

Collection of food intake data: a reappraisal of criteria for judging the methods

BY RENATO BORRELLI

Institute of Internal Medicine and Metabolic Disease, 2nd Medical School, Via Sergio Pansini 5, Naples 80131, Italy

(Received 22 May 1989 – Accepted 15 November 1989)

The relationship between diet and the development of chronic disease still remains a controversial area. One major difficulty is to obtain a valid estimate of habitual pattern and level of food consumption for each individual. There is, in fact, a voluminous and largely negative literature on the validity of dietary assessment methods. In the present paper the utility of the most frequently used dietary assessment method in epidemiological studies is discussed in terms of precision and accuracy.

Food intake: Epidemiology: Statistics

The relationship between diet and the development of chronic disease still remains a controversial area. In studying this relationship, the estimation of the habitual pattern of food consumption, the so called 'usual intake' of free-living individuals, is a crucial factor (Marr & Heady, 1986). One major difficulty, in fact, is obtaining a valid estimate of 'usual intake' of each individual. No current method has so far been able to yield precise and accurate quantitative amounts of food eaten and there is a voluminous and largely negative literature on the validity of dietary assessment methods (Block, 1982; Bingham, 1987).

However, it is generally thought that, for epidemiological purposes, the goal of dietary assessment methods is to rank individuals on the basis of their intake into broad categories, along the distribution of intake from very little to very much, rather than to obtain precise and accurate quantitative amounts of food eaten (Block, 1982). On the other hand, the conflicting results reported by different studies on this topic are largely attributed to the limitations of the methods used to collect information on food intake (Liu *et al.* 1978).

In order to have a better understanding of the judgement of dietary assessment methods for epidemiological studies, it appears useful to reappraise the criteria for judging a method.

Random error

It is well known that measurements of biological variables obtained from the same subject on different occasions, even under carefully standardized conditions, will not agree exactly. There is, in fact, a source of error denoted as random error (ϵ) which affects the measurements of the variables. So the measured value X of the i th-patient will be

$$[X]_i = [T]_i + \epsilon_i, \quad (1)$$

where $[T]_i$ is the true value of the i th-patient. As ϵ is a random type of error, the average value of repeated measurements of $[X]_i$ will tend to $[T]_i$.

Systematic error

The measurement of variables can also be affected by another source of error denoted as bias. In this case the observed values obtained by replicate measurements by the method under study on the same subject tend to a mean value which differs from the true value by

a difference called bias. This denotes the tendency of the method significantly to underestimate or overestimate true values. It is due to the imperfect specificity of the method (i.e. the sensitivity to the levels of substances (Z_i) other than that of interest).

The observed value of the i th-patient obtained by a method subject to bias and random error will be

$$[X]_i = A + B[T]_i + \epsilon_i + f([Z]_{i1} \dots [Z]_{in}), \quad (2)$$

where A , B and $f([Z]_{i1} \dots [Z]_{in})$ are the components of the bias. A and B are the constant and proportional components respectively. The non-specificity function $f(\cdot)$ describes the effect of the n interfering substances on the apparent concentration $[X]_i$ in the i th-patient. Usually it is quite complicated. However, Lawton *et al.* (1979), assuming that $f(\cdot)$ is essentially linear over a reasonable range of concentration about the mean levels, showed that model 2 can be reduced to

$$[X]_i = a^* + B[T]_i + \epsilon_i + e_i, \quad (3)$$

where $[X]_i$ is, in the case of dietary assessment method, the measured food intake of the i th-patient, and $[T]_i$ denotes the true intake. B is the proportional bias and a^* is the constant bias conditioned by the mean value (μ) of the n interfering substances

$$a^* = A + f(\mu[Z]_{j1} - \mu[Z]_{jn}).$$

a^* and B then denote the tendency of the method to underestimate or overestimate the true values.

ϵ_i is the random error and can be assumed to have a normal distribution with mean value equal to 0 and variance σ_ϵ^2 .

The error term (e_i) is

$$e_i = \sum_{j=1}^n y_j([Z]_{j1} - \mu[Z]_{j1}),$$

where y_j describes the function of the Z_j interfering factors. The error term (e_i) is the bias in excess of $a^* + B[T]_i$, due to the particular interfering factors that differ from the mean levels in the i th-patient. It can be assumed to be a variable with mean value equal to 0 and variance

$$\sigma_e^2 = \sum_{j=1}^n y_j^2 \sigma^2([Z]_{j1}),$$

where $\sigma^2([Z]_{j1})$ is the variance of the interfering factors.

