Reproducibility and Validity of an Expanded Self-Administered Semiquantitative Food Frequency Questionnaire among Male Health Professionals

Article *in* American Journal of Epidemiology · June 1992

DOI: 10.1093/oxfordjournals.aje.a116211 · Source: PubMed

American Journal of Epidemiology
Vol. 135, No. 10
Printed in U.S.A.

Reproducibility and Validity of an Expanded Self-Administered Semiquantitative Food Frequency Questionnaire among Male Health Professionals

Eric B. Rimm,[1] Edward L. Giovannucci,[2] Meir J. Stampfer,[1,2] Graham A. Colditz,[1,2]
Lisa B. Litin,[1] and Walter C. Willett[1,2,3]

The authors assessed the reproducibility and validity of an expanded 131-item semiquantitative food frequency questionnaire used in a prospective study among 51,529 men. The form was administered by mail twice to a sample of 127 participants at a one-year interval. During this interval, men completed two one-week diet records spaced approximately 6 months apart. Mean values for intake of most nutrients assessed by the two methods were similar. Intraclass correlation coefficients for nutrient intakes assessed by questionnaires one year apart ranged from 0.47 for vitamin E without supplements to 0.80 for vitamin C with supplements. Correlation coefficients between the energy-adjusted nutrient intakes measured by diet records and the second questionnaire (which asked about diet during the year encompassing the diet records) ranged from 0.28 for iron without supplements to 0.86 for vitamin C with supplements (mean r = 0.59). These correlations were higher after adjusting for week-to-week variation in diet record intakes (mean r = 0.65). These data indicate that the expanded semiquantitative food frequency questionnaire is reproducible and provides a useful measure of intake for many nutrients over a one-year period. *Am J Epidemiol* 1992;135:1114–26.

diet; epidemiologic methods; food; nutrition; nutrition surveys; questionnaires

Editor's note: *For a discussion of this paper and for the authors' response, see pages 1127 and 1133, respectively.*

Received for publication August 20, 1990, and in final form July 22, 1991.
[1] Department of Epidemiology, Harvard School of Public Health, Boston, MA.
[2] Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA.
[3] Department of Nutrition, Harvard School of Public Health, Boston, MA.
Reprint requests to Dr. Eric B. Rimm, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115.
Supported by research grant nos. HL 35464 and CA 40935 from the National Institutes of Health. Dr. Rimm was supported by National Institute of Environmental Health Services Research Service Award 5T32 E507069 from the Harvard School of Public Health.
The authors thank Doreen Hurd, Mary Johnson, Laura Sampson, Lisa Lomedico, Steve Stuart, Scott Taber, Susan Wu, Walkyria Paes de Almeida, and Sue-Wei Chiang for their assistance with compiling the data and preparing the manuscript.

Prospective studies of dietary etiologies of chronic disease become more feasible as better techniques for measuring diet are developed. Food frequency questionnaires are designed to measure average long-term diet rather than to provide a precise estimate of short-term intake (1). The low cost and ease of self-administration of a food frequency questionnaire facilitates use in large populations. Since 1979, we have been developing a semiquantitative food frequency questionnaire; an early 61-item version was shown to provide a reasonable measure of dietary intake among female nurses when compared with four one-week diet records (2, 3). A later version with 116 items was found to yield a similar degree of validity for assessing diet 3–4 years in the past (4). We have continued to refine and evaluate this method of dietary assessment.

The purpose of the present study was to assess the validity of nutrient intakes calculated from our expanded 131-item self-administered semiquantitative food frequency questionnaire completed by 51,529 male health professionals aged 40–75 years.

MATERIALS AND METHODS

Population

Men in this study are a subsample drawn from the Health Professionals Follow-up Study, a cohort of 51,529 male health professionals enrolled in a prospective study of dietary etiologies of heart disease and cancer (5, 6). The study population consists of dentists (57.6 percent), veterinarians (19.6 percent), pharmacists (8.1 percent), optometrists (7.3 percent), osteopathic physicians (4.3 percent), and podiatrists (3.1 percent). All cohort members completed a mailed self-administered 131-item semiquantitative food frequency questionnaire (described below) in 1986. During the following year, using the same age distribution from the cohort, a random sample (n = 323) of Boston area cohort members were invited to participate in a dietary validation study.

From this group, 43 (13 percent) could not be contacted after repeated telephone calls, 123 refused (38 percent), and 157 (49 percent) agreed to participate. The men who agreed to participate were asked to collect two one-week diet records over a one-year period, followed by a second semiquantitative food frequency questionnaire (figure 1). From the group of participants, 22 subjects were excluded because they did not have two complete one-week sets of diet records (n = 17) or did not complete a second food frequency questionnaire (n = 5). Subjects were also excluded from analysis if their daily energy intake calculated from either of the two semiquantitative food frequency questionnaires was outside the range of 800 to 4,200 kcal (n = 7) or their questionnaire had more than 70 items left blank (n = 1). These same restrictions for kilocalories and blank food items were applied in analyses of diet and disease in the Health Professionals Follow-up cohort (5, 6). The 127 participants (39 percent) included in this analysis had complete dietary information for both one-week diet records and both semiquantitative food frequency questionnaires.

The expanded semiquantitative food frequency questionnaire

The 131-item semiquantitative food frequency questionnaire is a refined and expanded version of a previously validated questionnaire (2). The original 61-item questionnaire was limited by size constraints and cost of manual data entry. Our expanded version of the questionnaire incorporated open-ended questions to identify specific brands of multiple vitamins, margarines, cooking oils, cold cereals, and other foods consumed at least once per week (4). In addition, we asked more detailed questions about several commonly eaten foods. For example, we asked separately about the frequency of eating cooked spinach (1/2 cup) and raw spinach (as in a salad) because these are typically consumed in different amounts.

