Again, the SUTVA is implicit in Rosenbaum's analysis and its violation might make the application of his results ineffective. For his medical examples, the SUTVA is not satisfied if the model only specifies a single control response, R_C rather than an R_{C_i} for the ith control group and, in fact, subjects in the first control group are exposed to a different kind of treatment than are those in the second control group. On the other hand, Rosenbaum shows that when such an assumption holds there is a clear benefit for the design and analysis of observational studies.

Because Rosenbaum focuses on the problem of assessing biases due to pretreatment differences, it is a little unfair to ask him to solve these other problems as well, but from the excellence of the present paper I am sure he is up to the task.

ADDITIONAL REFERENCES

HOLLAND, P. W., JAMISON, D. and RAGOSTA, M. (1979). *Computer-assisted Instruction and Compensatory Education: The ETS/LAUSD Study. Data Analysis: Fiscal year 1978.* Educational Testing Service, Princeton, N. J.

RAGOSTA, M., HOLLAND, P. W. and JAMISON, D. (1982). *Computer-assisted Instruction and Compensatory Education: The ETS/LAUSD Study. Final Report.* Educational Testing Service, Princeton, N. J.

RUBIN, D. B. (1986). What ifs have causal answers, discussion of "Statistics and causal inference," by P. Holland. *J. Amer. Statist. Assoc.* **81** 961–962.

Comment: The Use of Multiple Control Groups in Designed Experiments

Barry H. Margolin

I want to commend Dr. Rosenbaum on a most lucid presentation regarding the role of a second control group in an observational study. My contribution to this discussion is motivated by a remark of Cochran (1965), that in discussions of topics in observational studies, "it is relevant to indicate how the problem is tackled in controlled experimentation . . . ." With this in mind, I will discuss briefly and illustrate the roles that multiple control groups have played in designed experiments; these illustrations draw heavily upon my own research simply because they are readily accessible to me. Three distinct roles are discernible: (i) to detect the presence of unsuspected systematic effects; (ii) to determine whether there are hidden sources of extraneous random variability; and (iii) to assemble sufficient control data to permit a meaningful assessment of sampling model assumptions.

The earliest experiments whose design included features resembling multiple control groups appear to be uniformity trials in agricultural research (Cochran, 1937). These are agricultural experiments in which the land is divided into a number of plots of the same size. A single variety of the crop of interest is planted, although other factors such as fertilizer are kept constant from plot to plot, and the yield of each plot is observed. As Cochran (1937) noted, the primary purpose of a uniformity trial is to study the effects of amalgamation of the original plots into "larger plots of various sizes and shapes" and "to provide information on the optimum size and shape of plot" with regard to experimental error. As such, a uniformity trial is viewable as an experiment with a control group but no treatment group.

The analogy with multiple control groups becomes clearer, however, when one observes that uniformity trials are also conducted to validate the applicability of tests of significance that are based on analysis of variance (ANOVA). As Cochran (1937) writes,

> "A preliminary requirement for the application of the analysis of variance to be possible is that the experimental design used should be chosen at random from a set of designs such that, in the absence of any treatment effect, the average treatment mean square over the set should equal the average error mean square. . . . The further question arises: how good an approximation to the tabulated z distribution is generated by the process of randomization used? There again the question may be tested from uniformity trial data."

When any particular design is imposed on the uniformity trial data, which involve no true treatments, the results hopefully appear as if they derived from a

Barry H. Margolin is Head, Statistical Methodology Section, National Institute of Environmental Health Sciences, P. O. Box 12233, Research Triangle Park, North Carolina 27709.