Criteria for judging the methods

When a method is subject to random and systematic errors, it is appropriate to evaluate two characteristics, the precision and the accuracy (Westgard *et al.* 1974). The precision estimates to what extent a method gives similar results on different occasions. The accuracy determines the tendency significantly to overestimate or underestimate true values and identifies the confounding effects of other factors.

Precision

The precision can be evaluated by taking replicate measurements by means of components of variance analysis (Gardner & Heady, 1973). The precision, in fact, depends on the random error variance (usually called within-subject variance) (σ_w^2):between-subject variance (σ_b^2) ratio. The between-subject variation represents the differences between individuals in the true values, i.e. the variance of T .

As the estimation of the confidence bands of the components of variance is sensitive to departures from normality, and the within-subject variation may depend on the level of

intake, the nutrient intake must be appropriately transformed in order to achieve normal distribution and to stabilize the variance (Armitage & Berry, 1987). The Box-Cox method (Box & Cox, 1964) seems to be suitable for finding the appropriate transformation. This method is based on the family of transformations

$$\begin{aligned} (X + \delta)^\lambda & \text{ when } \lambda \neq 0 \\ \ln(X + \delta) & \text{ when } \lambda = 0, \end{aligned}$$

where X is the nutrient intake, involving the unknown parameters λ and δ . This extended version, involving the addition of a value δ before the power transformation (λ), appears preferable because some transformations (such as the logarithmic one) can only be applied to non-zero nutrient intake, while it is possible to have a null intake for some nutrients on some days.

Accuracy

The accuracy depends on the magnitude of the bias (a^* , B , σ_e^2). This can be evaluated by comparing the method under study with a reference method (Lawton *et al.* 1979). Unfortunately, up to now, no reference dietary assessment method on free-living individuals has been devised.

Assessing usual intake

The most frequently used methods in epidemiological research are diet record, diet history and 24 h recall. The diet record method requires weighing and recording the amounts of food eaten during a varying period of time (usually 7 d or less). Several modifications have been applied to this method (duplicate portion, estimated weights, etc). The diet history method, originally attributed to Burke (1957) but modified by subsequent investigators, consists of an extensive interview designed to elicit the usual diet. This method is time-consuming, and requires a trained interviewer and the cooperation of the interviewee. The 24 h recall method consists of an interview about the food eaten during the day preceding the interview.

The results of comparisons between one method and another have never been consistent within the same individual (Bingham, 1987). In particular, while diet history seems to overestimate food consumption, 24 h recall seems to underestimate it when compared with records of food intake. Bingham (1987) reported that, of a total of thirteen studies, eleven have shown that diet history gave estimates of food consumption significantly different (higher in seven, lower in four) from diet record, while of a total of twenty-four studies, seventeen have shown that 24 h recall gave estimates of food intake significantly different (lower in thirteen, higher in four) from diet record. The large standard error of differences obtained in these studies also showed that the ranking of individuals on the basis of their intake was not maintained by the different methods. Naturally, it is unknown whether the diet history method exaggerates food consumption or the diet record method underestimates it.

Bingham (1987) also described the potential sources of error in the methods most frequently used to assess dietary intake of individuals. Each method is associated with different sources of error. These errors may be random (ϵ), systematic (a^* , B) or due to interfering factors (e), as described in model 3. These types of error affect statistical analyses in different ways.

Impact of the random error on statistical analysis

If systematic (a^* , B) and random bias (e) are negligible in the dietary assessment method being used, model 3 can be reduced to

$$[X]_i = [T]_i + \epsilon_i.$$

In investigating the relationship between dietary intake and disease, it has been shown (Gardner & Heady, 1973; Liu *et al.* 1978; Beaton *et al.* 1983) that large measurement errors (σ_e^2) relative to between-subject variation (σ_T^2) will artificially attenuate the observed correlation coefficient. In fact, if Y is the biochemical variable of interest with variance σ_y^2 , and X is the dietary intake of some nutrient measured with error σ_e^2 (the within-subject variance), the variance of X will be σ_T^2 (the between-subject variance) plus σ_e^2 . Thus the observed correlation coefficient will be

