Diabetologia (2023) 66:614–617
https://doi.org/10.1007/s00125-022-05862-8

FOR DEBATE

Missing the forest-plot for the trees

Deirdre K. Tobias [1,2]

Received: 7 November 2022 / Accepted: 5 December 2022 / Published online: 14 January 2023

Abstract

Systematic reviews and meta-analyses are methods increasingly used in biomedical research since their introduction in the 1970s. They serve to replace other non-systematic and cherry-picked narrative reviews, which are highly variable in their approach, structure and content. Their increase in popularity parallels the increase in overall scientific output, and when properly conducted, systematic reviews can contribute highly impactful summaries of a fast-growing evidence base. Meta-analyses offer statistical summaries, called forest plots, which similarly provide a powerful synopsis unachievable by individual studies. Thus, it is not difficult to imagine why systematic reviews are published more often. Should scientists be concerned by the accelerated output of research, from systematic reviews or other? If quantity comes at the expense of quality, then yes, of course; but should important manuscripts be rationed out otherwise? A new scientific technique can seem scary at first, especially to the researcher who is unfamiliar with its application or uncertain of its validity. In that case, we should become familiar with new and popular methods, and understand their strengths and limitations. There is a rightful place for systematic reviews and meta-analyses among respectable research tools. Importantly, however, despite standard operating procedures and best practices, the quality of systematic reviews today is highly variable, warranting serious concerns for quantity exceeding quality. Therefore, the appropriate response should be to instil researchers with an appreciation for the complexity of conducting and interpreting a systematic review and meta-analysis, to create more knowledgeable authors, reviewers and editors, who collectively will improve, rather than dismiss, these important scientific contributions.

Keywords Meta-analysis · Systematic review

Introduction

There is no argument that systematic reviews and meta-analyses have increased both in number and in proportion of overall published literature. For example, there has been a steady rise in overall PubMed citations that include the word ‘diabetes’ (Fig. 1). Among them, the per cent with ‘meta-analysis’ reached an all-time high in 2022, now representing 3.9% of ‘diabetes’ citations (a 78-fold increase since 1988!). This ‘For Debate’ perspective addresses two questions that have been raised: (1) why have systematic reviews and meta-analyses risen in popularity; and (2) should we care?

✉ Deirdre K. Tobias
dtobias@bwh.harvard.edu

[1] Division of Preventive Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA

[2] Nutrition Department, Harvard TH Chan School of Public Health, Boston, MA, USA

‘In the past, meta-analyses were virtually non-existent’ [1]. Indeed, the first seminal work on the systematic review and meta-analysis framework was published by Dr Archibald Cochrane in 1972 (the same year biostatistician Sir David Cox put forth his now ubiquitous survival analysis model) [2]. Dr Cochrane, a Scottish clinician, promoted the value of having a standardised way of synthesising large bodies of evidence on a given clinical research question to optimise and disseminate critical knowledge in our resource-limited healthcare settings [3]. When feasible, a meta-analysis would then be conducted to statistically integrate the individual studies, generating the signature summary figure known as a forest plot [4]. The Cochrane Collaboration, a now well-known authority on systematic review and meta-analysis methods, was created in 1993, building on Dr Cochrane’s premise that ‘well-designed evaluations provide information that is essential for improving policies and decisions in health care and research’ [3].

In just 50 short years, systematic reviews and meta-analyses have been credited with significantly accelerating the use of life-saving therapies [5, 6] and served as the basis

