

## **Epidemiological Studies in Nutrition: Utility and Limitations**

**GEORGE G. RHOADS AND NATALIE KURINIJ**

*Epidemiology Branch, Prevention Research Program, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892*

There is considerable evidence that dietary habits contribute to the causation of the major chronic diseases of the industrialized world. The important influence of alcohol on cirrhosis, of sucrose and fluoride on dental caries and of dietary saturated fats and cholesterol on coronary heart disease is well accepted. Moreover, there is evidence that a substantial proportion of cancers have nutritional causes (1); diets rich in fats and animal protein may increase the risk of some cancers whereas dietary vitamin A, selenium and fiber are believed to have a protective effect. Diet may also contribute to the pathogenesis of a number of other prominent diseases such as stroke, diabetes and diverticulitis.

Historically, descriptive epidemiological observations have provided important clues for the unraveling of many of the known diet-disease relationships. Through careful observation of high risk populations and description of their diets, investigators such as Lind (2) and Goldberger (3) were led to suspect nutritional causes of scurvy and pellagra. Reversal of these conditions by dietary supplementation clearly demonstrated their etiology and provided important guidance to the later biochemical investigation of ascorbic acid and niacin.

In more recent years, epidemiology has provided leads to the nutritional etiology of degenerative chronic diseases. For instance, the observation that coronary disease was uncommon in Eskimos, despite their high fat diet, led Bang and Dyerberg to postulate that  $\omega$ -3 fatty acids from fish oils in the Eskimo diet were responsible for the low plasma lipid levels and prolonged bleeding times observed in this population (4). Although the biological effects of  $\alpha$ -linolenic acid (18:3n-3) have been studied for many years, only recently has this information been used to define a mechanism of action of  $\omega$ -3 fatty acids in the prevention of heart disease. Similarly, recent statistical associations from large complex surveys between calcium intake and blood pressure and between alcohol consumption and breast cancer have stimulated research in these areas.

The descriptive epidemiological methods used to link vitamin-deficient diets to clinical syndromes and to

link other dietary characteristics to degenerative chronic disease are the same. They depend on several lines of evidence such as comparisons of diet and life-style in population groups with high and low disease rates, and associations between time trends in consumption of certain foods and trends in cause-specific mortality. Sometimes dietary hypotheses have gained favor over genetic ones because of changing disease rates in migrant populations that experience a substantial change in diet.

Of course, this descriptive evidence does not carry the force of the elegant experiments of Lind (2) and Goldberger (3). The slow development of the major chronic diseases, their multifactorial origin and their irreversible nature have precluded the simple confirmatory experiments that were used in identifying nutrient deficiency diseases. Randomized trials have been devised to test nutritional hypotheses in special circumstances, but practical considerations often demand compromises in design. Thus the Lipid Research Clinics Coronary Primary Prevention Trial (5) provided evidence that a lowering of serum cholesterol would reduce coronary risk, but for practical reasons a drug, cholestyramine, was used as the method of intervention instead of diet alone. Currently, the National Cancer Institute is conducting a number of other trials. These include feasibility studies for two large studies of the effect of a low fat diet ( $\leq 20\%$  of energy) on reduction of breast cancer in high risk women and the decrease of recurrence of breast cancer in diagnosed cases (6). Although these are focused in special groups of women with very high risk, they will, if successful, provide considerable focus for the study of diet in human mammary carcinogenesis.

It is often not feasible or not ethical to carry out direct experiments to test diet-disease hypotheses. In such instances, common practice is to mount detailed observational studies of individuals to see whether the correlations between diet and disease noted among population groups can be reproduced at the individual level. Such studies have been notably successful in demonstrating the risks of smoking, hypertension, radiation

and a host of other factors. But attempts to test diet-disease hypotheses through comparisons of individual diets *within* populations have mostly been negative or difficult to reproduce.

Conflicting results among these studies may stem from several sources, but there appear to be two overriding limitations. First, variation in diet within a homogeneous population is substantially less than the variation between populations. Thus, in comparing Japan and the United States, there are smokers and non-smokers, drinkers and nondrinkers and promiscuous and monogamous persons in both populations. But there is almost no one in the United States who consumes a Japanese diet or in Japan who consumes an American one (7). The modest variation in diet composition within a population group results in a statistical inability to differentiate individuals by nutritional factors. Second, dietary measurements have poor reliability. Those with the greatest face validity, 24-h recall and diary methods, usually cover too short a period, while those that ask for "usual" intake require the subject to make an educated guess. A useful measure of reliability is the correlation coefficient between two measures of the same risk factor conducted a few months apart. Ideally this should be  $r = 1.0$ . For a highly reproducible measure such as height, it regularly exceeds 0.9. For well-established disease risk factors such as cigarettes per day, serum cholesterol or blood pressure, repeat measures usually correlate at a level of 0.7 or better. But for most dietary methods the values range downward from 0.6, sometimes to appalling levels (8). This type of reproducibility measure is affected both by real changes in diet as well as by measurement error; both compromise epidemiological studies, which usually seek to establish an average intake. Of course, reproducibility does not guarantee validity, but a method that does not yield reproducible results cannot provide a useful measure of disease risk.

