VIEWPOINT

Meta-Analyses Can Be Credible and Useful
A New Standard

John P. A. Ioannidis, MD, DSc
Departments of Medicine and Health Research and Policy, Stanford University School of Medicine, Stanford, California; and Meta-Research Innovation Center at Stanford, Stanford University, Stanford, California.

Corresponding Author: John P. A. Ioannidis, MD, DSc, Departments of Medicine and Health Research and Policy, Stanford University School of Medicine, 1265 Welch Rd, MSOB X306, Stanford, CA 94305 (jioannid@stanford.edu).

Overview

Carefully done meta-analyses constitute a major advance compared with expert opinion and nonsystematic attempts at summarizing, synthesizing, and integrating information. Meta-analyses serve many fields in summarizing an increasing stream of data and, for clinical purposes, streamlining information for decision making. However, there are flaws and caveats that threaten the validity and utility of meta-analyses.[1]

Most of these efforts are retrospective exercises that try to piece together fragments of information from multiple completed studies. They depend on information that is already published (or at least retrievable) with all the accompanying errors and biases and rarely correct these problems. For example, 8 meta-analyses of imaging in unipolar depression have reached inconsistent conclusions because they have used different studies with diverse protocols and methods that are difficult to standardize post hoc and different errors that other meta-analyses may or may not correct for. A meta-analysis should systematically probe, detect, dissect, and highlight major errors and biases (instead of sanctifying flawed studies by their inclusion). Careful bias scrutiny alone can be a major service to a field. However, it is always tempting to take a shortcut to talk about summary effects and forget about the deficiencies of the evidence.

The Existing Evidence

Bias is inflated when meta-analyses are done by authors and/or sponsors with financial or other conflicts of interest. Authorship by company employees and/or sponsorship by companies was the strongest risk factor for reporting no caveats for antidepressants among a body of 185 meta-analyses on antidepressants for depression published between 2007 and 2014.[2] Allegiance bias may be an equivalent problem for evidence on psychotherapies.[3] Conflicted meta-analyses compound the distortion that exists in the publication process of primary studies, including a publication bias against negative results, the selective reporting of negative trials as positive,[4] and other spins (eg, changing the analysis plan or the focus of interpretation) that lead to more favorable results and interpretations.

Of approximately 20 000 meta-analyses performed annually,[1] well over 1000 have relevance for mental health. In an empirical survey using stringent criteria for labeling meta-analyses,[5] 7% pertained to mental and behavior disorders. Most of these meta-analyses look only at published data and circumscribed, small fractions of the evidence that might be relevant for the question of interest. For example, in therapeutics research, of 822 network meta-analyses of clinical trials published until May 2015, only 39 pertained to mental health. Moreover, to my knowledge, there are few meta-analyses in mental health that have been able to use individual-level data. In a database of 829 meta-analyses with individual-level data published until 2012,[6] only 52 (6.3%) pertained to mental and behavioral disorders. Most of them either pertained to nontherapeutic questions (eg prognostic, biomarker, imaging, and association studies) that mostly had no clinical relevance or lacked systematic searches. I found only 7 meta-analyses that addressed therapeutic questions and had performed a systematic search to retrieve trials, and 4 of these dealt with only a single drug. What is more common is pooling projects done by the industry in which a few trials on a specific sponsored drug have their individual-level data combined. Typically, these pooling exercises work as marketing efforts reassuring, by default, that the drug is effective and safe. Typically, they do not perform a systematic review for the assessed drug, nor do they consider drugs from other competitors. While meta-analyses with individual-level data have become more common in other popular applications of the method beyond therapeutics—in particular for imaging studies—most imaging meta-analyses still depend on published group-level data and examine narrow questions.

Therefore, almost all meta-analyses with systematic searches depend on published group-level data and/or examine small fragments of the evidence space. However, when there are dozens of therapeutic options, as is the case with antidepressants or antipsychotics, a meta-analysis focusing on 1 agent has limited use because it says nothing about the relative benefits and harms of this agent compared with other competitors. Similarly, when data are combined from only a few imaging studies, the emerging picture will be uncertain and possibly misleading depending on what factors have shaped data availability.

The few network meta-analyses published to date in the field offer a wider view of the evidence by considering multiple treatments and understanding better their relative benefits and harms. However, those efforts are still in an exploratory phase and may disagree on the final conclusions. For example, network meta-analyses on antidepressants for depression have yielded radically different conclusions on the relative ranking of various antidepressants. Differences may be caused by variability in the eligibility criteria, efforts (or lack of efforts) to overcome publication and other selective reporting bias, the choice of outcomes, and analytical methods, among other factors.