The questionnaire was designed to classify individuals according to levels of average daily intake of selected nutrients over the past year. Dietary variables estimated by the questionnaire included calories, total fat, saturated fat, monounsaturated fat, polyun-

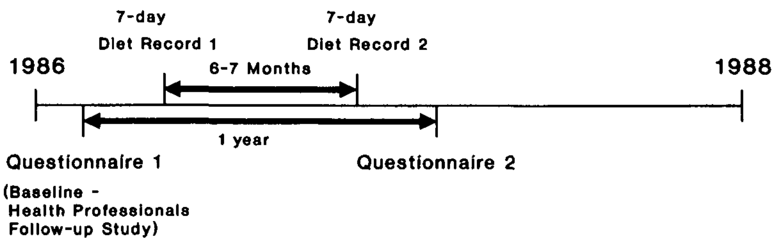

FIGURE 1. Time sequence of diet validation study conducted among 127 male health professionals living in the Boston, Massachusetts area aged 40–70 years.

saturated fat, animal and vegetable fat, total carbohydrate, sucrose, dietary and crude fiber, protein, vitamin A, pre-formed vitamin A, carotene, vitamins B1 (thiamin), B2 (riboflavin), B6, B12, C, D, and E, folate, calcium, phosphorous, magnesium, iron, sodium, potassium, zinc, caffeine, and alcohol. We included food items that explained the largest amount of variance in nutrient intake between subjects during a series of pilot studies (2) and in an analysis of 5,400 days of diet recording (7). In addition, those foods that made a large contribution to total absolute nutrient intake were included. Also, some food items were included because a hypothesis of interest related to a food rather than a nutrient (e.g., cabbage and mushrooms).

For 127 of the 131 food items, a serving size was indicated, using "natural" units (such as one apple, a glass of milk, a slice of bread) whenever possible, or otherwise using typical serving sizes based on analyses of diet records (8) (e.g., 1/2 cup of string beans or 4-6 oz (114-170 g) of meat as a main dish). Estimated portion sizes were used for four foods. Subjects were asked to indicate for each food how often, on average, they had used the amount specified during the past year. Nine multiple choice responses were possible, ranging from never or less than once per month to six or more times per day.

The food composition data base used to calculate nutrient values is based primarily on US Department of Agriculture publications (9, 10) and is continually supplemented by other published sources and personal communications from laboratories and manufacturers.

Nutrient intakes were calculated from the questionnaire by multiplying a weight assigned to the frequency of use (where once per day is equal to one) by the nutrient composition for the portion size specified for each food or vitamin supplement. Nutrients were summed across all foods and vitamins to obtain a total nutrient intake for each individual. Calculations also included specific nutrients from open-ended questions pertaining to regular use of multiple vitamin supplements, cold breakfast cereals, margarines, other foods not previously reported on the questionnaire, and types of fat or oil used for frying, cooking, and baking.

Dietary records

The first week of diet records were completed over a 6-month period and the second week of diet records were completed approximately 6 months after the first, thereby capturing seasonal variability within the two administrations of the questionnaire (figure 1). Each participant was given a dietetic scale for weighing portion sizes. A dietician (L.L.) provided detailed instructions to each participant for weighing and recording all foods consumed. After each participant returned a one-week diet record, the dietician reviewed entries and telephoned the participants to resolve any ambiguities. Diet records were analyzed using CBORD (11), a nutrient software package which uses the ESHA nutrient data base (12) (based primarily on the US Department of Agriculture publications (9, 10). For reported foods that were not in the data base, recipes were obtained from the participants and appropriate amounts of the component foods were entered. The diet record data base provides information on 31 nutrients. Many additional nutrients are computed for the questionnaire, including specific fatty acids, alcohol and caffeine, but are not included in this report because they were not part of the ESHA data base.

Analysis

Means and standard deviations were calculated on total nutrient intakes from both food frequency questionnaires and from the two one-week diet records. Most nutrient distributions were skewed toward higher values, therefore all variables were log (natural) transformed before analysis. Because all nutrients, as measured by the questionnaire, were correlated with energy intake, it is essential to remove the variation due to energy intake and its associated measurement error before assessing the validity of individual nutrients (1, 13). Adjustment for energy is based on a priori biologic considerations that a larger, more physically active person will require a higher caloric intake, which will also be associated with a higher absolute intake of all nutrients. This same absolute intake may have a different effect on a smaller, less active person; therefore, by energy adjusting, we are able to examine the composition of the diet given the individuals energy requirements. Energy adjustment removes some of the variation in diet (see table 5), but is necessary if we plan to use energy adjusted nutrients in studying the dietary etiologies of disease. To add an intuitive sense to the residual, we added a constant, the expected nutrient value at the caloric intake of 2,000 kcal. Pearson correlations were used to compare the two dietary assessment methods for both unadjusted and energy-adjusted nutrients. Because the questionnaire measures diet in the preceding year, correlation coefficients relating the first questionnaire to the diet records will tend to underestimate the validity of the questionnaire. The second questionnaire measures dietary intake over the one-year interval during which the dietary records were collected; however, the process of recording intake could possibly heighten dietary awareness and artificially increase the accuracy of the questionnaire. Using correlations for both questionnaires provides a minimum and maximum estimate of validity. Spearman and Pearson correlation coefficients did not appreciably differ for all nutrients, therefore parametric statistics are presented in this report and non-parametric statistics have been placed with the National Auxiliary Publications Service (NAPS) (see NAPS document no. 04904 for 3 pages of supplementary material. Order from NAPS, c/o Microfiche Publications, P.O. Box 3513, Grand Central Station, New York, NY 10163–3513. Note: remit in advance, in US funds only, $7.75 for photocopies or $4.00 for microfiche). For brevity, coefficients, standard errors, and intercept estimates from regressing dietary intake from the food frequency questionnaires on dietary intake from the dietary records are not presented, but are also available from NAPS.