What does the future hold? As laboratory methods become more sophisticated it may be possible to use biochemical measures to support or replace dietary assessment for some additional nutrients. This could potentially increase the accuracy of assessment and make epidemiological studies of individuals more feasible. Unfortunately, practical laboratory methods for estimating fiber, total energy and proportion of energy from the various macronutrients do not appear likely in the near future. Moreover, many nutrient levels in blood or serum are physiologically controlled or interact with other nutrients and drugs so that they do not correlate well with intake. Progress can still be expected in this area, but it may be mainly limited to selected vitamins and minerals.

A second approach is to try to improve the quality of future epidemiological studies. This probably means deemphasizing the role of case-control studies in which subjects with an illness are asked what they used to eat one or more years ago. This retrospective assess

ment compounds the measurement problems and has led to conflicting results. In prospective studies, the diets of individuals are characterized before the onset of disease. Unfortunately, large numbers of subjects are required to obtain enough cases of future illness. A prospective study of 90,000 nurses in whom diet was ascertained through a well-validated, mailed food-frequency questionnaire serves as an example of the kind of undertaking possible (9).

Perhaps the prospective approach could be improved by using food diary methods that, as noted above, have greater face validity than do food-frequency methods. Some years ago Liu et al. (10) pointed out that stable estimates of nutrient intake could be obtained with this approach if records were obtained for an adequate number of different days. It might be possible to mount a prospective study of diet and chronic disease that takes advantage of this principle by obtaining food intake histories for multiple days on a large number of healthy individuals. An effort to recruit a diverse study population, perhaps by oversampling vegetarians or other special groups, might well be worthwhile. It seems clear that continuing interest in the relationship between diet and chronic disease in both the lay and professional communities will ensure a long future for nutritional epidemiology. The challenge is to make this effort as productive as possible through careful epidemiological study design and active collaboration with laboratory and clinical scientists.

#### LITERATURE CITED

1. DOLL, R. & PETO, R. (1981) Quantitative estimates of avoidable risks of cancer in the United States today. *J. Natl. Cancer Inst.* 66: 1191-1308.
2. MCCOLLUM, E. V. (1957) *A History of Nutrition. The Sequence of Ideas in Nutrition Investigations*, pp. 252-264, Houghton Mifflin, Boston.
3. TERRIS, M. (1964) *Goldberger on Pellagra*, Louisiana State Univ. Press, Baton Rouge.
4. BANG, H. O. & DYERBERG, J. (1981) Personal reflections on the incidence of ischaemic heart disease in Oslo during the Second World War. *Acta Med. Scand.* 210: 245-248.
5. LIPID RESEARCH CLINICS PROGRAM (1984) The Lipid Research Clinics coronary prevention trial results. II. The relationship of reduction in incidence of coronary heart disease to cholesterol lowering. *J. Am. Med. Assoc.* 251: 365-374.
6. GREENWALD, P., CLIFFORD, C., BUTRUM, R. R. & IVERSON, D. C. (1987) Feasibility studies of a low-fat diet to prevent or retard breast cancer. *Am. J. Clin. Nutr.* 45: 347-353.
7. TILLOTSON, J. L., KATO, H., NICHAMAN, M. Z., MILLER, O. C., GAY, M. L., JOHNSON, K. G. & RHOADS, G. G. (1973) Epidemiology of coronary heart disease and stroke in Japanese men living in Japan, Hawaii, and California: methodology for comparison of diet. *Am. J. Clin. Nutr.* 26: 177-184.
8. RHOADS, G. G. (1987) Reliability of diet measures as chronic disease risk factors. *Am. J. Clin. Nutr.* 45: 1073-1079.
9. WILLETT, W. C., STAMPFER, M. J., COLDITZ, G. A., ROSHER, B. A., HENNEKENS, C. H. & SPEIZER, F. E. (1987) Dietary fat and the risk of breast cancer. *N. Engl. J. Med.* 316: 22-28.
10. LIU, K., STAMLER, J., DYER, A., MCKEEVER, J. & MCKEEVER, P. (1978) Statistical methods to assess and minimize role of intra-individual variability in obscuring relationships between dietary lipids and serum cholesterol. *J. Chronic Dis.* 31: 399-418.