

Research Techniques Made Simple: Workflow for Searching Databases to Reduce Evidence Selection Bias in Systemic Reviews

Laurence Le Cleach, M.D, associate professor, faculty¹, Elizabeth Doney, information specialist, faculty², Kenneth A Katz dermatologist MD, faculty³, Hywel C Williams, DSc, professor, faculty², Ludovic Trinquart, PhD, biostatistician, faculty⁴

¹Department of Dermatology, AP-HP, Hôpitaux Universitaires Henri Mondor; Université Paris-Est; EA EpiDermE, INSERM, Créteil, France

²Cochrane Skin Group, The University of Nottingham, Centre of Evidence-Based Dermatology, School of Medicine, University of Nottingham, UK

³Department of Dermatology, Kaiser Permanente, San Francisco, California, US

⁴Cochrane France; INSERM U1153 METHODS team, Paris, France

E Doney: elizabeth.doney@nottingham.ac.uk,
KA Katz: kenneth.katz@gmail.com,
HC Williams: Hywel.Williams@nottingham.ac.uk,
L Trinquart: ludovic.trinquart@aphp.fr

Corresponding author: Laurence Le Cleach,
Service de Dermatologie
Hôpital Henri Mondor
51 avenue du Maréchal de Lattre de Tassigny
94010 Créteil
laurence.lecleach@free.fr
Phone: +33 1 49 81 25 12
Fax : +33 1 49 81 25 08

Conflict of interests: None to declare

Abstract

Clinical trials and basic-science studies without statistically significant results are less likely to be published than studies with statistically significant results. Systematic reviews and meta-analyses that omit unpublished data are at high risk of distorted conclusions. Here, we describe methods to search beyond bibliographical databases to reduce evidence selection bias in systematic reviews. By searching conference proceedings, unpublished studies may be identified. Moreover, clinical trial registries – databases of planned and ongoing trials–, and regulatory agency websites – such as the European Medicine Agency (EMA) and the United States Food and Drug Administration (FDA) – may provide summaries of efficacy and safety data. Primary and secondary outcomes are pre-specified in trial registries, thus allowing the assessment of outcome reporting bias by comparison with the trial report. The sources of trials data and documents are still evolving, with ongoing initiatives promoting broader access to clinical study reports and individual patient data. There is currently no established methodology to ensure that the multiple sources of information are incorporated. Nonetheless, systematic reviews must adapt to these improvements and cover the new sources in their search strategies.

Introduction

As highlighted in related Research Techniques Made Simple articles, reporting bias remains one of the greatest threats to the validity of systematic reviews (2016; Abuabara et al., 2012).

To obtain a fair assessment of the effects of an intervention, systematic reviews of interventions for skin diseases should use stringent efforts to include all relevant evidence. An exhaustive search of trials is the most important step in systematic review methodology to reduce evidence selection bias. Yet many published articles labeled as "systematic reviews" search only a fraction of the evidence by limiting the search to one or two convenient databases.

In this article, we describe a workflow for searching sources beyond bibliographical databases (Figure). These techniques will be useful for systematic reviewers for planning an optimal search strategy and for readers of systematic reviews to judge whether suboptimal methods to identify trials may have introduced bias.

First, find the published trials

For a systematic review of dermatological interventions, the least one can do is to make every effort to identify all published randomized trials. Searching the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, and EMBASE will likely allow the researcher to find the majority of published trials. CENTRAL is particularly important to search because it offers a concentrated source of reports of randomized trials. Other specialized bibliographical databases may be relevant to specific topics (Online Appendix 1). Searching bibliographical databases should follow the methodological principles for information retrieval (Lefebvre et al., 2011). In particular, search equations should seek increased sensitivity and use ad hoc filters to identify randomized trials (such as the Cochrane

Highly Sensitive Search Strategies or filters listed at

<https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home>). Such a search should

be complemented by screening the reference lists of all selected trials and by searching for

previous systematic reviews on the same topic and screening the lists of selected trials.

