The influence of survey duration on estimates of food intakes – relevance for food-based dietary guidelines

Joyce Lambe* and John Kearney

Institute of European Food Studies, Trinity College, Dublin 2, Ireland

Development of food-based dietary guidelines relies on correct identification and interpretation of current dietary intake patterns. The process involves assessing which foods discriminate between those with desirable and undesirable intakes of particular nutrients. It is therefore important that those involved in this task are aware of any source of variability in the underlying food consumption data that cannot be attributed to true differences between individuals. A short survey duration can introduce a high degree of within-person variation and not reflect usual or more long-term food or nutrient intakes. As survey duration increases, the % consumers identified by a survey increases and the intakes among consumers only decrease while the total population intakes do not change. A short survey duration may also lead to misclassification of individuals into high and low consumers. Survey duration is therefore an issue that should be considered in the interpretation of dietary data when developing food-based dietary guidelines.

Survey duration: Food intakes: % Consumers

Development of food-based dietary guidelines involves the classification of individuals by their food and nutrient intakes to assess which foods discriminate between those with desirable and undesirable intakes of specific nutrients, such as fat and fibre, which have been linked to particular health outcomes. However, as outlined elsewhere in this supplement (Gibney, 1999), this process is not a complicated one and presents many challenges to the experts in terms of defining desirable and undesirable intakes and identifying 'discriminating' foods. While formulating scientifically sound principles for developing food-based dietary guidelines, it is extremely important that there is an awareness of the nature and magnitude of error that exists in the underlying food consumption data and how this may influence any analyses undertaken. This paper looks at survey duration, a major source of variability in food consumption surveys, which influences estimates of food intake and classification of individuals.

Survey duration and within-person variance

Survey duration exerts an effect on estimates of food intake because of the high level of within-person variation present as a result of the day-to-day variation in the composition of the self-selected diet (Beaton *et al.* 1979). In general, the within-person variation has been shown to exceed the between-person variation when nutrient intake estimates are considered (Basiotis *et al.* 1987; Beaton *et al.* 1979).

If the total variance (s_{obs}^2) of a distribution is defined as the sum of the between-person (s_b^2) variance and the within-person (s_w^2) variance divided by the number of replicate measures (k) (as described below), then the only way to diminish the impact of s_w^2 on s_{obs}^2 is by increasing the number of replicate measures (Sempos *et al.* 1991).

$$s_{obs}^2 = s_b^2 + s_w^2/k$$

Where the effects of day-to-day variation have been reduced, the data provide better estimates of the usual intakes of individuals.

Documented effects of a short survey duration include misclassification of individuals, inaccurate correlation coefficients, reduced power in testing for differences, and overestimation of prevalence of high and low intakes (Liu *et al.* 1978; Beaton, 1982; Sempos *et al.* 1991). The effects that are perhaps most relevant for the development of food-based dietary guidelines will be discussed in more detail in the following sections.

Survey duration, misclassification of individuals and relevance for formulation of food-based dietary guidelines

One of the first steps in investigating the development of effective food-based dietary guidelines is to look at prevailing patterns of food and nutrient intake. The intakes can be

---

*Corresponding author: Joyce Lambe, fax +353 1 670 9176, email iefs@iefs.ie

collapsed into quantiles (e.g. tertiles, quartiles, quintiles), and those in the lowest quantile of fat and saturated fat and those in the highest quantile of fibre and fruit and vegetables can be considered as the individuals with the most desirable intakes. Differences in the food consumption patterns between those in the highest and lowest quantiles can then be examined. To avoid drawing erroneous conclusions about the characteristics and food intakes of those classified into the upper and lower quantiles, it is important to be aware of sources of variability in the data, such as survey duration, which may lead to misclassification of individuals.