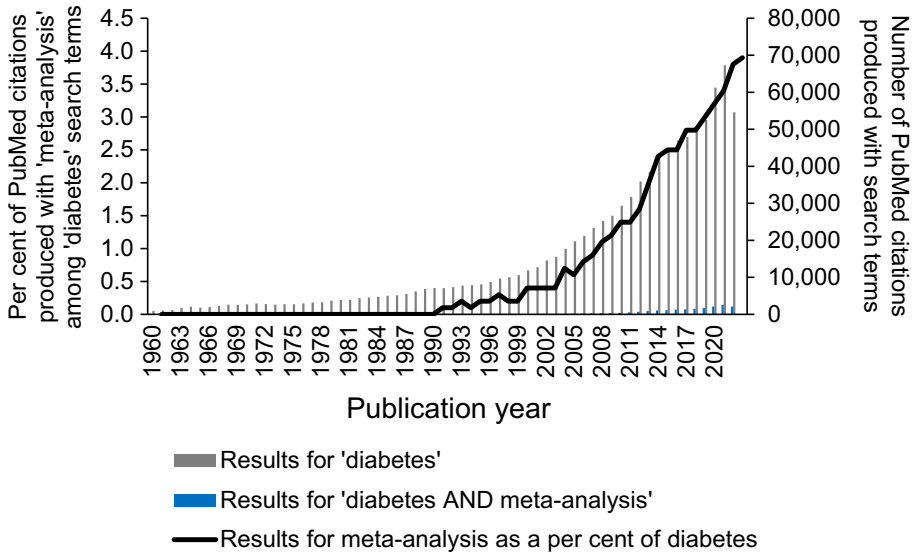

Fig. 1 Briefly, I queried the National Library of Medicine (PubMed) database to ascertain the number of citations per year that are retrieved for two crude, illustrative, search strategies: (1) 'diabetes AND meta-analysis' and (2) 'diabetes'. The publication dates ranged from 1 January 1960 to 31 December 2022 and the search was conducted on 28 October 2022. I then calculated the annual per cent of search 1 citations over the denominator of search 2 citations ('meta-analysis AND diabetes'/'diabetes') to estimate the trend proportion of diabetes literature containing reference to 'meta-analysis'. As seen in the figure, citations for 'meta-analysis AND diabetes' come on the scene with three citations in 1988 and reach a high of 2591 citations in 2021. However, there is also an increase in the total number of 'diabetes' citations, from 920 in 1960 to 67,305 in 2021

for clinical guidelines and public health recommendations [7–9]. Why then would an increase in systematic review papers be something to be concerned about?

Why are forest plots flourishing?

There are endless possible explanations for the dramatic rise in overall research outputs in recent decades, such that quantifying their individual contribution would be impossible. In his 'For Debate', Dr Bonora alleges the proliferation of systematic reviews and meta-analyses, in particular, is because investigators and journal editors have caught on to their citability, and are the result of underlying, even nefarious, motives of journal editors and researchers to boost impact factors and *h*-indexes [1].

There is no evidence that today's scientists are any more motivated by a quick citation boost via 'the least possible effort' [1] than past generations. It would also be difficult to support the assertion that such authors are disproportionately more common among systematic reviews and meta-analyses, compared with any other manuscript type. Thus, it is disheartening that a senior scientist would characterise authors of systematic reviews as exploitive of others' hard work and downgrade their conduct from 'scientist' to 'novelist', based not on evidence, but speculation.

If anything, this highlights a deeper concern that there are scientists who prioritise personal success above integrity. It is alarming when the allure for the 'shortcut to success' manifests as falsified laboratory data, patient negligence and other forms of research misconduct; this is arguably more detrimental to the progress of science than an influx of duplicative, 'nothing to add' papers [1]. However, without actual evidence that systematic review and meta-analysis papers serve as a conduit of research misconduct, we should avoid stigmatising generalisations.

An alternative explanation for the increase in systematic reviews is simply because they have been successfully marketed as an essential research tool. Systematic reviews can benefit virtually all clinical research domains and offer a rigorous alternative to the bias of eminence-based narrative reviews. A systematic review and meta-analysis can be conducted anywhere by anyone, with few resources required. Enhancements such as digital print and Open Access have lowered barriers to dissemination of scientific papers. A meta-analysis can overcome the low number of endpoints and imprecision of individual studies. A systematic review can also highlight paucity of evidence and areas of uncertainty for a given hypothesis. Further, given the massive increase in original scientific output, a rise in systematic reviews is warranted to meaningfully synthesise the growth and diversity of research. For scientists that, like Dr Bonora, find it 'difficult to keep up' with the number of original papers, well-done and timely systematic reviews should be a welcome sight.

Should we deforest-plot the literature?

In his 'For Debate', Dr Bonora calls for strict journal publishing limits and the exclusion of systematic reviews and meta-

analyses from impact rating metrics [1]. He postulates that without such restrictions, the allure of the quick success will continue to be at the grave expense of impairing the development of 'real scientists' [1]. Yet, there is no data to support that learning how to conduct a systematic review deters from developing proficiency in other research methodologies. Further, early career researchers would be at a disservice to be taught that major medical advances are made predominantly one paper at a time, through original experiments and data. Rather, meaningful scientific progress comes as several lines of evidence accumulate over years, and sometimes even whole careers. If a systematic reviewer can spend less than one year with their laptop and a spreadsheet to efficiently inform state-of-the-art scientific progress, why should they reconsider?