Next, find the unpublished trials

About 50% of clinical trial results that are presented at meetings and congresses remain unpublished (Scherer et al., 2007). As a consequence, conference abstracts should be searched in order to identify trials with unpublished results. Data reported in conference abstracts may be not be reliable as full publication, as abstracts may contain preliminary results and may not contain sufficient information to assess methodological quality. However, abstracts allow documenting the existence of unpublished trials (more specifically, their number and sample size), and unpublished outcomes. It allows statistical analysis to gauge the sensitivity of the systematic review conclusions to the non-dissemination of these trials.

Some databases index conference proceedings. However, there is currently no centralized registry of abstracts from all conferences. Systematic reviewers most frequently hand search or electronically search abstracts made available by the corresponding societies (e.g. American Academy of Dermatology, European Society for Dermatological Research, Society for Investigative Dermatology, Japanese Society for Investigative Dermatology) through journal supplements or on their websites. The Cochrane Skin Group has hand searched and added to its Specialized Register 42 journals and 28 conference proceedings (Appendix 1).

The value of trial registries for identifying missing outcome data

Clinical trial registries - databases of planned and ongoing trials - have become essential sources for identifying unpublished trials. In 2005, the International Committee of Medical Journal Editors (ICMJE) stated that to be considered for publication, trials need to have been

registered in a public ICMJE-approved registry before the beginning of enrollment.

Systematic review authors can search the World Health Organization International Clinical Trials Registry Platform Search Portal, which gathers records of trials registered on 16 data providers including clinicaltrials.gov and the European Union (EU) Clinical Trials Register. Besides institutional registries, pharmaceutical companies have also developed clinical trial registries. When a relevant completed trial is identified but no published article can be matched, the systematic review authors can contact the trialists or sponsors to inquire about the trial status and ask for results. Some authors have even suggested that only prospectively registered trials should be included in meta-analyses because the risk of bias with any other form of trials is too great (Roberts et al., 2015).

Trial registries also allow identifying unreported outcomes since the primary and secondary outcomes are documented in each trial record. In cases of publication, one can compare the reported outcomes to the registered outcomes and assess selective outcome reporting bias, i.e. when negative outcomes remain unreported (Nankervis et al., 2012). An example of outcome reporting bias is the Multicenter Selective Lymphadenectomy Trial -1 study that sought to determine whether wide excision followed by sentinel node biopsy and immediate lymphadenectomy for nodal metastases is better than wide excision followed by nodal observation for melanoma. The trial produced much valuable data, yet the primary outcome of overall survival which was identified in the original trial registration was never published in the “final” report. Derivation of overall survival data from the study report suggested no overall survival increase for sentinel biopsy plus selective lymphadenectomy (Williams, 2015).

Clinical trial registries may also contain summary trial data. In clinicaltrials.gov, the results of applicable clinical trials as defined by Section 801 of the Food and Drug Administration Amendments Act are required to be posted and the results of many other trials are also posted

voluntarily. For systematic reviewers, it is therefore crucial to use clinicaltrials.gov to find trial results, in particular safety information. The European Medicines Agency (EMA) has also enacted a proactive publication of summary results through the EU Clinical Trials Register. Some pharmaceutical companies have also developed their own clinical trial result databases.

The untapped data buried in regulatory agency websites

Regulatory agencies, such as the Food and Drug Administration (FDA) and the EMA, also offer access to additional data through the pharmaceutical companies' approval applications (Online Appendix 2). The FDA provides a searchable catalog of approved drug products. These unpublished trial data are directly usable for systematic reviews and their inclusion can result in modification of the conclusions. In a re-analysis of 41 meta-analyses based on published data only, the addition of unpublished FDA trial data changed the outcome to a lower treatment effect in 46.3% of meta-analyses, did not change the estimate in 7.4%, and changed the outcome to a larger treatment effect in 46.3% (Hart et al., 2012). The EMA publishes European Public Assessment Reports for every medicine application, whether it has been granted or refused a marketing authorization. A comparison of FDA and EMA data for 27 drugs has shown that detailed data on efficacy and harms were available; the information were easier to find on the EMA than on the FDA website, however more data on harms were available on the latter (Schroll et al., 2015).

