



# Are systematic reviews and meta-analyses still useful research? We are not sure

Morten Hylander Møller<sup>1\*</sup>, John P. A. Ioannidis<sup>2</sup> and Michael Darmon<sup>3,4</sup>

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

In a traditional systematic review and meta-analysis, the best available evidence is sought, systematically identified, critically appraised and synthesised, in order to try answer some clinical or research question [1]. Systematic reviews have long been considered at the top of the evidence-based medicine study hierarchy, and the number of published systematic reviews and meta-analyses has increased significantly over the past 20 years [2]. An assessment using stringent criteria a decade ago suggested that at least 2500 new systematic reviews are published each year [2]. Since then, their rate of production has increased approximately 10-fold [3]. About a third of them include quantitative syntheses (meta-analyses) [3].

## Upside of systematic reviews

Systematic reviews in theory help inform clinicians and researchers on patient-important issues, and they try to provide a transparent outline of the balance between the benefits and harms of healthcare behaviour and interventions [3]. They also provide balanced and transparent evidence for patients, relatives and policymakers, and they can be the starting point for trustworthy clinical practice guidelines [4].

Systematic reviews and meta-analyses are also increasingly being used in the context of research programs, in which they form the first step and are used to identify and highlight uncertainties and unanswered research areas where additional knowledge is warranted [5]. This

valuable information is used by many, including funders wishing to support new research, and by journal editors in prioritising research findings [2].

Summarising data from individual trials increases the statistical power and the precision of the effect estimates, as data from individual trials are often underpowered with high risk of type 1 and 2 errors [6]. This is especially important when assessing adverse events, as most individual trials are powered to detect benefits but not adverse events [7]. Consequently, meta-analyses have the potential to reveal important adverse event associations which may require further verification in the future. Carefully done systematic reviews and meta-analyses have also a unique opportunity to measure and try to understand sources of heterogeneity and to offer a broader view of the evidence.

## Downside of systematic reviews

Major challenges and threats to the validity and interpretation of systematic reviews and meta-analyses include lack of systematic and transparent conduct and reporting, poor methodological quality of the included studies, risk of random errors, unrecognized and unaccounted statistical and clinical heterogeneity, data dredging in non-predefined statistical analyses, and lack of assessment of the overall quality of evidence.

Lack of systematic and transparent conduct and reporting of systematic reviews is still common, despite efforts at prompting registration and careful reporting. Many systematic reviews are often flawed, redundant and/or misleading [3].

The methodological rigour of many published trials is not adequate, and most suffer from high risk of systematic errors (bias), which in turn increases the risk of overestimating benefit and underestimating harm [8].

\*Correspondence: [mortenhylander@gmail.com](mailto:mortenhylander@gmail.com)

<sup>1</sup> Department of Intensive Care 4131, Copenhagen University Hospital, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark  
Full author information is available at the end of the article

For contrasting viewpoints, please go to <https://doi.org/10.1007/s00134-018-5066-3> and <https://doi.org/10.1007/s00134-018-5102-3>.

Many statistical analyses are inadequately conducted and reported, and even changed after the completion of a trial [9]. Meta-analyses may not necessarily remove these biases.

Meta-analyses are at further risk of random errors due to repetitive testing and spurious findings [10]. Trial sequential analysis approaches [11] and other methods may help confer the right amount of uncertainty about the conclusiveness of the results.

Summarising data from various sources increases the risk of heterogeneity. Although heterogeneity may be estimated statistically, currently available approaches, such as the  $Q$  test and  $I^2$  estimates, remain imperfect, weak in detecting heterogeneity and likely give a false sense of reassurance when failing to demonstrate significant heterogeneity [12].

Overall, the quality of evidence of many systematic reviews and meta-analyses is low because of risks of bias, indirectness, imprecision, inconsistency and publication bias [13]. This has significant impact on the validity of the findings.