Importantly, a blanket wariness of systematic reviews and meta-analyses disregards the meaningful impact that this tool can have. Dr Bonora argues that because systematic reviews are built on other scientists' 'struggles', they cannot be considered innovative, and are borderline, if not overt, plagiarism [1]. Ironically, a prerequisite to conducting a meta-analysis is that at least two studies have been conducted testing the same hypothesis; in fact, Cochrane guidelines set the minimum at ten studies before a meta-analysis is considered reasonable [10]. It is difficult to reconcile the idea that an author of the tenth clinical trial is innovative, but the systematic reviewer is redundant.

We all acknowledge how critical replication is, but even 'original research' is eventually unoriginal. Thus, because of their ability to efficiently update accumulating evidence, almost in real-time, systematic reviews and meta-analyses have the unique ability to identify when a critical mass of knowledge has been reached on a research topic, saving subsequent duplicative efforts. In this sense, systematic reviews serve as gatekeepers, not spigots, of redundancy.

While not all researchers will be motivated to master this skillset, those that are should be encouraged to develop their proficiency in conducting systematic reviews as they would any other research method. But like any research instrument, systematic reviews and meta-analyses can be problematic and even damaging in the wrong hands. Thus, rather than forbid the tool, let us teach how to properly wield it, and encourage researchers to not only learn best practices, but to maintain the healthy scepticism needed to innovate and advance the technique.

Garbage in, garbage out

For decision-makers and stakeholders of evidence-based medicine, among the most pressing concerns facing systematic reviews and meta-analyses is not the 'unfair' contribution to *h*-index calculations. Rather, the worry that comes with widespread use is the drastic variability in quality. Even for manuscripts that are decidedly important additions to the literature, which all research papers should be beholden to justify, there is no guarantee that their conduct and output will be valid or useful. This reality is often underappreciated, evident by the fact that systematic reviews and meta-analyses continue to be positioned at the top of evidence-rankings as though they, by default, generate the infallible final say on the matter.

For example, the quality of systematic reviews can be greatly diminished by reviewer missteps, including typographical and conversion errors, insufficient literature search strategies, inappropriate meta-analytic methods, pooling of duplicate study populations, retroactive inclusion or exclusion of studies to achieve a preferred result, and more. Authors may also fail to appropriately consider potential error, bias and quality of the underlying studies. Similarly, inattention to the degree and potential drivers of between-study heterogeneity (i.e. when studies are statistically combined despite reporting on different hypotheses and/or having different effect estimates) plagues many low-quality meta-analyses.

Additional concerns for the validity of systematic reviews and meta-analyses pertain to the limitations of the individual studies themselves. A meta-analysis generates a weighted average that includes all flaws, errors and biases of the primary research studies. It cannot overcome poor intervention adherence, exposure and outcome measurement error, residual and unmeasured confounding bias, selection bias, losses to follow-up, publication bias and missing data. The seriousness of these limitations depends entirely on the underlying studies, which are also often immensely variable in the quality of their conducting and reporting.

Other issues are more complex and may be less straightforward to improve, including reviewers' ability to identify errors, potential biases and sources of between-study heterogeneity. Further, weighing the potential impact of these in the interpretation of the certainty and validity of the overall evidence requires biostatistical, study design and subject-matter expertise.

Thus, the skillset and expertise required to properly undertake a systematic review and meta-analysis should not be underestimated. Fortunately, as with any scientific tool, many of these shortcomings can be remedied with greater attention to rigour and better standardisation of the approach. Notable improvements include innovations in systematic review software, freely available analysis code and journals increasingly requiring prospective protocol registration.

As such, attempts to improve and standardise the systematic review process have likely reduced errors and improved overall quality. However, having a framework that is too rigid may have also led to certain unintended consequences through a false sense of 'one size fits all'. Despite the step-by-step standardisation of screening, data extraction and statistical analysis, authors may still arrive at erroneous or over-stated conclusions.