The benefit of searching regulatory agency websites is exemplified in studies on use of imiquimod cream for molluscum contagiosum. In a Cochrane review published in 2009, the one published trial comparing imiquimod to placebo in 23 patients showed a relative risk of 3.67 (95% confidence interval [CI] from 0.48 to 28.0) for complete clearance of lesions. However, 3 industry-sponsored unpublished trials were included in a FDA's publicly available review (Papadopoulos, 2007). These 3 trials randomized a total of 827 patients.

When added to the published trial, the pooled relative risk was 0.93 (95% CI from 0.73 to 1.19) suggesting that imiquimod is ineffective in that indication.

Lastly, health technology assessment agencies, through their requests to industry, may have access to unpublished data and make them publicly available by publishing benefit assessment dossiers online (Online Appendix 1).

Limitations of statistical diagnosis or correction for bias

A comprehensive search is even more important when considering that no statistical method allows complete documenting or excluding of reporting bias in a systematic review with certainty. Asymmetry of the funnel plot may reveal that smaller trials give different findings from larger trials. But funnel plot asymmetry has several possible causes, in particular heterogeneity, and its presence or absence cannot be equated with the presence or the absence of reporting bias. Moreover, many statistical methods have been introduced to detect or adjust for reporting bias. But, their use is inappropriate in most meta-analyses because of too few trials or excessive heterogeneity (Ioannidis, 2008).

Potential challenges to handle with multiple sources of data

Comprehensive searching adds to the resources needed to complete the systematic review, but searching some sources may not always yield additional evidence. Among 114 systematic reviews that searched FDA documents, unpublished data was available from the FDA for 17%. (McDonagh et al., 2013) The extent and depth of the search strategy might be adapted according to the review question and context. For example, in a systematic review of a drug for an unapproved indication, searching the FDA documents is unlikely to provide unpublished evidence. But attributes of reviews that will most likely benefit from searching additional sources such as the FDA are still unknown.

Another challenge is that multiple reports for the same trial may be identified and discrepancies for results can exist between different sources (Hartung et al., 2014). Systematic review authors then have to link all reports of the same trial together, and decide and describe clearly which report is to be chosen as the primary source of information. Although there is no established consensus, an order of priority may be pre-specified. For instance, FDA-prepared documents may be considered as more reliable than journal articles. In fact, FDA statistical reviewers reanalyze raw data, while journal articles may be affected by selective reporting of a subset of statistical analyses based on the results.

The way forward

Trial registration is now a legal requirement in the United States, EU, and many countries, but compliance is far from perfect. Enhanced transparency is encouraged by the alltrials.net campaign, an initiative of several organizations such as Cochrane, the BMJ and the Centre for Evidence-Based Medicine calling for registration and reporting of results of all clinical trials. Another project, OpenTrials.net, will aggregate information from a wide variety of existing sources to provide a comprehensive picture of all the data and documents available for all trials. One key source of trial data are clinical study reports, which are prepared by trial sponsors and transmitted to regulators. These documents are still infrequent, but are becoming increasingly publicly available through requests to the EMA and FDA. Moreover, the goal to obtain reporting transparency will be reached as prominent journals continue to establish clear requirements for making trial data available (Taichman et al., 2016). The clinicalstudydatarequest.com and the Yale University Open Data Access websites allow researchers to request access to individual patient data and supporting documents from industry-sponsored clinical trials. Moreover, the European Medicines Agency policy has released guidance on the publication of clinical data for medicinal products. This policy has

entered into force in 2015 for the publication of clinical reports but, in a later stage, it will also concern the publication of individual patient data.

