International Journal of Epidemiology
Vol. 26, No. 1(Suppl. 1)

Printed in Great Britain

The Malmö Food Study: Validity of Two Dietary Assessment Methods for Measuring Nutrient Intake

ELIO RIBOLI,* SÖLVE ELMSTÅHL, RODOLFO SARACCI,*,† BO GULLBERG,** AND FOLKE LINDGÄRDE‡**

Riboli E (International Agency for Research on Cancer, 150 cours Albert-Thomas, 69372 Lyon Cédex 08, France), Elmståhl S, Saracci R, Gullberg B and Lindgärde F. The Malmö Food Study: Validity of two dietary assessment methods for measuring nutrient intake. *International Journal of Epidemiology* 1997; **26 (Suppl. 1):** S161–S173.

Background. Nutritional epidemiology relies largely on dietary assessment methods for the estimation of the 'exposure' variables which may be related to disease risk.

Methods. This paper describes a methodological study conducted in Malmö, Sweden, to compare nutrient intake—estimated by two alternative dietary assessment methods—with a reference method consisting of 18 days of weighed food records. The two candidate methods were an extensive food frequency questionnaire with portion size to be estimated from a booklet of 120 sets of photos (method A) and a method involving the combination of a shorter questionnaire and a two-week food record (method B).

Results. In absolute values, both methods overestimated nutrient intake by 20–40%, with method B closer to the reference for most nutrients. Both crude and energy-adjusted correlations between A-reference and B-reference were of the order of 0.50–0.60 for energy, energy-providing nutrients and most vitamins and minerals. Correlations were in the same range for most of the 14 fatty acids considered in the analyses. Protein intake, estimated from the analyses of urinary nitrogen on 6–8 repeated 24-hour urine collections per subject, was almost identical to the reference method values. Correlation between nitrogen-derived values and dietary measurement was 0.75.

Conclusions. Overall, the study indicated that both methods A and B had good ranking validity compared to the reference and that in most cases the combined method (B) performed slightly better than the extensive food frequency method (A).

Keywords: diet, nutrients, validity, dietary methods, urinary nitrogen, nutritional surveys, Sweden, EPIC

Research on the relation of diet to the aetiology of cancer and other chronic diseases has proved to be one of the most challenging areas of modern epidemiology. The difficulty of measuring diet during the time period relevant to disease aetiology and with a validity and precision sufficient to detect underlying relations has been widely discussed.[1,2]

As compared to the empirical approach often used in epidemiological studies on diet and cancer during the 1970s, there has been a growing awareness in recent years of the limitations of using dietary questionnaires of unknown validity and reproducibility and, as a consequence, a new interest in the methodological development of dietary questionnaires tailored for epidemiological studies.

Some of these questionnaires have been tested for their relative validity and reproducibility in so called 'validation' studies.[3–5] In these methodological studies the measurements of current usual diet during a given time period for a group of subjects were compared to the measurements obtained for the same subjects with a method of (relatively) much higher validity but too cumbersome to be used for thousands and even tens of thousands of subjects in large epidemiological cohort studies.

The Malmö diet study was originally designed in 1984 in order to test two alternative methods of measuring diet, to be used eventually in the Malmö Diet and Cancer Study. The design of the study was conceptually based on the idea of comparing three 'families' of measurements: the two candidate methods, the dietary method of reference (weighed food records) and biomarkers of diet from blood and urine samples. A succinct description of the study and partial results were reported in the proceedings of a symposium held in Malmö in 1991.[6]

This paper reports the results on the relative validity of the measurements of 16 main nutrients and 12 fatty

* International Agency for Research on Cancer, 150 cours Albert-Thomas, 69372 Lyon Cédex 08, France.
** Department of Community Health Sciences, Lund University, S-21401 Malmö, Sweden.
† National Research Council, via Trieste 41, 56100 Pisa, Italy.
‡ Department of Medicine, Lund University, Malmö General Hospital, S-21401 Malmö, Sweden.

Randomized groups	Dietary method	Reference dietary method (3-day weighed food records – 6 times)						Dietary method	No. of subjects
		1	2	3	4	5	6		
1	A1	→	→	→	→	→	→	A2	57
		1	2	3	4	5	6		
2	A1	→	→	→	→	→	→	B2	50
		1	2	3	4	5	6		
3	B1	→	→	→	→	→	→	B2	54
		1	2	3	4	5	6		
4	B1	→	→	→	→	→	→	A2	45
Time in months	0	1–2	3–4	5–6	7–8	9–10	11–12	12–13	

A: 250-question food frequency questionnaire
B: 130-question food frequency questionnaire combined with a two-week food record

FIGURE 1 *General outline of the study protocol*

acids obtained by the two candidate methods as compared to those obtained with the weighed food records. In addition, measurements of proteins from the dietary assessment methods were compared to the protein intake estimated from urinary nitrogen excretion. Results on reproducibility, which have been published,[7] showed correlation coefficients of the order of 0.50–0.90 for nutrients and foods.

The main reason for including two alternative methods in the study was to compare an extensive food frequency questionnaire—with food portions estimated using photos—to a new method based on the combination of 14 days of (non-weighed) food records with a shorter quantitative food frequency questionnaire with closed and open questions. This decision seemed justified to the authors by the potential interest in using information on 'actually consumed' food, as can be provided by food records, which was expected to overcome some of the limitations of the more traditional food frequency approach. Moreover, there was an interest in testing the performances of a food frequency questionnaire developed close to the maximum length still compatible with reasonable compliance. These approaches were somewhat at variance with the prevailing tendency of the 1980s of trying to simplify as far as possible the dietary questionnaires designed for diet and cancer studies.

MATERIALS AND METHODS

Study Design

To estimate the validity of the two dietary assessment methods it was decided to sub-divide the study subjects randomly into six groups. The two methods which, for brevity, are called Method A (extensive food frequency questionnaire) and Method B (combined food record-food frequency method) are described below.

Groups 1 to 4 were set up to estimate validity by comparing method A or B, administered at the beginning and at the end of the 12-month study period, with the reference method. While groups 1 and 3 filled in twice either method A or B, the subjects in groups 2 and 4 switched method from A to B or from B to A. This provided a kind of 2×2 squared design, the aim of which was to estimate whether the concordance between methods would be affected by the order in which dietary assessment methods are administered in validation studies (Figure 1).

