covariate can be ignored without concern for bias. If the responses in the two control groups are not similar, then care must be taken since there is evidence that the effect of exposure or treatment depends on variables not controlled for in the design and not measured for entry into the analysis.

How could this be done in the example of the study of stainless steel workers? Many times such studies contain comparisons with other workers in the same plant or with local, state or national population rates after adjusting for age. In comparisons within the same plant, cancer rates may be affected by substances in the environment other than nickel and chromium. The principle of control by systematic variation presented by Rosenbaum would lead to the formation of more than one control group distinguishable by different types and levels of exposure within the workplace environment, but not including nickel or chromium. For instance, one group could consist of office personnel. Another could consist of workers exposed to heat and dust in a plant that produces steel without nickel or chromium as ingredients.

In the example concerning the evaluation of burn care, the potential for survival may not be fully depicted by objective measures of the severity of the burn and by information on age and comorbidity that reflect the health and resilience of patients. Therefore it would be appropriate to include hospitals in the control, as well as treatment group, which treat a variety of types of patients. Possibilities include an inner city hospital that treats many indigent patients, a military hospital and a research hospital.

It may seen obvious that multiple control groups would be informative and desirable in these examples. Yet many observational studies use only one control group. Rosenbaum's review focuses attention on the benefit to be gained from multiple control groups and develops useful principles for their design and interpretation.

ADDITIONAL REFERENCES

CORNELL, R. G. (1984). Mortality patterns among stainless steel workers. In *Nickel in the Human Environment* (E. W. Sunderman, Jr., ed.) 65–71. International Agency for Research on Cancer, Oxford Univ. Press, N. Y.

CORNELL, R. G., FLORA, J. D. and ROI, L. D. (1983). The statistical evaluation of burn care. *Undergraduate Mathematics Applications Project J.* **4** 85–115.

WOLFE, R. A., ROI, L. D., FLORA, J. D., FELLER, I. and CORNELL, R. G. (1983). Mortality differences and the speed of wound closure among specialized burn care facilities. *J. Amer. Med. Assoc.* **250** 763–766.

Comment

Norman Breslow

The major conclusions of this paper should come as no surprise to biostatisticians and epidemiologists involved in the applications of statistical methods and concepts to clinical and observational studies in public health and medicine. For the most part, the arguments lend mathematical precision to principles of study design and interpretation that are well-known and widely used. Few epidemiologists involved with cohort studies, for example, would take issue with the statement that "... if, after adjustment for (the covariables), the two control groups differ with respect to the response ... then ... at least one of the control groups is not comparable to the treated group." The principles of "control by systematic variation" and "bracketing" seem intuitively clear and are evidently well-established in the social science literature. Although it is reassuring to note that these principles are confirmed by the randomization paradigm developed by Rubin and Rosenbaum, it would perhaps be of even greater interest to see examples where the paradigm led to novel insights and new methods of study design and analysis. Recent work of Robins (1986) on causal inference in occupational cohort studies represents a start in this direction.

Turning specifically to the field of case-control studies, the major conclusion is that inference regarding the validity of a covariable-adjusted relative risk estimated for a particular exposure is strengthened if one can demonstrate equality in the covariable-adjusted odds ratios that contrast the exposures in two or more control groups. This confirms the insights of Hill (1971): "If a whole series of control groups, e.g., of patients with different diseases, gives much the same answer and only the one affected group differs, the evidence is clearly much stronger than if the affected group differs from merely one other group." In their pioneering work on lung cancer and smoking, Hill and Doll in fact utilized two control groups. One consisted of patients with diagnoses other than cancer who were

---

Norman Breslow is Professor and Chairman, Department of Biostatistics, University of Washington, Seattle, Washington 98195.

matched to the cases for age, sex, hospital and date of diagnosis. The other consisted of patients with cancers of the stomach, colon and rectum who were not matched for age and sex but who were notified to the investigators in the same way as the cases so that they could be interviewed without prior knowledge of the diagnosis. Doll (1959) remarks that this second control group served to test for the specificity of the relationship between lung cancer and smoking and for the presence of what today are known as selection bias and information bias. French workers in the field (Schwartz and Denoix, 1957) utilized four separate control groups for their age- and hospital-matched lung cancer cases, namely patients with (i) cancer elsewhere in the respiratory or digestive tracts, (ii) other cancers, (iii) noncancerous diseases, and (iv) trauma from accident or injury. They found roughly equivalent smoking habits among the last three control groups and substantially more smoking among the cases. The first control group occupied an intermediate position. These results enhanced the validity of the case-control comparison and raised the question of a possible link between smoking and certain other cancer sites. We now know that cancers of the oral cavity and upper respiratory tract are related to both tobacco and alcohol consumption.