Liu *et al.* (1978) explain that if the within-person variation is large and the number of replicate measures per individual is small, the observed average value (x) may be quite different from the true mean (μ) and the probability of misclassification may therefore be quite large. Alternatively, if the variance is small and the number of replicate measures is large, then the probability of misclassification will be small. In addition to the variance and the number of replicate measures, the probability of misclassifying an individual also depends on the distance between the true mean and the cut-off point. Freudenheim *et al.* (1987) used 6844 food records collected from 106 Wisconsin women over a 3-year period to look at the extent of misclassification of nutrient intakes of individuals using 1, 2, 3 and 7 days of recording. Table 1 is taken from this paper by Freudenheim *et al.* and shows how increasing the number of records per subject improved classification into the lowest and highest quintiles, relative to usual intake as determined by the mean of the nutrient from all diet records submitted by that subject over the 3-year period (minimum of 37 records). Classification of subjects into high and low quintiles of intake was not accurate with a 1 d record (e.g. only 57% of subjects were correctly classified into the lowest quintile of calcium intake); 3 d and 7 d records improved accuracy considerably, depending on the within-person variation of the nutrient of interest. Table 2 contains data on the %energy from fat from the Dietary and Nutritional Survey of British Adults (Gregory *et al.* 1990) to illustrate this further. Only 62·4 % of subjects classified into the highest tertile of %energy from fat by 7 d records were also classified as high consumers based on 1 d intakes. Of the 37·6 % of subjects who were misclassified, 9·0 % were grossly misclassified into the lowest tertile of intake based on 1 day. Extending the survey from 1 day to 3 days showed a marked improvement, increasing the percentage of subjects similarly classified into the highest tertile from 62·4 to 74·2 %. The percentage of subjects grossly misclassified was reduced from 9·0 to 2·0 %.

Table 1. Percentage of subjects who were correctly classified into the lowest and highest quintiles of intake of selected nutrients based on 1, 2, 3 and 7 days compared to 'usual' intake based on 37–72 days of recording

(data from Wisconsin women aged 35–65 years, 1979–1982)*

	Lowest quintile				Highest quintile			
	1 day	2 day	3 day	7 day	1 day	2 day	3 day	7 day
Calcium	57	52	67	76	57	62	62	67
Kcals	57	67	76	71	57	57	67	62
Vitamin A	43	48	48	62	43	33	43	52
Vitamin C	43	48	48	48	67	71	76	76

*From Freudenheim *et al.* (1987).

Table 2. Percentage of subjects similarly classified into tertiles of %energy from fat (%en fat) by 1 day and 3 days of dietary records compared with 7 days, based on data from the Dietary and Nutritional Survey of British Adults*

(*n* = 2197)

		7 days		
		Lowest tertile (%energy fat < 36 %)	Middle tertile (%energy fat 36–41 %)	Highest tertile (%energy fat > 41 %)
1 day	Lowest tertile (%energy fat < 34 %)	63·4	27·6	9·0
	Middle tertile (%energy fat 34–42 %)	27·5	44·1	28·6
	Highest tertile (%energy fat > 42 %)	9·2	28·4	62·4
	Total (n)	732	733	732
3 days	Lowest tertile (%energy fat < 36 %)	76·5	21·4	2·0
	Middle tertile (%energy fat 36–41 %)	20·6	55·5	23·8
	Highest tertile (%energy fat > 41 %)	2·9	23·1	74·2
	Total (n)	732	733	732

*Gregory *et al.* 1990.

Table 3. Influence of survey duration on % consumers, mean consumer only intakes and total population intakes for selected foods from the Dietary and Nutritional Survey of British Adults* ($n = 2197$)

	Apples			Carbonated beverages		
	1 day	3 days	7 days	1 day	3 days	7 days
% consumers	21	35	48	19	38	54
Mean consumer only intakes (g/d)	122	66	45	339	188	125
Mean total population intakes (g/d)	26	23	22	65	71	68

*Gregory *et al.* 1990.

Obviously, misclassification has implications for the development of food-based dietary guidelines because, unless individuals are classified correctly, differences in dietary patterns between quantiles may not in fact reflect the differences between the diets of those individuals with the most and least desirable current intakes. Knowledge of the within- and between-person variance ratios may help in interpreting data in surveys in which the basic dietary methodology is a single 1 d record (Beaton *et al.* 1979). In his review of the possible use of food consumption surveys to estimate exposure to additives, however, Löwik (1996) points out that while statistical procedures are available for adjusting the within-person variance of an observed distribution, caution is called for in the usage of such procedures.

