Validating the 24-hour recall method as a dietary survey tool[1]

ELISABETH E. I. LINUSSON[2], DIVA SANJUR[3], and EUGENE C. ERICKSON[4]
Cornell University, Ithaca, New York

SUMMARY

Purpose was to evaluate the validity of the 24-hour recall method in a specific population, 86 lactating women in Cornell Medical Center, the New York Hospital. The major hypothesis was that 24-hour recall is a valid dietary survey method for obtaining reliable information on food consumed. This was tested by comparing the quantity of food consumed (by weighing) with food recalled during an interview. Two aspects of the hypothesis were considered: First, the quantity of food recalled for each of 14 food groups, and second, the number of food items recalled. Regression analyses between recall and actual intake showed for all food groups a tendency to overestimate actual intake when consumption was low and underestimate it when it was high. The validity coefficient, ranging from 0.28 (salad) to 0.72 (breakfast cereal), showed a low correlation for most foods. Also the r^2 varied from 0.08 (salad) to 0.52 (breakfast cereal), indicating considerable unobserved errors, varying with food groups. Great variations were observed in the biases. In eight of the food groups the biases were significantly large to invalidate the method. In general, underestimation was greater than overestimation. The recall method appears fairly accurate for qualitative estimation of average for a population group but not highly valid for ascertaining quantity of food consumed.

---

1. This paper is based on a thesis submitted to the Graduate School by Elisabeth Linusson in partial fulfillment of the requirements for a degree of Master of Science in Nutrición.
2. Former graduate student.
3. Associate instead and Assistant Profesor of Human Nutrition and Food.
4. Associate Profesor of Developmental Sociology.

Recibido: 23-8-73.

INTRODUCTION

An important part of nutrition work is to gain valid knowledge of what foods and how much people eat - as well as when, where, and why. The source of information is often either a dietary survey as part of a nutrition status survey, or a food consumption survey per se.

In evaluating the nutritional status of a population group. not only is coordination between the different parameters such as clinical, dietary, biochemical and anthropometric tests needed, but also some assurance that indicators at each level have been periodically assessed for their level of validity. On this respect, several researchers agree that the remarkable inaccuracy of data from dietary surveys in our field studies has been disconcerting (1-4).

In assessing short-term past food intake of a population group, the 24-hour recall method is the most commonly used indicator. It is a rapid, inexpensive, convenient, and fairly accurate method for obtaining a mean estimate of intake for a population group (5, 6). However, the most crucial disadvantage of recall is the limitation of memory, whatever the purpose of a study may be (7, 8). Human errors in the interview also have an important effect on the validity of this technique, as well as the limited observation period of a single day's recall. Often communication problems may result in inaccurate dietary recalls. The respondent may think that the interviewer is not familiar with a particular food or feels that the food is such a common part of the diet that it is not necessary to report it (9). Other shortcomings of this method are the inability of some subjects to identify kinds of food and, in obtaining quantitative information - their inability to estimate correctly the amounts of food consumed.

Yet, in spite of methodological limitations, the 24-hour recall may continue to be the most practical and useful survey tool for assessing dietary intake of population groups. The growing trends in urbanization and modernization, accompanied by changes in life styles (increasing eating-out patterns, changes in food purchasing practices, attitudinal resistance to strangers doing "research"), constitute barriers to

the utilization of more accurate but elaborate and time consuming food weighing techniques at the household level. (10).

Thus, in view of the above, the present study was undertaken with the objectives of evaluating under a particular set of circumstances the validity of the 24-hour recall method as a tool in assessment of food consumption, and to determine whether this ability to recall varies with the different food groups consumed.

The Concept of Validity

In testing validity of a method, the basic analysis of validity must concern its *meaningfulness*. Does the tool measure what is supposed to measure. Validity, however, is not a fixed characteristic; it varies with people, time and environment. Thus, periodic appraisal of the validity is necessary. A requirement for validity is reliability or reproducibility. Yet, a measurement can be reliable without being valid (11, 12, 13, 14). There are many forms of validity. A distinct line exists between *internal* and *external* validity. The former asks whether a difference exits at all in any given comparison (15) while external validity is a question of generalization (15, 16). In addition, *convergent* and *discriminant* validity have been used in dietary surveys (7). For the purpose of this study, convergent validity will show to what degree data from the 24-hour recall method agree with the data from the control method, i.e., the actual food intake, or whether the two separate methods are reasonably independent of each other. The discriminant validity, will indicate whether the subjets remember each food equally well, or if a particular food is a consistently forgotten item.