$$r_o = \frac{\text{covariance}(X, Y)}{((\sigma_T^2 + \sigma_e^2) \times \sigma_y^2)^{0.5}},$$

while the true correlation p is

$$p = \frac{\text{covariance}(T, Y)}{(\sigma_T^2 \times \sigma_y^2)^{0.5}}.$$

As an example, assume (just for this example!) that the dietary protein (T) (mean 62.6 g/d, variance 342) and serum prealbumin (Y) (mean 260.6 mg/l, variance 36.2) are measured without any error. The correlation coefficient between these two variables is

$$r_{T,Y} = \frac{24.3}{(342 \times 36.2)^{0.5}} = 0.218.$$

This represents the true correlation between dietary protein and serum prealbumin in this population. Assuming that the dietary method was, instead, subject to a random error (ϵ) normally distributed with mean 0 and variance 407 (this was generated by a pseudo-random number algorithm generator by SPSS computer program (Nie *et al.* 1975)), the following correlation coefficient will be obtained:

$$r_{X,Y} = \frac{26.5}{(757 \times 36.2)^{0.5}} = 0.16.$$

In practice statistical theory allows us to derive an estimate of the true correlation coefficient if the precision of the method is known, i.e. if σ_e^2 and σ_T^2 are known (Rosner & Willett, 1988; Borrelli *et al.* 1989). In fact, assuming that measurement errors in X are independent of Y , we can obtain an unattenuated correlation coefficient by multiplying the observed correlation (r)

$$r = \frac{\text{covariance}(T, Y)}{((\sigma_T^2 + \sigma_e^2) \times \sigma_y^2)^{0.5}} \quad \text{by} \quad \left(1 + \frac{\sigma_e^2}{\sigma_T^2}\right)^{0.5}.$$

In the previous example, the true correlation can be obtained by multiplying $r_{X,Y}$ by 1.4.

The variances σ_e^2 and σ_T^2 can be estimated by taking replicate measurements by means of the components of variance analysis after adequate transformation (Gardner & Heady, 1973; Armitage & Berry, 1987).

The ratio of these variances can also be used in other statistical analyses, such as multiple regression analysis and the proportional hazards model in epidemiological studies (Liu, 1988, 1989).

Impact of the bias on statistical analysis

If the bias (a^* , B , σ_e^2) is, instead, relatively large, there will be a biased estimation of σ_T^2 with unpredictable consequences on the correlation coefficient.

In the previous example, the correlation coefficient between X and Y would have been

0.189 (and σ_T^2 1188), if a proportional bias B of 1.5 had been present in measuring X in addition to the random error, or it would have been 0.107 if the proportional bias B had been 0.5 (and σ_T^2 497).

It appears, then, that the correct ranking of individuals into broad categories on the basis of their intake as a criterion of acceptability of a dietary assessment method for epidemiological studies seems rather inadequate. More important becomes the evaluation of the random error (σ_e) and the bias (a^* , B , σ_e^2) in dietary assessment methods. The magnitude of the bias will influence the acceptability of the dietary assessment method, leading to a biased estimation of the between-subject variance σ_T^2 . The magnitude of the within-subject variance σ_e is not as relevant, provided it is estimated, because it is possible to correct it.

Thus, we can conclude that, for epidemiological purposes, dietary intake assessment methods have to be accurate (small values of a^* , B and σ_e^2) rather than precise (small value of σ_e^2), provided that the precision is evaluated by taking replicate measurements.

Sources of errors

Diet record. The main source of random variation (ϵ) in the diet record method is the daily variation. In fact, it has been shown that each individual is liable to vary considerably from day to day in his or her level of food consumption, especially for some nutrients, such as cholesterol or vitamin A. For this reason, many epidemiologists thought that only a poor estimate of the individual's 'usual intake' is obtained by recording food intake over a period of a few days (Gardner & Heady, 1973; Liu *et al.* 1978; Marr & Heady, 1986). However, as previously described, the problem of the large daily variation in the pattern and level of food consumption in individuals can be overcome in the diet record method. Indeed the unattenuated correlation coefficient between nutrient intake (measured with random error σ_e^2) and disease can be obtained provided that σ_e^2 be estimated by taking replicate measurements.

Another source of random variation (ϵ) is seasonal variation. The limited information currently (McHenry *et al.* 1945; Black *et al.* 1983) available to estimate the magnitude of this error seems to indicate that seasonal variation is negligible compared with daily variation (Cole & Black, 1983).

