



Are systematic reviews and meta-analyses still useful research? Yes

Djillali Annane^{1*}, Roman Jaeschke^{2,3} and Gordon Guyatt³

© 2018 Springer-Verlag GmbH Germany, part of Springer Nature and ESICM

Systematic reviews (SR) are commonly defined as “a summary of studies addressing a clear question, using systematic and explicit methods to identify, select, and critically appraise relevant studies, and to collect and analyse data from them” [1]. One of the essential tenets of evidence-based medicine is that optimal care requires up-to-date, rigorous summaries of evidence [2]. Without such summaries, clinicians and patients are vulnerable to unrepresentative samples of the evidence, misinterpreted and biased estimates of benefits and harms of interventions. For instance, consider the evolution of expert views regarding thrombolysis for acute myocardial infarction (AMI) [3]. An SR found 33 trials evaluating streptokinase in AMI patients, 25 suggesting mortality reduction and six reaching conventional levels of statistical significance. A cumulative meta-analysis found that by 1973—after eight trials (2432 patients)—results showed a mortality reduction with streptokinase. None of the subsequent trials changed the direction or magnitude of the odds ratio of dying. Yet, 34,542 AMI patients continued to be exposed to placebo in subsequent trials, including so-called mega-trials. The reason for the conduct of all the unnecessary trials was that experts, without the benefit of an SR and meta-analyses, could not grasp the big picture of the benefits of thrombolytic therapy, forcing trialists to conduct study after study until the message finally got through. In the meanwhile, patients died unnecessarily.

The fundamental standards for trustworthy clinical practice guidelines is that they rest on a foundation of

a systematic summary of the highest quality evidence; among those offering standards for trustworthy guidelines, the need for such summaries is not in dispute [4, 5]. Moreover, those who have written compellingly regarding waste in medical research have emphasized the needless duplication and misguided research directions when carrying out new studies not preceded by systematic summaries of actual knowledge [6].

We consider these tenets indisputable: optimal clinical practice, optimal guidelines to inform clinical practice and avoiding research waste; all require systematic summaries of evidence. Indeed, though they were first introduced in medical science [7], other fields, including social [8] and business [9] sciences, have recognized the crucial role of SR.

The issue then arises of how to produce the most efficient, useful and informative SR. Arguing that the problems of duplication and poor conduct of SR cast doubt on their usefulness—indeed, their necessity—is as facile and imprudent as arguing that the profusion of poorly conducted randomized trials fundamentally challenges their usefulness.

In the 30 years since authorities suggested initial standards for the conduct of SR [10], their methodology has advanced enormously, including insights into the importance of appropriately chosen eligibility criteria and of partnership of methodologists and clinicians, greater sophistication in assessing risk of bias in individual studies, insights into use of optimal statistical models, and the advent of network meta-analysis for simultaneous consideration of multiple interventions. The GRADE system is the most recent and possibly most relevant development.

The GRADE methodology for rating the quality (otherwise known as certainty or confidence) of the evidence has been now adopted by over 100 organizations worldwide including the Cochrane Collaboration, the World

*Correspondence: Djillali.annane@aphp.fr

¹ Service de médecine intensive réanimation, hospital Raymond Poincaré (APHP), laboratory of Infection & Inflammation U1173 INSERM/UVSQ; Dean of the School of medicine Simone Veil, University Versailles SQY-Paris Saclay, 104 boulevard Raymond Poincaré, 92380 Garches, France
Full author information is available at the end of the article

For contrasting viewpoints, please go to <https://doi.org/10.1007/s00134-017-5039-y> and <https://doi.org/10.1007/s00134-018-5066-3>.



Health Organization, UpToDate, and most critical societies. That methodology, extensively described for clinicians and other users [11], and for SR authors, guideline developers and health technology assessment practitioners [12], allows the classification of evidence as high, moderate, low or very low quality considering study design, risk of bias, precision, consistency, directness, publication bias and magnitude of effect. The detailed guidance for making quality judgements provides a transparent, straightforward approach that has facilitated a new rigour for the SR process. Not surprisingly, the widespread adoption of GRADE has not ushered in an era of uniformly exemplary SR, conducted efficiently and with successful efforts to avoid duplication. Systematic reviews continue to share a number of problems with other forms of research—poorly designed question, questionable population/interventions/outcomes, the conduct full of lapses in rigour both at the level of individual studies and of SR itself, and non-transparent academic and financial conflicts of interests.