METHODOLOGY

The investigation, as well as the pretesting phase of it, was carried out at the Lying-In Hospital of Cornell Medical Center, the New York Hospital, during the Spring of 1972. The specific sample population consisted of 86 women hospitalized postpartum, who were in a healthy state. The sample subjects were all from approximately the same social stratum

and whose age ranged from 26 to 39 years old. The selection of the respondents was done on a daily basis, including women who had a living newborn (2 or 3 day-old) child, recovered from operational delivery remaining in the hospital 2 or more days, and who were on a regular diet.

The null hypothesis —that the 24-hour recall is a valid method for determining food intake among women— was tested by comparing the 24-hour recall with weighing of all prepared food, served and left over.

The weighing method served as the control criterion for validating the 24-hour recall method. The day prior to the 24-hour recall interview every prepared individual food item for each subject and meal was measured (household and/or institutional measurements), weighed on Hanson dietetic scales — models 1440 (500 gm.) and 1460 (1000 gms. capacity) and recorded before being placed on the patient's tray. After the meal each subject's tray was checked and recorded by measuring and weighing everything left over. The time relationship between actual intake and recall interview was 24-hours. This took place in the respondent's room and required an average of 23 minutes. Four or five respondents were studied per day. They were unaware of being included in a survey. Memory aids like household measuring cups, glasses, bowls, spoons, a drawing of different-sized pieces of meat, and one plastic model of a 1½ inch square piece of cornbread were used. The subjects attempted to report all the kinds of food consumed and estimated the quantity for each item. All recall information was included in the study material except in-between meals. After the recall of food intake, the women were asked about their food habits. Information on sociocultural, economic, and health characteristics was obtained from the hospital chart. The weighing method was done by one of the authors, who also did the personal recall interviewing.

Analyses

The data were tabulated from each sample, and estimates were converted into grams, cored, verified, and puched into IBM cards for use with data processing equipment (IBM 360). The recall information was analyzed and interpreted in units of food groups, keeping in mind the following criteria: pattern

in which the subjects expressed their food habit answers; the type of grouping based on content of certain nutrients and often used in dietary surveys; and the realistic food grouping for the type of hospital menu and according to their similarity in use - i.e.:

1 = Meat main dishes (meat, fish, poultry, egg)
2 = Dairy products (milk, butter, cheese)
3 = Vegetables (cooked)
4 = Fruits (fresh and all breakfast fruits)
5 = Salads (mixed fruit and vegetables)
6 = Breakfast cereals (hot and cold)
7 = Starches (potato, rice, spaghetti, macaroni)
8 = Breads (all kinds, including sweet rolls, muffins, crackers)
9 = Soups (including clear broth and creamed soups)
10 = Desserts (pie, cake, pudding, ice cream, sherbet, fruits (no fresh)
11 = Beverages (hot and cold, except milk)
12 = Combined main dishes (spaghetti and meat sauce, meat and vegetable, stew, cheeseloaf, etc.)
13 = Sweets (sugar, jelly, honey, candy)
14 = Miscellaneous (dressing, ketchup, mustard, gravy, sauce, pickles, potato chips, olives)

Two aspects of the hypothesis were considered:

1) *Quantity of food recalled for each of 14 food groups.* In determining the pattern and degree of validity, the statistical methods used for these kind of data were linear regression, Pearson's product moment correlation, and scattergram. A multitrait-multimethod matrix designed after Campbell and Fiske, and Becker *et al* (17, 7) was used for interpretation of the validity. The latter authors illustrate its application on data collected by Huenemann and Turner (5). From the regression analysis, bias and unobserved errors were calculated. The symmetry in the deviation of recall from actual intake was calculated.

2) *Percentages of food items recalled per day.*—The data were classified according to frequency distribution into low, medium and high, i.e., $\leq$ 70%, 71-85%, or $\geq$ 86% of correct recall respectively.

