

Weak Associations in Nutritional Epidemiology: The Importance of Replication of Observations on Individuals

W A VAN STAVEREN, J BUREMA, P DEURENBERG AND M B KATAN

Van Staveren W A (Wageningen Agricultural University, Department of Human Nutrition, PO Box 8129, 6700 EV Wageningen, The Netherlands), Burema J, Deurenberg P and Katan M B. Weak associations in nutritional epidemiology: the importance of replication of observations on individuals. *International Journal of Epidemiology* 1988, 17 (Suppl): 964-969.

Day-to-day variation within an individual's dietary intake may obscure associations between diet and disease. We calculated the extent of misclassification of subjects by dietary polyunsaturated to saturated fatty acid (P/S) ratio for 1, 3, 7 and 19 survey days per person. In 59 young women diet was estimated 19 times over a period of two and a half years by a 24-hour recall method. The dietary P/S ratio was compared with the P/S ratio in adipose tissue biopsies, which was used as the reference. The proportion of subjects properly classified increased with the number of survey days. Between three and seven recalls appeared to be most cost effective. The effect of misclassification on the observed odds ratio in a case-control study was calculated for a hypothetical distribution of exposure for cases and controls.

For studies on the relationship between diet and disease, methods are necessary which can adequately assess the relative magnitude of the habitual food intake of an individual.¹ The large day-to-day variation within an individual's diet has been considered the major reason for the inability to find significant associations between an individual's intake of a food or nutrient and the disease in question.²⁻⁷ Unless repeated surveys are made of the same individuals, inter-individual variation will be overestimated and intra-individual variation will be underestimated. In general, one 24-hour recall or a one-day record is not appropriate to adequately estimate the habitual intake of an individual.

Several authors have demonstrated that correlations between diet and disease will rise if more surveys per person are performed.⁵⁻⁷ We have shown this improvement in a study in which fat tissue biopsies were taken from 59 young Dutch women after a period of two and a half years in which their food intake was estimated 19 times by a 24-hour recall method.⁸ This study was part of a larger survey on the seasonal variation of food intake. In the present paper, the extent of misclassification as a function of the number

of 24-hour recalls per subject ($k = 1, 3, 7$ and 19 days) will be evaluated by comparing the dietary polyunsaturated fatty acid to saturated fatty acid (P/S) ratio calculated from a certain number of recalls with the P/S ratio in adipose tissue. The latter is then considered to be an objective estimate of the P/S ratio of the habitual diet, and in this article serves as the reference method.

Due to misclassification, moderate associations tend to be weakened, thus resulting in underestimation of the effect. In addition, weak associations are prone to produce false negative conclusions. We calculated the effect of misclassification on the relative risk estimate in case-control studies according to the procedures of Marshall.⁹

MATERIALS AND METHODS

Methods have been described extensively earlier.⁸ For this paper the essentials will be summarized.

Subjects

Out of a study group of 123 subjects, 59 women (aged 32-35 years) were willing to undergo a fat tissue biopsy procedure. Data of two of these 59 women were discarded for the present analysis because more than one dietary recall was missing. Mean body weight of the group was stable during the observation period; the mean body mass index was 21.6 kg/m² (SD 1.7) at the end of the study.

Wageningen Agricultural University, Department of Human Nutrition,
PO Box 8129, 6700 EV Wageningen, The Netherlands.

Diets and Chemical Analysis

The interviews were conducted by trained dietitians during home visits between August 1981 and December 1983.

The first 14 interviews were carried out at monthly intervals and the last five interviews were obtained at intervals of 2–3 months. For the present analyses we have used for $k = 1$, the interview in month 11, for $k = 3$, the interviews in months 11, 12 and 13, and for $k = 7$, those in months 2, 6, 9, 11, 12, 13 and 14. Complete data for all these interviews were available for 57 subjects. For all 19 interviews combined, 12 women missed one interview.

