Systematic reviews: let's keep them trustworthy

I. Garcia-Doval[1,2], E. van Zuuren[3], F. Bath-Hextall[4], J. R. Ingram[5]

1.- Research Unit, Fundación Piel Sana AEDV. Madrid, Ferraz 100, 1º izda, Madrid, Spain.
2.- Complexo Hospitalario Universitario de Vigo, Vigo. Meixoeiro sn 36200, Vigo, Spain.
3.- Department of Dermatology, Leiden University Medical Centre, Albinusdreef 2, 2333ZA Leiden, the Netherlands.
4.- Centre for Evidence Based Healthcare, School of Health Sciences, Faculty of Medicine and Health Sciences, The University of Nottingham, Nottingham, U.K.
5.- Department of Dermatology & Wound Healing, Division of Infection and Immunity, Cardiff University, University Hospital of Wales, 3rd Floor Glamorgan House, Heath Park, Cardiff CF14 4XN, U.K.

E-mail: ignacio.garcia.doval@sergas.es

http://orcid.org/0000-0002-6881-5260 (I. Garcia-Doval)
https://orcid.org/0000-0002-4780-0182 (E.J. van Zuuren)
http://orcid.org/0000-0001-6583-029X (F. Bath-Hextall)
http://orcid.org/0000-0002-5257-1142 (J.R. Ingram)

Conflicts of interest: none

Funding: non

Acknowledgement: We thank Hywel Williams for his review and comments.

The advent of systematic reviews is relatively new and represents an important tool to summarise the best evidence for studying the effects of healthcare interventions. Without systematic reviews, the effectiveness of therapies (or lack of benefit) can remain unrecognised for many years. The Cochrane logo (www.cochranelibrary.com) provides a good example: it shows the forest plot of a systematic review describing how corticosteroids given to women who are about to give birth prematurely can save the life of the newborn child, an important summary of the evidence available at the time. Systematic reviews as a publication type has increased over the last decade. Initially concerned with therapeutic interventions, there are now many more types, from “umbrella” reviews (reviews of reviews) to reviews of observational studies and case reports. The methods of the review will depend on the question being asked.

Systematic reviews carry an aura of infallibility but are highly dependent on the methods used and the quality of information that they summarise. A well performed systematic review of high quality randomised controlled trials (RCTs) is usually assigned to the top of the evidence pyramid for assessing the effects of interventions. However, a poorly performed systematic review that only includes a subset of all relevant evidence with little attempt to rate the quality of included studies can produce false conclusions. The review’s position in the evidence pyramid is anchored by the quality of its included studies. For example, a systematic review of case reports is likely to be misleading due to publication bias affecting the included studies, in which reports showing benefit of a new intervention are more likely to be published than those showing no benefit. Cochrane reviews are generally more methodologically rigorous than non-Cochrane reviews in dermatology[1], but are limited to studies of interventions and diagnostic test accuracy. These are the types of reviews with most established methods[2]. The Joanna Briggs Institute has also developed methodology for qualitative systematic reviews. However, reviews of other types of questions are less well

developed and authors have less guidance to ensure a high quality review. The BJD aims to publish only excellent systematic reviews of any type.

An advantage of systematic reviews over narrative reviews is that their methods should be clearly described and reproducible. This is the objective of reporting guidelines such as PRISMA or MOOSE, ensuring complete description of the methods so that the reader can judge the quality of the review[3,4]. In a similar way to registering or publishing the protocol for an RCT, the planned methodology for a systematic review should be registered beforehand on the International Prospective Register of Systematic Reviews (PROSPERO) so that reviewers and readers can judge whether the authors did what they said they would do. Cochrane reviews protocols are also published before the review starts. Several tools exist for undertaking critical appraisal and methodological quality assessment of systematic reviews. Although none has become universally accepted, the AMSTAR tool is probably the most commonly used quality assessment tool for systematic reviews of RCTs[5]. A tool for observational studies is being developed.

As in any study, the main concerns that can lead to false results in systematic reviews are chance, confounding and bias. These can arise in the original studies and in the process of producing the systematic review and can be amplified in the review if steps are not taken to mitigate them.

Taking into account the effect of chance is probably the least problematic. It is achieved using statistics and tools such as meta-analysis, provided the included studies are sufficiently similar to combine in such a way – which is always a matter of judgement. It is important to ensure that meta-analysis does not hide, but describes and helps explain heterogeneity of results.