There are about 11 SR published daily in the field of medicine [13] with high risk of redundancy, flawed methodology, conflict of interest-driven biases and misinterpretation of evidence [14]. Journals should refuse to publish SR not meeting rigorous standards, including duplicate assessment of eligibility and risk of bias,

explanation of heterogeneity and consideration of conflict of interest. Application of GRADE in further use of SR reinforces adherence to optimal standards. Stemming the flow of poor quality misleading research will require a change in academic medical culture. As long as job security, promotion and recognition require high volumes of publications, academics will put a premium on volume rather than quality, and there will always be journals that accommodate that culture. Editors should no longer consider SR for publication without an open access protocol such as on PROSPERO [15]. Fortunately, there are many rigorous SR and guides for clinicians and other users to distinguish the credible from the fundamentally flawed [16]. The alternative to using systematic reviews is simply not available (Fig. 1).

Author details

¹ Service de médecine intensive réanimation, hospital Raymond Poincaré (APHP), laboratory of Infection & Inflammation U1173 INSERM/UVSQ; Dean of the School of medicine Simone Veil, University Versailles SQY-Paris Saclay, 104 boulevard Raymond Poincaré, 92380 Garches, France. ² Department of Medicine, McMaster University, Hamilton, Canada. ³ Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, Canada.

Acknowledgements

We are grateful to Solène PITTION-ROSSILLON, department of education, university of Versailles SQY, for her help in drawing Figure 1.

Received: 7 February 2018 Accepted: 15 February 2018
Published online: 16 April 2018

References

1. Systematic review. GET-IT glossary. <http://www.getitglossary.org/>. Accessed 22 Jan 2018
2. Guyatt G, Jaeschke R, Wilson M, Montori V, Richardson S (2014) What is evidence-based medicine. In: Guyatt G, Meade MO, Cook DJ, Rennie D (eds) *Users' guides to the medical literature: a manual for evidence-based clinical practice*, 3rd edn. McGraw-Hill, New York
3. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC (1992) Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 327:248–254
4. Institute of Medicine (2011) *Clinical practice guidelines we can trust*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/13058>
5. Brouwers M, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna S, Littlejohns P, Makarski J, Zitzelsberger L, For the AGREE Next Steps Consortium (2010) AGREE II: advancing guideline development, reporting and evaluation in healthcare. *CMAJ* 182:e839–e842
6. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, Schulz KF, Tibshirani R (2014) Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 383:166–175
7. Cochrane AL (1979) 1931–1971: a critical review, with particular reference to the medical profession. *Medicines for the year 2000*. Office of Health Economics, London, pp 1–11
8. Littell Julia H (2006) Systematic reviews in the social sciences: a review. *Evid Policy* 2(4):535–537
9. Tranfield D, Denyer D, Smart P (2003) Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *Br J Manag* 14:207–222
10. Oman A, Guyatt GH (1991) Validation of an index of the quality of review articles. *J Clin Epidemiol* 44:1271–1278
11. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, GRADE Working Group (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336:924–926
12. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A (2011) GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol* 64:380–382
13. Bastian H, Glasziou P, Chalmers I (2010) Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 7(9):e1000326
14. Ioannidis JP (2016) The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 94:485–514
15. Booth A, Clarke M, Dooley G, Ghera D, Moher D, Petticrew M, Stewart L (2012) The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev* 1:2
16. Murad MH, Montori VM, Ioannidis JP, Jaeschke R, Devereaux PJ, Prasad K, Neumann I, Carrasco-Labra A, Agoritsas T, Hatala R, Meade MO, Wyer P, Cook DJ, Guyatt G (2014) How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *JAMA* 312:171–179