Ideally, for comparison with the questionnaire, diet records would be maintained for the preceding year because nutrient intakes vary considerably from day to day (14, 15). As a compromise, dictated by cost and compliance, we used average daily nutrient intakes from two different weeks separated by approximately 6 months. Two weeks of diet records are adequate to estimate individual intake for many nutrients (14, 16, 17), although within-person variation in weekly intake can attenuate correlations between the questionnaire and diet records. Therefore, we used the within- and between-person components of variation in diet record intake (treating the two one-week averages as random units of observation) to "deattenuate" (see footnote, table 3) Pearson correlation coefficients (18–20).

The purpose of this report is to quantify measurement error rather than test hypotheses, therefore *p* values are not presented.

RESULTS

We compared the means for average daily nutrient intakes from the 14 days of diet recording and from questionnaires 1 and 2 for the 127 men included in the validation analysis (table 1). The questionnaire measurements of total energy intake were similar (within 7 percent) to intake assessed by diet records. Intake of most nutrients was slightly lower when measured by questionnaire compared with diet records. However, total vitamin A intake was overestimated by questionnaire as were intakes of many vitamins with supplements. This overestimation of intake by the food frequency questionnaire was primarily caused by extreme values for a few individuals. Apart from vitamin supplements, 75 percent of mean values for nutrients (not energy-adjusted) measured by questionnaire were within 15 percent of the diet record values. In addition, average intakes for men who were asked, but did not participate in the validation study, were within 10 percent of men who did participate, with the exception of retinol without supplements (12 percent lower) and iron

TABLE 1. Mean (± standard deviation) absolute daily nutrient intakes estimated by two one-week diet records and from questionnaires completed by 127 Boston area health professionals and 49,999 male health professionals from the United States.

	Diet record (n = 127)	Questionnaire 1 (n = 127)	Questionnaire 2 (n = 127)	HPFS cohort (n = 49,999)
Calories	2,167 ± 462	2,092 ± 594	2,014 ± 610	1,967 ± 620
Total fat (g)	79.8 ± 25.3	70.7 ± 24.9	67.9 ± 23.8	70.7 ± 28.0
Saturated fat (g)	27.5 ± 10.7	24.4 ± 10.1	23.6 ± 9.6	24.8 ± 10.7
Polyunsaturated fat (g)	16.4 ± 5.1	14.1 ± 4.8	13.4 ± 4.7	13.6 ± 5.9
Monounsaturated fat (g)	29.0 ± 9.8	26.0 ± 10.4	24.9 ± 9.3	26.2 ± 11.1
Cholesterol (mg)	342.3 ± 132	305.2 ± 124	287.1 ± 128	322.9 ± 175
Carbohydrates (g)	257.8 ± 69	264.4 ± 103	252.5 ± 96	234.6 ± 88
Protein (g)	92.0 ± 17.8	87.7 ± 24.5	84.9 ± 25.5	87.4 ± 29.6
Dietary fiber (g)	19.2 ± 8.1	24.1 ± 12.2	21.6 ± 7.9	22.6 ± 10.9
Total vitamin A (IU)	10,440 ± 7,668	15,030 ± 9,909	14,069 ± 7,677	15,312 ± 13,147
Without supplements	9,246 ± 7,383	13,635 ± 9,052	12,454 ± 6,719	12,388 ± 11,143
Retinol (IU)	3,261 ± 3,330	3,489 ± 3,447	3,534 ± 3,977	4,265 ± 4,507
Without supplements	2,131 ± 1,494	2,369 ± 2,250	2,274 ± 2,580	2,349 ± 2,069
Carotene (IU)	7,179 ± 7,178	11,265 ± 8,992	10,180 ± 6,342	10,254 ± 11,066
Vitamin B1 (mg)	6.04 ± 14.8	3.27 ± 4.5	3.34 ± 5.4	4.55 ± 9.3
Without supplements	1.82 ± 0.59	1.54 ± 0.54	1.45 ± 0.44	1.43 ± 0.64
Vitamin B2 (mg)	6.20 ± 14.7	3.60 ± 4.3	3.68 ± 5.1	4.94 ± 9.0
Without supplements	2.08 ± 0.68	1.94 ± 0.67	1.86 ± 0.61	1.93 ± 0.81
Vitamin B6 (mg)	6.01 ± 14.3	5.38 ± 16.0	4.06 ± 7.9	8.27 ± 23.2
Without supplements	2.31 ± 0.84	2.25 ± 0.75	2.14 ± 0.72	2.19 ± 0.83
Vitamin B12 (g)	14.1 ± 25.7	12.3 ± 9.0	11.9 ± 8.7	13.2 ± 19.0
Without supplements	7.3 ± 4.10	10.3 ± 7.1	9.89 ± 6.7	9.79 ± 9.6
Folate (g)	383.5 ± 186	407.3 ± 208	407.2 ± 226	450.8 ± 281
Without supplements	324.7 ± 126	349.6 ± 118	336.4 ± 108	333.8 ± 157
Vitamin C (mg)	296.7 ± 491	332.3 ± 362	343.7 ± 418	401.7 ± 448
Without supplements	138.5 ± 65.7	182.0 ± 101	188.1 ± 119	160.5 ± 95.0
Vitamin E (mg)	46.5 ± 96.5	56.7 ± 136.0	48.2 ± 123.0	86.4 ± 184.3
Without supplements	15.0 ± 5.92	10.2 ± 5.8	9.2 ± 3.1	9.5 ± 6.3
Calcium (mg)	814.7 ± 280	888.3 ± 403	838.9 ± 346	923.8 ± 485
Without supplements	796.0 ± 268	818.6 ± 325	804.4 ± 311	831.2 ± 408
Iron (mg)	20.1 ± 7.7	17.2 ± 8.4	17.5 ± 13.7	19.1 ± 13.5
Without supplements	17.9 ± 5.4	15.6 ± 5.9	14.2 ± 4.5	15.0 ± 7.1
Magnesium (mg)	357.4 ± 111	366.7 ± 115	350.9 ± 108	367.9 ± 133
Phosphorous (mg)	1,411 ± 334	1,336 ± 428	1,268 ± 387	1,348 ± 478
Sodium (mg)	3,176 ± 1,255	2,624 ± 1,193	2,677 ± 1,277	2,893 ± 1,372
Potassium (mg)	3,229 ± 816	3,564 ± 1,119	3,424 ± 1,052	3,423 ± 1,170
Zinc (mg)	14.9 ± 8.2	15.5 ± 12.1	14.4 ± 11.1	20.5 ± 21.5

without supplements (13 percent lower) (data not shown).

