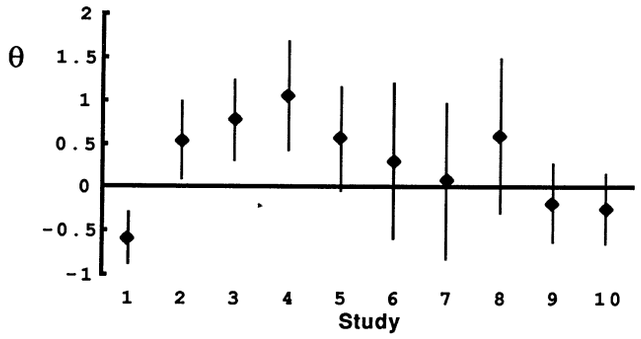

FIG. 1. *Estimates* (±2 *SE*) *for each study from Table* 4.

clear that the hypothesis of a common value of θ for every study is ruled out. Studies 2, 3 and 4 definitely have higher effect sizes than average, studies 1, 9 and 10 are definitely lower than average, although the other four studies fall somewhere in the middle. That granted, just what does the subsequent analysis, based on the assumption of equal θ's, illustrate? It is hard to tell whether the features of the method being discussed, such as the behavior of the likelihood contours, are typical or just due to a poorly fitting model. Furthermore, the standard errors presented in Section 4 depend on the model being appropriate.

Can the authors' selection model be extended to handle with case where the ten "true values" of θ have been drawn from a superpopulation with mean μ and variance σ^2? Perhaps it can, especially if the statistician has, and is willing to use, prior information about the distribution of true values and about the mechanisms governing the selection bias. In their conclusion, the authors raise the question of design issues for meta-analyses. What possible design issues can arise if the meta-analysis uses such a simplified model? Once one admits a component of variance for between study variation, the trade-off between making many smaller studies or fewer large studies begins to get interesting. If, in addition, one uses other characteristics differentiating the studies to build a hierarchical prior distribution for the θ's, then design considerations can become paramount, as discussed in DuMouchel and Harris (1983) and DuMouchel and Groër (1987).

ADDITIONAL REFERENCE

DUMOUCHEL, W. and GROËR, P. (1987). A Bayesian methodology for scaling radiation studies from animals to man. Presented at the 26th Hanford Life Sciences Symposium, October 1987. *Health Physics*. To appear.

Rejoinder

Satish Iyengar and Joel B. Greenhouse

Our objectives in writing this paper were to illustrate a practical application of selection models as a technique for sensitivity analysis in meta-analysis, to develop further statistical methodology for the file drawer problem and to identify statistical and practical issues related to the theory and practice of meta-analysis. We are indebted to the discussants on several accounts. Each of them has made fundamental contributions to the selection model or meta-analysis literature, and the roots of this paper are found in these earlier works. Furthermore, in their comments they have suggested modifications and alternative approaches that are likely to improve the methods discussed as well as the general practice of meta-analysis.

Professor Hedges and Professors Rosenthal and Rubin suggest that the issue of publication bias is overemphasized. Hedges believes that the related problem of "reporting bias" where studies test many hypotheses and report sufficient statistics only for results that achieve statistical significance is more widespread. Rosenthal and Rubin point out that the usual file drawer problem portrays a rather extreme view of publication bias and in fact empirical research shows that "neither nonsignificant nor unpublished means unretrievable." Although both of these points are well taken, nevertheless, Dickersin, Chan, Chalmers, Sacks and Smith (1987) report that "the results of published RCTs (randomized clinical trials) are more likely to favor the new therapy than are the results of unpublished RCTs . . . " and conclude that " . . . it seems likely that bias against the publication of 'negative' results does exist." As we note in our paper, with the general interpretation of the weight function as a model for the selection mechanism, both reporting and retrieval bias can be treated as special cases of the general methodology. Finally, we hope that as authors and editors of journals begin to adopt guidelines for reporting statistical studies, such as those suggested by Bailar (1986), the problems of reporting bias might diminish.

The empirical results concerning publication bias presented by Rosenthal and Rubin are interesting but