The portion sizes of the foods most frequently used were checked by the dietitian by weighing; other foods were estimated in household measures. The food consumption data were converted into nutrient intakes using the 1981 edition of the computerized Dutch Nutrient data bank.¹⁰

Three months after the food consumption study had been completed, subcutaneous adipose tissue samples were collected during a home visit. The samples were taken from the buttock as described by Beynen and Katan.¹¹ Methyl esters of the component fatty acids were analysed by gas-liquid chromatography.¹²

STATISTICAL METHODS

Subjects were classified according to tertiles of dietary P/S intake, viz high, intermediate and low intake of polyunsaturated fatty acids relative to saturated fatty acids. This was compared with a similar classification into three categories by P/S ratio of the adipose tissue. Subjects for whom the recalls resulted in the same classification (high, intermediate, or low) as the biopsy were considered to be correctly classified by dietary intake data; others were taken as misclassified.

The probability p of misclassification into an adjacent category was assumed to be identical for all pairs of adjacent categories (Table 1). Since it was not realistic in the present case to assume that

misclassification was restricted to adjacent categories, the model by Marshall⁹ was extended so as to allow for more extreme misclassifications. Thus, the probability q of misclassification of a subject in one extreme category to the opposite extreme category was also estimated from the data.

The probabilities p and q were estimated from the observed relative frequencies (column percentages) in the off-diagonal cells of a 3x3 table. The mean of the two relative frequencies associated with classifications into the opposite extreme category was used as an estimate of q , and relative frequencies in the other four off-diagonal cells were averaged in order to obtain an estimate of p .

Pearson correlation coefficients were calculated to examine the impact of the number of survey days on the association between the P/S ratio in the diet and in the adipose tissue.

RESULTS

The composition of the diet as determined from all 24-hour recalls has already been described elsewhere.^{8,13} The dietary pattern was typical for a Western population and was similar to that found in earlier studies of young Dutch adults.¹⁴

For the present paper, the dietary P/S ratio was calculated from one, or the average of three, seven, or nineteen 24-hour recalls. As an example, Table 2 presents the cross classification according to estimated (dietary) and reference (adipose tissue) level of exposure when the dietary P/S ratio is calculated, for each subject, as the average of seven 24-hour recall measurements. In case of perfect agreement all subjects should be in the diagonal cells. The mean relative cell frequency that estimates q was equal to $(5/19 + 0/19)/2 = 0.132$ (Table 3). Similarly, 0.237 was the mean of four relative frequencies estimating p , viz 9/19, 5/19, 4/19 and 0/19. As the number of measurements increased, so did the proportions of subjects classified correctly. There was one exception, in that there happened to be one more misclassification

TABLE 1 Model for the expected relative cell frequencies (proportions). For a given reference level of exposure, p is the probability of misclassification according to estimated level of exposure into an adjacent category*, and q is the probability of misclassification into the opposite extreme category*

Estimated level of exposure: dietary P/S ratio	Reference level of exposure: adipose tissue P/S ratio		
	High	Intermediate	Low
High	$1-p-q$	p	q
Intermediate	p	$1-2p$	p
Low	q	p	$1-p-q$

* Adjacent or opposite extreme category, ie in the same column.

TABLE 2 Cell frequencies in a cross classification according to tertiles of dietary P/S ratio, calculated as an average of seven measurements, and of adipose tissue P/S ratio ($N=57$)

Estimated level of exposure: dietary P/S ratio	Reference level of exposure: adipose tissue P/S ratio		
	High	Intermediate	Low
High	10	9	0
Intermediate	4	10	5
Low	5	0	14
Total	19	19	19

TABLE 3 Estimates of probabilities of misclassification, for specified number of observations used

Number of dietary measurements averaged	Probability of misclassification	
	Adjacent category (p)	Opposite category (q)
	1	0.382
3	0.316	0.132
7	0.237	0.132
19*	0.211	0.158

* For 12 of the 57 subjects, only 18 measurements were available.

into the opposite extreme category using the mean of 19 recalls compared with the mean of 7 or 3 measurements only; this was considered to be due to chance.