Systematic reviews must adapt to these improvements and cover the multiple new information sources in their search strategies. Conference proceedings, clinical trial registries, regulatory agency reviews, and health technology assessment reports contain unpublished evidence that can be essential in resolving publication bias and selective outcome reporting.

References

2016. Assessing Risk of Bias in Systematic Reviews. *J Invest Dermatol*.
- Abuabara K, Freeman EE, Dellavalle R: 2012. The role of systematic reviews and meta-analysis in dermatology. *J Invest Dermatol* 132:e2.
- Hart B, Lundh A, Bero L: 2012. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ* 344:d7202.
- Hartung DM, Zarin DA, Guise JM, McDonagh M, Paynter R, Helfand M: 2014. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. *Ann Intern Med* 160:477-483.
- Ioannidis JP: 2008. Interpretation of tests of heterogeneity and bias in meta-analysis. *J Eval Clin Pract* 14:951-957.
- Lefebvre C, Manheimer E, Glanvill J: 2011. Chapter 6: Searching for Studies, in: Higgins J, Green S eds., *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0* (updated March 2011), The Cochrane Collaboration.
- McDonagh MS, Peterson K, Balshem H, Helfand M: 2013. US Food and Drug Administration documents can provide unpublished evidence relevant to systematic reviews. *J Clin Epidemiol* 66:1071-1081.
- Nankervis H, Baibergenova A, Williams HC, Thomas KS: 2012. Prospective registration and outcome-reporting bias in randomized controlled trials of eczema treatments: a systematic review. *J Invest Dermatol* 132:2727-2734.
- Papadopoulos E: 2007. Clinical Review NDA 20723 Aldara, Imiquimod 5% Cream.
- Roberts I, Ker K, Edwards P, Beecher D, Manno D, Sydenham E: 2015. The knowledge system underpinning healthcare is not fit for purpose and must change. *BMJ* 350:h2463.
- Scherer RW, Langenberg P, von Elm E: 2007. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev*:MR000005.
- Schroll JB, Abdel-Sattar M, Bero L: 2015. The Food and Drug Administration reports provided more data but were more difficult to use than the European Medicines Agency reports. *J Clin Epidemiol* 68:102-107.
- Taichman DB, Backus J, Baethge C, et al.: 2016. Sharing Clinical Trial Data: A Proposal From the International Committee of Medical Journal Editors. *Ann Intern Med* 164:505-506.
- Williams HC: 2015. Place your bet and show us your hand. *Br J Dermatol* 173:1104-1105.

Summary points

- Trials without statistically significant results are less likely to be published than trials that show apparent differences (publication bias). Moreover, trial outcomes that do not support the use of the new treatment are less likely to be published than those that do support its use (outcome reporting bias).

- Systematic reviews and meta-analyses that omit unpublished data are at high risk of biased conclusions. To increase their validity, systematic reviews should rely on a thorough search for published and unpublished trials.

- The Cochrane Central Register of Controlled Trials, MEDLINE, and EMBASE should be searched for published trials.

- Sources for finding unpublished trials have expanded recently. Conference proceedings, clinical trial registries, regulatory agency reviews, and health technology assessment reports should be searched for unpublished trials.

- A limitation is that there is no standard methodology yet to decide which sources of unpublished trials to search and how to search them.

Multiple choice questions

1) Which of the following would result in publication bias?

- a. Trials with negative results were not published and could not be selected in the systematic review
- b. Trials with statistically significant results were cited more often by subsequent articles, increasing the likelihood of being selected in the systematic review
- c. Trials were published in languages other than English and could not be selected in the systematic review
- d. Trials were published more than once, increasing the likelihood of the trial being selected in the systematic review
- e. All of the above

2) Searching beyond bibliographical databases for a systematic review potentially reduces

- a. Publication bias
- b. Validity of the systematic review
- c. Outcome reporting bias
- d. Labor intensity of the search
- e. A and C

3) The sources to search for published trials include

- a. MEDLINE only
- b. the Cochrane Central Register of Controlled Trials
- c. the Cochrane Database of Systematic Reviews
- d. EMBASE
- e. B C and D

