In this issue, Bingham et al. present analyses comparing nutrient intakes estimated by a 1-week diet record (DR) or food frequency questionnaire (FFQ) to a series of biomarkers and incidence of coronary heart disease (CHD). Such triangulation analyses can provide valuable evidence on the relative validity of different dietary assessment methods and are thus of considerable interest. Unfortunately, the results provided by Bingham et al. are difficult to interpret due to the design of the study and their choice of statistical models.

In the design of a study in which multiple dietary assessment methods are compared with biomarkers, the temporal relationships are critical. Underlying this is the concept that, for epidemiologic applications, the dietary methods are meant to represent intake over an extended period, usually many months or years, not just the few days or weeks around the time the data are collected. Thus, if the biomarker is sensitive to recent intake (true for most urinary and plasma measurements, but less true for biomarkers using adipose or nail tissue) and is collected close in time to a measure of short-term intake, such as a DR or several 24-h recalls, the correlations will tend to be exaggerated. This is a form of correlated error that can result from fluctuations in diet due to seasonality or random variation that will be reflected in both measurements (see Chapter 6). These issues are different for a FFQ, which typically asks about diet over the past year, and should therefore reflect dietary intake over a much longer period, not just recent diet. In the cross-sectional analyses by Bingham et al., the FFQ, DR and biomarkers were all collected in close temporal proximity, which would tend to bias the correlations in favour of the DR. Their argument in the discussion that the FFQ was collected within 30 days of the biomarkers misses the critical point. The solution to this bias is to collect the biomarker at some time remote from the collection of the dietary intakes assessments, either before or later. Fortunately, Bingham et al. have done this. In an earlier, although smaller, validation study, plasma vitamin C...
levels were measured both at the time of FFQ and DR collection and 1 year earlier (Michels and Day, presented at the 2001 meeting of the Society for Epidemiologic Research). When appropriately adjusted for energy intake, the correlation between vitamin C intake and plasma vitamin C was stronger for DR than for FFQ when the biomarker was measured at the time of dietary assessment, but when the biomarker was measured 1 year earlier the association was virtually identical for DR and FFQ. Also, in the present article, the associations with plasma vitamin C measured 3–4 years after dietary assessment were similar for FFQ and DR (the authors note that the associations were weakened by increases in vitamin C intake, but this should affect the validity of both dietary methods similarly). These findings suggest that long-term intake of energy-adjusted vitamin C is measured similarly well by a 1-week DR and a FFQ.

The choice of statistical models also clouds the interpretation of findings presented by Bingham et al. When assessing intake of a macronutrient in an energy-adjusted model, as in an isocaloric-controlled feeding study, decisions need to be made about the other macronutrient(s) for which this will be substituted. In studies of dietary fat, it is typical to compare total or a specific type of fat with the same number of calories for carbohydrate or a mix of the other macronutrients in the diet. However, in the models for total fat used by Bingham et al., carbohydrate was included as a covariate with total energy and alcohol, meaning that calories from fat were compared with the same number of calories from protein, i.e. the model evaluated the effect of calories from fat substituting for protein. This is an unusual comparison and may explain why they failed to see the expected positive association between total fat and HDL cholesterol and the inverse association between total fat and triglyceride levels that has been seen consistently in controlled feeding studies and in an earlier validation study of a FFQ that we conducted. The results for fat intake compared with intake of carbohydrate or other calories would be of considerable interest because of the strong interest in the association between dietary fat and breast cancer. Also, Bingham et al. have argued that DRs are superior to FFQs based on a stronger association between dietary fat and a small number of incident cases of breast cancer when using a DR than when using an FFQ (although in that analysis the FFQ data were adjusted for energy intake but the DR data were not). It is unfortunate that they did not provide these results despite repeated requests in the review process. Because the results for fat are usually reciprocal with those for carbohydrate, the fact that they found similar associations for the FFQ and DR when comparing carbohydrate intake and HDL cholesterol suggests that fat was measured similarly well by the two methods, but it would have been better to evaluate this directly. Similar problems affect their other models in both the cross-sectional and longitudinal analyses of CHD. For example, saturated fat is compared with other types of fat plus protein, and in their analyses of isoflavones, the data are not adjusted for standard demographic and risk factors or energy intake.

In their analyses predicting CHD, intakes of vitamin C and of fruits and vegetables assessed by DR appear to be more strongly predictive than by FFQ, but this is only one component of diet, and the data that would be more valuable if other dietary factors previously associated with CHD, such as trans- and specific-polyunsaturated fatty acids and cereal fibre, were also included in these comparisons. Also, the results for vitamin C predicting CHD, interpreted to be a marker of fruit and vegetable intake, appear to be inconsistent with the similar prediction of plasma vitamin C levels.

Despite the biases and obscure models, the findings do not show a consistent advantage for one method. For carbohydrate and dietary fibre intakes, the coefficients for DR and FFQ are similar in both the cross-sectional and longitudinal analyses, and for sodium, potassium and vitamin C the coefficients for DR are only modestly stronger. As the authors note, they previously showed that correlations for N-3 fatty acids were similar for DR and FFQ. The authors dismiss earlier literature comparing different dietary methods with biomarkers on the basis of having fewer than 500 subjects. However, these studies were of sufficient size to be informative, and they overall suggest similar associations comparing FFQs with DRs or multiple 24-h recall; for some nutrients one method appears a bit better and for other nutrients the converse was seen. Although having large numbers is always nice, 12,000 subjects is overkill for these cross-sectional comparisons and other investigators should not be deterred from conducting such studies with smaller sample sizes. Most importantly, a large sample size cannot overcome a flawed study design.

Studies of relative validity using biomarkers can be useful to improve our understanding of different dietary assessment methods, and more studies are desirable because the relative validity can vary by dietary factor and among populations. At this time, the overall evidence suggests that one method is not consistently better than the other, and decisions whether to use diet DRs, FFQs or multiple 24-h recalls will be influenced by multiple considerations. One consideration is that week-to-week variation for foods, which are also of interest in most dietary studies, tends to be greater than for nutrients, with the result that the relative validity for DRs is likely to be lower. Cost and subject burden also differ greatly; in a pilot study in the Nurses’ Health Study, only about 40% of participants who had completed an FFQ were willing to complete a 1-week DR, and in a small pilot from the AARP cohort conducted 8–9 years...
after baseline ~14% of participants provided a DR (A Schatzkin, personal communication). Also, the lower cost and burden of the FFQ means that it can be repeated many times; for example, now eight times in the Nurses' Health Study. This is important in a long-term study, in part because using the average of multiple measurements can reduce random error, and the repeated measures can account for real changes in the food supply and dietary choices. For example, we have found that correlations between the average of three FFQs and multiple DRs over a 6-year period can be as high as 0.8–0.9.11 Because of the possibility that some associations might be better detected with one dietary method than another, it is useful that not all studies use the same method. Still, given the available evidence and practical considerations, most epidemiologists will continue to use carefully designed, appropriately adapted and calibrated FFQs to measure dietary intake until superior methods are documented. As alternative dietary assessment methods are developed, triangulation comparisons with existing methods and biomarkers can be valuable. However, appropriate temporal relationships among these measures are essential for valid comparisons.

Conflict of interest: None declared.

References