Another source of error derives from weighing and recording the foods. If there were no subgroups of the population that are particularly liable to weigh their foods incorrectly, this type of error could be considered random.

Unfortunately errors, which are likely to be systematic rather than random, do occur in the diet record method. These, in fact, seem to affect subgroups of the population in different ways. For instance, Prentice *et al.* (1986) showed that obese subjects are likely to give a distorted or wrong estimation of habitual food consumption because they often go deliberately on a diet or incompletely report their intake when being studied. Marr (1971) observed that the tedious task of weighing food could itself limit the variety of food consumed and the number of meals or snacks eaten during the observation period. So body mass index or cooperation level of the subject seem to affect the measurement of food intake and could represent confounding factors. Furthermore it is likely that people in a low socio-economic class do not keep records as well as better educated subjects (Morgan *et al.* 1987). In addition, the presence of disease could influence the ability to keep records (Hebert & Miller, 1988). Members of low socio-economic groups could overestimate intake of 'status' foods such as meat. The intake of foods such as condiments, nuts and seeds are difficult to estimate. The intake of beverages (especially beer and wine) is probably under-reported by those consuming large amounts.

It is important to point out that day-to-day variation, seasonal variation, and so on may

not be strictly random. There could in fact be differences in intake between weekend and working days, seasonal changes, etc. However, these can be overcome at the stage of sampling by mixing day of the week (or seasonal period) (Cole & Black, 1983).

Diet history. The diet history approach has been favoured because it attempts to estimate directly the 'usual intake' of each individual, eliminating at a stroke the daily variation. The diet history method is probably also subject to random errors, but the correction for a low precision cannot easily be applied to the diet history method because the estimation of σ_e^2 is more difficult. Indeed, a repeated interview, apart from being expensive, could measure a real change in habitual dietary pattern rather than random error.

The diet history method is also subject to bias. This is mainly due to errors and omissions of quantities of food eaten because the ability of individuals to estimate directly their own true mean for dietary intake is not a fixed capacity but largely depends on memory, the ability of the interviewer and other factors (Liu *et al.* 1978). If a^* , B and σ_e^2 are relatively large, the estimate of food intake will be invalidated. However, as no reference method is actually available, these cannot be estimated.

I do not subscribe to the recommendation of some authors (Bazzare & Yuhas, 1983; Karkeck, 1987) that diet history should be used for retrospective investigations of the relationship between dietary intake and incidence of disease, and diet record used when the actual food intake is to be measured. In fact, unbiased retrospective estimates of food consumption are probably unobtainable due to lack of memory, poor cooperation, etc. Furthermore, it is likely that the occurrence of disease modifies dietary intake. In any event, there is at present no way to demonstrate the validity of a retrospective dietary intake assessment.

24 h recall. The usefulness of the 24 h recall method is even more questionable if we consider that it is subject to the same bias as diet history, that one day's intake does not represent the usual intake of each individual and that it is not possible to estimate the daily variation.

Conclusion

I therefore conclude that for epidemiological purposes recorded intake over a few days seems to be the preferable technique to estimate the usual intake, but the magnitude of the possible bias should be evaluated.

Unbiased retrospective estimates of food consumption are probably unobtainable.

The lack of a reference method makes it difficult, if not impossible, to evaluate the possible bias. There is, therefore, a need to develop new approaches which are free of bias for assessing food intake. The 'biological markers' of food consumption (biochemical indices which reflect the nutrient intake) are a fertile field, but considerable efforts are needed to develop them for the majority of nutrients (James *et al.* 1981; Bingham, 1987). In particular, the lack of bias and random error in these new techniques has to be demonstrated. Once their acceptance as 'gold standard' techniques has been established, they could be used to evaluate the bias of the dietary method with a population subsample. Alternatively, they could be used to replace the conventional methods.

REFERENCES

- Armitage, P. & Berry, G. (1987). *Statistical Methods in Medical Research*. Oxford: Blackwell Scientific Publications.
- Bazzare, T. L. & Yuhas, J. A. (1983). Comparative evaluation of methods of collecting food intake data for cancer epidemiology studies. *Nutrition and Cancer* **5**, 201-214.
- Beaton, G., Milner, J., McGuire, V., Feather, T. E. & Little, J. A. (1983). Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins and minerals. *American Journal of Clinical Nutrition* **37**, 986-995.