RESULTS AND DISCCUSSION

Assessment of Biases in the 24-hour Recall Method.

There are always variations between two measurements. Both the weighing method (actual food intake) and the 24-hour recall method have degrees of accuracy, because a measuring instrument cannot be expected to be error-free, i. e., the 24-hour recall method cannot be one hundred percent valid. It is thus important to know the error of measurement, and it is important that it falls within prescribed limits.

The mean difference between the recalled and actual amounts of food eaten was observed first. This difference will be called the *bias*. If the 24-hour recall method were perfect, then the bias would be zero. The following statistical procedure was used to determined the size of bias in the 14 food groups:

The standard error of the bias is

$$\sqrt{\frac{\sigma_y^2 \quad \sigma_x^2 \quad - \quad 2\rho_{xy}\,\sigma_x\,\sigma_y}{n}}$$

where σ x and σ y refer to the standard deviations of actual and recalled measurements. Estimates of these quantities are available from the regression analysis, so we can test departure from zero for the biases.

The degree of validity of quantity of recalled food intake was evaluated by determination of size of systematic errors in each of 14 food groups.

Table 1 shows the size of biases in these food groups. The size of the bias is significant when the 24-hour recall method is used to estimate the actual food intake for vegetables, salads, breads, soups, desserts, and sweets at the 0.01 percent level and for fruits and combined main dishes at the 0.05 per cent level. The errors were significant in light of the various food groups which may suggest that the 24-hour recall method may be a poor predictor of the actual amount of food eaten. The biases show that the errors in interview technique and in the conversion of amount of food eaten (from household measurements to grams) contribute to the recallability in all eight

TABLE 1

BIASES IN THE 24-HOUR RECALL METHOD FOR 14 FOOD GROUPS

Food Group		Bias [1]	Bias as a Percent of Actual Intake	Standard Error	Z
1	Meat main dishes	-14.30	-9.5	7.53	-1.90 NS
2	Dairy products	-33.10	-5.7	27.32	-1.2. NS
3	Vegetables	-25.95	-22.5	8.48	-3.06 **
4	Fruits	-11.25	- 2.4	6.61	-1.70 *
5	Salads	-41.01	-53.0	7.08	-5.80 **
6	Breakfast cereals	+9.81	+11.9	6.82	+1.43 NS
7	Starches	-16.01	-17.8	8.90	-1.80 NS
8	Breads	-14.01	-19.9	4.26	-3.29 **
9	Soups	-41.38	-51.8	9.08	-4.56 **
10	Desserts	-43.87	-30.8	10.05	-4.37 **
11	Beverages	-28.71	-5.1	24.30	-1.19 NS
12	Combined main dishes	-15.27	-29.2	6.02	-2.54 *
13	Sweets	-16.07	-49.1	2.30	-6.79 **
14	Miscellaneous	-3.32	-12.4	3.05	-1.09 NS

* Significant at the 0.05 level where Z = ±1.96

** Significant at the 0.01 level where Z = ±2.58

NS = Not significant

1. Expressed as mean difference in grams between the recall and actual intake.

foods groups. Small biases were also found in the recall of breakfast cereals, dairy products, meat main dishes, starches, beverages, and miscellaneous foods, but these errors are not significant and thus do not invalidate the 24-hour recall method.

The Pattern of Relationship between Actual and Recall Food Intake

The type of relationship between actual and recalled food intake is shown in Table 2. It shows that for all food groups the "2" values, or intercepts, are positive and the "b" values,

TABLE 2

FITTED REGRESSION LINES AND CORRELATIONS BETWEEN ACTUAL AND RECALLED FOOD QUANTITIES FOR 14 FOOD GROUPS

Food Groups		r*	r^2	$1-r^{2**}$	b	a
1	Meat main dishes	0.57	0.33	0.67	0.47	65.0
2	Dairy products	0.69	0.47	0.53	0.73	124.1
3	Vegetables	0.53	0.28	0.72	0.58	23.1
4	Fruits	0.55	0.30	0.70	0.45	14.4
5	Salads	0.28	0.08	0.92	0.24	17.8
6	Breakfast cereals	0.72	0.52	0.48	0.89	18.7
7	Starches	0.44	0.19	0.81	0.45	34.0
8	Breads	0.48	0.23	0.77	0.47	23.0
9	Soups	0.30	0.09	0.91	0.31	13.9
10	Desserts	0.36	0.13	0.87	0.48	30.6
11	Beverages	0.67	0.45	0.55	0.73	120.6
12	Combined main dishes	0.69	0.48	0.52	0.60	5.5
13	Sweets	0.40	0.16	0.84	0.45	2.0
14	Miscellaneous	0.41	0.17	0.83	0.82	1.4

