Meta-analysis: A tool for better understanding of medical research

V. Raman Kutty

Quite often, published results of studies addressing the same questions provide different or even conflicting answers. This is true of many fields such as psychology and education, but particularly of epidemiology and medicine. Review articles provide partial solutions to the problem, but cannot be said to be completely free of subjectivity. But the statistical technique of meta-analysis is designed to provide a scientific way out.

Meta-analysis has become quite popular as a method. A scan of the medical literature using computerized search (MEDLINE-Compact Cambridge) yielded 52 meta-analyses published in 1990 alone. Meta-analysis has been defined as the 'practice of using statistical methods to combine outcomes of a series of different experiments/investigations'[1]. The purposes of meta-analysis can be: (i) to increase statistical power for primary endpoints or subgroups, (ii) to resolve uncertainty when reports disagree, (iii) to improve estimates of effect size, (iv) to answer questions not posed at the start of the individual trials[2]. Thus meta-analysis is a research strategy in which a study is itself the unit of observation. Just as in other experiments/observations, the meta-analyst has to guard against bias in many forms.

Epidemiological studies can be broadly categorized on the basis of methodology into hypothesis-testing studies and effect-estimating studies. The first variety attempts to prove or reject a certain pre-stated hypothesis; the second variety is more precise in the sense that it attempts to estimate the magnitude of certain relationships such as risk ratios or risk differences within a stated confidence interval. Both these types of studies are amenable to meta-analysis, even though effect estimation should be the ultimate objective in most meta-analytic approaches. This is because, to say that a certain effect is of a particular degree of magnitude is a stronger statement than to say that the effect is present.

Meta-analyses have addressed such important research questions in recent times as the importance of physical exercise in reducing risk of coronary artery disease[3], the value of propranolol in preventing variceal bleeding[4], the use of CT scan to stage lung cancer[5], and the use of the McGill questionnaire to measure pain[6]. A common approach is to use meta-analysis in integrating the results of several randomized controlled trials of a new therapeutic intervention or drug.

The steps in meta-analysis, as in any other epidemiological study, have been defined as (i) formulating specific goals, (ii) locating and evaluating the published literature, (iii) choosing a common metric of outcome, (iv) deciding on a statistical model for the analysis ('fixed effects' versus 'random effects'), (v) statistical analysis, and (vi) interpretation. As can be seen, statistical analysis is only one part of the whole programme, and usually only elementary statistical techniques are used. Choosing a common outcome measure by which to combine several studies can often prove to be the trickiest part of meta-analysis. The question of the statistical model is decided by the analyst in the light of existing knowledge in the field. There are mainly two approaches, viz. (i) the fixed-effects model, in which the several outcomes are assumed to be the only ones possible, and (ii) the random-effects model, in which the observed outcomes are assumed to be among several possible.

A considerable amount of bias is likely to occur at the collection-of-literature stage. Studies vary enormously in their range of quality. To provide equal weight to all would be to bias the analysis in favour of poorly defined studies. It has been advised that, after collection of the literature, an initial qualitative assessment of the literature, conducted blind to the published outcomes, should categorize the papers based on the methods section into unacceptable, acceptable and good[7]. As one can imagine, there is scope for considerable degree of difference of opinion in weighting studies. In further analysis, the papers should be weighted according to their merits and a pooled measure of outcome derived.

An important bias to be remembered is the 'publication bias', i.e. the fact that 'positive' results are more likely to be published, especially with regard to randomized controlled trials. Moreover, in a survey of 71 negative trials, Freeman *et al.*[8] have shown that in many studies the sample size was too small to detect a 25% difference between the groups at a 90% probability level. It is quite possible that, if all the studies had been pooled together and subjected to meta-analysis, probably the outcome and interpretation would have been quite different.

One of the interesting outcomes of this approach of looking afresh at published data has been the description of the 'regression–dilution bias'. In studies relating adverse outcomes to continuous variables such as blood pressure and serum cholesterol, it is common in epidemiology to employ regression techniques and describe a regression equation. MacMahon and others have shown that such studies, due to unbiased (random) fluctuations of the measurement process, are likely to systematically underestimate the regression slope, and therefore the magnitude of the effect of the variable on the outcome[9]. In other words, if we take this bias into consideration, blood pressure and cholesterol have a worse effect on the probability of coronary events than was previously thought. They also suggest some statistical 'fixes' for this.

Meta-analysis, for all the advantages conferred by the computer search, should not be thought of as an easy task. The success of this technique depends on the quality of the primary studies, on there being enough of them, and on the possibility of finding a universal common metric[10]. When a

'meta-meta-analysis' of 86 meta-analyses of randomized controlled trials using a scoresheet of 23 items divided into six major criteria (study design, combinability, control of bias, statistical analysis, sensitivity analysis and applicability of results) was done, only 28% were found to have addressed all six items[2].

1 Laird, N. M. and Mosteller, F., *Int. J. Technol. Assessment in Health Care*, 1990, **6**, 5.

2. Sacks, H. S., Bernier, J., Reitman, D., Ancona-Berk, V. A and Chalmers, T. C., *New Engl. J. Med.*, 1987, **316**, 450.
3. Berlin, J. A. and Colditz, G. A., *Am. J. Epidemiol.*, 1990, **132**, 612.
4. Hayes, P. C., Davis, J. M., Lewis, J. A. and Bouchier, I. A., *Lancet*, 1990, **336**, 153.
5. Dales, R. E., Stark, R. M. and Raman, S., *Am. Rev. Resp. Dis.*, 1990, **141**, 1096.
6. Wilkie, D. J., Savedra, M. C., Holzemer, W. L., Tesler, M. D. and Paul, S. M., *Nurs. Res.*, 1990, **39**, 36.
7. Yach, D., *South Afr. Med. J.*, 1990, **78**, 94.
8. Freeman, J. A., Chalmers, T. C., Smith, H. (Jr) and Kubler, R. R., *New Engl. J. Med.*, 1978, **299**, 690.
9. MacMahon, S., Peto, R., Cutler, J., Collins, R., Sorlie, P., Neaton, J., Abbott, R., Godwin, J., Dyer, A. and Stamler, J., *Lancet*, 1990, **335**, 765.
10. Graney, M. J. and Engle, V. F., *J. Gerontol. Nurs.*, 1990, **16**, 16.

V. Raman Kutty is in the Sree Chitra Tirunal Institute for Medical Sciences and Technology, Thiruvananthapuram 695 011.