Confounding is not a big problem in reviews of RCTs because randomization, if done properly, decreases confounding. However, for reviews of observational studies, controlling for confounding is difficult. Important confounders should be pre-defined and taken into account in the review, which can be difficult to achieve if they have not been measured in all studies. Even after diligent adjustment, residual confounding due to partial control for known confounders or the existence of unknown confounders can cause the results of a review to deviate from the truth.

Bias is perhaps the most important issue to assess. Bias can be present in the original papers or may be introduced in the process of creating the systematic review. There are tools such as the Cochrane risk of bias tool that can be used to evaluate risk of bias in the original papers, mostly designed for use with RCTs or the ROBINS-I tool[6] for non-randomized studies of interventions. Risk of bias figures give a useful general overview of potential bias within all the included studies, although it can be difficult to use them to determine the influence of bias on a particular result. The Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach is a standard method to make recommendations on the basis of evidence, and includes tools to assess the quality of evidence per outcome[2]. ConQual[7] and CerQual[8] have the same purpose of assessing the confidence that we can place in qualitative reviews findings. Assessment of risk of bias introduced during the process of review is key in the process of peer reviewing systematic reviews. New tools such as ROBIS have been developed with this aim[9].

An important source of bias in the creation of a review is study selection. Authors should be as certain as possible that they have included all relevant studies in their review. The main concern is to ensure that studies that could modify the results of the review are not missed. These are usually studies with negative results that remain unpublished, or published in harder to reach journals. Performing a comprehensive literature search is not easy, especially when looking for observational studies. As a letter to the editor by Grindlay in this

issue highlights[10], seeking the help of a trained information specialist is very useful and has been associated with more sensitive searches[11].

One additional frequent problem encountered by systematic reviews is outcome measure instrument heterogeneity. Outcomes measured in the source studies may be different, unvalidated and hence hard to compare and at many times not very useful for patients or clinicians. Ongoing efforts to improve homogeneity of outcomes in dermatology, such as developing core outcomes for skin conditions in the CSG-COUSIN initiative[12], are a step forward in this area for future clinical trials, provided that the recommendations for core outcome sets are widely disseminated and used.

Keeping systematic reviews trustworthy is important to avoid research waste and to help us cope with information overload[13]. From the BJD perspective, methodological peer-review of these papers should be rigorous and ideally include an information specialist, and an expert in the methods of systematic reviews. The proliferation of tools to assess the quality of information and reviews is an indication of how difficult this can be. We suggest that reviewers use tools such as AMSTAR or ROBIS when evaluating the quality of systematic reviews. Referees should check that all limitations and differences between protocol and review are acknowledged, and that recommendations are based on the evidence present in the review. Looking ahead, protocol registration in PROSPERO may in the future become a requirement for publication of non-Cochrane systematic reviews in the BJD.

References:

1 Collier A, Heilig L, Schilling L *et al.* Cochrane Skin Group systematic reviews are more methodologically rigorous than other systematic reviews in dermatology. *The British journal of dermatology* 2006; **155**: 1230-5.

2 Cochrane Handbook for Systematic Reviews of Interventions. In: (Higgins J, Green S, eds): The Cochrane Collaboration, 2011. Available at: www.handbook.cochrane.org. (last accessed 10 May 2017).

3 Moher D, Liberati A, Tetzlaff J *et al.* Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.

4 Stroup DF, Berlin JA, Morton SC *et al.* Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-12.

5 Shea BJ, Hamel C, Wells GA *et al.* AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009; **62**: 1013-20.

6 Sterne JA, Hernan MA, Reeves BC *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016; **355**: i4919.

7 Munn Z, Porritt K, Lockwood C *et al.* Establishing confidence in the output of qualitative research synthesis: the ConQual approach. *BMC Med Res Methodol* 2014; **14**: 108.

8 Lewin S, Glenton C, Munthe-Kaas H *et al.* Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med* 2015; **12**: e1001895.

9 Whiting P, Savovic J, Higgins JP *et al.* ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016; **69**: 225-34.

10 Grindlay D. Letter in this issue (to be completed).

11 Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC *et al.* Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. *J Clin Epidemiol* 2015; **68**: 617-26.

12 Core Outcomes Set Initiative (CSG-COUSIN). Available at: http://skin.cochrane.org/core-outcomes-set-initiative-csg-cousin (last accessed 10 May 2017).

13 Page MJ, Shamseer L, Altman DG *et al.* Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. *PLoS Med* 2016; **13**: e1002028.