20424241
News	Lancet. 2010 Apr 24;375(9724):1423. No abstract available. Erratum in: Lancet. 2010 Jul 31;376(9738):332.	20425882
News	JAMA. 2010 Jul 21;304(3):257-8. doi: 10.1001/jama.2010.961. No abstract available. Erratum in: JAMA. 2010 Sep 22,304(12):1330.	20639553
News	JAMA. 2010 Aug 18;304(7):729-31. doi: 10.1001/jama.2010.1062. No abstract available. Erratum in: JAMA. 2010 Oct 13;304(14):1557.	20716730
News	Lancet. 2010 Oct 23;376(9750):1377-8. No abstract available. Erratum in: Lancet. 2011 Mar 5;377(9768):812.	20979325

	JAMA. 2010 Dec 15;304(23):2576-7. doi: 10.1001/jama.2010.1738. No abstract available. Erratum in:	
News	JAMA. 2011 Feb 14;305(5):470.	21156940
Late erratum	Heart. 2009 Dec;95(24):2023-8. doi: 10.1136/hrt.2009.170241. Epub 2009 Aug 6. Erratum in: Heart. 2012 Jul;98(14):1108.	19666460
Late erratum	BMJ. 2009 Sep 7;339:b3354. doi: 10.1136/bmj.b3354. Review. Erratum in: BMJ. 2013;346:f151.	19736282
Late erratum	Circulation. 2009 Sep 22;120(12):1056-64. doi: 10.1161/CIRCULATIONAHA.108.827121. Epub 2009 Sep 8. Erratum in: Circulation. 2014 Jul 29;130(5):e43.	19738141
Late erratum	Lancet. 2009 Oct 3;374(9696):1160-70. doi: 10.1016/S0140-6736(09)61558-0. Epub 2009 Sep 14. Erratum in: Lancet. 2014 Jan 25;383(9914):308.	19758692
Late erratum	Heart. 2010 May;96(9):662-7. doi: 10.1136/hrt.2009.177162. Epub 2009 Sep 23. Erratum in: Heart. 2014 Feb;100(4):350.	19778920
Late erratum	BMJ. 2009 Oct 26;339:b4248. doi: 10.1136/bmj.b4248. Erratum in: BMJ. 2016 May 27;353:i3039.	19858174
Late erratum	JAMA. 2009 Nov 11;302(18):1985-92. doi: 10.1001/jama.2009.1568. Erratum in: JAMA. 2012 Nov 21;308(19):1973.	19903919
Late erratum	N Engl J Med. 2009 Dec 3;361(23):2241-51. doi: 10.1056/NEJMoa0909938. Epub 2009 Nov 17. Erratum in: N Engl J Med. 2018 Aug 16;379(7):697.	19920051
Late erratum	Circulation. 2009 Dec 15;120(24):2470-7. doi: 10.1161/CIRCULATIONAHA.109.867929. Erratum in: Circulation. 2012 Sep 11;126(11):e173.	19948974
Late erratum	Arch Intern Med. 2010 Jan 25;170(2):136-45. doi: 10.1001/archinternmed.2009.492. Erratum in: JAMA Intern Med. 2015 Mar;175(3):470.	20101008
Late erratum	Lancet. 2010 Feb 13;375(9714):555-62. doi: 10.1016/S0140-6736(10)60101-8. Epub 2010 Jan 29. Erratum in: Lancet. 2014 Oct 25;384(9953):1504.	20116841
Late erratum	BMJ. 2010 Mar 11;340:c1040. doi: 10.1136/bmj.c1040. Erratum in: BMJ. 2016 May 18;353:i2823.	20223872
Late erratum	J Am Coll Cardiol. 2010 Apr 6;55(14):e27-e129. doi: 10.1016/j.jacc.2010.02.015. No abstract available. Erratum in: J Am Coll Cardiol. 2013 Sep 10;62(11):1039-40.	20359588
Late erratum	BMJ. 2010 May 17;340:c2220. doi: 10.1136/bmj.c2220. Erratum in: BMJ. 2016 Nov 29;355:i6423.	20478960
Late erratum	BMJ. 2010 May 25;340:c2265. doi: 10.1136/bmj.c2265. Erratum in: BMJ. 2016 May 27;353:i3041.	20501583
Late erratum	BMJ. 2010 Jun 17;340:c3187. doi: 10.1136/bmj.c3187. No abstract available. Erratum in: BMJ. 2015;350:h3495.	20558517
Late erratum	Eur J Heart Fail. 2010 Jul;12(7):721-9. doi: 10.1093/eurjhf/hfq095. Erratum in: Eur J Heart Fail. 2013 Mar;15(3):360.	20576835
Late erratum	BMJ. 2010 Jun 25;340:c3134. doi: 10.1136/bmj.c3134. No abstract available. Erratum in: BMJ. 2019 Aug 23;366:l4910.	20581031
Late erratum	BMJ. 2010 Aug 25;341:c4130. doi: 10.1136/bmj.c4130. No abstract available. Erratum in: BMJ. 2018 Sep 26;362:k4080.	20739363
Late erratum	Lancet. 2010 Oct 30;376(9751):1476-84. doi: 10.1016/S0140-6736(10)60960-9. Epub 2010 Sep 7. Erratum in: Lancet. 2016 Aug 6;388(10044):564. Dosage error in published abstract; MEDLINE/PubMed abstract corrected; Dosage error in article text.	20825986
Late erratum	Circulation. 2010 Oct 19;122(16 Suppl 2):S298-324. doi: 10.1161/CIRCULATIONAHA.110.970996. No abstract available. Erratum in: Circulation. 2013 Nov 5;128(19):e393.	20956253
Late	J Am Coll Cardiol. 2010 Dec 7;56(24):2006-20. doi:	21126642

erratum 10.1016/j.jacc.2010.05.059. Erratum in: J Am Coll Cardiol. 2016 Apr 5;67(13):1660.