Subjects in groups 5 and 6 were asked to complete twice, 1 year apart, either method A or B but not to keep weighed food records. The results of this component of the study were reported in a paper on reproducibility of dietary measurements.[7]

Subject Recruitment and Participation

A list of 887 residents (half men, half women) of the town of Malmö aged 50–69 was randomly extracted from the computerized population registry. All subjects were invited by letter to attend the general health screening which was ongoing at the Institute of Preventive Medicine and were asked to volunteer to take part in a study on diet.

Of the 887 invited, 563 (64%) attended the health screening in September-October 1984. After 19 subjects had been excluded for medical reasons or language problems, 544 people were randomized to groups 1 to 6; 75% of those assigned to groups 1 to 4 accepted, while the corresponding figure for groups 5 and 6 was 84%. Of those in groups 1 to 4 who started the first

3-day period of diet recording, 18% dropped out during the 1-year period and did not complete the study. The dropout rate was higher in groups 5 and 6; of those who filled in questionnaire A or B at the start, 24% did not return 1 year later when they were invited to fill in the same questionnaire a second time.

Methods for Measuring Diet

Extensive food frequency questionnaire with portion size estimation (method A). Method A, which was designed to cover the whole diet over the preceding 12-month period, contained questions on frequency of consumption and usual portion size for 350 foods. It was 50 pages long, self-administered, and was structured by major groups of foods, prepared dishes and beverages. Subjects were asked to indicate frequency of usual consumption per week or month, seldom or never. Frequency was open and not in pre-defined categories.

The portion size was estimated from a 78-page booklet with 122 sets of photographs or drawings of foods and dishes in varying portions (usually four). For each food the corresponding series of pictures in the booklet was indicated by a number in the questionnaire, and portions were indicated by alphabetical letters A, B, C or D going from the smallest to the largest. The subjects, in groups of 1–4 people, were instructed by a dietician on how to fill in the questionnaire in the hospital where health screening took place. It took most of the subjects 1–2 hours to complete the questionnaire. Approximately 1 week later a dietician had a personal meeting with each subject to check whether the questionnaire was filled in satisfactorily.

Combined food record-food frequency method (method B). Method B involved the combination of a much shorter food frequency questionnaire (130 items) and a 2-week food record. The questionnaire did not cover any 'hot meals', i.e. lunch and dinner, and, to assess these, a 2-week food record was used. The subjects were instructed to record detailed information on what was eaten at each 'hot meal'. When subjects returned the questionnaire and the food records 2–3 weeks later, the dietician recorded the usual amounts of each item consumed by the subject. The initial instructions were given to study subjects in small groups, and the checking was carried out individually when the subjects returned the questionnaires.

The reference method, to which methods A and B were compared, consisted of weighed food records kept over six 3-day periods, evenly distributed over 1 year (Table 1). The periods were selected so that each subject recorded the food once every 2 months and so that the 18 days of records for each subject included two Tuesdays, Wednesdays and Thursdays, and three of each of the other days of the week. Thus, the subject was asked to weigh and record everything he/she ate or drank for a total of 18 days. For each recording period the dietician had two individual meetings with the subject, the first to provide instructions and the second to check the food records. Information for the validity study (groups 1, 2, 3 and 4) was collected by four dieticians, while the repeatability study (groups 5 and 6) was run by only one dietician. Each dietician followed 'her' subjects from the first health screening, and all through the second screening. In groups 1, 2, 3 and 4, however, another dietician carried out the final A or B measurement because the participant's 'usual' dietician was suspected of knowing too much about the person's food habits from the reference method to make a 'blind' measurement.

In order to minimize inter-interviewer variation and its potential effects on the study results, all the dieticians were trained together, quality checks were performed several times during the study, and each dietician in the validity study had approximately 25% of the subjects in each group (1 to 4).

Food coding and nutrient calculations. The data were coded using the Swedish Food Data Base[8] prepared by the National Food Administration, which provides figures on the content of 34 nutrients for about 1500 food items, drinks and recipes. The food data collected were coded by dieticians other than those doing the field work. The concordance of the coding was checked on several occasions (e.g. by two dieticians coding the same material); only a very low proportion of minor discrepancies were found, and these were solved by the

TABLE 1 *Mean and standard deviation of age, height, weight and quetelet index for the subjects randomized in the four study groups*

Study groups:		1	2	3	4
Dietary method:		A1 + A2	A1 + B2	B1 + B2	B1 + A2
No. of subjects		57	50	54	45
Age	Mean	62.7	61.2	60.3	61.3
	± SD	5.7	5.1	5.8	5.5
Height	Mean	168.2	168.8	169.9	167.0
	± SD	10.7	8.1	9.7	8.6
Weight	Mean	73.3	71.7	71.5	72.0
	± SD	12.7	8.1	11.0	12.0
Quetelet	Mean	25.9	25.1	24.7	25.8
	± SD	3.8	3.4	2.4	3.9

head nutritionist. For the purposes of the statistical analyses presented in this paper, vitamin intakes were restricted to vitamins contained in foods and did not include vitamin supplements. This information was collected, and will be used in connection with measurements of vitamins in blood.

Collection and Analysis of Urine and Blood Samples

Twenty-four hour urine samples were collected from a subsample of the subjects included in the methodological study. Each subject was asked to collect 24-hour urines eight times during the 12-month study period. To simplify the task and make it less cumbersome for the volunteers, the 8 days were in most cases subdivided into four periods of 2 days each, but the urines of the two consecutive days were kept in separate bottles.

Samples were collected using standard procedures. The subjects were provided with a bag, plastic bottles and cups. They were instructed by the dieticians to keep a careful record of their urine collection, which included the time of start and end of the collection, and a self-evaluation of the completeness of collection at each micturation (certainly complete, possibly incomplete, certainly incomplete). When the subjects brought in their samples, the dieticians went over with them the reports on completeness of urine collection over the 24-hour period, and a score of completeness was then computed on a scale of 1 to 10 for the whole 24-hour period. This was used to classify 24-hour urine samples into 'probably or certainly complete' and 'probably or certainly incomplete'. Statistical analyses were conducted on the two subgroups separately.