Survey duration, influence of food intake parameters and relevance for formulation of food-based dietary guidelines

The best means of identifying the specific foods or patterns of intake that currently contribute to achieving desirable intakes of specific nutrients or food groups is an issue that will need to be resolved by those developing food-based dietary guidelines (Gibney, 1999). Regardless of the method chosen, however, dietary guidelines are formulated for the whole population, and therefore it is the mean total population intake of a food that will form the basis of the assessment. Differences in mean total population intakes between those in the upper and lower quantiles of nutrient intakes may be a result of differences in the % consumers in each group or similarities in % consumers but differences between the intakes among consumers. While group means for the total population intakes are uninfluenced by survey duration once the sample size is large enough, this is not the case for the values for % consumer and intakes among consumers only. As survey duration increases, the % consumers of foods increases and, as a consequence of the increased % consumers and introduction of non-consuming days among consumers, there is a decrease in the mean intakes among consumers only. Table 3 considers selected foods (apples and carbonated beverages) from the database of the Dietary and Nutritional Survey of British Adults (Gregory *et al.* 1990) to illustrate the effect of survey duration on the various parameters of food intake. As survey duration increases from 1 to 7 days, the % consumers of apples more than doubles and the % consumers of carbonated beverages almost trebles. The increase in % consumers is paralleled by a dramatic decrease in mean consumer only intakes, with the 7 d value for both foods being only approximately one-third of the 1 d value. In contrast to the changes in % consumers and mean consumer only intakes, the mean total population intakes remain relatively stable.

These data show that short survey durations will not provide very reliable estimates of more usual or long-term food consumption patterns, particularly for rarely consumed foods. Awareness of the impact of survey duration is important when making judgements about whether strategies should involve increasing the % consumers of a food(s) or increasing the amount of that food(s) among those who are already consuming it. It may become particularly relevant if food-based dietary guidelines are to incorporate some degree of quantification.

Discussion and conclusions

A variety of dietary survey methodologies have been used by the member states of the European Union for the measurement of the food and nutrient intakes of their populations. Such data should help to form a logical basis for setting dietary targets and help to assess how best to achieve those targets. As Beaton (1994) has pointed out, dietary intake cannot be estimated without error and probably never will be. However, we can improve our use and interpretation of dietary data if we gain a greater understanding of the nature and magnitude of the errors that occur in dietary data and consider the impact of these errors in the context of the questions that we are posing. Survey duration affects estimates of % consumers, mean and high consumer only intakes of foods, and also classification of individuals as high or low consumers of foods or nutrients. All of these factors are relevant for the development of food-based dietary guidelines. Analysis of dietary data for this purpose should therefore be qualified with information on the number of days on which the estimates were made and its likely impact.

References

Basiotis PP, Welsh SO, Cronin FJ, Kelsay JL & Mertz W (1987) Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. *Journal of Nutrition* **117**, 1638–1641.

Beaton GH (1982) What do we think we are estimating? In *Proceedings of the symposium on dietary data collection, analysis and significance* pp 36–48 [VA Beal and MJ Laus, editors]. Massachusetts Agricultural Research Station, University of Massachusetts, Amherst, MA. Research Bulletin No. 675.

Beaton GH (1994) Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. *American Journal of Clinical Nutrition* **50** (suppl.), 253S–261S.

Beaton GH, Milner J, Corey P, McGuire V, Cousins M, Stewart E, de Ramos M, Hewitt D, Grambsch PV, Kassim N & Little JA (1979) Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *American Journal of Clinical Nutrition* **32**, 2456–2559.

Freudenheim JL, Johnson NE & Wardrop RL (1987) Misclassification of nutrient intake of individuals and groups using one-, two-, three-, and seven-day food records. *American Journal of Epidemiology* **126(4)**, 703–713.

Gibney MJ (1999) Development of food-based dietary guidelines: a case-study of fibre intake in Irish women. *British Journal of Nutrition* **81** (suppl.), S151–S152.

Gregory J, Foster K, Tyler H & Wiseman M (1990) *The Dietary and Nutritional Survey of British Adults*. London: HM Stationery Office.

Liu K, Stamler J, Dyer A, McKeever J & McKeever P (1978) Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *Journal of Chronic Diseases* **31**, 399–418.

Löwik MRH (1996) Possible use of food consumption surveys to estimate exposure to additives. *Food Additives and Contaminants* **13(4)**, 427–441.

Sempos C, Looker A & Johnson C (1991) The importance of within-person variability in estimating prevalence. In *Monitoring Dietary Intakes* pp 99–109 [I MacDonald, editor]. New York: Springer Verlag.