- Bingham, S. A. (1987). The dietary assessment of individuals: methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviews* **57**, 705-742.
- Black, A. E., Ashby, D. R., Day, K. C., Bates, C. J. & Paul, A. A. (1983). Analytical versus food table values for vitamin C foods: the effect on calculated vitamin C intake of elderly subjects. *Human Nutrition: Applied Nutrition* **37A**, 9-22.
- Block, G. (1982). A review of validations of dietary assessment methods. *American Journal of Epidemiology* **115**, 492-505.
- Borrelli, R., Cole, T. J., Di Biase, G. & Contaldo, F. (1989). Some statistical considerations on dietary assessment methods. *European Journal of Clinical Nutrition* **43**, 453-463.
- Box, G. E. P. & Cox, D. R. (1964). An analysis of transformation. *Journal of the Royal Statistical Society B* **26**, 211-252.
- Burke, B. S. (1957). The diet history as a tool in research. *Journal of the American Dietetic Association* **23**, 1041-1046.
- Cole, T. J. & Black, A. E. (1983). Statistical aspects in the design of dietary surveys. In *The Dietary Assessment of Populations. MRC Environmental Epidemiology Unit Scientific Report no. 4*, pp. 5-7. Southampton: Southampton General Hospital.
- Gardner, M. & Heady, J. (1973). Some effect of within person variability in epidemiological studies. *Journal of Chronic Diseases* **26**, 781-795.
- Hebert, J. R. & Miller, D. R. (1988). Methodological considerations for investigating the diet-cancer link. *American Journal of Clinical Nutrition* **47**, 1068-1077.
- James, W. P. T., Bingham, S. A. & Cole, T. J. (1981). Epidemiological assessment of dietary intake. *Nutrition and Cancer* **2**, 203-212.
- Karkeck, J. M. (1987). Improving the use of dietary survey methodology. *Journal of the American Dietetic Association* **87**, 869-871.
- Lawton, W. H., Sylvestre, E. A. & Young-Ferraro, B. J. (1979). Statistical comparison of multiple analytic procedures, applications to clinical chemistry. *Technometrics* **21**, 397-409.
- Liu, K. (1988). Measurement error and its impact on partial correlation coefficient and multiple linear regression analyses. *American Journal of Epidemiology* **12**, 864-874.
- Liu, K. (1989). Consideration of and compensation for intraindividual variability in nutrient intake. In *Epidemiology, Nutrition and Health*, pp. 87-97 [L. Kohlmeier and E. Helsing, editors]. London: E. Smith-Gordon.
- Liu, K., Stamler, J., Dyer, A., McKeever, J. & McKeever, P. (1978). Statistical methods to assess and minimize the role of intraindividual variability in obscuring the relationship between dietary lipid and serum cholesterol. *Journal of Chronic Diseases* **31**, 399-418.
- McHenry, E. W., Ferguson, H. P. & Gurland, J. (1945). Sources of error in dietary surveys. *Canadian Journal of Public Health* **36**, 355-361.
- Marr, J. W. (1971). Individual dietary surveys, purposes and methods. *World Review of Nutrition and Dietetics* **13**, 105-164.
- Marr, J. W. & Heady, J. A. (1986). Within and between person variation in dietary surveys, number of days needed to classify individuals. *Human Nutrition: Applied Nutrition* **40A**, 347-364.
- Morgan, K. J., Johnson, S. R., Rizek, R. L., Reese, R. & Stampley, G. L. (1987). Collection of food intake data: an evaluation of methods. *Journal of the American Dietetic Association* **87**, 888-896.
- Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K. & Bent, D. H. (1975). *SPSS Statistical Package for the Social Sciences*. New York: McGraw-Hill.
- Prentice, A. M., Black, A. E., Coward, W. A., Davies, H. L., Goldberg, G. R., Murgatroyd, P. R., Ashford, J., Sawyer, M. & Whitehead, R. G. (1986). High levels of energy expenditure in obese women. *British Medical Journal* **292**, 983-987.
- Rosner, B. & Willett, W. C. (1988). Interval estimates for correlation coefficient corrected for within person variation: implications for study design and hypothesis testing. *American Journal of Epidemiology* **127**, 377-386.
- Westgard, J. O., Carey, R. N. & Wold, S. (1974). Criteria for judging precision and accuracy in method development and evaluation. *Clinical Chemistry* **20**, 825-832.