In this issue of Mayo Clinic Proceedings, Blauwet et al report on the relative frequency with which clinical studies in cardiology are published with the inclusion of sex-specific findings. They found that in a sample of recent articles, a minority (37% in general medical journals and 23% in cardiology journals) reported sex-specific results. They noted some progress but a continuing gap in knowledge about treatment efficacy in women due in part to the persistent tendency not to report sex-specific results in cardiovascular treatment trials. They acknowledged the possible lack of power but maintained that this should not deter the reporting of sex-specific results because the reporting would at least facilitate the possibility of analyzing results for women in meta-analytic approaches.

From a statistical point of view, this issue presents a dilemma. On the one hand, we are often cautioned in regard to reporting subgroup specific results from clinical trials because of the danger of overinterpretation of apparent (but not real) differences between the findings in different subgroups and because of the fact that, often, trials are not designed with enough power to make self-contained inferences specific to each subgroup. Sex-specific reporting would seem to be another example of this perilous practice. On the other hand, sex is a variable that is well-known a priori to influence many diseases and treatment responses, including those of cardiovascular disease. Is it not, therefore, always better to present more detailed (ie, sex-specific) information, even if it is not precise?

The answer to that last question is less obvious than it might appear. The decision to present subgroup specific results tends to carry the implication that the subgroup specific results are more accurate than the “overall results,” which may or may not be true. Inaccuracy of an estimate of a treatment effect has 2 components—bias and variance. Bias is the inaccuracy caused by ignoring the variability of the “true” treatment effect from one sex (or subgroup) to the other. Variance is the inaccuracy caused by sampling error, ie, by smaller sample size. Consider 2 possible estimates of the effect of a treatment on women: (1) $E_1$, the effect based on the analysis of the women only, and (2) $E_2$, the “overall” effect based on men and women. The estimate $E_1$ might be said to be “unbiased,” in that it would be perfectly accurate in a very large sample. However, it may have a large variance—women may contribute the minority of patients, or even more to the point, the minority of events. On the other hand, the estimate $E_2$ will tend to have less variance but may be biased. In any given situation, the “optimal” choice between $E_1$ and $E_2$ is unknown. Indeed, one can argue that the truly best estimate of the treatment effect in women is some compromise between $E_1$ and $E_2$.

Therefore, one may be getting a very precise estimate of the “average” male-female treatment difference, but the same issue of heterogeneity that accompanies any meta-analysis potentially clouds the interpretation here as well.

Finally, any attempt to perform a sex-stratified meta-analysis in a context in which only a third of studies report sex-specific results may be seriously compromised by a version of the “publication bias” issue. That is, one has to be worried about whether sex-specific results were examined and not reported in some or many of the studies that did not report sex-specific results. To the extent that this is the case, one could be obtaining a very selective set of sex-specific results on which to base sex-specific conclusions. This consideration clearly lends support to the authors’ suggestion that sex-specific results should be routinely reported because the less selective the reporting, the better the meta-analysis.

Thus, this issue of sex-specific reporting emphasizes the different scientific criteria that would be used for reporting results in a single study, if it were to be considered a self-contained study, vs the criteria for reporting the same results if the study were merely to be considered as 1 of a population of studies to be the subject of meta-analysis. It is not easy to reconcile these 2 different points of view, but it is certainly worth some thought and attention toward reconciling them.

Kent R. Bailey, PhD
Division of Biostatistics
Mayo Clinic College of Medicine
Rochester, Minn

Address correspondence to Kent R. Bailey, PhD, Division of Biostatistics, Mayo Clinic College of Medicine, 200 First St SW, Rochester, MN 55905 (e-mail: baileyk@mayo.edu).

© 2007 Mayo Foundation for Medical Education and Research

See also page 166