EDITORIAL

Are systematic reviews and meta-analyses still useful research? No

Sylvie Chevret[1]* , Niall D. Ferguson[2,3] and Rinaldo Bellomo[4,5,6]

"Facts are stubborn, but statistics are more pliable" Mark Twain

Introduction
Whatever the intervention of interest for a given clinical condition or public health problem, it is likely that previous studies have been performed to address it. It is therefore logical that such prior knowledge should be understood and analysed in order to inform decisions and help plan, design and justify future studies. Unfortunately, heterogeneity of prior to knowledge in terms of settings, teams and patient features creates uncertainty about its validity, robustness, significance, relevance and implications for both modern knowledge and future investigations. Moreover, it creates opportunities for evidence distortion driven by confirmation, financial and academic biases among many.

Faced with such uncertainty, even as far back as 1904, Karl Pearson [1] proposed the use of formal techniques to combine data from different studies to examine the preventive effect of serum inoculations against enteric fever. Since then, the need to systematically assess prior knowledge has become even more seductive, given the current global research output with 20,000 journals publishing more than 2 million articles per year, often with unclear or conflicting results. This task, however, is challenging, no matter how theoretically desirable. In particular, most modern literature remains heterogeneous in quality, design, patient characteristics, diagnoses and outcomes. In addition, many studies involve small numbers, are at

high risk of type I and type II errors, and lack blinding. In response to such limitations, systematic reviews and meta-analyses (SRMAs) have been promoted as the "path to scientific salvation" by providing the statistical magic that will allow clinicians, public health officials or investigators to determine the possible efficacy or harm of a given intervention [2], despite the use of data from poor quality, unblinded, heterogeneous trials. Moreover, even if a systematic review does not necessarily involve quantitative statistical pooling (meta-analysis), it is frequently argued that it will provide important synthetic information of prior knowledge and identify significant (and allegedly previously unnoticed) knowledge gaps. Such SRMAs have become particularly fashionable in the era of evidence-based medicine.

Evidence-based medicine and the evolution of SRMAs
The advent of the evidence-based medical and healthcare paradigm in the 1990s has provided strong impetus for the use of SRMAs, placing them at the top of the hierarchical pyramid of evidence [3]. In particular, the Cochrane Collaboration was established in 1993 to support the generation and dissemination of SRMAs [4]. Since then, SRMA specialists have attempted to develop standards on how to conduct and report such SRMAs and created tools to assess the risk of bias or the quality of SRMAs [5]. Although such frenetic corrective statistical activity implies that SRMAs carry the same flaws as those of the data they dredge, SRMAs have proliferated [6], so that the concept of synthesizing data from two or more SRMAs has gained traction in healthcare research—yielding SRMAs of SRMAs, also named "overviews" [7] or "umbrella views" [8]. New methods for dealing with multiple SRMAs published on the same

*Correspondence: sylvie.chevret@univ-paris-diderot.fr
[1] Biostatistics Unit, Paris Diderot University, AP-HP, Saint Louis hospital, Paris, 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/s00134-018-5102-3.

Springer

Table 1 The 10 main issues of systematic reviews and meta-analyses (SRMAs)

1. They do not discover anything unknown
2. They do not carefully consider each study's characteristics, populations, design; studies performed perhaps 30 years ago are presented together with studies performed a year ago
3. They do not accurately predict the results of subsequent pivotal trials
4. They often come to trivial conclusions; if all studies go in one direction, a meta-analysis cannot possibly provide any useful information
5. They can use the same studies and come to different conclusions
6. They are proliferating in response to the availability of statistically packages rather than actual clinical need
7. The techniques used are not empirically validated to predict the correct final assessment; for instance, the weighting given to a study on the basis of study patient numbers is not empirically validated
8. They are often influenced by a single major trial
9. They often come to conclusions based on poor evidence creating dogma where none should exist
10. They are presented as the pinnacle of evidence when, in fact, they are derivative, flawed information analysis

topic area and methods for displaying outcome data in overviews have been proposed [9]. In addition, network meta-analyses now allow the inclusion of both directly and indirectly relevant evidence from individual studies to describe the relative benefits (or harms) of a range of interventions even in the absence of head-to-head comparisons [10]. Lastly, updating the evidence by continuous or trial sequential meta-analysis (TSA) for timely decision-making has been proposed for SRMAs [11], and cumulative network meta-analyses [12]. Many believe that such approaches will lead to improvements in the quality of SRMAs and to more reliable estimates of intervention effects.

Irrespective of the SRMA technique, all of the above developments will now ensure that a veritable tsunami of SRMAs will drown clinicians, public health officials and investigators in perpetuity. Indeed, their production has now reached "epidemic proportions" [13]. Thus, we are in the middle of an ideological, publication-fuelled bubble which, similar to the alchemists of centuries ago, promises to make gold out of clods of earth. Faced with this onslaught, it is legitimate to ask the key question of whether SRMAs are "useful research".

