

Collecting data for meta-analysis – Part II

[Print](#)

Ideally, a meta-analysis should include every single completed study on a given research question, irrespectively of the results, published or not. Well, you can never be sure if you found everything. And most likely you will not, especially because we all have only a limited time to collect the data. Best strategy is to make an efficient use of existing search tools – online databases like [Scopus](#), [Web of Knowledge](#), [GoogleScholar](#) and many other more or less specialized ones. Starting from the one that you expect to have the best coverage of your research topic. Spending some time just exploring which search terms and combinations of search terms give best set of “hits” both in numbers but also accuracy – you don’t want to waste time reading thousands of completely irrelevant papers. Later, repeating the search using other databases just to fish out the papers you initially missed.

It is well worth to thoroughly read key review papers, not only to learn more about the topic, but also to check out the papers the review is referring to. Also, while using online search engines, it is usually quite easy to do the reverse (which other published papers the paper is citing) and forward (in which other papers the given paper was cited later on) searches of the references.

And there will come the point when the rate of finding suitable publications (study inclusion/exclusion criteria is a separate and important problem to think about) gets really low. Then, you may assume you got most of them. Does it matter if you don’t find all of them? In my opinion, it will not matter much. But only under two conditions: first, your search and data-entry were not biased (e.g. you did not include into the dataset only these results that showed expected/desired effect). Second, assuming that your sample of publications is random, you have tested the dataset for [publication bias](#).

Unfortunately, many authors of meta-analytic papers do not test for publication bias of their datasets, instead reassuring the readers and reviewers that their search was so exhausting that they didn’t need to do so. Or they simply don’t mention this potential problem at all.

Fortunately for these who want to do a good job analysing their data, there are couple of simple methods developed for testing for the presence of publication bias. There is also one easy method that allows adjusting for publication bias – so called trim-and-fill method ([Duval and Tweedie, 2000](#)). This method is implemented in [R](#) package [meta](#) as function [trimfill](#), in [SAS](#) as macro PubBias, in [STATA](#) as METATRIM, it is also included in [Comprehensive Meta Analysis](#) (CMA) and [MIX](#) software. Out of these programs only R is free – and the others are quite pricey (not an excuse for not doing proper analyses, just learn to use R!). However, the trim-and-fill method is not really a cure to the problem of publication bias, if it exists. It is more like a sensitivity analysis where you check whether you conclusion still holds after re-filling the dataset with simulated “not-published” values. For more details refer to the original paper ([Duval and Tweedie, 2000](#)) and a newer one ([Peters et. al 2007](#)).

Details

Written by Malgorzata (Losia) Lagisz

Category: General

Created: 09 June 2010

Last Updated: 27 August 2015

Hits: 1567976