strength of the evidence against a null hypothesis. The fail-safe sample size can likewise be considered as measuring the weight of evidence. Thus, $n(0)$ is the number of *hypothetical* studies, conducted under the null hypothesis, that need to be added to the database to just offset a significant result. Here "offset" is used in the sense that the updated test statistic should have its expected value, *conditional on the published studies*, just equal to the critical point.

The authors' equation (4) modifies the foregoing in two respects: (i) the hypothetical studies are replaced by actual and unpublished studies and (ii) a weight function is made available for modeling the publication bias. Now we have difficulty interpreting the numerator of (4) as a conditional expectation. Granted the values reported in the published studies are noninformative for the unpublished studies, but the number k of published studies, when coupled with the weight function, is informative and should be reflected in the conditional expectation. Thus, does $n(\alpha)$ in equation (4) refer to actual studies (in file drawers) or to hypothetical studies?

As an alternative, we suggest using the weight function to estimate the number of unpublished studies. For example, letting k and k_0 be the number of published and unpublished studies and $N = k + k_0$,

$$k_0 \approx N \int (1 - \omega(x)) f(x)\, dx,$$

$$k \approx N \int \omega(x) f(x)\, dx,$$

and k_0 can be estimated as the solution of

$$\frac{k_0}{k} = \frac{E[1 - \omega(X)]}{E[\omega(x)]}. \tag{3}$$

Note that the righthand side of (3) reduces to a $(1 - \alpha)/\alpha$ in case of a dichotomous weight function. Also, $\omega(x)$ must have a phenomenological interpretation as a probability. It will not do to replace ω by a scalar multiple or to use an unbounded ω.

The solution of (3) is sensitive to the weight function, and one may prefer to look upon the estimate of k_0 as shedding light upon the reasonableness of the chosen ω instead of upon the "true" significance of the published results.

ADDITIONAL REFERENCES

GLYNN, R. J., LAIRD, N. M. and RUBIN, D. B. (1986). Selection modeling versus mixture modeling with nonignorable nonresponse. In *Drawing Inferences from Self-Selected Samples* (H. Wainer, ed.). Springer, New York.

HANSEN, M. H. and HURWITZ, W. N. (1946). The problem of nonresponse in sample surveys. *J. Amer. Statist. Assoc.* **41** 517–529.

LITTLE, R. J. A. (1983). The nonignorable case. In *Incomplete Data in Sample Surveys* (W. G. Madow, I. Olkin and D. B. Rubin, eds.). Academic, New York.

PETO, R. (1987). Why do we need systematic overviews of randomized trials? *Statist. Med.* **6** 233–240.

RUBIN, D. B. (1977). Formalizing subjective notions about the effect of nonresponse in sample surveys. *J. Amer. Statist. Assoc.* **72** 538–543.

SIMES, R. J. (1987). Confronting publication bias: A cohort design for meta-analysis. *Statist. Med.* **6** 11–29.

Comment

M. J. Bayarri

"Meta-analysis, like rock and roll, is here to stay" claim the authors of this interesting and stimulating paper, and they are right. Similar experiments are conducted and replicated, providing informaiton about the same unknown quantity, and statisticians have to face the challenge of providing methods for pooling this information. In a sense, the problem is similar to that of combining a set of expert opinions. Unfortunately, results from experiments are not, in general, expressed as distributions of the unknown quantity. If they were, then not only would the publication bias due to statistical significance be greatly reduced, but also the techniques for combining probability distributions would be available (for an excellent summary and comprehensive annotated bibliography about these techniques see Genest and Zidek, 1986). Moreover, meta-analysis is usually based on results that got published in the scientific literature. Due to the overabuse of hypothesis testing as a statistical methodology and to the overappreciation of statistical significance, it does not come as a surprise that publications are highly biased toward studies showing statistically significant results. This publication bias should be taken into account when carrying out a

M. J. Bayarri is Titular Professor at the University of Valencia (Spain). Her mailing address is Departamento de Estadística e IO, Facultad de Matemáticas, Av. Dr. Moliner 50, 46100 Burjasot, Valencia, Spain.