Following the example of Marshall,⁹ we demonstrate the impact of error in exposure assessment on the distortion of the odds ratio in a case-control study for a hypothetical distribution of true exposure level for cases and controls (Table 4). Assuming non-differential misclassification, the probabilities of misclassification, p and q, as derived from this study in non-diseased subjects, do also apply to cases. According to the model of Table 1, the expected number of cases that will be observed in the category of high dietary P/S ratio can be calculated as $15(1-p-q) + 30(p) + 60(q)$, which is equal to 24.5 if p and q are estimated from Table 2. Similarly, $15(p) + 30(1-2p) + 60(p) = 33.6$ cases can be expected to fall in the middle category, and 47.0 in the category of low P/S ratio. On the other hand, applying the same formulae to the distribution of controls leaves the numbers per category unchanged in this particular example, eg $35(1-p-q) + 35(p) + 35(q) = 35$. Thus, the odds ratio for intermediate dietary P/S ratio turned out to be equal to $(33.6 \times 35) / (35 \times 24.5) = 1.37$, rather than 2.0. Similar calculations can be made for other values of p and q (Table 3). As can be seen in Table 5, the odds ratios are always biased towards unity.

The extent of such a distortion obviously depends upon the extent of misclassification. Our results of Table 2 were applied to the hypothetical distribution of Table 4. Marshall⁹ has shown that for various such hypothetical distributions, all associated with the same odds ratios, the resulting reduction of the observed odds ratio is similar. In Table 5, the weakening effect of exposure misclassifications is reflected by smaller estimates of the odds ratio for fewer days of observation.

The unexpectedly high proportion of misclassifications into non-adjacent categories for the average of 19 measurements caused a departure from

monotonically increasing observed odds ratios for an increasing number of measurements. Such a distortion was not found in the Pearson correlation coefficients between dietary and adipose tissue P/S ratio, which ranked from 0.25 when one measurement was used to 0.58 for the average of 19 measurements (Table 6). A correlation coefficient of 0.58 means that one third of the variation between individuals in the P/S ratio in fat tissue is accounted for by the dietary measurement.

DISCUSSION

We used both the fatty acid composition of the adipose tissue and food consumption data obtained during one or more days of observation as indicators of the P/S ratio of the habitual food consumption. Our results show the effect of a large day-to-day variation in food intake on the correlation coefficient and on the extent of misclassification when the P/S ratio in the adipose tissue is used as the reference. Increasing the number of survey days from one to three, and even from three to seven per person produced decisive improvement,

TABLE 4 Hypothetical distribution of subjects in a case-control study on dietary P/S ratio and disease

True exposure	Number of subjects		Odds ratio
	Cases	Controls	
High	15	35	1.0
Intermediate	30	35	2.0
Low	60	35	4.0

TABLE 5 Odds ratios in the hypothetical case-control study of Table 4, if the exposure (P/S ratio in diet) is estimated on the basis of 1, 3, 7 or 19 24-hour recalls per subject

Number of dietary measurements averaged	Odds ratio by estimated dietary P/S ratio		
	High P/S	Intermediate	Low P/S
1	1.0	1.23	1.39
3	1.0	1.35	1.74
7	1.0	1.37	1.92
19	1.0	1.31	1.84
Odds ratio according to true P/S ratio	1.0	2.0	4.0

TABLE 6 Observed correlation coefficients of P/S ratio in adipose tissue and P/S ratio in the diet, for specified number of observations used

Number of dietary measurements averaged	Observed correlation coefficient
1	0.25
3	0.46
7	0.54
19	0.58

but a further increase to 19 survey days hardly seemed to be cost effective. The proportion of subjects misclassified into the opposite extreme category was actually higher with 19 survey days than with 3 and 7 survey days. Although this is probably just due to chance, one may conclude that in a new sample the improvement upon going from 7 to 19 will not be impressive. Similar results have been obtained by other authors. It appears that for the assessment of the intake of energy and the macronutrients, three to seven days of recording is cost effective in many populations.^{7,15,16}

Causes of Poor Correlations Between Dietary and Fat Tissue P/S Ratio

Weight changes. Even with nineteen 24-hour recalls, about two-thirds of the variance of the P/S ratio in the adipose tissue could not be explained. In addition, more than 10% of the subjects were misclassified into the opposite extreme categories, which causes a strong downward bias of relative risk estimates in a case-control study. This may be due partly to differences in fatty acid metabolism between subjects. We have demonstrated a difference in the relation of dietary and adipose tissue P/S ratio between women who lost or gained more than 3 kg of body weight during the survey period and women with a more stable weight. In the subgroup of 32 women with a stable weight, the correlation coefficient after 19 months was not 0.58 but 0.75.⁸

Measurement errors. Unexplained variation might also be due to errors made in recalling the types and amounts of foods eaten, identification of ingredients in ready-to-eat foods, and shortcomings of the food composition table used to calculate the fatty acid content of foods.