Multiple case or control groups are not uncommon in epidemiological research. One design for the study of occupational cancer compares the histories of exposure to a number of chemical agents among patients with a variety of cancer sites (Siemiatycki, Richardson, Gérin, Goldberg, Dewar, Désy, Campbell and Wacholder, 1986). A major inferential problem is posed by the large multiplicity of odds ratios associating each pair of cancer sites with exposure to each agent. Strategies for causal inference based on an extension of Rosenbaum's ideas could well prove useful in deciding which of the observed associations should be followed up in future research. Another common design is to use one control group consisting of a random sample of the population from which the cases arose (population controls), usually matched or stratified by age and sex, and another consisting of patients hospitalized in the same institutions as the cases but with different diagnoses (hospital controls). In principle, the population controls within each age/sex stratum constitute an unbiased sample vis-à-vis the exposures and thus formally satisfy the conditions needed for "X-adjustability." However, because the control interviews are conducted at home rather than in hospital by someone who knows that the interviewee is not a case, there may be concern about information bias. Interviews with hospital controls, especially if conducted with potential cases before the diagnosis is established for certain, would not be subject to this problem. However, the hospital controls may lack the sampling unbiasedness of the population controls due to the fact that the catchment area of each hospital changes with diagnosis or because a particular control diagnosis is itself related to the exposure in question. Epidemiologists usually are elated if they can demonstrate similar relative risks using these two control groups. However, Rosenbaum's remarks regarding the hazards of concluding too much from the demonstration of equivalence between partially comparable control groups in cohort studies would appear to apply here also.

Several research teams have shown explicitly how data from case-control studies with several case and/or control groups may be analyzed using the logistic model for polytomous outcomes (Cox, 1970, Section 7.5). Their results provide a useful extension of the Mantel-Haenszel techniques used by Rosenbaum, especially when the exposure and covariables are multiple and continuous. Thomas, Goldberg, Dewar and Siemiatycki (1986) and Dubin and Pasternack (1986) consider simple random or stratified samples for which the unconditional logistic analysis suffices. Liang and Stewart (1987) develop the corresponding conditional analysis of matched data. These methods are well-suited for testing the equality of covariable-adjusted relative risks among various subsets of the control and/or case groups and thus for formalizing some of the proposed tests for X-adjustability.

ADDITIONAL REFERENCES

Cox, D. R. (1970). *The Analysis of Binary Data.* Chapman and Hall, London.

Doll, R. (1959). Retrospective and prospective studies. In *Medical Surveys and Clinical Trials* (L. J. Witts, ed.) 71–98. Oxford Univ. Press, London.

Dubin, N. and Pasternack, B. S. (1986). Risk assessment for case-control subgroups by polychotomous logistic regression. *Amer. J. Epidemiol.* **123** 1101–1117.

Hill, A. B. (1971). *Principles of Medical Statistics*, 9th ed. Oxford Univ. Press, New York.

Liang, K. Y. and Stewart, W. F. (1987). Polychotomous logistic regression methods for matched case-control studies with multiple case or control groups. *Amer. J. Epidemiol.* **125** 720–730.

Robins, J. M. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *J. Math. Modelling* **7** 1392–1512.

Schwartz, D. and Denoix, P. F. (1957). L'enquête française sur l'étiologie du cancer broncho-pulmonaire. Rôle du tabac. *La Semaine des Hôpitaux de Paris* **33** 3630–3643.

Siemiatycki, J., Richardson, L., Gérin, M., Goldberg, M., Dewar, R., Désy, M., Campbell, S. and Wacholder, S. (1986). Associations between several sites of cancer and nine organic dusts: results from an hypothesis-generating case-control study in Montreal, 1979–1983. *Amer. J. Epidemiol.* **123** 235–249.

Thomas, D. C., Goldberg, M., Dewar, R. and Siemiatycki, J. (1986). Statistical methods for relating several exposure factors to several diseases in case-heterogeneity studies. *Statist. Med.* **5** 49–60.