Urine samples were analysed for urinary nitrogen in Professor Isaksson's Department of Clinical Nutrition, University of Gothenburg, by the Kjelldahl method. Protein values (g/day) were computed from nitrogen output by means of the formula:

$$\text{Protein in diet (g)} = (\text{Ng} + 2) \times 6.25$$

where 2 indicates the average extra-urinary losses of nitrogen (in grams, through faeces, sweat and skin cell exfoliation), and 6.25 the standard nitrogen factor for protein in diet.[9]

During the study period, blood samples were collected four times (once every 3 months) from 60 subjects. Results of the biochemical analyses will be presented in a separate report.

Statistical Methods

Data on nutrient intake for each day of food recording (reference method) and for the usual daily intake estimated by methods A and B were first analysed for quality control in terms of distributions, outliers, absolute and relative values. Outliers were systematically checked by returning to the original records, and corrections were made, when appropriate, in the reference method. Errors detected in methods A and B were not corrected if they were due to a subject's misinterpretation of the questionnaire, as this was considered to be part of the characteristics of the methods to be tested, while errors due to data input or management were corrected whenever detected.

Nutrient data were analysed using SAS statistical procedures for calculation of means, standard deviations, quintiles and Pearson's correlation coefficients. Energy-adjusted correlation coefficients were calculated following the regression residual method described by Willett and Stampfer.[10] The correlations were also computed adjusting for age but they are not presented; as might be expected from the narrow age range (50–69) they were identical to those adjusted for energy only.

RESULTS

Average age and standard deviation for age, height, weight and Quetelet's index (Body Mass Index) for the 206 subjects (101 men and 105 women) included in the final analysis are shown in Table 1 for each of the four study groups included in the validity study. The four groups did not differ significantly for any of these variables.

The average intake of energy, energy-providing nutrients, and some vitamins and minerals estimated by different dietary assessment methods are reported in Tables 2a and 2b for men and women, respectively.

The dietary patterns estimated by the reference method are typical of those expected in a middle-aged/elderly group of men and women (aged between 50–69) with a rather sedentary lifestyle.[11]

Both methods A and B overestimated the intake of energy-providing nutrients and therefore of total energy. The two methods also provided substantially larger estimates of between-person variability than the reference method, as shown by the larger standard deviations of the means for all nutrients. The overestimation was more pronounced for A1 and B1 than for A2 and B2. These results may suggest either that A2 and B2 were more appropriate measurements of the diet during the 12 months covered by the reference method or that the subjects had 'learned' during the study period to report their diet more correctly. While this issue is of general interest for the interpretation both of the present study and of similar studies, and will be reconsidered in the discussion, it is of interest to note that the lower overestimation observed for A2 and B2 is independent of the order in which method A or B was administered.

TABLE 2a *Daily intake of selected nutrients estimated by either method A (extensive food frequency questionnaire) or method B (combined food records-food frequency) compared to the reference method. A1 and B1 at start of 12-month study period. A2 and B2 at the end.*

		Men							
No. of subjects		Groups 1 & 2 57		Groups 3 & 4 44		Groups 1 & 4 46		Groups 2 & 3 55	
		A1	Ref	B1	Ref	A2	Ref	B2	Ref
Energy (kcal)	Mean:	3093	2348	3005	2323	2892	2340	2771	2335
	± SD	950	451	1127	461	933	437	721	470
Protein, % energy	Mean:	13	14	14	14	14	14	13	14
	± SD	2	2	2	2	2	2	2	2
Fat, % energy	Mean:	40	39	40	40	39	40	41	39
	± SD	8	5	7	5	7	5	7	5
Carbohydrates, % energy	Mean:	45	44	44	43	45	43	44	44
	± SD	7	5	8	6	8	5	7	6
Alcohol, % energy	Mean:	2	3	3	3	2	3	2	3
	± SD	3	3	4	4	3	4	2	2
Protein (g)	Mean:	100	80	103	79	97	80	90	80
	± SD	29	14	44	16	29	14	22	16
Fat (g)	Mean:	138	102	131	104	126	103	126	102
	± SD	56	22	48	23	53	20	40	24
Cholesterol (mg)	Mean:	446	390	460	379	407	309	395	381
	± SD	170	106	176	115	128	98	130	120
Carbohydrates (mg)	Mean:	349	262	337	253	327	254	307	261
	± SD	119	62	154	68	120	64	100	65
Sugar (g)	Mean:	62	47	52	43	54	46	49	45
	± SD	41	25	27	24	25	23	35	26
Alcohol (g)	Mean:	9	10	10	10	9	11	8	8
	± SD	10	9	13	12	9	12	8	8
Retinol (µg)	Mean:	2213	1384	2162	1529	2203	1345	2133	1533
	± SD	1176	638	1369	754	1689	593	1263	758
β-carotene (µg)	Mean:	3653	1943	2587	2024	2665	1958	2775	1995
	± SD	2454	1074	1548	1265	1648	1215	2023	1115
α-tocopherol (µg)	Mean:	12 342	8622	11 183	8676	10 800	8368	11 485	8878
	± SD	8993	3266	6501	2904	9686	3116	6924	3093
Ascorbic acid (mg)	Mean:	98	61	82	61	81	58	76	64
	± SD	47	35	38	33	38	34	41	34
Folate (µg)	Mean:	308	220	294	208	287	209	267	220
	± SD	96	53	128	52	109	53	89	52
Calcium (mg)	Mean:	1300	921	1352	916	1104	921	1096	918
	± SD	570	267	910	344	461	305	447	302
Zinc (µg)	Mean:	12 667	10 372	13 738	10 464	12 558	10 356	12 072	10 459
	± SD	3764	1970	6307	2215	3728	1909	3132	2211
Selenium (µg)	Mean:	41	37	40	34	39	35	36	37
	± SD	14	12	14	9	14	9	11	12
Fibre (g)	Mean:	27	18	25	17	25	17	25	19
	± SD	11	5	11	5	12	5	11	6