* All r's significant at .01 level or less.
** Unobserved errors.

or regression coefficients, are between zero and one. The line regression shows that in all food groups the fitted regression line falls above the origin when $X = 0$ and that they have slopes ranging from 0.24 (salads) to 0.89 (breakfast cereals), see Figure 1.

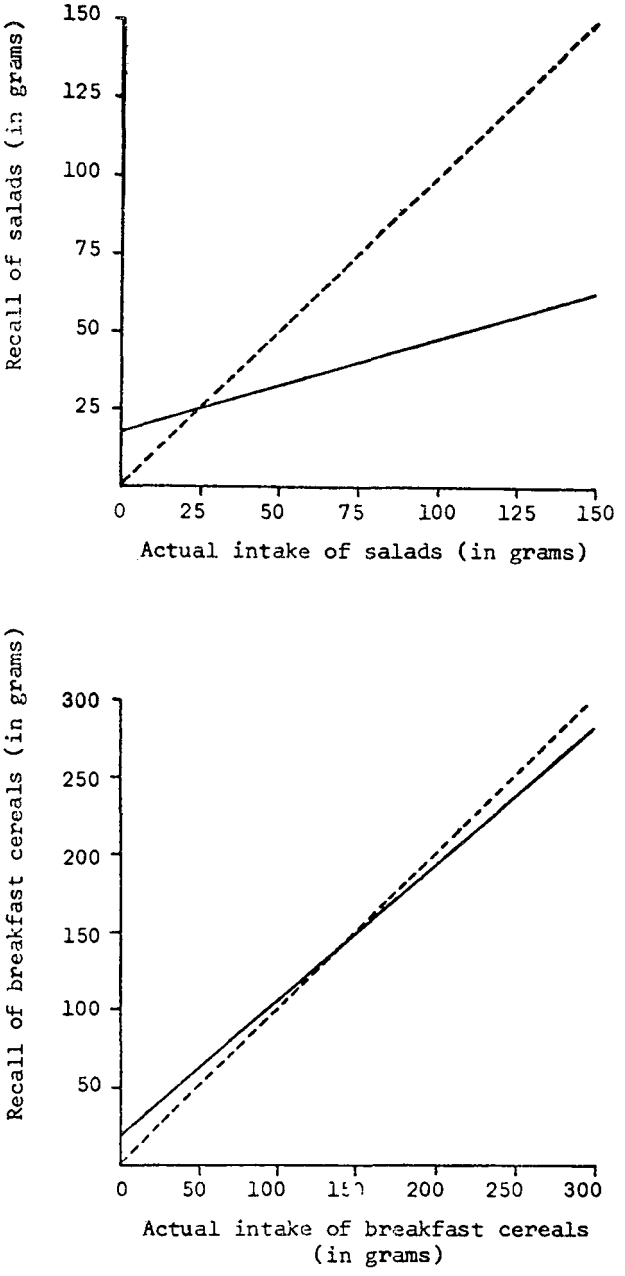

Figura 1. Fitted linear regressions between actual and recalled food intake.

Table 2 also shows that the food groups with "b" close to one also have the highest correlations. The food groups nearest the ideal slope generally have the least error variation due to faulty recall.

Assessment of Unobserved Errors in the 24-Hour Recall

Assuming that there is a linear association between recall and actual intake, then the unobserved errors [$(1-r^2) \times 100\%$] are the percentage of error which the recall method could not predict. Assessment of unobserved errors is shown in Table 2. The complementary estimate of unobserved errors is $r^2 \times 100\%$, which is an estimate of the amount of variation in actual intake that is accounted for by knowing the recalled intake for a given sample. Thus for the food groups salads, soups, desserts, sweets, miscellaneous, and starches, this error, not explained by recalled per se, ranges from 92 to 81 per cent. The figures are somewhat better for the other food groups. For example, breakfast cereals and dairy products have 48 and 53 per cent error variation respectively. Scatter diagrams of the relation between recalled and actual food intake showed no obvious departures from linearity for any of the 14 food groups.