Although not significantly different, mean daily intake of most vitamins was higher for the whole cohort (n = 49,999 valid questionnaires after exclusions) compared with responses from questionnaires 1 or 2 from people in the validation study. Because daily intakes of vitamins without supplements were almost identical, differences were explained entirely by less supplement use among validation study participants. This is consistent with the 20 percent lower rate of current vitamin use in cohort members from Massachusetts compared with cohort members from the rest of the United States.

We computed intraclass correlations for unadjusted and energy-adjusted daily nutrient intakes between the average intake measured by each of the one-week diet records and between intakes assessed by the two food frequency questionnaires (table 2). The comparisons between unadjusted nutrient intake from each one-week diet record indicate a high degree of reproducibility, ranging from r = 0.50 with vitamin E with-

TABLE 2. Intraclass correlations (*r*) and coefficients of variation* for two one-week diet records and two semiquantitative food frequency questionnaires (FFQ) completed by 127 male health professionals ages 40 to 75 years†

	Record 1 vs Record 2						FFQ 1 vs. FFQ 2	
	Unadjusted			Energy adjusted			Unadjusted	Energy adjusted
	r	CV_b%	CV_w%	*r*	CV_b%	CV_w%	*r*	*r*
Calories	0.77	2.6	1.4				0.65	
Total fat	0.75	6.8	3.9	0.72	4.5	2.7	0.66	0.52
Saturated fat	0.79	11.4	6.0	0.80	8.8	4.4	0.69	0.66
Polyunsaturated fat	0.58	10.0	8.4	0.44	6.6	7.5	0.59	0.38
Monounsaturated fat	0.76	9.5	5.6	0.71	6.7	4.2	0.65	0.56
Cholesterol	0.66	6.0	4.3	0.63	5.2	4.0	0.67	0.60
Carbohydrates	0.79	4.7	2.4	0.81	2.9	1.4	0.63	0.62
Protein	0.64	3.8	2.8	0.61	3.0	2.3	0.64	0.47
Dietary fiber	0.77	12.6	6.9	0.78	8.6	4.6	0.60	0.62
Total vitamin A	0.59	7.2	6.0	0.58	7.1	6.0	0.62	0.60
Without supplements	0.56	8.8	7.8	0.46	5.5	6.0	0.63	0.60
Retinol	0.69	11.2	7.4	0.69	10.9	7.3	0.57	0.56
Without supplements	0.49	7.5	7.6	0.44	6.6	7.5	0.56	0.55
Carotene	0.56	10.0	8.7	0.56	9.8	8.7	0.66	0.65
Vitamin B1	0.89	92.9	33.0	0.88	92.3	34.3	0.70	0.71
Without supplements	0.74	49.2	29.2	0.65	36.6	26.9	0.61	0.46
Vitamin B2	0.90	79.5	26.6	0.89	78.1	27.4	0.68	0.69
Without supplements	0.74	40.0	23.6	0.69	30.1	19.8	0.68	0.62
Vitamin B6	0.87	70.8	27.4	0.86	70.3	28.2	0.71	0.72
Without supplements	0.71	40.3	26.0	0.66	34.4	24.8	0.67	0.67
Vitamin B12	0.67	30.8	21.7	0.67	30.6	21.8	0.65	0.61
Without supplements	0.43	18.3	21.0	0.41	17.4	20.9	0.66	0.61
Folate	0.74	7.0	4.2	0.73	6.8	4.1	0.57	0.54
Without supplements	0.73	5.9	3.5	0.67	4.9	3.4	0.65	0.63
Vitamin C	0.79	15.3	7.8	0.79	15.2	7.8	0.80	0.79
Without supplements	0.67	15.8	11.0	0.63	7.8	5.9	0.69	0.64
Vitamin E	0.83	29.8	13.3	0.83	29.5	13.6	0.72	0.73
Without supplements	0.50	12.0	12.0	0.44	9.7	10.9	0.47	0.38
Calcium	0.68	4.5	3.0	0.62	3.3	2.6	0.64	0.54
Without supplements	0.67	4.3	3.1	0.58	3.2	2.6	0.70	0.59
Iron	0.74	10.5	6.2	0.75	9.8	5.6	0.58	0.57
Without supplements	0.70	8.6	5.6	0.65	6.4	4.8	0.54	0.35
Magnesium	0.75	4.4	2.6	0.75	3.7	2.1	0.69	0.72
Phosphorous	0.72	2.9	1.8	0.70	2.0	1.3	0.67	0.57
Sodium	0.62	3.6	2.8	0.50	2.5	2.5	0.72	0.53
Potassium	0.75	2.8	1.7	0.65	1.9	1.4	0.65	0.60
Zinc	0.74	13.7	8.0	0.77	13.7	7.4	0.65	0.64

* The coefficient of variation (CV%) is the standard deviation/mean × 100. The standard deviation is calculated using ANOVA to separate the between- and within-person variance components.

† All data are $\log_e$ transformed.

out supplements to $r = 0.90$ for vitamin B2. The intraclass correlations measuring the reproducibility of the unadjusted nutrients from the semiquantitative food frequency questionnaire spaced one-year apart range from 0.47 for vitamin E without supplements to 0.80 for vitamin C with supplements. Adjustment of the nutrient intake for total energy intake before testing reproducibility did not appreciably alter the correlations.