TABLE 2b *Daily intake of selected nutrients estimated by either method A (extensive food frequency questionnaire) or method B (combined food records-food frequency) compared to the reference method. A1 and B1 at start of 12-month study period. A2 and B2 at the end.*

		Women							
No. of subjects		Groups 1 & 2 50		Groups 3 & 4 55		Groups 1 & 4 56		Groups 2 & 3 49	
		A1	Ref	B1	Ref	A2	Ref	B2	Ref
Energy (kcal)	Mean:	2039	1634	1886	1663	1959	1670	1768	1625
	± SD	460	277	572	315	566	283	537	312
Protein, % energy	Mean:	14	14	14	14	14	14	14	14
	± SD	3	2	2	2	3	2	2	2
Fat, % energy	Mean:	39	39	39	39	39	39	38	39
	± SD	6	4	5	4	7	4	7	4
Carbohydrates, % energy	Mean:	46	45	45	45	45	44	46	45
	± SD	6	5	5	5	7	5	7	5
Alcohol, % energy	Mean:	1	2	2	2	1	2	1	2
	± SD	1	3	2	2	2	3	1	2
Protein (g)	Mean:	73	58	65	58	69	58	62	58
	± SD	19	11	16	10	19	9	19	12
Fat (g)	Mean:	88	71	83	73	87	73	76	70
	± SD	27	13	32	17	37	16	29	14
Cholesterol (mg)	Mean:	337	289	299	291	308	289	288	292
	± SD	101	94	92	80	103	87	106	87
Carbohydrates (mg)	Mean:	236	185	216	187	221	187	307	185
	± SD	60	41	65	40	59	38	68	43
Sugar (g)	Mean:	39	37	37	37	38	37	37	36
	± SD	22	16	17	14	20	14	31	16
Alcohol (g)	Mean:	3	5	4	5	3	6	2	4
	± SD	3	7	6	5	4	7	3	5
Retinol (µg)	Mean:	1701	1009	1228	927	1260	895	1163	1048
	± SD	1198	690	888	521	781	474	865	724
β-carotene ((µg)	Mean:	4464	2741	3047	2237	3704	2442	3043	2516
	± SD	3111	1942	1837	1648	3059	1819	2650	1815
α-tocopherol (µg)	Mean:	10 036	7178	8965	7601	9655	7783	8303	6962
	± SD	5873	2417	5056	2899	9979	2697	7620	2611
Ascorbic acid (mg)	Mean:	104	78	82	67	94	72	84	73
	± SD	53	43	38	33	53	38	42	38
Folate (µg)	Mean:	276	193	224	183	244	189	209	186
	± SD	95	58	91	52	198	49	67	61
Calcium (mg)	Mean:	940	672	882	725	920	792	820	693
	± SD	456	185	344	180	367	163	420	206
Zinc (µg)	Mean:	9371	7415	8385	7390	8936	7425	8091	7575
	± SD	2705	1531	2084	1366	2809	1273	2264	1622
Selenium (µg)	Mean:	34	30	29	28	30	28	28	30
	± SD	11	12	9	9	10	10	10	12
Fibre (g)	Mean:	22	15	19	15	19	15	18	15
	± SD	8	5	6	4	8	4	6	5

TABLE 3a *Over- or underestimation of dietary methods A1 and B1 as compared to the reference method. The values in the table were calculated as ([A1-Ref]/Ref) × 100 for each subject and expressed as the average percentage of the reference value. (A1 and B1 were administered at the start of the 12-month study period)*

Nutrient	Men		Women		All subjects	
	A1	B1	A1	B1	A1	B1
No. of subjects	57	44	50	55	107	99
Energy	33	29	27	14	30	21
Protein, % energy	–8	0	0	0	–4	0
Fat, % energy	3	0	0	0	1	0
Carbohydrates, % energy	2	2	2	0	2	1
Protein	25	30	27	14	26	21
Fat	35	27	28	15	32	21
Cholesterol	17	26	23	8	20	16
Carbohydrates	36	33	30	16	33	24
Sugar	51	43	11	3	32	21
Alcohol	14	186	7	–21	11	75
Retinol	79	61	96	51	87	56
Carotene	120	76	87	89	104	83
Tocopherol	40	31	41	17	41	23
Ascorbic acid	90	59	45	36	69	46
Folate	44	42	47	28	45	34
Calcium	42	52	41	23	42	36
Zinc	24	31	28	16	26	23
Selenium	16	22	23	11	19	16
Fibre	49	44	46	32	48	38

TABLE 3b *Over- or underestimation of dietary methods A1 and B1 as compared to the reference method. The values in the table were calculated as ([A1-Ref]/Ref) × 100 for each subject and expressed as the average percentage of the reference value.*

Nutrient	Men		Women		All subjects	
	A1	B1	A1	B1	A1	B1
No. of subjects	57	44	50	55	107	99
Saturated						
Total	37	30	28	14	33[a]	21[a]
10:0, Capric acid	41	43	32	15	37	28
12:0, Lauric acid	57	40	39	27	49	32
14:0, Myristic acid	37	35	28	12	33	22
16:0, Palmitic acid	35	29	29[b]	13[b]	32[c]	20[c]
18:0, Stearic acid	41	29	28[a]	15[a]	35[d]	21[d]
20:0, Arachidic acid	10	1	5	–5	8	–2
Monounsaturated						
Total	35	26	24	15	30	20
16:1, Palimitoleic acid	19	35	12	12	16	23
18:1, Oleic acid	28	23	17	14	23	18
Polyunsaturated						
Total	38	29	38	22	38	25
18:2, Linoleic acid	39	31	39	23	39	26
18:3, Linolenic acid	34	29	27	16	31	22
20:4, Arachidonic acid	33	31	31	18	32	24
20:5, Timnodonic acid	24	30	5	31	15	31
22:5, Eicosapentaenoic acid (EPA)	28	22	17	15	23	18
22:6, Eicosahexaenoic acid (DHA)	15	22	11	16	13	19
P/S ratio	2	4	8	8	5	6
Mono+poly/sat. ratio	17	14	13	11	16	12

[a,b,c,d] Indicate statistically significant ($P < 0.05$) differences in overestimation (A > B).