The Degree of Relationship - The Convergent Volidity

The degree of linear association between actual and recall food intake for the 14 food groups was calculated by the Pearson's product moment correlation coefficient (r), i.e., the validity coefficient shown in the multitrait-multimethod matrix of Table 3. In this table, the 14 food groups represent traits. The values in the diagonal are the validity coefficients and they all seem to meet the requirement outlined by Campbell and Fiske (17)[1]. In evaluating the convergent validity of this

1. Validity coefficients should be significantly different from zero and sufficiently large to encourage further examination of validity. Secondly, a validity diagonal value should be higher than the values lying in its column and row in the triangles on both sides of the diagonal. That is, a validity value for a variable should be higher than the correlations obtained between that variable and other variables having neither trait nor method in common (i.e., the coefficients in the two triangles on both sides of the diagonal represent discriminant validity). This requirement may seem so minimal and so obvious as to not need stating, yet an inspection of the literature shows that is frequently not met, and may not be met even when the validity coefficients are of substantial size.

TABLE 3

CORRELATION MULTITRAIT-MULTIMETHOD MATRIX FOR EVALUATION OF CONVERGENT VALIDITY OF THE 24-HOUR RECALL METHOD

Food Groups	24-Hour Recall Method	Weighing Method (Actual Intake)													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Meat main dishes	1	.57	.31	.08	.06	.00	.18	.08	.14	-.07	.10	-.16	-.14	.12	.25
Dairy products	2	.09	.69	.36	.19	.12	.25	-.05	.17	.09	.16	-.11	.30	-.16	.17
Vegetables	3	.06	.24	.53	.37	.16	.11	-.03	.15	.18	-.01	.16	.20	.07	.15
Fruits	4	.03	.15	.41	.55	.21	.22	-.15	.15	.08	.13	.00	.15	-.13	-.09
Salads	5	-.12	.10	-.02	.13	.28	-.11	.08	.01	.17	.09	-.03	.13	-.19	.04
Breakfast cereals	6	.26	.21	.14	.00	.08	.72	.08	.26	-.08	.06	.08	-.12	.26	.19
Starches	7	.08	.03	-.06	.06	.16	.15	.44	.34	-.00	.06	-.06	-.22	.25	.13
Breads	8	.06	.25	.11	.07	.27	.07	.14	.48	.13	.35	.03	.12	-.02	..14
Soups	9	-.23	-.03	.01	.08	.13	-.18	.08	.17	.30	.03	.20	.03	-.05	.12
Desserts	10	-.07	.10	-.05	.13	-.01	.08	.05	.05	.14	.36	.28	.19	.18	.05
Beverages	11	-.15	-.06	.10	.23	-.00	-.08	.06	.08	.03	.20	.67	-.05	.39	.18
Combined main dishes	12	-.21	.14	.08	.10	.11	-.10	-.25	.11	.14	.14	.11	.69	-.17	-.05
Sweets	13	.09	.03	-.12	.05	.04	.11	-.02	.00	-.12	.07	.18	-.17	.40	.04
Miscellaneous	14	.07	.29	.14	.30	.13	.12	.16	.26	.04	.05	.14	-.08	.12	.41

matrix, we note that only four food groups have validity coefficients larger than 0.60. These are breakfast cereals (0.72), dairy products (0.69), combined main dishes (0.69), and beverages (0.67). The lowest values of "r" obtained were for salads (0.28), soups (0.30), and desserts (0.36). The r's for sweets, miscellaneous, and starches were also low, i.e., below 0.45. The validity coefficients for the breads, vegetables, fruits, and meat groups were medium-low to medium.

Over and Underestimation - Extent of Symmetry in the Deviation.

The symmetry in the deviation is an important component of a measurement which must be considered in our aim to standardize the method. Knowledge of an acceptable range around the mean is likewise needed. This is particularly significant, since the 24-hour recall is an estimating tool and analysis of dietary data will rely heavily on the way the deviations are interpreted.

