Comments on Statistical Issues in November 2012

Commentary

Yong Gyu Park

Department of Biostatistics, The Catholic University of Korea College of Medicine, Seoul, Korea

In this section, we explain the statistical models and evaluation of heterogeneity among studies in meta-analysis, which appeared in the articles titled, "A comparison of efficacy and safety of nonsteroidal anti-inflammatory drugs versus acetaminophen in the treatment of episodic tension-type headache: a meta-analysis of randomized placebo-controlled trial studies" by Yoon et al.¹⁾ and "Use of proton pump inhibitor and risk of colorectal cancer: a meta-analysis of observational studies" by Ahn et al.²⁾, both published in September 2012.

Statistical Models in Meta-analysis

It is often the case that more than one investigation is performed to study a particular research question, often by different research groups. In some instances, results are seemingly contradictory, with some research groups reporting significant differences for a particular finding while other research groups report no significant differences.

The issue is, what is the appropriate way to combine evidence across all the studies so as to reduce sampling error and increase the power of the investigation, and in some instances to resolve the inconsistencies among the study results? The (statistical) method for accomplishing this is called meta-analysis.

It is controversial among research workers whether a fixed- or random-effects model should be used when performing a meta-analysis. Under a fixed-effects model, the between-study variance is ignored in computing the study weights and only the within-study variance is considered. In a random-effects model, both between- and within-study variation are considered. If the between-study variation is substantial relative to the within-study

variation, then larger studies will be given proportionally more weight under a fixed-effects model than under a random-effects model. Hence, the pooled estimates under these two models may also be different.

It is argued that if there is substantial variation among the study results, then one should investigate the source of the heterogeneity (e.g., different study designs, etc.) and not report an overall pooled estimate. Others insist that between-study variation should always be considered in meta-analysis. A reasonable compromise might be to check for significant heterogeneity among study-specific estimates and use the decision rules as follows (Table 1).³⁾

Evaluate the Heterogeneity among Studies

Cochran's Q-test⁴⁾ and Higgins' I²-statistic⁵⁾ are frequently used to evaluate the statistical heterogeneity among study results. Cochran's Q-test provides us with a P-value which can be used directly to make a decision about the heterogeneity among studies. However, the small number of studies used in most meta-analyses gives us very low statistical power, thus the P-value of Cochran's Q-test usually shows no significance even though the heterogeneity among studies seems quite large. So we often use a significance level of 0.1 instead of a conventional 0.05 in meta-analysis. Higgins' I²-statistic is derived from Cochran's Q-statistic and defined as the proportion of variation due to the heterogeneity from total variation among studies.

$$I^2 = ([Q - df] / Q) \times 100 (\%)$$

where $df = k - 1$, k = number of studies.

Table 1. Criteria for choosing statistical models.

P-value for heterogeneity	Model
$P \geq 0.5$	Fixed-effects model
$0.05 \leq P < 0.5$	Random-effects model
$P < 0.05$	Do not report pooled estimate; Assess sources of heterogeneity

Table 2. Degree of heterogeneity and values of I^2 .

Value of I^2	Degree of heterogeneity
$I^2 < 25\%$	Low
$25\% < I^2 < 75\%$	Moderate
$I^2 > 75\%$	Severe

Values of I^2 are between 0% and 100% as shown in table 2, and 0% represents no heterogeneity among studies (In this case, both fixed- and random-effects models have the same results). Negative values are treated as 0% and $I^2 = 50\%$ is used as a criterion for choosing a statistical model.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

1. Yoon YJ, Kim JH, Kim SY, Hwang IH, Kim MR. A comparison of efficacy and safety of non-steroidal anti-inflammatory drugs versus acetaminophen in the treatment of episodic tension-type headache: a meta-analysis of randomized placebo-controlled trial studies. *Korean J Fam Med* 2012;33: 262-71.
2. Ahn JS, Park SM, Eom CS, Kim S, Myung SK. Use of proton pump inhibitor and risk of colorectal cancer: a meta-analysis of observational studies. *Korean J Fam Med* 2012;33:272-9.
3. Rosner B. *Fundamentals of biostatistics*. 5th ed. Pacific Grove: Duxbury; 2000.
4. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;10:101-29.
5. Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J Health Serv Res Policy* 2002;7:51-61.