In practice, the tendency is the same for the subjects in groups 1 and 3 (A-Ref-A or B-Ref-B) as for those in groups 2 and 4 (A-Ref-B, B-Ref-A).

When intake values, at the individual level, obtained by methods A1 or B1 were compared to those obtained by the reference method (Table 3a), there was a statistically significant overestimation for method A of energy, protein, carbohydrates, cholesterol, retinol, tocopherol, folate, calcium and zinc among women. Among men, the overestimation was also more pronounced for method A than for method B, but the two methods did not differ significantly. Overall, method B showed a less pronounced overestimation of intakes than method A, for both energy-providing nutrients and for vitamins and minerals. It is of interest to note that the differences in estimated intake of vitamin C were mainly due to overestimation of vegetable consumption by both methods A and B, and the differences in calcium intake were mainly due to overestimation of dairy products.

In recent years there has been growing interest in investigating the role of the fatty acid composition of dietary fats in addition to total fat intake, and to go beyond the simple distinction between saturated, monounsaturated and polyunsaturated fats. Table 3b shows the average percentage over- or underestimation of the intakes of specific fatty acids estimated by either method A or B compared to the reference method. It can be seen that overestimation was higher for method A than for method B, particularly among women. The difference in overestimation was statistically significant for 16:0 among women and for 16:0, 18:0 and total saturated fat in both sexes combined. It should be noted that the P/S ratio was only very slightly overestimated (2–8%). In fact, the overestimation was very similar at the individual level for the intake of both polyunsaturated and saturated fatty acids, and the ratio was almost not affected.

Crude and energy-adjusted Pearson correlations between the reference method and either method A or B are reported in Tables 4a and 4b. To correct for skewness

TABLE 4a *Pearson correlation coefficients, crude and energy-adjusted, between daily nutrient intakes measured either by method A and reference or by method B and reference. Correlations were computed on log-transformed variables*

	Men							
	Crude correlation				Energy-adjusted correlation			
	A1-Ref	B1-Ref	A2-Ref	B2-Ref	A1-Ref	B1-Ref	A2-Ref	B2-Ref
Energy	0.55	0.59	0.56	0.55	–	–	–	–
Protein	0.56	0.52	0.55	0.49	0.54	0.53	0.49	0.54
Fat	0.59	0.51	0.63	0.51	0.65	0.50	0.62	0.64
Cholesterol	0.56	0.44	0.59	0.56	0.54	0.46	0.59	0.56
Carbohydrates	0.57	0.66	0.60	0.64	0.65	0.60	0.60	0.66
Sugar	0.63	0.61	0.49	0.67	0.62	0.57	0.50	0.60
Alcohol	0.67	0.75	0.85	0.78	0.72	0.74	0.86	0.80
Retinol	0.24	0.37	0.55	0.27	0.37	0.28	0.58	0.39
Carotene	0.41	0.29	0.48	0.53	0.46	0.30	0.42	0.48
Tocopherol	0.68	0.59	0.70	0.60	0.77	0.45	0.70	0.65
Ascorbic acid	0.59	0.56	0.58	0.66	0.64	0.58	0.62	0.64
Folate	0.48	0.61	0.47	0.66	0.53	0.66	0.42	0.75
Calcium	0.63	0.56	0.74	0.64	0.58	0.55	0.75	0.70
Zinc	0.46	0.52	0.41	0.52	0.41	0.57	0.25	0.58
Selenium	0.53	0.48	0.53	0.45	0.53	0.44	0.56	0.46
Fibre	0.58	0.69	0.54	0.72	0.61	0.72	0.60	0.74

95% confidence interval for a sample size = 50: –0.09–0.46 for r = 0.20; 0.01–0.54 for r = 0.30; 0.13–0.62 for r = 0.40; 0.25–0.69 for r = 0.50; 0.38–0.76 for r = 0.60; 0.52–0.82 for r = 0.70; 0.67–0.88 for r = 0.80.

TABLE 4b *Pearson correlation coefficients, crude and energy-adjusted, between daily nutrient intakes measured either by method A and reference or by method B and reference. Correlations were computed on log-transformed variables*

	Women							
	Crude correlation				Energy-adjusted correlation			
	A1-Ref	B1-Ref	A2-Ref	B2-Ref	A1-Ref	B1-Ref	A2-Ref	B2-Ref
Energy	0.28	0.54	0.48	0.55	–	–	–	–
Protein	0.56	0.24	0.41	0.46	0.64	0.40	0.60	0.53
Fat	0.22	0.56	0.35	0.55	0.60	0.58	0.37	0.69
Cholesterol	0.59	0.43	0.41	0.69	0.68	0.52	0.58	0.71
Carbohydrates	0.51	0.63	0.66	0.64	0.46	0.65	0.43	0.70
Sugar	0.68	0.77	0.57	0.74	0.67	0.73	0.49	0.74
Alcohol	0.79	0.82	0.82	0.83	0.73	0.81	0.80	0.78
Retinol	0.52	0.32	0.40	0.69	0.52	0.29	0.44	0.72
Carotene	0.68	0.51	0.55	0.71	0.68	0.50	0.58	0.70
Tocopherol	0.57	0.77	0.66	0.72	0.70	0.74	0.69	0.83
Ascorbic acid	0.77	0.64	0.78	0.73	0.78	0.65	0.85	0.71
Folate	0.59	0.40	0.69	0.73	0.69	0.47	0.71	0.75
Calcium	0.62	0.57	0.58	0.72	0.63	0.65	0.58	0.73
Zinc	0.54	0.28	0.39	0.45	0.47	0.31	0.51	0.44
Selenium	0.61	0.15	0.52	0.41	0.60	0.22	0.59	0.44
Fibre	0.61	0.50	0.69	0.63	0.67	0.60	0.69	0.69

95% confidence interval for a sample size = 50: –0.09–0.46 for r = 0.20; 0.01–0.54 for r = 0.30; 0.13–0.62 for r = 0.40; 0.25–0.69 for r = 0.50; 0.38–0.76 for r = 0.60; 0.52–0.82 for r = 0.70; 0.67–0.88 for r = 0.80.