Table 2 also contains the coefficients of variation for both the between- and within-person variation using the average daily intake from each one-week diet record. The ratio of the within- to between-person variation illustrates the degree to which total variation in measured dietary intake from diet records is accounted for by within-subject (week-to-week) variation compared with between-subject variation. For example, using one week as the unit of measurement, within-person variation is only 33 percent of between-person variation for unadjusted total fat intake (using the square of the coefficients of variation from table 2). For other nutrients, such as polyunsaturated fat and vitamin B12 without supplements, within-person variation accounts for a much larger proportion of total variation. These fluctuations, which can be due to day-to-day or seasonal factors, cause measurement error in the diet record, the standard for comparison in this study. The effects of this error, which may attenuate correlations between the semiquantitative food frequency questionnaire and the diet records, can be reduced by increasing the number of dietary record measurements or by statistically reducing the bias in correlation coefficients due to within-person variability (see below).

To assess the validity of the questionnaire, we compared unadjusted and energy-adjusted nutrient intakes estimated from each administration of the questionnaires with the 14-day average from diet records (table 3). Correlations between questionnaire 1 and the 14-day average from the diet records were slightly lower than correlations comparing questionnaire 2 with the diet records. Questionnaire 2 provided reasonably good correlations for energy-adjusted saturated fat (0.71) and cholesterol (0.67), and moderate correlations for energy-adjusted total vitamin A with supplements (0.41) and protein (0.38). For vitamins B1, B6, and E (with supplements), Spearman correlations were about 20 percent lower than Pearson correlations (data not shown). However, Spearman and Pearson correlations were similar for all other nutrients. Partial correlations between nutrients calculated from the food frequency questionnaire and from diet records, controlling for age and body mass index, were almost identical to correlations presented in table 3. For both questionnaires, correlations for nutrients from individual weeks (records 1 and 2) were similar, indicating little effect of the proximity in time to diet recording (data not shown). The columns labeled "de-attenuated" in table 3 list the correlations between nutrient intake from questionnaires and the average daily nutrient intake from the diet records after correcting for the effects of within-person variation (20). The de-attenuated correlation coefficients between questionnaire 2 and diet records range from 0.37 for polyunsaturated fat to 0.92 for vitamin C with supplements (mean = 0.65). The last column of table 3 contains regression coefficients (b) derived from models when the questionnaire is used to predict intake estimated by diet records. These coefficients can be used to correct measures of bivariate association when the food frequency questionnaire is used in etiologic studies (21, 22). The corresponding standard errors and intercept terms are also available from NAPS, as noted above.

Comparisons of extreme quintiles are often used to examine risk of disease. Table 4 summarizes the joint classifications for nutrient intake using quintiles from the questionnaire by nutrient intake using quintiles from 14 days of diet records. For both questionnaire and diet records, separate quintile cutpoints were established from their respective distributions of nutrient intake. Using energy-adjusted saturated fat as an example, 88 percent of subjects in the lowest quintile of intake from the diet records were in the bottom two quintiles of saturated fat as calculated from the questionnaire. We observed no gross misclassification of subjects from the first quintile of saturated fat intake measured by diet records to the fifth quintile of intake measured by the questionnaire. In those subjects whose intake was in the highest quintile, only one person (4 percent) was grossly misclassified by the questionnaire into the lowest quintile of intake. Thus, only minimal misclassification exists between extreme quintiles of intake. On average, for each nutrient, only 4 percent (1 person) were grossly misclassified into extreme quintiles.

Data from the diet records can also be used to estimate the absolute levels of intake defined by quintiles of the food frequency

TABLE 3. Pearson correlation (*r*) and regression coefficients (*b*) between semiquantitative food frequency questionnaires and the average of two one-week diet records* calculated for unadjusted and energy-adjusted nutrients†

Nutrient	Questionnaire 1			Questionnaire 2			Regression coefficient§ *b*
	Unadjusted *r*	Adjusted *r*	De-attenuated‡ *r*	Unadjusted *r*	Adjusted *r*	De-attenuated‡ *r*	
Calories	0.27		0.29	0.40		0.43	0.28
Total fat	0.42	0.48	0.53	0.52	0.61	0.67	0.60
Saturated fat	0.52	0.59	0.63	0.63	0.71	0.75	0.77
Polyunsaturated fat	0.29	0.30	0.38	0.33	0.29	0.37	0.27
Monounsaturated fat	0.46	0.56	0.61	0.53	0.62	0.68	0.61
Cholesterol	0.57	0.59	0.67	0.62	0.67	0.76	0.63
Carbohydrates	0.40	0.62	0.65	0.48	0.69	0.73	0.63
Protein	0.22	0.28	0.32	0.25	0.38	0.44	0.21
Dietary fiber	0.45	0.59	0.63	0.49	0.64	0.68	0.86
Total vitamin A	0.45	0.52	0.61	0.35	0.41	0.48	0.61
Without supplements	0.49	0.56	0.71	0.41	0.48	0.61	0.46
Retinol	0.52	0.53	0.59	0.67	0.68	0.75	0.62
Without supplements	0.46	0.47	0.60	0.44	0.43	0.55	0.34
Carotene	0.50	0.55	0.65	0.48	0.54	0.64	0.62
Vitamin B1	0.66	0.67	0.69	0.81	0.83	0.86	0.97
Without supplements	0.34	0.47	0.53	0.35	0.53	0.60	0.68
Vitamin B2	0.67	0.68	0.70	0.83	0.85	0.88	1.03
Without supplements	0.45	0.60	0.66	0.46	0.48	0.53	0.46
Vitamin B6	0.77	0.78	0.81	0.80	0.82	0.85	1.02
Without supplements	0.52	0.61	0.68	0.52	0.65	0.73	0.85
Vitamin B12	0.37	0.37	0.41	0.49	0.50	0.56	0.63
Without supplements	0.29	0.27	0.35	0.40	0.40	0.52	0.32
Folate	0.45	0.48	0.52	0.61	0.71	0.77	0.78
Without supplements	0.47	0.56	0.62	0.50	0.63	0.70	0.84
Vitamin C	0.72	0.74	0.79	0.84	0.86	0.92	0.94
Without supplements	0.50	0.57	0.65	0.64	0.68	0.77	0.66
Vitamin E	0.74	0.74	0.78	0.86	0.87	0.92	0.73
Without supplements	0.34	0.40	0.49	0.28	0.31	0.42	0.44
Calcium	0.38	0.42	0.48	0.51	0.53	0.61	0.42
Without supplements	0.44	0.45	0.51	0.52	0.53	0.60	0.46
Iron	0.32	0.38	0.41	0.42	0.50	0.54	0.37
Without supplements	0.28	0.41	0.46	0.28	0.28	0.32	0.30
Magnesium	0.49	0.62	0.67	0.50	0.66	0.71	0.73
Phosphorous	0.45	0.57	0.63	0.43	0.57	0.63	0.55
Sodium	0.40	0.39	0.48	0.44	0.49	0.60	0.42
Potassium	0.41	0.56	0.63	0.49	0.65	0.73	0.62
Zinc	0.50	0.59	0.63	0.56	0.66	0.71	0.60