Why SRMAs are not useful research

There are several reasons why SRMAs are not useful research. First, they are not new research at all; rather they are simply summaries and statistical transformations of existing research based on empirically untested assumptions. Such analytical processes are often managed by SRMA professionals who, in some sense, academically and somewhat symbiotically benefit from the years of effort expended by investigators and research teams conducting clinical trials. Second, SRMAs do not consider and cannot estimate the biological plausibility, reproducibility and external validity of their results [14]. Third, they have limited predictive ability for the results

of subsequent randomized controlled trials (RCTs), often producing a misleading assessment of knowledge [15]. Fourth, understanding prior knowledge in a way that is useful requires a systematic and reproducible approach (as is the case for all science), which delivers reproducible outcomes. SRMAs often do not deliver such reliable outcomes; SRMAs are sometimes published reporting different conclusions from seemingly almost identical sources. Fifth, the most common conclusions of SRMAs are typically unhelpful and include sentences such as "better studies are needed" or "there is insufficient evidence" (Table 1). These kinds of conclusions are often also reported after the non-pivotal underpowered RCTs, for whose poor level evidence the SRMAs are supposed to compensate. Assessment of such prior knowledge by experts could have easily come to the same conclusion in one line. In this regard, the concerns associated with SRMAs are similar to those associated with low quality research. Unfortunately, SRMAs sometimes seem to amplify the reach of such low quality work rather than represent a correction for their shortcomings.

Conclusions

The synthesis of prior knowledge and the assessment of its quality are essential to scientific progress. SRMAs offer a (often flawed) means of delivering such synthesis and assessment but are not useful research themselves because they do not provide novel information or deliver the results of experiments. SRMAs are often based on poor studies and low quality primary evidence and thus cannot deliver useful and robust information or insights (garbage in = garbage out). SRMAs often confer a whiff of legitimacy to that which should be dismissed and are used as tools to support specific agendas. Finally, SRMAs promote misleading views among readers by showing that a single-centre, unblinded study of 20 patients belongs to the same forest plot as a 2000-patient,

multicentre, double-blind RCT. Any rational consideration of the consequences of such an approach must inevitably lead to the logical conclusion that, in general, SRMAs are not useful research and can never be a substitute for reasoned and careful assessment of the literature.

Author details
[1] Biostatistics Unit, Paris Diderot University, AP-HP, Saint Louis hospital, Paris, France. [2] Interdepartmental Division of Critical Care Medicine, Departments of Medicine and Physiology, Institute for Health Management, Policy, and Evaluation, University of Toronto, Toronto, ON, Canada. [3] Division of Respirology, Department of Medicine, University Health Network and Sinai Health System, Toronto, ON, Canada. [4] Department of Intensive Care, Austin Hospital, Heidelberg, Australia. [5] School of Medicine, University of Melbourne, Parkville, Australia. [6] Australian and New Zealand Intensive Care Research Centre, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia.

References
1. Pearson K (1904) Report on certain enteric fever inoculation statistics. Br Med J 2:1243–1246
2. Becker LA, Oxman AD (2008) Overviews of reviews. In: Higgins JPT, Green S (eds) Cochrane handbook for systematic reviews of interventions. Wiley, Chichester, p 607–631
3. Rosenberg WM (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. JAMA 268:2420–2425
4. Herxheimer A (1993) The Cochrane Collaboration: making the results of controlled trials properly accessible. Postgrad Med J 69(817):867–868
5. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C (2007) Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 7:10
6. Chandler J, Hopewell S (2013) Cochrane methods–twenty years experience in developing systematic review methods. Syst Rev 20(2):76. https://doi.org/10.1186/2046-4053-2-76)
7. Hartling L, Chisholm A, Thomson D, Dryden DM (2012) A descriptive analysis of overviews of reviews published between 2000 and 2011. PLoS One 7:e49667
8. Salleh S, Thokala P, Brennan A, Hughes R, Booth A (2017) Simulation modelling in healthcare: an umbrella review of systematic literature reviews. Pharmacoeconomics 35(9):937–949
9. Crick K, Wingert A, Williams K, Fernandes RM, Thomson D, Hartling L (2015) An evaluation of harvest plots to display results of meta-analyses in overviews of reviews: a cross-sectional study. BMC Med Res Methodol 15:91
10. Alhazzani W, Alshamsi F, Belley-Cote E, Heels-Ansdell D, Brignardello-Petersen R, Alquraini M, Perner A, Møller MH, Krag M, Almenawer S, Rochwerg B, Dionne J, Jaeschke R, Alshahrani M, Deane A, Perri D, Thebane L, Al-Omari A, Finfer S, Cook D, Guyatt G (2018) Efficacy and safety of stress ulcer prophylaxis in critically ill patients: a network meta-analysis of randomized trials. Intensive Care Med 44(1):1–11
11. Biau DJ, Boulezaz S, Casabianca L, Hamadouche M, Anract P, Chevret S (2017) Using Bayesian statistics to estimate the likelihood a new trial will demonstrate the efficacy of a new treatment. BMC Med Res Methodol 17(1):128
12. Nikolakopoulou A, Mavridis D, Egger M, Salanti G (2016) Continuously updated network meta-analysis and statistical monitoring for timely decision-making. Stat Methods Med Res 1:962280216659896
13. Ioannidis JPA (2016) The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q 94(3):485–514
14. Bellomo R, Bagshaw SM (2006) Evidence-based medicine: classifying the evidence from clinical trials–the need to consider other dimensions. Crit Care 10(5):232
15. LeLorier J, Grégoire G, Benhaddad A, Lapierre J, Derderian F (1997) Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med 337(8):536–542