TABLE 5 *Pearson correlation coefficients, crude and energy-adjusted for daily intake of specific fatty acids, measured by method A and reference or by method B and reference. Correlations were computed on log-transformed variables*

Fatty acids	Men				Women			
	Crude correlation		Energy-adjusted correlation		Crude correlation		Energy-adjusted correlation	
	A1-Ref	B1-Ref	A1-Ref	B1-Ref	A1-Ref	B1-Ref	A1-Ref	B1-Ref
Saturated								
Total	0.56	0.55	0.68	0.56	0.28	0.54	0.63	0.68
10:0, Capric acid	0.59	0.59	0.71	0.65	0.56	0.24	0.68	0.64
12:0, Lauric acid	0.56	0.52	0.62	0.41	0.22	0.56	0.52	0.49
14:0, Myristic acid	0.57	0.51	0.73	0.59	0.59	0.43	0.66	0.64
16:0, Palmitic acid	0.63	0.44	0.64	0.47	0.51	0.63	0.66	0.66
18:0, Stearic acid	0.67	0.67	0.52	0.51	0.68	0.77	0.61	0.69
Monounsaturated								
Total	0.41	0.75	0.50	0.59	0.52	0.32	0.58	0.66
16:1, Palmitoleic acid	0.20	0.37	0.66	0.51	0.68	0.51	0.56	0.45
18:1, Oleic acid	0.68	0.29	0.47	0.58	0.57	0.44	0.57	0.63
Polyunsaturated								
Total	0.59	0.39	0.72	0.26	0.57	0.77	0.69	0.64
18:2, Linoleic acid	0.48	0.59	0.77	0.23	0.77	0.64	0.69	0.68
18:3, Linolenic acid	0.63	0.56	0.58	0.22	0.59	0.40	0.79	0.58
20:4, Arachidonic acid	0.46	0.61	0.70	0.55	0.62	0.57	0.54	0.44
20:5, Timnodonic acid	0.53	0.56	0.49	0.24	0.54	0.28	0.69	0.38
22:5	0.58	0.52	0.38	0.37	0.61	0.15	0.65	0.40
22:6, Docosahexaenoic acid	0.69	0.48	0.50	0.20	0.61	0.50	0.70	0.27
P/S ratio	0.73	0.44	0.72	0.41	0.75	0.77	0.74	0.74
Mono+poly/sat. ratio	0.58	0.51	0.65	0.21	0.57	0.62	0.65	0.68
Average correlation[a]	0.55	0.52	0.59	0.40	0.58	0.52	0.62	0.56

[a] Does not include P/S and Mono+poly/sat ratios.

95% Confidence Interval for a sample size = 50: –0.09–0.45 for r = 0.20; 0.01–0.54 for r = 0.30; 0.13–0.62 for r = 0.40; 0.25–0.69 for r = 0.50; 0.38–0.76 for r = 0.60; 0.52–0.82 for r = 0.70; 0.67–0.88 for r = 0.80.

of the distributions, all values of nutrient intake were log transformed. Average energy-adjusted correlations are of the order of 0.53–0.57 for method A and 0.52–0.69 for method B, depending on subgroup and sex. Most of the correlations were similar for methods A and B.

It should be noted that correlations for fat intake were relatively high for method B (0.50–0.69), while for method A the crude correlation for fat among women was quite low (0.22). As usual, the highest correlations were found for alcohol intake (0.67–0.86).

Overall, the means of the energy-adjusted correlations for method B are 0.40 and 0.56 for men and women, respectively, and 0.59 and 0.62 for men and women, respectively, for method A. The detailed values for fatty acid composition available from the Swedish food tables made it possible to estimate the validity of the measurements of specific fatty acids which may be of particular interest in investigations on cancer or cardiovascular disease and diet (Table 5). The correlations for the essential fatty acids (linoleic and linolenic fatty acids) were higher for method A (range 0.58–0.79) than for method B (range 0.22–0.68). The correlations for long chain polyunsaturated fatty acids were also generally lower with method B than with method A.

For all variables shown in Tables 4 and 5 we also computed rank order (Spearman) correlations, crude and energy-adjusted, and obtained results very similar to those provided by Pearson correlations. The Spearman correlations are therefore not shown.

Table 6 examines the concordance between methods from a different point of view, namely that of the correspondence between quartile distribution obtained with each method as compared to the reference. For simplicity, the table reports the proportion of subjects in the lowest or in the highest quartile of intake by the

TABLE 6 *Cross-classification of subjects by quartile of nutrient intake for reference method and methods A or B. The figures indicate the proportion of subjects in the lowest quartile on the reference method who were also classified in the lowest quartile (correctly) or in the highest quartile (incorrectly) by methods A or B (men and women combined)*

Nutrient	Lowest quartile on reference method (%)						Highest quartile on reference method (%)					
	Lowest quartile on A1	Second lowest quartile on A1	Highest quartile on A1	Lowest quartile on B1	Second lowest quartile on B1	Highest quartile on B1	Highest quartile on A1	Second highest quartile on A1	Lowest quartile on A1	Highest quartile on A1	Second highest quartile on B1	Lowest quartile on B1
Energy	42	31	8	50	33	0	70	19	4	72	16	0
Protein	54	31	4	38	46	13	63	22	4	60	20	0
Fat	42	33	4	54	21	4	70	26	4	72	20	4
Saturated	46	31	8	63	17	4	67	30	0	56	36	4
Monounsaturated	46	15	4	63	21	4	67	26	4	72	20	8
Polyunsaturated	46	31	4	58	29	4	59	30	7	60	20	0
Cholesterol	54	19	12	54	21	8	59	39	4	60	16	4
Carbohydrates	58	27	4	50	38	0	56	26	4	68	28	0
Sugar	54	27	0	54	33	8	63	19	7	50	29	8
Alcohol	73	19	4	75	13	4	67	33	0	82	24	0
Retinol	54	27	12	42	38	4	41	30	7	48	24	16
β-carotene	50	31	4	42	29	13	52	26	0	56	16	8
α-tocopherol	50	31	0	71	25	0	67	11	11	56	36	0
Ascorbic acid	46	27	4	54	21	4	63	26	0	46	38	8
Folate	54	27	4	58	25	8	44	22	11	48	36	8
Calcium	65	15	8	58	21	8	56	33	4	56	32	0
Zinc	46	31	4	46	33	8	59	22	4	60	20	4
Selenium	50	31	4	50	21	13	59	19	7	36	36	16
Fibre	54	27	0	58	21	4	63	22	4	60	24	0
Average per cent	53	28		48	23		60	23		60	23	
	81		4	81		6	83		5	83		5

reference method who were correctly classified in the same or in the adjacent quartile by methods A or B, and the proportion of subjects grossly misclassified.

