(1985). *Evaluation of Short-Term Tests for Carcinogens: Report of the International Programme on Chemical Safety Collaborative Study on In Vitro Assays.* North Holland, Amsterdam.

COCHRAN, W. G. (1937). Catalogue of uniformity trial data. *J. Roy. Statist. Soc. Suppl.* **4** 233–253.

MARGOLIN, B. H. (1985). Statistical studies in genetic toxicology: A perspective from the U. S. National Toxicology Program. *Environ. Health Perspectives* **63** 187–194.

MARGOLIN, B. H., RESNICK, M. A., RIMPO, J. Y., ARCHER, P., GALLOWAY, S. M., BLOOM, A. D. and ZEIGER, E. (1986). Statistical analyses for in vitro cytogenetic assays using Chinese hamster ovary cells. *Environ. Mutagen.* **8** 183–204.

Comment

Richard G. Cornell

Rosenbaum's review provides a logical framework for thinking about the value of more than one control group in observational studies and contains useful discussion of the implications for the design and interpretation of investigations in which randomization to control and treatment groups is not feasible.

I found his discussion of case-control studies particularly interesting. Controls in these studies are "noncases" selected for comparison with a group of "cases" known to have a particular disease or other condition. Controls and cases are compared with respect to the extent of exposure to potential causative agents or with respect to other background variables. Rosenbaum emphasizes that a comparison of the histories of two or more control groups provides a check of the assumptions that underlie the estimation of the effect of exposure after covariance adjustment. The example he presents on the extent of exposure to sunlight of cataract cases and controls involves three control groups with other eye conditions. The use of multiple control groups enabled him to conclude that adjustment for age and sex is not sufficient for unbiased estimation of the effect of sunlight exposure on the prevalence of cataracts.

This example serves as a prototype for the interpretation of other case-control studies with more than one control group and a warning of possible undetected bias in case-control studies with only one control group. More importantly, it serves as a reminder that it is best to select more than one control group when the ideal control group cannot be formed through randomization. This allows a check on assumptions that cannot be attained through randomization and yet are crucial to conclusions on the effect of exposure after taking covariates into account.

Rosenbaum also refers to an example of a study in which the groups to be compared are exposure groups and the outcome is the prevalence of coronary thrombosis, with subclasses within each exposure group defined by covariates. Observational studies of this type are common in epidemiology and medicine. One example is a study, described by Cornell (1984), of cancer rates relative to exposure to the environment in steel plants, which produce stainless steel. One purpose of this study was to see if there is evidence of an increase in lung cancer rate attributable to the nickel and chromium used in stainless steel production. Exposure groups were formed by area worked within a plant. Comparisons were made after adjusting for age.

Another example is the comparison of survival rates for burn victims using registry data grouped by hospital, and subsequently by the speed of wound closure attained by the burn care practice in a hospital. Again, age is an important covariate for comparisons of burn survival. So is burn severity as measured by the extent of full thickness burn. A model that takes these variables as well as other demographic and severity variables into account for purposes of estimation, prediction and evaluation is discussed by Wolfe, Roi, Flora, Feller and Cornell (1983) and presented in detail by Cornell, Flora and Roi (1983).

These examples are typical of observational studies in epidemiology and medicine in that morbidity or mortality rates are compared between groups formed by exposure categories or type of treatment. Comparisons are made within categories defined by covariates or after adjustment for covariates. Common types of covariates are demographic variables, such as age and sex, and initial severity measures.

Rosenbaum gives guidance with respect to the design and analysis of such studies. He says that it is desirable to select two control groups in such a way that a possibly relevant, but unobserved, covariate has different distributions in the two control groups, and then to check to see if the responses in the two control groups are similar. If they are, then the unobserved

Richard G. Cornell is Professor of Biostastistics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109.