* All data are $\log_e$ transformed.

† The energy-adjusted correlations between dietary methods use the residuals from regressing each nutrient on the total calories as measured by the semiquantitative food frequency questionnaire or diet records.

‡ The de-attenuated correlation coefficient is calculated using the ratio of the within- to between-person variance (table 2) measured from the weekly averages for the 2 weeks of diet records. The formula for this corrected correlation is calculated as: $\rho_c = \rho_o\sqrt{1 + [(\sigma_w^2/\sigma_b^2)/n]}$, where ρ_o is the observed correlation between the energy-adjusted nutrients (except for energy itself) from the food frequency questionnaire and diet records, σ_w^2 is the within-person variation, and σ_b^2 is the between-person variation and n is the number of replicate measurements. For our calculations, $n = 2$ represent each week of diet records.

§ The regression coefficient (*b*) is calculated from regressing the average of energy-adjusted nutrient intake from two one-week diet records on the energy-adjusted nutrient calculated from the second food frequency questionnaire.

questionnaire. In table 5, each nutrient has three columns representing the dietary method used to define the quintiles. Column 1 represents the "true" intake measured by the two one-week diet records, column 2 represents the intake measured directly by the semiquantitative food frequency questionnaire, and column 3 is the "true" intake (from the diet records) using quintiles defined by the food frequency questionnaire.

TABLE 4. Comparison of nutrient intakes from the second semiquantitative food frequency questionnaire with the mean of two one-week diet records based on cross-classification of quintiles, calculated from energy-adjusted nutrient intake

	Lowest quintile* on diet record			Highest quintile* on diet record		
	Lowest quintile† on questionnaire (%)	Lowest 2 quintiles† on questionnaire (%)	Highest quintile† on questionnaire (%)	Highest quintile† on questionnaire (%)	Highest 2 quintiles† on questionnaire (%)	Lowest quintile† on questionnaire (%)
Calories	28	56	0	36	60	12
Total fat	60	80	4	40	68	8
Saturated fat	60	88	0	52	76	4
Polyunsaturated fat	32	64	8	28	52	12
Monounsaturated fat	56	80	4	48	68	0
Cholesterol	56	88	4	64	84	4
Carbohydrates	64	76	0	52	76	0
Protein	44	64	4	32	60	12
Dietary fiber	36	68	0	56	80	8
Total vitamin A	32	56	12	56	80	4
Without supplements	40	60	4	40	68	12
Retinol	64	84	4	60	84	0
Without supplements	48	68	8	44	64	8
Carotene	40	60	4	44	76	12
Vitamin B1	52	64	0	80	88	1
Without supplements	52	72	0	44	60	0
Vitamin B2	56	80	0	76	84	0
Without supplements	48	76	0	44	72	0
Vitamin B6	52	76	0	76	88	4
Without supplements	56	84	0	48	88	0
Vitamin B12	40	60	8	52	92	0
Without supplements	52	68	12	40	72	0
Folate	56	88	0	52	76	0
Without supplements	48	76	0	40	76	0
Vitamin C	48	80	0	62	80	4
Without supplements	56	80	0	52	76	0
Vitamin E	60	80	4	40	68	8
Without supplements	44	68	8	32	40	16
Calcium	48	72	4	44	68	0
Without supplements	36	68	4	48	76	8
Iron	44	60	0	48	62	8
Without supplements	32	48	8	36	56	0
Magnesium	48	72	4	44	68	4
Phosphorous	52	80	0	56	68	8
Sodium	36	60	8	52	80	4
Potassium	52	68	0	52	60	0
Zinc	32	52	0	52	68	12

* Quintiles defined using the nutrient distribution from the 14 days of diet records.
† Quintiles defined using the nutrient distribution from the second food frequency questionnaire.

Thus, column 3 represents the measurable variation in nutrient intake after accounting for the error from the questionnaire. As an example, total fat intake measured by the two one-week diet records ranged from a median of 51 g in the lowest quintile to a median of 112 g of fat in the highest quintile. When we used the semiquantitative food frequency questionnaire to measure total fat intake in the same population, the range was from a median of 40 g in the lowest quintile to 97 g in the highest quintile. However, column 3 illustrates that, because of regression toward the mean due to measurement error in the questionnaire, the actual fat intake in categories defined by the questionnaire ranges from an average of 65 g in the lowest quintile to 94 g in the highest quintile.