The proportion of subjects correctly classified in the same highest or lowest quartile as by the reference method ranged between 42 and 73% for method A and between 38 and 75% for method B for different nutrients. By adding also the proportion of subjects in the highest and lowest quartiles by the reference method who were classified in the adjacent second highest or second lowest quartile by method A or B, the figures ranged between 66 and 100% for method A, and 71 and 96% for method B. On average, 81–83% of subjects in the lowest or in the highest quartile by the reference method were classified in the lowest or highest quartile or in the adjacent one by methods A or B. The proportion of subjects grossly misclassified in the opposite highest or lowest quartile was on average 4–6%, and was very similar for both methods and both sexes.

Protein intake estimated from urinary nitrogen output was estimated for 63 subjects who completed at least six 24-hour collections. However, results are presented for 53 subjects who had the highest score for completeness of urinary collection. Their average protein intake estimated by urinary nitrogen output was 70.8 g per day, while for the 12 subjects who reported incomplete collection it was 56.4 g per day. This large difference indicates that subjects carefully reported the information on incompleteness of urine collection on the *ad hoc* questionnaire and interview.

The mean protein intakes estimated from the urinary nitrogen by the methods tested (reference, A1, A2, B1 and B2) are reported in Table 7. The mean values from urinary nitrogen and the reference method are very close (difference: +1.2 g), and the Pearson correlation between the protein intakes estimated by the reference method and those derived from urinary nitrogen was 0.75, indicating good concordance, both in mean absolute

TABLE 7 *Mean daily intake of protein estimated from urinary nitrogen (average of 6–8 × 24-hour urines) and from the reference method, and methods A and B in 53 subjects (24 women, 29 men)*

	Women		Men	
	No.	Mean ± SD	No.	Mean ± SD
Protein calculated from urinary nitrogen	24	62.2[a] ± 10.1	29	78.0[a] ± 15.3
Reference (food records)	24	61.0[a] ± 10.2	29	76.8[a] ± 13.3
Method A1	12[b]	75.6[c] ± 13.8	16[b]	95.1[c] ± 29.2
Method B1	12	71.4[c] ± 11.9	13	93.5[c] ± 36.1
Method A2	15	67.3[c] ± 15.2	13	88.7[c] ± 25.1
Method B2	9	74.0[c] ± 12.6	16	89.9[c] ± 19.0

[a] Paired t-test between food record value and protein calculated from urinary nitrogen: $P = 0.60$.
[b] The numbers of subjects indicated refer to subsets of the 24 women and 29 men who completed either method A or method B on different occasions.
[c] Paired t-test versus protein calculated from urinary nitrogen: $P = < 0.05$.

values and subject ranking, as shown by the plot in Figure 2. On the other hand, both methods A and B overestimated protein intake, compared to the nitrogen-derived values, by +8.2% to 23.8% depending on the subgroup of study subjects.

DISCUSSION

The present study was designed to test the validity of two dietary assessment methods developed for a large epidemiological study in Sweden. Both methods were designed with the aim of collecting as much detailed information as possible on a large variety of foods while at the same time avoiding making the methods so cumbersome that subjects' compliance would be too low.

The purpose of dietary measurements within a prospective cohort study on diet and chronic diseases is to estimate the usual intake of foods and nutrients in individual subjects who will then be followed up, and for whom the occurrence of a given disease will be investigated in relation to their diet at baseline.

The main difficulty in evaluating the validity of a dietary questionnaire relates to the very nature of what the questionnaire is intended to measure. In fact, while it would be relatively easy to evaluate the validity of a method designed to measure diet over a short period of time (e.g. a few days), it is much more complex to evaluate the validity of a measurement of usual diet referring to a long period of time such as one or more years. The difficulty is obviously related to the practical absence of a real 'gold standard' which could be used as measurement of reference. In fact, any of the methods normally used as reference in 'validity' studies might themselves be affected by under- or overestimation biases as well as by some degree of random error.

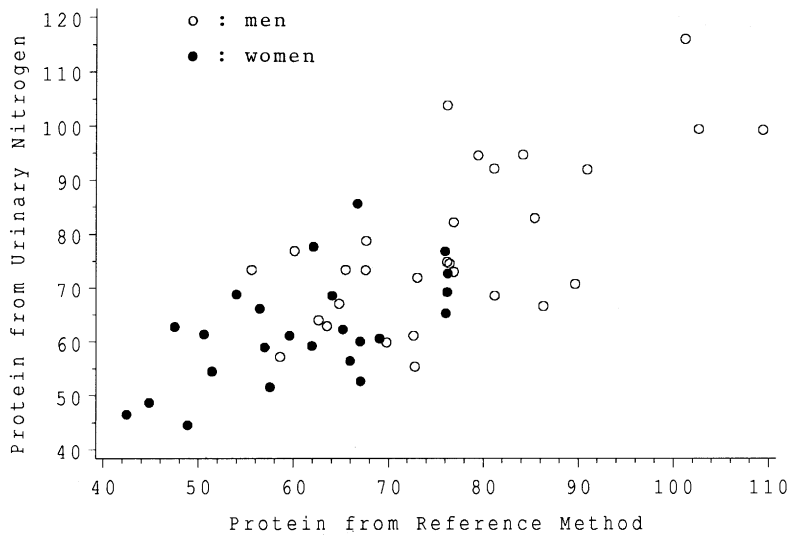

FIGURE 2 *Malmö Food Study—Plot of the mean daily intake of protein estimated in 53 subjects by:*
x = mean urinary nitrogen output measured in 6–8 × 24-hour urine collections per subject
y = mean protein intake estimated from 18 days of weighed food records

In the present study it was decided to use as reference method a series of 18 days of weighed records spread over six periods of 3 days repeated every 2 months. The rationale for this design was to collect dietary data over a sufficient number of days in order to minimize random errors due to day-to-day variations in food intake and, at the same time, to cover seasonal variations throughout a 1-year period. It is important to note that this is the first of the recent generation of 'validation' studies of dietary assessment methods in which, in addition to weighed food records, urine and blood samples were collected from the study subjects on repeated occasions during the study period in order to measure objective biochemical markers of nutrient intake. The results reported on protein provide strong support for the validity of the measurements obtained with 18-day weighed food records in the present study and, at the same time, add value to the relatively good performance observed for the two candidate methods.