The extent of symmetry in the deviation of the recalled food consumed is shown by Figures 2 and 3, demonstrating the total under and overestimation respectively within the 14 food groups. A comparison of these figures show that the overall tendency to *underestimate* is much greater than it is to overestimate.

In table 4 above the deviations of the average estimation show the range around the mean of under and over estimation.

OVERVIEW

Reports in the literature, as well as the findings hereby discussed seem to raise critical questions concerning the validity of the 24-hour recall when ascertaining quantity of food consumed. Similarly, there seems to be considerable agreement among investigators regarding the usefulness of this method when qualitative aspects of the diets are to be assessed. Present tendencies in eating-out patterns, as well as changes in life styles and attitudes toward strangers doing "research" have also raised some critical issues among nutritionists regarding the urgent need for more practical, though less accurate, short methods of dietary appraisals.

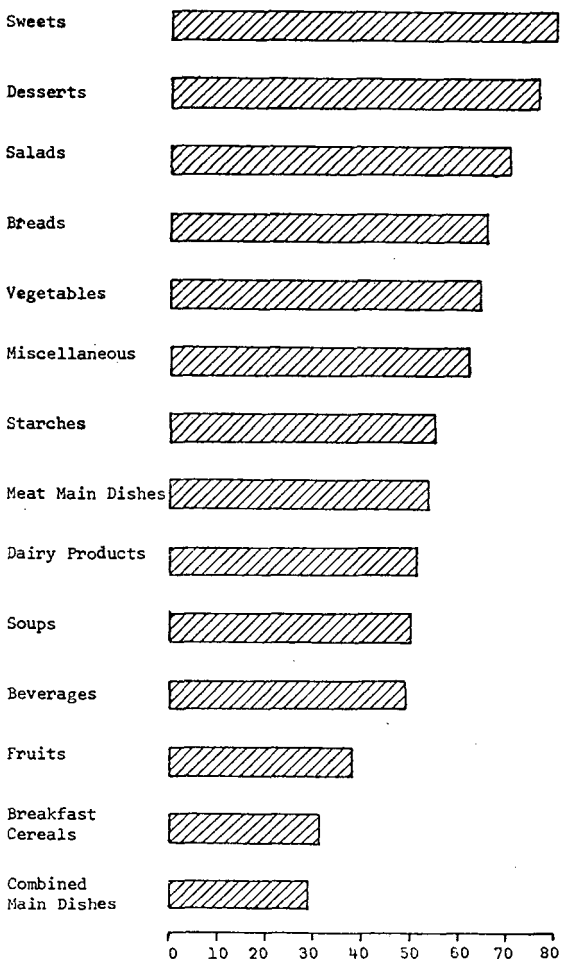

Figure 2. Actual percentage of underestimation of quantity of food.

It is felt that what is crucial prior to determining the choice of a dietary method is a clear definition of any study objectives. If quantitative dietary information is the clear need (as in the case of balance studies, or some epidemiological studies), then the choice of appropriate methods to employ is also clear - for the correct estimation of the quantities of food consumed, there is almost no other alternative but to use one technique, i.e. *direct weighing of food.* Conversely, if