D. PubMed IDs for included records (n = 451).

19303630	19699857	19815268	19952036	20164481	20375410	20538127	20696973	20956206
19491172	19709732	19819004	19959221	20171405	20375417	20538140	20702548	20956221
19560810	19710486	19819542	19959589	20171823	20381674	20546887	20723641	20956222
19564629	19710495	19822821	19962480	20172927	20381676	20547907	20728210	20956224
19571281	19712806	19822822	19995815	20177011	20382326	20548004	20733095	20956225
19571284	19713419	19828525	19995818	20177041	20386001	20548009	20733105	20956226
19571295	19716960	19833707	20018364	20177042	20386013	20550610	20800156	20956227
19584113	19717184	19837741	20019324	20179132	20388896	20551152	20801500	20956228
19586656	19717844	19841457	20026784	20179285	20393179	20552718	20802247	20956231
19596718	19720936	19843903	20026787	20181637	20399327	20554986	20805436	20956233
19597050	19721018	19843906	20032318	20181970	20399974	20557927	20816131	20956261
19597073	19724045	19850208	20032324	20181977	20400552	20558370	20816531	20956487
19602691	19726774	19850702	20034972	20185494	20404381	20569726	20816540	20961244
19602713	19733768	19850703	20040556	20188410	20409500	20569762	20818854	20961248
19605831	19738093	19853696	20040562	20189239	20410499	20569843	20818856	20974996
19605834	19738096	19853705	20042483	20193826	20410513	20570343	20818865	20979469
19616680	19741227	19853906	20061357	20197539	20413028	20570344	20818888	21035838
19616701	19744618	19854499	20065199	20200386	20417052	20570952	20818894	21041579
19616706	19752383	19856489	20068215	20202522	20417778	20571020	20818903	21045098
19620161	19755696	19861667	20071021	20206760	20417837	20573924	20818911	21047223
19620164	19755703	19861668	20075331	20211292	20417856	20581392	20820036	21059976
19620503	19759377	19864673	20083621	20212177	20421558	20587560	20822810	21060079
19625712	19761931	19878986	20083678	20215609	20424254	20587585	20825316	21062666
19633103	19762075	19880172	20083825	20219749	20427809	20587587	20825319	21067805
19638651	19765502	19887653	20083830	20220183	20430113	20594588	20837277	21070923
19641202	19765507	19892392	20085991	20220184	20434591	20595256	20837826	21073557
19641205	19766755	19897665	20089582	20220188	20434763	20605201	20837833	21080835
19647866	19770180	19906418	20089974	20223871	20435194	20609688	20837908	21083382
19657122	19770388	19913903	20089975	20228402	20439577	20609950	20841544	21083386
19657123	19770391	19914707	20096451	20228403	20445182	20609967	20851244	21091062
19657127	19770412	19917887	20100987	20228404	20452521	20620713	20854837	21095277
19661138	19773328	19917888	20102915	20231535	20458084	20625013	20854959	21098446
19665633	19773565	19920001	20102979	20233780	20460620	20625026	20855784	21098448
19667304	19776405	19920235	20103014	20233822	20463332	20638121	20855802	21098771
19667307	19776408	19920237	20103758	20237344	20463338	20638563	20855804	21105788
19671655	19776409	19920272	20107216	20308616	20463342	20639569	20858879	21105794
19671886	19778665	19920274	20107219	20308643	20472173	20643261	20863951	21111891
19671905	19783291	19920276	20109866	20338500	20472938	20644016	20870200	21119084
19671910	19797279	19920278	20117433	20338943	20472939	20647196	20879894	21121834

19683630	19797281	19922995	20124228	20346417	20478405	20647204	20888527	21126614
19683638	19797284	19923152	20130261	20346815	20478865	20650362	20889988	21139111
19683639	19797287	19923169	20142454	20356621	20479152	20654732	20921445	21142535
19686772	19799931	19926002	20145228	20357284	20494435	20660401	20924012	21142537
19690317	19801099	19926021	20145232	20362720	20497980	20660809	20924013	21146205
19690345	19805767	19932790	20145236	20362730	20501927	20673995	20933357	21149752
19691156	19808819	19933514	20152220	20362757	20505173	20679583	20937757	21149763
19692693	19808820	19933955	20152233	20368645	20505178	20688208	20937973	21156948
19695364	19812054	19934408	20152562	20368665	20513599	20691825	20942670	21156954
19695457	19812404	19940300	20152563	20371785	20525992	20692531	20942673	21175315
19699853	19812406	19942100	20154049	20375406	20528971	20692693	20952767	21175316
								21175318