As meta-analyses are conducted more broadly across diverse research domains, there is an increased requirement for nuance and subject-matter knowledge to appropriately combine and interpret results. For example, environmental and behavioural exposures are not often amenable to the randomised, placebo-controlled study designs which meta-analyses were originally developed for. The systematic review and meta-analytic techniques have evolved since their inception, expanding to accommodate non-randomised research, dose–response analyses, bias detection and quality assessment methods, and formal investigation of between-study heterogeneity. Similar innovations are seen in the wide variety of checklists and tools available for evaluating evidence quality and bias, for various study designs and healthcare domains. However, it's important for the reader to understand that many of the methodological decisions made by the review authors are not necessarily as 'systematic' as the name might imply. For example, poorly defined research questions and study exclusion criteria [11], meta-analysis weighting scheme [12] and choice of study-level bias assessment tool [13] have led to striking differences in reviewers' conclusions.

Conclusions

Overall, there has certainly been an increase in systematic reviews and meta-analyses over the past half-century. The response should not be to forcibly prevent their publication, but rather educate to develop better widespread proficiency. The problem isn't the tool itself, it's the misuse of the tool. If a systematic review and meta-analysis is sloppily conducted, poorly interpreted, redundant and contributes nothing toward our understanding of the hypothesis, then of course its publication should be questioned; but the same should be said for every type of manuscript. However, restricting this tool altogether would be a disservice to rigorous scientific progress. By instilling this generation of researchers with, at the very least, an appreciation for the complexity of conducting and interpreting a systematic review and meta-analysis, we will create more knowledgeable peer-reviewers and keen journal editors to properly scrutinise and improve, rather than dismiss, this critical and growing evidence base.

Authors' relationships and activities The author declares that there are no relationships or activities that might bias, or be perceived to bias, their work.

Contribution statement DT is the sole contributor to this paper.

References

1. Bonora E (2022) The 'scientist', the 'analyst' and the 'novelist': science or metrics? Diabetologia. https://doi.org/10.1007/s00125-022-05808-0
2. Cochrane AL (1972) Effectiveness and efficiency: random reflections on health services. Nuffield Trust
3. Chalmers I (1993) The Cochrane collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. Ann N Y Acad Sci 703:156–163; discussion 163-165. https://doi.org/10.1111/j.1749-6632.1993.tb26345.x
4. Lewis S, Clarke M (2001) Forest plots: trying to see the wood and the trees. BMJ 322(7300):1479–1480. https://doi.org/10.1136/bmj.322.7300.1479
5. Smith ML, Glass GV (1977) Meta-analysis of psychotherapy outcome studies. Am Psychol 32(9):752–760. https://doi.org/10.1037//0003-066x.32.9.752
6. Shah HM, Chung KC (2009) Archie Cochrane and his vision for evidence-based medicine. Plast Reconstr Surg 124(3):982–988. https://doi.org/10.1097/PRS.0b013e3181b03928
7. Grundy SM, Stone NJ, Bailey AL et al (2019) 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation 139(25):e1082–e1143. https://doi.org/10.1161/CIR.0000000000000625
8. Slutsky J, Atkins D, Chang S, Sharp BAC (2010) AHRQ series paper 1: comparing medical interventions: AHRQ and the effective health-care program. J Clin Epidemiol 63(5):481–483. https://doi.org/10.1016/j.jclinepi.2008.06.009
9. US Department of Agriculture, Agricultural Research Service (2016) Nutrient Data Laboratory. USDA National Nutrient Database for Standard Reference, Release 28 (Slightly revised). Version Current: May 2016. Available from http://www.ars.usda.gov/nea/bhnrc/mafcl. Accessed 22 October 2022
10. Higgins JPT, Thomas J, Chandler J et al (2022) Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022). Cochrane. Available from www.training.cochrane.org/handbook. Accessed 22 October 2022
11. Satija A, Malik VS, Willett WC, Hu FB (2017) Meta-analysis of red meat intake and cardiovascular risk factors: methodologic limitations. Am J Clin Nutr 105(6):1567–1568. https://doi.org/10.3945/ajcn.117.153692
12. Malik VS, Willett WC, Hu FB (2009) Sugar-sweetened beverages and BMI in children and adolescents: reanalyses of a meta-analysis. Am J Clin Nutr 89(1):438–439; author reply 439-440. https://doi.org/10.3945/ajcn.2008.26980
13. Tobias DK, Wittenbecher C, Hu FB (2021) Grading nutrition evidence: where to go from here? Am J Clin Nutr 113(6):1385–1387. https://doi.org/10.1093/ajcn/nqab124