Responder bias. In addition, a responder bias towards the mean of the population, which has been called the flat slope syndrome or the phenomenon of 'talking a good diet', may mask the association between a diet factor and disease.¹⁷

A bias away from the mean, to extreme values, may also occur. In the same study as described above we investigated the validity of reports of extremely high or low energy intake.¹³ Data obtained from those who reported a very low energy intake and those who reported a very high energy intake were inconsistent with the lack of changes in their body weight (Table 7). In five subjects who reported a very low energy intake we examined whether the inconsistency between their reported mean energy intake and the weight development was caused by a disturbed metabolic rate. Table 8 shows that in these subjects the

TABLE 7 Mean daily energy intake, body weight, body height, body mass index (BMI) and change in body weight over a period of 14 months of young adult Dutch women with different levels of energy intake

Energy intake	No.	Energy intake* (kcal/day)	Body weight (kg)	Body height (cm)	BMI (kg/m ²)	Change in body weight† (kg)
Very low (< 1675 kcal)	24	1461	67.3	168	24.0	0.0
Low (1675–2000 kcal)	25	1856	62.3	166	22.7	0.1
Intermediate (2001–2155 kcal)	25	2089	59.6	166	21.6	1.2
High (2156–2345 kcal)	24	2238	58.9	166	21.2	0.9
Very high (> 2345 kcal)	25	2619	60.4	169	21.3	0.8
Total group	123					
— Mean		2056	61.7	167	22.1	0.6
— SD		412	8.5	6.0	2.7	2.2

* Estimated from 14 monthly 24-hour recalls.

† Difference between the last three and the first three measurements.

TABLE 8 Body weight, height, change in weight over a period of 14 months and RMR of five young adult women. Energy intake as assessed by fourteen 24-hour recalls

Subject	Body weight (kg)	Body height (cm)	Weight change (kg)	RMR (kcal/24 hour)	Energy intake (kcal/24 hour)
A	80.0	174	-0.3	1859	1035
B	65.1	168	3.8	1502	1282
C	62.8	174	0.3	1500	845
D	66.9	166	4.0	1563	1162
E	109.6	170	2.0	1866	1611

mean resting metabolic rate as determined by indirect calorimetry was more than the reported mean total daily energy intake, which should have resulted in a loss instead of the observed gain in body weight in four of these five subjects. This consistent and probably unconscious underestimation of intakes is very hard to overcome and seems to be related to obesity and restrained eating patterns.¹⁸

Variation between individuals. The rather low correlation coefficient of the dietary P/S ratio with the adipose tissue P/S ratio and the high proportion of misclassifications are also due to the fact that we studied a homogeneous group. A larger between-individual variation would increase the correlation coefficients and decrease the extent of misclassification.⁶ If the intake of food or nutrients is related to a disease, the study population may be expected to have a larger variation in intake of energy and the various nutrients than the present study group. Multicentre studies of representative groups of

subjects with different food habits and living in areas at different disease risks have shown a number of associations of disease with diet. Such studies are becoming more and more popular, but a drawback is that they are expensive. In such situations nutritional epidemiology is inevitably faced with two opposing interests. On the one hand the method chosen must be sufficiently precise to be able to classify individuals correctly in quantities of intakes; on the other hand the method must be feasible for use in large numbers of subjects so as to accumulate sufficient numbers of individuals in the groups of cases and controls.

Amount of Dietary Observations

The number of people to be enrolled in a study and the number of days for which dietary intake should be recorded depend upon the objectives and the items of interest in the study. Rational methods exist for the determination of optimal combinations of sample size and number of replicates. It has been shown that the daily records of 24-hour recalls per person cannot be considered as adequate.^{6,7,16,19} For specific nutrients with a large day-to-day variation such as dietary fibre and vitamin A, short questionnaires have been developed for a rough estimate of the intakes, yet accurately enough to discriminate between 'high' and 'low' consumers.^{20,21} These questionnaires may be used in addition to a repeated dietary recall or record method.

