This issue of the Journal features a diverse set of papers on dietary epidemiology. Although these papers might not help us decide which party buffet items to select in the coming days of the holiday food season, they do represent a cross section of the current status of the field of dietary epidemiology. This field includes a rapidly growing number of studies using variations of the food frequency method of dietary assessment and, at the same time, studies that continue to critically examine the validity of that very same method. The food frequency validation studies in this issue (1–3) and their accompanying commentaries (4, 5) are noteworthy contributions to the process of building a consensus about the proper use of the food frequency method of dietary estimation.

There is perhaps no other epidemiologic discipline that has attracted as much public attention and, at the same time, as much scientific criticism as has dietary epidemiology. That is because the exposure is both of immediate interest to the public and notoriously difficult to measure. A small preliminary study suggesting an effect of the frequency of broccoli consumption on breast cancer survival might make headlines when even more definitive findings from studies on other important topics would not. Publicity surrounding variations in findings from studies on other important topics would not. Publicity surrounding variations in findings from studies of dietary epidemiologic studies has led to both scientific and public cynicism about the field. Some excess publicity can be blamed on epidemiologists, but most has been a consequence of the high public appetite for information about diet and health. As we have learned the sometimes difficult lessons that popular hypotheses about diet and disease may well be null, the validity of dietary measurements has been openly questioned.

Experimental nutritional scientists, who have been trained to consider nutrients as precisely quantitated factors, often scoff at dietary epidemiologic methods. When nutritional scientists ponder epidemiologic findings relating long-term diet to disease based on dietary estimates from such questions as “How often do you eat peas?” their skepticism is understandable. The disagreement between experimental nutritional scientists and nutritional epidemiologists appears to result from not only their different standards of precision for exposures but also their use of different endpoints. Although experimental nutrition can precisely measure nutrient exposures, their outcomes, such as disease in laboratory animals or short-term biomarkers in human volunteers, are often only tangentially related to the human diseases of interest. In contrast, epidemiologic dietary studies relate the relevant disease endpoints to the relevant dietary exposures, but the measures of diet are quite crude according to the standards of experimental nutrition science.

Epidemiologists have also been skeptical about the ability of food-frequency self-reports to support etiologic inference. There has been concern about not only random measurement errors in food frequency reports but also biases in reporting diet related to disease or disease-linked factors. This concern comes in part from the small size of associations observed in dietary epidemiology, often with relative risks of the order of 1.5–1.9. The challenge in dietary epidemiology is that, because of imperfect dietary assessment, it is difficult to know whether a relative risk of 1.5 should be interpreted as an indicator of a much larger relative risk that has been diluted by random measurement error, or whether it is an overestimation of an even smaller relative risk that has been inflated by bias or confounding.

Do we yet have enough understanding of the validity of the types of food frequency dietary assessments used in epidemiologic studies to begin to put these controversies behind us? How good is good enough, and how bad is too bad for dietary measures? Many previous comparisons of alternative methods of estimating diet have shown essentially the same findings as are shown by the current study reported by Dr. Subar et al.: that food frequency measures fill the glass of validity about half full. Studies comparing
food frequency measures with repeated dietary recalls generally show, as was observed by Subar et al., correlations of the order of 0.4–0.7 (1). The study by Subar et al. is in many ways the strongest such study conducted to date. It is large, it includes both genders, it is based on a representative sample of the general population, and it compares reference diet measures with three different food frequency questionnaires, two in widespread use in various formats and a newly developed questionnaire. In their accompanying commentaries, both Dr. Willett (4) and Dr. Block (5) conclude that food frequency methods are sufficiently valid for etiologic studies. I agree. That biomarkers of nutrient intake, such as circulating triglycerides, vitamin C, or carotenoids, also correlate to a similar degree with food frequency estimates of both macronutrients (2) and micronutrients (3) further strengthens the conclusion that food frequency reports, though seemingly crude measures, do indeed substantially reflect the truth of nutrient intake. It is important to note that, despite its greater length and theoretical cognitive advantages, the new questionnaire tested by Subar et al. yielded energy-adjusted correlations with the reference diet measures that were quite similar to the other older and shorter questionnaires. The observation by Dr. Willett is noteworthy that there may be a “ceiling of validity” for food frequency questionnaires somewhere near correlations of 0.7 (4). Further refinements and improvements in dietary questionnaires may not substantially improve our current capabilities using existing questionnaires.

If we can now accept that food frequency measures are sufficiently valid to support etiologic studies, as I think we can, it is important to also now further consider the validity of what is done with those measures in the phase of data analysis. The question of the interpretation of caloric intake measures is answered differently by Drs. Willett and Block, who present conflicting views of the principal purpose of dietary measures (4, 5). Whether we regard food frequency reports more as representations of dietary composition (4) or as estimates of absolute intake (5) has important implications for how we analyze dietary data. This difference in biologic inference of analyses conducted with or without caloric adjustment is but one of many questions about how to analyze complex dietary data. Many hypotheses can be tested in many ways. Because any given food contains many different nutrients, any given nutrient comes from many foods, and foods can be considered as being in many types of groups, we can choose analyses that are simple or with multivariate adjustments for other foods, and we can now also choose to describe food frequency reports within constellations of complex dietary patterns (6). Even if we accept that food frequency measures are sufficiently valid for supporting causal inference, as I do, we must still be aware that threats to validity do not end with the collection of the food frequency reports.

REFERENCES