4) The sources to search for unpublished trials include

- a. clinicaltrials.gov
- b. alltrials.net
- c. Drugs@FDA
- d. Proceedings to the American Academy of Dermatology Annual Meeting

e. A, C and D

5) Some limitations of sources of unpublished trials are

- a. Clinical trial registries include ongoing and completed trials and potentially posted trial results
- b. Reviews obtained from regulatory agencies typically lack sufficient detail to assess the risk of bias for a trial
- c. Conference abstracts are not restricted by treatment type (pharmacological and non pharmacological)
- d. Searching conference abstracts, clinical trial registries, regulatory and health technology assessment agency website is burdensome

e. B and D

Answers:

1) Only a. corresponds to publication bias (the whole trial results are made or not made publicly available according to the nature and direction of the results); b. corresponds to citation bias (citation or non-citation of a trial report, depending on the nature and direction of the results); c. corresponds to language bias (the publication of trial results in a particular

language) and d. corresponds to duplicate publication bias (multiple publication of trial results). e. All of the above correspond to reporting biases

2) a. and c. Searching sources such as conference proceedings, clinical trial registries, or regulatory agency websites may allow identifying trials with unpublished results, thus reducing publication bias; in particular, primary and secondary outcomes are pre-specified in clinical trial registries, thus allowing reducing outcome reporting bias.

3) b. c. and d. Searching MEDLINE only is insufficient as many relevant published trials are indexed in other databases such as EMBASE and CENTRAL . Searching c. would allow identifying previous systematic reviews on the same topic and screening the lists of selected trials.

4) a., c. and d correspond respectively to a clinical trial registry, a regulatory agency website, and a conference proceedings repository. d. corresponds to the website of an initiative calling for registration and reporting of results of all clinical trials

5) b. and d. Reviews obtained from regulatory agencies typically include few details about the trial methodology itself but they can be complemented with information from the trial protocol or a journal article. Another limitation is that comprehensive searching adds to the resources needed to complete the systematic review

Figure: Summary Workflow for Searching Databases in a Systemic Review

A Cochrane systematic review about oral antiviral therapy for prevention of genital herpes outbreaks in immunocompetent and nonpregnant patients is used as an illustration of this workflow (Le Cleach et al. Cochrane Database Syst Rev. 2014;8:CD009036. doi: 10.1002/14651858.CD009036.pub2). Details can be found in Online Appendix 1 and a tutorial to search FDA drug approval packages and EMA public assessment reports can be found in Online Appendix 2

Workflow

Example from Cochrane review CD009036 on genital herpes

A. First, find the published trials

- 1) Bibliographical databases (CENTRAL, MEDLINE, EMBASE)
- 2) Reference lists of all selected trials
- 3) Systematic review databases (Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects)

"We searched a range of bibliographical databases, including the Cochrane Central Register of Controlled Trials (CENTRAL); the specialized registers of the Cochrane Infectious Diseases Group, the Cochrane Skin Group, and the Cochrane Sexually Transmitted Infections Group; MEDLINE, EMBASE, and LILACS, with no restriction on language or date. [...] We screened the reference lists of all selected trials.

B. How to find the unpublished trials

- 1) Conference abstracts
- 2) Clinical trial registries and results databases
- 3) Contact trialists and sponsors
- 4) Industry trial registries and results databases

"We contacted the main authors in the field to identify any additional published or unpublished data. We searched the proceedings of the following conferences: European Congress International Union Against Sexually Transmitted Infections (IUSTI), [...]. We contacted the pharmaceutical companies [...] and searched the clinical trial results database of each company to [...] identify ongoing and unpublished trials. We also searched the search portal of the World Health Organization International Clinical Trials Registry Platform."

C. The untapped data buried in regulatory agency websites

- 1) Regulatory agency online databases
- 2) Health technology assessment agencies

"We searched reviews submitted to the Food and Drug Administration (FDA) for drug registration."