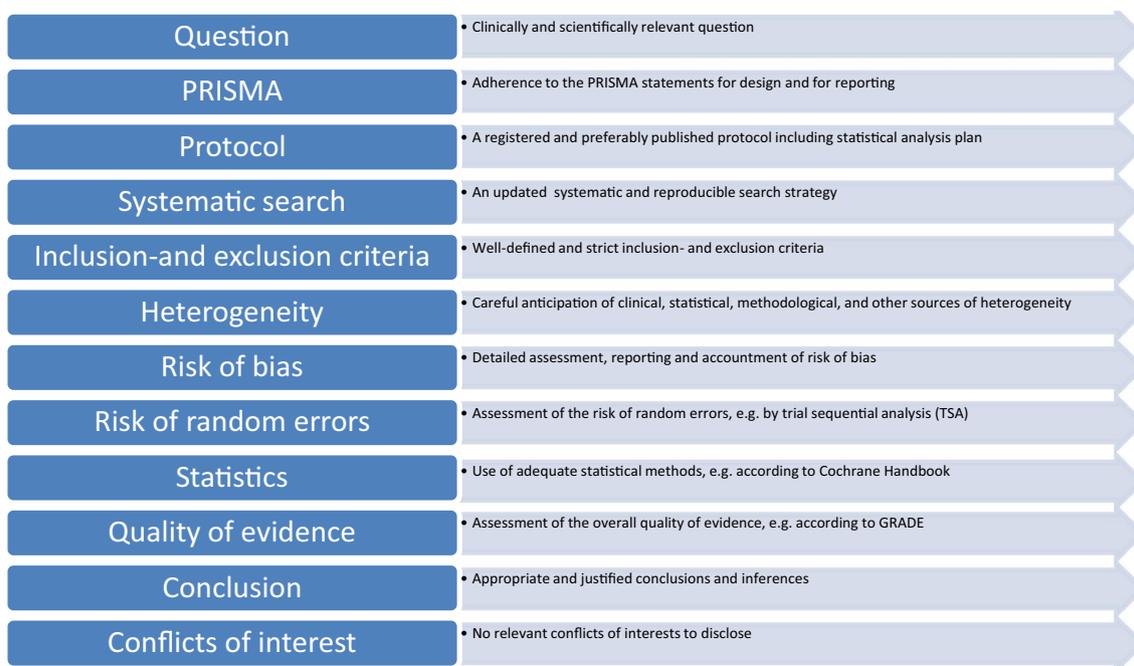
### Future of systematic reviews

Threats to the validity and interpretation of systematic reviews are common, but they can all be overcome to some extent with careful design and execution of the review (Fig. 1). For those that cannot be overcome (e.g. the poor quality of the primary data), a good systematic

review can still be helpful if it recognises and carefully maps the limitations of the evidence.

Trustworthy systematic reviews and meta-analyses with accurate and reliable conclusions should be prepared systematically and transparently [3]. The preferred reporting items for systematic reviews and meta-analyses (PRISMA) statements provide guidance on how to design [14] and report a systematic review [2]. These guidelines, although necessary, may not be considered as sufficient, as no guidance on design and reporting can make a meta-analysis useful if it does not deal with a useful, important question to start with (Fig. 1).

Conventional retrospective systematic reviews are limited by the inherent limitations of the primary data that they try to unearth and combine. Systematic reviews of systematic reviews (overview of systematic reviews or umbrella reviews) and network meta-analyses are offering opportunities for even broader views of the evidence [15], and when carefully conducted, they can further enhance the utility of evidence syntheses. Individual patient-data meta-analyses are also becoming increasingly popular and the increasing availability of raw data may make them the default standard in the future, while currently they are still limited by data availability and much higher requirements in terms of effort and resources. As a way to move forward, meta-analyses may need to be planned and designed prospectively. In this concept, the research agenda of multiple trials may



**Fig. 1** Key items in a trustworthy systematic review and meta-analysis

deserve to be designed with the explicit expectation that all trials will contribute to a continuously updated meta-analysis. This may allow reaping the most useful results, as primary data and meta-analyses become synchronised.

## Conclusion

An evaluation of the landscape of current systematic reviews and meta-analyses suggests that many of them are focused on unimportant questions, many are redundant and unnecessary, a sizeable proportion are flawed beyond repair, and eventually only about 3% of them are both well done and clinically useful [3]. So, if this literature is to be seen in its bulk, it represents mostly a waste of no utility. However, there is a minority of systematic reviews and meta-analyses that can be extremely useful. The main challenge is how to enhance this minority and, if possible, make it even the majority among future systematic reviews and meta-analyses.

## Author details

<sup>1</sup> Department of Intensive Care 4131, Copenhagen University Hospital, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark. <sup>2</sup> Meta-Research Innovation Center at Stanford (METRICS) and Departments of Medicine, Health Research and Policy, and Biomedical Data Science, Stanford University School of Medicine, Stanford, CA, USA. <sup>3</sup> AP-HP, Saint Louis Hospital, Medical Intensive Care Unit, Paris, France. <sup>4</sup> Medicine University, Paris 5 University, Paris, France.

Received: 18 December 2017 Accepted: 21 December 2017

Published online: 16 April 2018

## References

- 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
- Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 339:b2535
- Ioannidis JP (2016) The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 94:485–514
- Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336:924–926
- Scoglio D, Fichera A (2014) Establishing a successful clinical research program. *Clin Colon Rectal Surg* 27:65–70
- Ridgeon EE, Bellomo R, Aberlegg SK, Sweeney RM, Varughese RS, Landoni G, Young PJ (2017) Effect sizes in ongoing randomized controlled critical care trials. *Crit Care* 21:132
- Clarke M, Brice A, Chalmers I (2014) Accumulating research: a systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. *PLoS One* 9:e102670
- Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, Als-Nielsen B, Balk EM, Gluud C, Gluud LL, Ioannidis JP, Schulz KF, Beynon R, Welton NJ, Wood L, Moher D, Deeks JJ, Sterne JA (2012) Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 157:429–438
- Anand V, Scales DC, Parshuram CS, Kavanagh BP (2014) Registration and design alterations of clinical trials in critical care: a cross-sectional observational study. *Intensive Care Med* 40:700–722
- Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, Guyatt G, Devereaux PJ, Thabane L (2011) The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis—a simulation study. *PLoS One* 6:e25491
- Wetterslev J, Thorlund K, Brok J, Gluud C (2008) Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol* 61:64–75
- Ioannidis JP, Patsopoulos NA, Evangelou E (2007) Uncertainty in heterogeneity estimates in meta-analyses. *BMJ* 335:914–916
- Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH (2011) GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 64:401–406
- Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA, Group PP (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 4:1
- Mills EJ, Thorlund K, Ioannidis JP (2013) Demystifying trial networks and network meta-analysis. *BMJ* 346:f2914