STATISTICAL SIGNIFICANCE

Meta-analysis is widely used to combine the evidence from multiple studies on the efficacy of a treatment. While there is extensive research on the pooling methods, there is little to no research on the critical step of extracting data from published research papers where there is a lot of missing data.

Data extraction of published treatment effects

META-ANALYSIS. In the health sciences, meta-analyses provide stronger and more accurate evidence of the efficacy of a treatment. Ideally, a meta-analysis would pool the individual patient data from all studies. In practice this is not possible. In fact, less than one percent of all meta-analyses uses this approach. Instead, meta-analyses rely on data extracted from published literature. Unfortunately, published research papers often lack the necessary information to feed into a meta-analysis.

DATA EXTRACTION. While there has been an effort to promote better reporting practices, currently there is no established protocol on how to approach the problem of missing information in the published research papers. In most cases, published meta-analyses do not provide the details on how they addressed the missing data problem –data is simply manipulated into a convenient format without justification.

UNRELIABLE RESULTS. To get an idea of the degree to which poor reporting of summary statistics is a problem, a study found that only 35.5% of the estimates required for a meta-analysis were available. In addition, we re-analyzed a meta-analysis of mortality after one year between two treatments and found that 9 out of 10 studies had a missing data problem. Under these circumstances it is clear that the reliability of the resulting meta-analyses should be questioned.

RELIABLE RESULTS. We provide a novel methodology[1] that allows us to use information found in each published study to estimate the missing values. Since each study is different in terms of the amount and quality of the information it provides, the methodology allows modeling each study individually.

The resulting treatment effect has a larger variance compared to that of a meta-analysis that does not adjust for uncertainty from the missing data. It is clear that this difference may turn some significant into not significant results, dramatically changing conclusions and corresponding health implications.

Our methods would not be necessary if published studies provided better estimates and complete information of their follow-up issues. Unfortunately, the poor reporting of summary statistics will continue to prevail in the published literature, making our proposed methodology a requirement to any meta-analysis performed that does not include individual patient data.

[1]Bayesian Meta-Analysis Data Extraction.
Shemra Rizzo[2], Robert E. Weiss[2], Raj R. Makkar[3].
[2] Biostatistics department, University of California, Los Angeles.
[3]Heart Institute, Cedars-Sinai Medical Center, Los Angeles, California.