

Around meta-analysis (4): data sharing and vanishing data

[Print](#)

We have noticed that my three blog posts ([I](#), [II](#), [III](#)) about collecting data for meta-analysis, which were published in 2010, received lots of hits from the Internet users. For example, when you use combinations of words: meta-analysis collecting data, or similar, one of these blog posts will appear on the first results page. This encouraged me to continue the series, just with a bit broader scope.

Recently, I came across a short letter published in Nature about tampering with the rules of data sharing for currently published papers (Bonniolo and Vaccari 2012). This subject is quite worrying, but the neglect regarding what happens with the data from slightly "older" papers is of even more concern to me, as the amount of research results that are, or quickly will be, unavailable for re-examination is massive.



Obviously, not much can be done to get access to the datasets created so long ago that their authors finished their academic careers. However, when it comes to more recent data from the active researchers, the classical principle of sharing scientific results is not embraced widely and easily. Relying on the good will of the authors is often futile as direct data requests usually get no reply or result in refusal to share. Authors can also be difficult to reach – their contact details change frequently as they change jobs and institutions. Relatively few journals request that authors make their data available by depositing them online. Such journal policies are often not enforced, resulting in broken links to data (Evangelou et al. 2005), partial or insufficiently described datasets being deposited (Alsheik-Ali et al. 2011).

There are some costs and risks of data sharing, and, apart of a good feeling from helping others, there are not many benefits for scientists sharing their data, except increased citation rate for their original studies (Piwowar et al. 2007). Despite numerous calls from researchers in various disciplines, journals do not give any direct incentives for data sharing. They even do not provide extra space allowances in the main paper text for referencing primary studies associated with the re-analysed datasets (which can sometimes run into the hundreds in meta-analyses).

Scientific data should be freely available for re-use by other scientists and for evaluation by non-scientists, but we are still a long way from reaching this ideal.

References:

Bonniolo, G. & Vaccari, T. Publishing: Alarming shift away from sharing results. Nature 488,

157–157 (2012). [\[link\]](#)

Alsheikh-Ali, A. A., Qureshi, W., Al-Mallah, M. H. & Ioannidis, J. P. A. Public Availability of Published Research Data in High-Impact Journals. PLoS ONE 6, e24357 (2011). [\[link\]](#)

Evangelou, E., Trikalinos, T. A. & Ioannidis, J. P. A. Unavailability of online supplementary scientific information from articles published in major journals. FASEB J 19, 1943–1944 (2005). [\[link\]](#)

Piwowar, H. A., Day, R. S. & Fridsma, D. B. Sharing Detailed Research Data Is Associated with Increased Citation Rate. PLoS ONE 2, e308 (2007). [\[link\]](#)

Details

Written by Malgorzata (Losia) Lagisz

Category: General

Created: 22 November 2012

Last Updated: 03 February 2015

Hits: 1284109