The two dietary methods which were compared to the weighed food records were both new and had been designed to cover as wide a range as possible of foods commonly eaten in southern Sweden. The more original of the two was the combined food record/food frequency method (B) in which the information on food consumption is provided by a 'retrospective' quantitative questionnaire addressing relatively 'simple' foods (breakfast, snacks, sandwiches, beverages) while the 14-day food record is intended to capture the consumption of more complex foods such as hot dishes, casseroles and specific recipes.

The results of the analyses at the group level indicated that both methods A and B overestimated the intake of all nutrients with the exception of alcohol. Although weighed food records may underestimate total food intake for the simple reason that study subjects may forget to write down some of the foods they eat, further analyses of our dietary data in comparison to biomarkers, and particularly urinary nitrogen excretion, indicated that in this study the underestimation of the reference method was negligible while the overestimation of both methods A and B was considerable (+9–13 g/day).

Pearson correlation coefficients are the most immediate statistical indicator of the concordance in ranking between two measurements. They have the advantage of not being influenced by systematic over- or underestimation of one method with respect to the other but mainly by the concordance in ranking between methods.

In our study, the correlations found between either method A or B and the reference method are of the order of 0.5–0.6 for most of the nutrients, thus indicating fairly good concordance between the two candidate methods and the reference. These results indicate that the relative validity of the two dietary methods tested in our study rank with the best reported in previous studies.[4,5,12] in which weighed food records were used as reference method. Adjustment for energy did not significantly modify the correlation coefficients for most of the nutrients. The very modest effect of energy adjustment may be explained partly by the fact that the subjects in our study were rather homogeneous with respect to age, and possibly that the dietary assessment methods were relatively extensive and complete in terms of foods and energy-providing nutrients.

Regarding energy-providing nutrients it is worth noting the good correlations found for fat and for most of the fatty acids, particularly for method B. This is particularly important in view of the interest in investigating the role of the fat composition of diet in the aetiology of several types of cancer.

Finally, the analyses of the distribution by quartile indicated that about half of the subjects in the lowest quartile were correctly classified by either method A or B and about 80% of the subjects were correctly classified above or below the median (data not shown).

Overall, the results of our validity study indicate that both the extensive food frequency questionnaire and the new method combining food records plus a short food frequency questionnaire provided correlations with the reference method comparable with the highest observed in similar validation studies. However, method B provided more accurate estimates of absolute intake of several nutrients (including energy-providing nutrients), some fatty acids, vitamins and minerals of particular interest in research on cancer, cardiovascular diseases and diet. Considering also the interest in investigating dietary patterns as well as in adopting an open-ended method for collection of consumption data on specific foods, the investigators' preference went for the combined method. The prospective design of the project influenced this choice, as it gave weight to the value of a dietary method which is not based on predefined food lists and by means of which it is therefore possible to combine and to group specific foods according to interesting scientific hypotheses which may be put forward over the next decade, when it will be possible to harvest the results of current efforts.

ACKNOWLEDGEMENTS

This study was supported by the Swedish Medical Research Council (Grant K84-19X-7010-1) and by the International Agency for Research on Cancer (Collaborative Research Agreement DEB/85/43).

The authors acknowledge the contribution of Ms Eva Callmer who supervised the fieldwork during the data collection phase, and are grateful to the nutritionists who collected the dietary data and to the computer technicians in charge of data management and statistical analyses (Bertrand, Hémon and Geneviève Burnod).

REFERENCES

[1] Block G, Hartman A M. Issues in reproducibility and validity of dietary studies. *Am J Clin Nutr* 1989; **50:** 1133–38.

[2] Willet W. *Nutritional Epidemiology* (Monographs in Epidemiology and Biostatistics, Vol. 15). New York: Oxford University Press, 1990.

[3] 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.

[4] Pietinen P, Hartman A M, Haapa E *et al*. Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. *Am J Epidemiol* 1988; **128:** 667–76.

[5] Rimm E B, Giovannucci E L, Stampfer M J, Colditz G A, Litin L B, Willett W C. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol* 1992; **135:** 1114–26.

[6] Callmer E, Riboli E, Saracci R, Akesson B, Lindgärde F. Dietary assessment methods evaluated in the Malmö Food Study. *J Intern Med* 1992; **233:** 53–57.

[7] Elmståhl S, Gullberg B, Riboli E, Saracci R, Lindgärde F. The reproducibility of a novel diet history method and an extensive food frequency questionnaire. The Malmö Food Study. *Eur J Clin Nutr* 1996; **50:** 134–42.

[8] Statens Livsmedelsverk. *Livsmedelstabeller*. Stockholm: Liber Tryck A B, 1986.

[9] Isaksson B. Urinary nitrogen output as a validity test in dietary surveys. *Am J Clin Nutr* 1980; **33:** 4–12.

[10] Willett W C, Stampfer M J. Total energy intake: Implications for epidemiologic analyses. *Am J Epidemiol* 1986; **124:** 17–27.

[11] Berntorp K, Lindgärde F. Type 2 diabetes heredity and nutrient intake. A dietary history assessment in non-obese normoglycaemic men. *Eur J Clin Nutr* 1989; **43:** 347–53.

[12] Hankin J H, Wilkens L R, Kolonel L N, Yoshizawa C N. Validation of a quantitative diet history method in Hawaii. *Am J Epidemiol* 1991; **133:** 616–28.