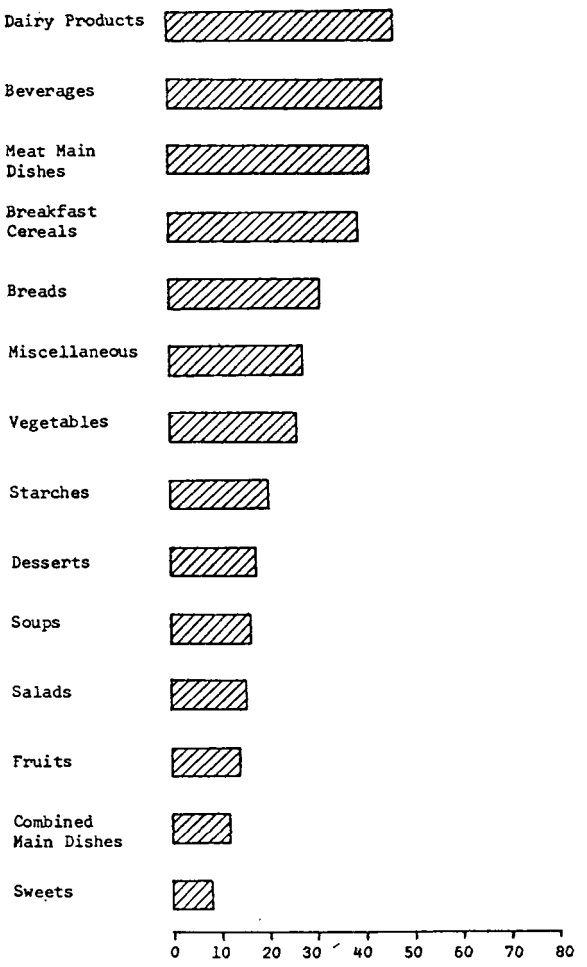

Figure 3. Actual percentage of Overestimation of quantity of food.

TABLE 4

AVERAGES RANGES OF DEGREE OF UNDER AND OVERESTIMATION FROM ACTUAL FOOD INTAKE IN 14 FOOD GROUPS AS ESTIMATED BY THE 24-HOUR RECALL METHOD (expressed in percentages).

Food Groups	III Under-Estimation ≥ -30%	II Medium -15-29%	I Good ±14%	II Medium +15-29%	III Over-Estimation ≥ +30%
1 Meat main dishes	25	17	26	7	25
2 Dairy products	29	6	38	8	19
3 Vegetables	48	11	14	8	19
4 Fruits	30	7	47	2	14
5 Salads	57	7	21	4	11
6 Breakfast cereals	15	11	35	8	31
7 Starches	39	10	29	2	20
8 Breads	42	9	22	6	21
9 Soups	46	2	34	8	10
10 Desserts	47	17	23	4	9
11 Beverages	20	19	31	7	23
12 Combined main dishes	19	5	65	0	11
13 Sweets	68	11	14	1	6
14 Miscellaneous	51	6	17	3	23

the purpose is to gather *trends in dietary patterns* from sizeable population groups - possibly to serve as baseline data in Applied Nutrition and Public Health programs - then the choice of the 24-hour recall method is a valid one.

The ultimate decision depends then on the type of study, sample size, feasibility and other important factors. The 24-hour recall method could be adequate for field longitudinal studies if surveys are carried out in an appropriate way.

Acceptable margins of error up to 10% have been reporter when the 24-hour recall method has been used, the acceptability emerging from the fact that the larger sample size

minimizes the size of the error (19). Moreover, what is an acceptable error will also be a function of the specific objectives of each study.

The data hereby reported seem to suggest that much caution needs to be exercised in using the 24-hour recall method for estimating quantitative food intake of groups in the population. However, in assessing the nutritional status of population groups in a developing country, where less variation in the food intake exists, this limitation might be lessened.

The circumstances under which the present validation of the 24-hour recall was conducted must be kept in proper perspective when examining the findings. The study was conducted in an institutional research setting which provided the outside criterion of control (food weighing method), a requirement for validation purposes. Nevertheless, this may very well not be the usual circumstances under which this methodological tool is usually applied in population studies. In field dietary research wide varieties of situations may arise, giving way to many "intervening" variables - some a function of intrapersonal/interpersonal variance ratio, rather than one of inter or intra-personal variance alone - over and beyond the dimension of "limitation of memory" most extensively tested in the present investigation.

RESUMEN

Valor del método recordatorio de 24 h como instrumento en encuestas dietéticas.