Choice of the Dietary Method

In this paper we have focused on the reproducibility of the 24-hour recall method with portion sizes checked by weighing. In a recent review article the 24-hour recall method had been found to be more prone to error and variation than other methods for dietary assessment.¹⁶ It was not the purpose of this paper to promote the 24-hour recall method for use in all nutrition surveys. The selection of a method should be based on several criteria, which have been described elsewhere.¹

CONCLUSION

The aim of this paper was to show how, when using a 24-hour recall method, the weakening effect of a large day-to-day variation in food consumption on the relative risk estimate can be reduced by replication of observations per person. However, even with a high number of survey days per person, a relationship between a purported dietary risk factor and disease can only be demonstrated if cases and controls differ substantially in intake of the dietary component under investigation.

REFERENCES

- 1 Cameron M E, Staveren W A van, eds. *Manual on methodology for food consumption studies*. Oxford, Oxford University Press, 1988.
- 2 Räsänen L. Validity and reliability of recall methods. In: Hautvast J G A J, Klaver W, eds. *The diet factor in epidemiological research*. Euronut Report 1, Wageningen, 1982 p92-9.
- 3 Chalmers F W, Clayton M M, Lorraine O, et al. The dietary record — how many and which days? *J Am Diet Assoc* 1952; 28: 711-7.
- 4 Keys A. Dietary survey methods in studies on cardiovascular epidemiology. *Voeding* 1965; 26: 464-83.
- 5 Liu K, Stamler J, Dyer A, McKeever J, McKeever P. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *J Chron Dis* 1978; 31: 399-418.
- 6 Beaton G H, Milner J, Corey P, et al. Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am J Clin Nutr* 1979; 32: 2546-59.
- 7 Sempos C, Johnson N E, Smith E L, Gillican C. Effects of intra-individual variation in repeated dietary records. *Am J Epidemiol* 1985; 121: 120-30.
- 8 Staveren W A van, Deurenberg P, Katan M B, et al. Validity of the fatty acid composition of subcutaneous fat tissue microbiopsies as an estimate of the long-term average fatty acid composition of the diet of individuals. *Am J Epidemiol* 1986; 123: 455-63.
- 9 Marshall J R, Priore R, Grahams S, Brasure J. On the distortion of risk estimates in multiple exposure level case-control studies. *Am J Epidemiol* 1981; 113: 464-73.
- 10 Hautvast J G A J. Ontwikkeling van een systeem om gegevens van voedings-enquêtes met behulp van de computer te verwerken. *Voeding* 1975; 36: 356-61.
- 11 Beynen A C, Katan M B. Rapid sampling and long-term storage of subcutaneous adipose-tissue biopsies for determination of fatty acid composition. *Am J Clin Nutr* 1985; 42: 560-3.
- 12 Metcalfe L D, Schmitz A A, Pekka J R. Rapid preparation of fatty acid esters for lipids for gas chromatographic analysis. *Anal Chem* 1966; 18: 514-5.
- 13 Staveren W A van, Deurenberg P, Burema J, Groot L C P G M de, Hautvast J G A J. Seasonal variation in food intake, pattern of physical activity and change in body weight in a group of young adult Dutch women consuming self-selected diets. *Int J Obesity* 1986; 10: 133-45.
- 14 Baecke J A H, Staveren W A van, Burema J. Food consumption, habitual physical activity and body fatness in young Dutch adults. *Am J Clin Nutr* 1983; 37: 278-86.
- 15 Freudenheim J L, Johnson N E, Wardrop R L. Misclassification of nutrient intake of individuals and groups using one, two, three and seven days of food records. *Am J Epidemiol* 1987; 126: 703-13.
- 16 Bingham S A. The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviews* 1987; 57: 701-42.
- 17 Madden J P, Goodman S Y, Guthrie H A. Validity of 24-hour recall. Analysis of data obtained from elderly subjects. *J Am Diet Assoc* 1976; 68: 143-7.
- 18 Strien T van. *Eating behaviour, personality traits and body mass*. Lisse, Swets & Zeitlinger, 1986.

19. Beaton G H, Milner J, McGuire V, *et al.* Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins and minerals. *Am J Clin Nutr* 1983; **37**: 986-95.
20. Räsänen L, Pietinen P. A short questionnaire method for evaluation of diets. *Prev Med* 1982; **11**: 669-76.
21. Willett W C, Sampson L, Stampfer M J, *et al.* Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 1985; **122**: 51-65.