TABLE 5. Median unadjusted and energy-adjusted nutrient intake for quintiles derived from two one-week diet records, the second semiquantitative food frequency questionnaire, and diet records using quintiles derived from the second semiquantitative food frequency questionnaire (data from 127 men aged 40–75 years)

Quintile	Diet record	Question-naire	Diet record using questionnaire quintiles	Diet record	Question-naire	Diet record using questionnaire quintiles	Diet record	Question-naire	Diet record using questionnaire quintiles
	Total fat (g)			% of calories from fat			Saturated fat (g)		
Unadjusted									
1	51	40	65	24.2	23.2	27.9	16	14	20
5	112	97	94	40.0	37.5	37.4	40	33	35
Energy adjusted									
1	57	51	64				17	17	18
5	94	83	87				34	30	33
	Polyunsaturated fat (g)			Cholesterol (mg)			Vitamin C with supplements (mg)		
Unadjusted									
1	11	8	13	203	156	264	72	96	89
5	23	19	18	505	471	426	709	902	644
Energy adjusted									
1	11	10	14	202	179	243	75	104	96
5	20	18	18	482	413	440	720	913	709

The variation between persons was much greater for intakes of cholesterol and vitamin C with supplements. The ranges of energy-adjusted nutrients listed in table 5 are slightly attenuated because variation due solely to calories has been removed.

In this study, we excluded improperly completed forms defined primarily by extreme values of total energy intake rather than on the number of blank food items. Informal examination of questionnaires from the total cohort suggested that forms with a high number of blanks appeared plausible and carefully completed. However, to determine whether the number of blanks was associated with the degree of measurement error, we calculated the correlations between the number of blank responses on the second food frequency questionnaire with the error, defined as the absolute value of nutrient intake measured by diet records minus nutrient intake measured by the food frequency questionnaire. The number of blanks was not significantly correlated with the error for any of the nutrients, and we found no appreciable differences between correlation coefficients from analyses with and without 26 men who had more than two food items blank on their questionnaire. Finally, the errors, as defined above, were not correlated with age, body mass index, or alcohol consumption ($r < 0.08$ for the average error over all macronutrients (saturated, monounsaturated, and polyunsaturated fat, protein, and carbohydrate), vitamins without supplements, minerals, and cholesterol and fiber).

DISCUSSION

We compared individual nutrient intakes estimated by a 131-item food frequency questionnaire with intake calculated from two one-week diet records collected approximately 6 months apart in a population of male health professionals. Correlations, after correcting for within-person variation assessed by dietary records, averaged 0.60 for macronutrients. As we observed previously, correlations between the two methods (2) were somewhat higher for the energy-adjusted nutrients (mean correlation = 0.54 for macronutrients) than for raw nutrients (mean correlation = 0.44 for the same nutrients). This results because the reduction of correlated measurement error for total calories and the macronutrients exceeds the reduction in between-person variation for nutrient intake caused by controlling for total energy intake (1).

Results from this study could be biased if men who agreed to participate in the validation study completed the first food frequency questionnaire differentially from subjects who were contacted, but refused to participate. However, for questionnaire 1 (completed by all men), the median number of blanks were quite low and almost identical between the validation study participants (median = 1) and nonparticipants (median = 2). In addition, the percent of calories from fat did not significantly differ between the participants (30.5 percent) and the nonparticipants (31.3 percent) in the validation study. Therefore, reasons for nonparticipation do not appear to be associated with major differences in diet or with the ability to complete a questionnaire.

Other validation studies of food frequency questionnaires (2, 4, 23–29) that have used diet records or 24-hour recalls for a comparison method have been reviewed elsewhere (1), with the exception of a few recent publications (30, 31). Differences in study populations and in the magnitude of the between-person variation in diets among these populations preclude formal comparisons of questionnaire validity across studies. Nevertheless, the correlations we observed in this study were, in general, as high or higher than those reported earlier, including those using questionnaires that were much longer (25), that were administered by interviewers (30), or that included pictures of foods to help estimate portion sizes (24). In particular, the average correlation coefficient for energy-adjusted nutrient intakes was 0.55 in our previous comparison of a 61-item self-administered food frequency questionnaire with 4 weeks of diet records among 173 women (2); for the same nutrients in this study, the mean de-attenuated correlation coefficient was 0.67. Although we did not directly compare interviewer administration with our self-administered, mailed format, the relatively high correlations seen in this study do not support the suggestion by Sobell et al. (15) that mailed dietary questionnaires may provide a low degree of validity.

In assessing the validity of a dietary measurement technique, an absolute standard does not exist, therefore it is desirable to compare two methods with uncorrelated errors to reduce the possibility of artificially inflating correlation coefficients (1). The food frequency questionnaire relies primarily on a person's ability to recall usual frequency of intake over the past year, whereas the diet record depends on a person's ability (and willingness) to weigh and record current diet rather than relying on memory. Completing two one-week diet records could have influenced responses to the second food frequency questionnaire, however, it is unlikely that the baseline food frequency questionnaire, which requires approximately 20–30 minutes to complete, would have influenced actual dietary consumption months later. Correlations comparing nutrients from the diet records with nutrients from the second questionnaire were only slightly higher than correlations for nutrients from the first questionnaire. This small difference may reflect some learning bias, but, alternatively, some difference would be expected because the second questionnaire represents the time period during which the diet records were collected.

The lower correlation between the food frequency questionnaire and diet records for polyunsaturated fat appears to be due equally to error in our standard (diet records) as to error with the food frequency questionnaire. Hunter et al. (32) compared, in our same population, polyunsaturated fatty acid intake calculated from two one-week dietary records and from the food frequency questionnaire with polyunsaturated fatty acid concentrations from subcutaneous fat aspirates. Correlations comparing the two dietary methods to actual polyunsaturated fat stores were very similar, thus indicating that both methods have similar degrees of error for this nutrient.