Se efectuó una evaluación del método recordatorio de 24 h en un grupo de 86 mujeres lactantes en el Cornell Medical Center del New York Hospital. La hipótesis que este método produce resultados confiables fue probada comparando la cantidad de alimentos consumidos y determinados por peso con la que se recordó durante una entrevista. Se consideraron los siguientes aspectos: 1º La cantidad de alimentos recordados de c/u de 14 alimentos y 2º el número de alimentos recordados. Al efectuar el análisis de regresión entre alimentos recordados y consumo real se evidenció una tendencia de una estimación alta en los casos de consumo bajo y de una estimación baja cuando el consumo era alto. El coeficiente de validez demostró una correlación baja para la mayoría de los alimentos, estaba entre 0.28 para la lechuga y 0.72 para cereales de desayuno. También variaba el valor de r^2 de 0.08 para lechuga hasta 0.52 para cereales de desayuno, indicándose así considerables errores que variaban con los gru-

pos de alimentos. Generalmente, era más frecuente la estimación baja que la alta. Parece que el método recordatorio es medianamente exacto para la estimación cualitativa de los valores medios para un grupo de población pero no muy válido para averiguar la cantidad de alimentos consumidos.

BIBLIOGRAPHY

1. McHenry, E. W., H. P. Gerguson and J. Gurland. "Sources of error in dietary surveys". **Canad. J. Publ. Health, 36: 355-361, 1945.**
2. Thomson, A. M. "Techniques and perspective in clinal and dietary studies in human pregnancy". **Nutr. Coc. Proc. 16: 45-51, 1957.**
3. Young, C. M., and M. F. Trulson. "Methodology for dietary studies in epidemiological surveys: II. Strengths and weaknesses of existing methods". **Amer. J. Pub. Health, 50: 803-813, 1960.**
4. Groover, J. E., L. Boone, P. C. Houk and S. Wolf. "Problems in the quantitation of dietary surveys". **J. Am. Med. Assoc. 201: 86-89, 1967.**
5. Huenemann, R. L., and D. Turner. "Methods of dietary investigation". **J. Am. Dietet. 18: 562-568, 1942.**
6. Young, C. M., F. W. Chalmers, H. N. Chuch, M. M. Clayton, L. O. Gates, G. C. Hagan, B. F. Steele, R. E. Tucker, A. W. Wertz and Foster. **Cooperative Nutritional Status Studies in the North-East Region: III. Contribution to Dietary Methodology Studies.** Massachusetts Agricultural Experimental Bulletin Nº 469, 1952.
7. Becker, B. G., B. P. Indik and A. M. Beeuwkes. **Dietary Intake Methodologies: A Review.** University of Michigan UMRI Project 03188, 1960.
8. Fidanza, F. "Sources of error in dietary surveys". Abstract, XIth Symposium of GEN., Warsaw, Poland, 1972.
9. Sanjur, D., J. Cravioto, A. van Veen, L. Rosales. "Infant Feeding and Weaning Practices in a Rural Preindustrial Setting: A Sociocultural Approach". **Acta Paediat. Scand. 200: 1-56, 1970.**
10. Sanjur, D., E. Romero and G. Nevill. **A Community Study of Food Habits Among Families Residing in East Harlem, New York.** Department of Human Nutrition and Food, Cornell University, Ithaca, N. Y., 1972, mimeographed).
11. Gulliksen, H. **Theory of Mental Tests.** New York, John Wiley & Sons, Inc., 1950.
12. Moore, F. E. "Committee on design and analysis of studies. Epidemiology of cardiovascular diseases methodology". **Am. J. Publ. Health, 50: 10-19, 1960.**
13. Amos, J. R., F. L. Brown and O. G. Mink. **Statistical Concepts,** New York: Harper & Row, 1965.

14. Suchman, E. A. **Evaluative Research**, New York: Russell Sage Foundation, 1967.

15. Webb, E. J., D. T. Campbell, R. D. Schwartz and L. Sechrest, **Unobtrusive Measures: Non-Reactive Research in the Social Sciences**, Chicago: Rand McNally & Company, 1966.

16. Campbell, D. T. and J. C. Stanley. **Experimental and Quasi-Experimental Designs for Research**, 7th Printing, Chicago: Rand McNally & Company, 1971.

17. Campbell, D. T. and D. W. Fiske. "Convergent and discriminant validation by the multitrait-multimethod matrix". **Psychol. Bull. 2: 56-59, 1959.**

18. Schorr, B. C., D. Sanjur and E. C. Erickson. "Teen-age food habits". **J. Am. Dietet. Assoc. 61: 415-420, 1972.**

19. Flores, M. "Metodología en encuestas alimentarias entre pre-escolares". **Arch. Latinoamer. Nutr. 22: 360-384, 1972.**