Improvements

Improvements in meta-analyses may stem from improvements in the primary studies that they synthesize and

improvement in the design, conduct, and reporting of meta-analyses themselves.

First, there is encouraging news that after many years of inertia, protocols and individual-level data from randomized clinical trials may start becoming more routinely available for use in reanalyses, meta-analyses, or other purposes. Similarly, several efforts try to promote the availability of raw data from other studies, such as brain imaging (eg NeuroVault [neurovault.org] and OpenfMRI [openfmri.org]). The use of individual-level data may become more convenient, and thus meta-analyses of individual-level data may offer advantages and bypass the shortcomings of partial, biased coverage of the evidence by the available raw data.[7]

Second, there are ongoing efforts to increase the proportion of registered clinical trials and the completeness of information provided on registration.

Third, constructive criticism on the poor standards of many mental health trials—eg, the use of short follow-up, nonrelevant, extremely diverse, and nonstandardized outcomes, and use of inappropriate statistical methods such as the last observation carried forward—may start improving the primary material that meta-analyses assess and combine. This also applies to improved methods and standardized protocols for other types of studies, such as imaging.

Fourth, the wider adoption of reporting standards for meta-analyses (including the Preferred Reporting Items for Systematic Reviews and Meta-analyses and its extensions for harms and networks) may couple with the adoption of standards for meta-analysis protocols (Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocols).

Fifth, the preregistration of systematic reviews and meta-analyses in registries such as PROSPERO may also enhance transparency. It should also become easier to update meta-analyses with new data and to consider systematic reviews as "living documents"[8] able to incorporate new evidence in real time. Otherwise, numerous meta-analyses on the same question are an avoidable redundancy.

Sixth, network meta-analyses should also become more inclusive and should be considered routinely, whenever appropriate, in therapeutic topics in which multiple intervention choices are available. Finally, systematic reviews and meta-analyses also need to strengthen their editorial independence and protection from conflicts of interest.

Conclusions

Primary studies and meta-analyses should eventually become more confluent and, whenever possible, they should coincide. For many topics of prognostic, biomarker, and association research, the consortia of multiple teams can work with pre-agreed protocols and statistical, clinical, and laboratory methods to generate data prospectively and integrate them in a single meta-analysis. To my knowledge, the consortium paradigm has always been very successful in omics fields and is immediately relevant to imaging and biomarker studies. For therapeutics research, meta-analyses may also be designed upfront with the intention to summarize a research agenda of multiple primary trials. The research agenda is constructed with the anticipation that all of the studies will provide full detailed data to an ongoing updated meta-analysis.[9] The meta-analysis update can also inform the need for future studies, as well as the sample size and comparisons that they should address.[10] Sufficient safeguards should be in place to guarantee the independence of these large living individual-level network data syntheses from conflicts of interest. Meta-analysis can become the new prototype of robust, primary, original research.

ARTICLE INFORMATION

Published Online: February 22, 2017. doi:10.1001/jamapsychiatry.2017.0035

Conflict of Interest Disclosures: None reported.

REFERENCES

1. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94(3):485-514.

2. Ebrahim S, Bance S, Athale A, Malachowski C, Ioannidis JP. Meta-analyses with industry involvement are massively published and report no caveats for antidepressants. *J Clin Epidemiol*. 2016;70:155-163.

3. Cuijpers P, Cristea IA. How to prove that your therapy is effective, even when it is not: a guideline. *Epidemiol Psychiatr Sci*. 2016;25(5):428-435.

4. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358(3):252-260.

5. 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(5):e1002028.

6. Huang Y, Mao C, Yuan J, et al. Distribution and epidemiological characteristics of published individual patient data meta-analyses. *PLoS One*. 2014;9(6):e100151.

7. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *BMJ*. 2012;344:d7762.

8. Page MJ, Moher D. Mass production of systematic reviews and meta-analyses: an exercise in mega-silliness? *Milbank Q*. 2016;94(3):515-519.

9. Ioannidis JPA, Karassa FB. The need to consider the wider agenda in systematic reviews and meta-analyses: breadth, timing, and depth of the evidence. *BMJ*. 2010;341:c4875.

10. Nikolakopoulou A, Mavridis D, Salanti G. Planning future studies based on the precision of network meta-analysis results. *Stat Med*. 2016;35(7):978-1000.