In addition to assessing the validity of the food frequency questionnaire, the results from a validation study can also be used to "calibrate" a questionnaire used in a larger cohort study to correct observed diet-disease associations for measurement error. Using the slope from regressing average nutrient intake from diet records on nutrients from a food frequency questionnaire, Rosner et al. (22) have described a linear approximation method to correct observed relative risk estimates. For example, if the observed relative risk in a cohort study of saturated fat and disease was 1.7, the corrected relative risk estimate would be 2.0, using, from the validation study, the regression coefficient of 0.77 for saturated fat (table 3) to compensate for the error from the food frequency questionnaire. A formula for a corrected confidence interval has also been reported (22).

The findings of this validation study can probably be directly generalized to well-educated adult US men. However, because the form is quite simple and earlier versions have performed reasonably in a variety of populations (1), it is likely that the validity would not differ greatly in other general US populations. Ideally, the expanded food frequency questionnaire should be validated for use in populations which differ substantially from the men presented in this study.

In summary, the 131-item semiquantitative food frequency questionnaire used in this population of male health professionals provided reasonably reproducible measurements for individuals over a one-year period and mean values for most nutrients were close to those obtained by dietary records. Although we did not directly compare this questionnaire with earlier, shorter versions, we observed a trend toward higher correlations with intakes measured by diet records than we found in previous validation studies. The degree of variation in nutrient intake and validity of the questionnaire indicate that important associations between diet and disease can be reasonably quantified in the larger cohort of men.

REFERENCES

1. Willett W. Nutritional epidemiology. New York: Oxford University Press, 1990.
2. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51-65.
3. Salvini S, Hunter DJ, Sampson L, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol 1989;18:858-67.
4. Willett WC, Sampson L, Browne ML, et al. The use of a self-administered questionnaire to assess diet four years in the past. Am J Epidemiol 1988; 127:188-99.
5. Grobbee DE, Rimm EB, Giovannucci E, et al. Coffee, caffeine, and cardiovascular disease in men. N Engl J Med 1990;323:1026-32.
6. Rimm EB, Giovannucci E, Willett WC, et al. Prospective study of alcohol consumption and risk of coronary disease in men. Lancet 1991;2:464-8.
7. Stryker WS, Salvini S, Stampfer MJ, et al. Contributions of specific foods to absolute intake and between-person variation of nutrient consumption. J Am Diet Assoc 1991;91:172-8.
8. Hunter DJ, Sampson L, Stampfer MJ, et al. Variability in portion sizes of commonly consumed foods among a population of women in the United States. Am J Epidemiol 1988;127:1240-9.
9. Adams CF. Nutritive value of American foods. Handbook no. 456. Washington, DC: US Department of Agriculture, 1975.
10. Consumer and Food Economics Institute. Composition of foods: Handbook 8 series. Washington, DC: US Department of Agriculture, 1976-1989.
11. The CBORD Group Inc. Version 3.0.3. Ithaca, NY: The CBORD Group Inc., 1988.
12. ESHA Research Database. Salem, OR: ESHA, 1988.
13. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. Am J Epidemiol 1986;124:17-27.
14. Beaton GH, 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.
15. Sobell J, Block G, Koslowe P, et al. Validation of a retrospective questionnaire assessing diet 10-15 years ago. Am J Epidemiol 1989;130:173-87.
16. Potosky AL, Block G, Hartman AM. The apparent validity of diet questionnaires is influenced by number of diet-record days used for comparison. J Am Diet Assoc 1990;90:810-13.
17. Nelson M, Black AE, Morris JA, et al. Between- and within-subject variation in nutrient intake from infancy to old age: estimating the number of days required to rank dietary intakes with desired precision. Am J Clin Nutr 1989;50:155-67.
18. Beaton GH, 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.
19. Liu K, Stamler J, Dyer A, et al. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. J Chronic Dis 1978;31:399-418.
20. Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hy-

pothesis testing. Am J Epidemiol 1988;127: 377–86.
21. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. Am J Epidemiol 1990;132:734–45.
22. Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. Stat Med 1989;8:1051–69.
23. Pietinen P, Hartman AM, 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;128:655–66.
24. Jain MG, Harrison L, Howe GR, et al. Evaluation of a self-administered dietary questionnaire for use in a cohort study. Am J Clin Nutr 1982;36:931–5.
25. Pietinen P, Hartman AM, Haapa E, et al. Reproducibility and validity of dietary assessment instruments. II. A qualitative food frequency questionnaire. Am J Epidemiol 1988;128:667–76.
26. Balogh M, Medalie JH, Smith H, et al. The development of a dietary questionnaire for an ischemic heart disease survey. Isr J Med Sci 1968;4: 195–203.
27. Hankin JH, Nomura AMY, Murphy NJ, et al. Reproducibility ofa dietary history questionnaire in a case-control study of breast cancer. Am J Clin Nutr 1983;37:981–5.
28. Stuff JE, Garza C, Smith EO, et al. A comparison of dietary methods in nutritional studies. Am J Clin Nutr 1983;37:300–6.
29. Hunt IF, Luke LS, Murphy NJ, et al. Nutrient estimates for computerized questionnaires vs. 24-hr. recall interviews. J Am Diet Assoc 1979;74: 656–9.
30. Block G, Hartman AM, Naughton D. A reduced dietary questionnaire: development and validation. Epidemiology 1990;1:58–64.
31. Larkin FA, Metzner HL, Thompson FE, et al. Comparison of estimated nutrient intakes by food frequency and dietary records in adults. J Am Diet Assoc 1989;89:215–23.
32. Hunter DJ, Rimm EB, Sacks FM, et al. Comparison of measures of fatty acid intake by subcutaneous fat aspirate, food frequency questionnaire, and diet records in a free-living population of US men. Am J Epidemiol 1992;135:418–27.