



Calibration of the Dietary Questionnaire for a Multiethnic Cohort in Hawaii and Los Angeles

Daniel O. Stram,¹ Jean H. Hankin,² Lynne R. Wilkens,² Malcolm C. Pike,¹ Kristine R. Monroe,¹ Sohee Park,¹ Brian E. Henderson,¹ Abraham M. Y. Nomura,² Maj E. Earle,² Faye S. Nagamine,² and Laurence N. Kolonel²

The performance of the dietary questionnaire used in a multiethnic cohort study in Hawaii and Los Angeles was assessed in a calibration substudy that compared diet reported from the questionnaire with three 24-hour dietary recalls. For the calibration substudy, subjects from each of eight subgroups defined by sex and ethnic group (African-American, Japanese-American, Latino, and White) were chosen randomly from among the cohort members, and each participant's previous day's diet was assessed by telephone recall on three occasions over approximately 2 months. After completing the three 24-hour recalls, each calibration subject was sent a second questionnaire; 1,606 persons completed three recalls and a second questionnaire (127 to 267 per ethnic-sex group). This report describes correlation coefficients and calibration slopes for the relation between the 24-hour recalls and second questionnaire values for a selected set of macro- and micronutrients, as absolute intakes, nutrient densities, and calorie-adjusted nutrients. In all subgroups, estimates of the correlation between the questionnaire and 24-hour recalls were greater after energy adjustment (average correlations ranged from 0.57–0.74 for nutrient densities and from 0.55–0.74 for calorie-adjusted nutrients) than when absolute nutrient values were used (average range 0.26–0.57). For absolute nutrient intakes, the correlations were greatest for Whites, somewhat lower for Japanese-Americans and Latinos, and lowest for African-Americans. After energy adjustment, the difference between subgroups were diminished, and the correlations were generally highly satisfactory. *Am J Epidemiol* 2000;151:358–70.

calibration; diet surveys; epidemiologic methods; ethnic groups; nutrition surveys; questionnaires

Disease risk estimates derived from cohort or case-control studies of diet are generally smaller than are corresponding risk estimates from ecologic analyses (1, 2). However, while ecologic studies suffer from inadequate control of confounders and other well-known limitations, estimates of risk within cohorts can be substantially attenuated by measurement errors in assessing individual diets (3). Measurement error is particularly a problem in cohort studies that include subgroups for whom the questionnaire may perform differently, because this complicates inter-group comparisons. A calibration substudy, in which data from diet history questionnaires are compared with information from a second source that is

assumed to provide an unbiased estimate, allows for correction of risk estimates for measurement error. We report here on the results of a calibration substudy for a recently established cohort study that includes several different ethnic groups. The two principal purposes of conducting the calibration study were: 1) to provide information that will ultimately be used for correction of risk estimates obtained from analysis of nutritional risk factors as cancer incidence is observed prospectively in this new multiethnic cohort; and 2) to evaluate the performance of the questionnaire, by correlational analysis, for each ethnic group under study.

The Hawaii-Los Angeles Multiethnic Cohort Study of diet and cancer is based on the prospective follow-up of over 215,000 adults who were aged 45–75 years at enrollment and who completed a mailed survey instrument that included a quantitative food frequency questionnaire (QFFQ) designed to assess typical food intake of individual males and females in five principal ethnic groups—African-Americans, Japanese-Americans, Latinos, Native Hawaiians, and Whites—and who lived in Hawaii and Los Angeles, California. This paper includes four of the five groups. Native Hawaiians were added to the cohort

Received for publication April 7, 1998, and accepted for publication December 8, 1998.

Abbreviations: BMI, body mass index; EPIC, European Prospective Investigation into Cancer and Nutrition; QFFQ, quantitative food frequency questionnaire.

¹ Department of Preventive Medicine, School of Medicine, University of Southern California, Los Angeles, CA.

² Cancer Etiology Program, Cancer Research Center of Hawaii, University of Hawaii, Honolulu, HI.

Reprint requests to Dr. Daniel O. Stram, Department of Preventive Medicine, University of Southern California School of Medicine, CHP-220, Los Angeles, CA 90033-9987.

after the other groups; hence, their calibration data are not as yet complete. The cohort includes substantial numbers of members from each ethnic group. This adds power to the statistical assessment of the relation between diet and cancer, by augmenting the variability of individual dietary habits within ethnic group with the variability between ethnic groups in mean consumption of nutrients and particular food items. Primary issues arising in the use of this cohort to assess cancer risk include the comparability of risk estimates across ethnic groups and the extent to which between-ethnic group differences in cancer rates are explained by the between-ethnic group differences in diet.

The standard approach to the validation and calibration of a dietary questionnaire compares the questionnaire intakes with daily intakes for a representative sample of subjects. In this calibration study, we compared the QFFQ intakes with intakes from three 24-hour telephone recalls of the previous day's diet. The statistical method used in this comparison is that of regression-calibration (3–5), in which the average of the 24-hour recall intakes is modeled as a function of the QFFQ intake.

The correlation between nutrient intakes from the 24-hour recalls and nutrients derived from the QFFQ is used as a standard measure of validity of the questionnaire. Under the assumption that each of the 24-hour recalls provides an unbiased estimate of a single day's diet, a regression-calibration equation can be used as a prediction equation for imputing an estimated true consumption given the QFFQ value. These imputed values, when used in an analysis of the relation between disease and nutrient intake, can eliminate (or at least substantially reduce) certain types of measurement error biases that would otherwise affect the risk estimates (3–5).

The QFFQ we used was developed for use in the multiethnic cohort and includes many ethnic food items. Of particular interest in this study is the comparison of the performance of the QFFQ in the comparatively unstudied populations of African-Americans, Japanese-Americans, and Latinos compared with its performance in Whites. Because we anticipated that our QFFQ might perform differently by ethnic group and sex, we designed our calibration study with enough subjects from each ethnic-sex group to provide well-estimated correlations and regression-calibration equations for each ethnic-sex subcohort. The general criteria used for determining the number of subjects was that the variability in the calibration slope estimate should not add importantly to the variance of corrected risk estimates ultimately obtained from the main study.

MATERIALS AND METHODS

Design of the QFFQ

Our objective was to produce a self-administered diet history applicable across the ethnic groups. Particular attention was paid to wording and format to ensure that the QFFQ could be understood and completed by persons with less than a high school education.

We developed the QFFQ as a modification of the face-to-face interview method used for many years by our group in Hawaii. In this method, usual frequencies and amounts of foods consumed during the past year are obtained, using photographs showing three portion sizes to estimate quantitative intakes of most foods. To develop the self-administered modification of this method, we initially collected 3-day measured food records from about 60 men and women, aged 45–75 years, from each ethnic group in Hawaii and Los Angeles. These records were utilized to identify food items for inclusion into the questionnaire. For each ethnic group, the percent contribution of the individual foods was computed for each dietary component of major interest (e.g., fat, dietary fiber, vitamin A, carotenoids, and vitamin C). The foods were compiled in a list sorted in order of importance, and the first set of foods contributing at least 85 percent of the nutrient intake was selected. Because foods that met this criteria for any ethnic group were included, the questionnaire actually accounts for much more than 85 percent of the intakes of the major nutrients. Lastly, specific food items uniquely associated with the traditional diets of a particular group irrespective of their contribution to intake were included (e.g., ham hocks for African-Americans, tofu and salted fish for Japanese-Americans, and tamales for Latinos). Similar foods with comparable nutrient composition were grouped into food items.

We chose a single questionnaire for all ethnic groups rather than separate questionnaires for three reasons: 1) The instrument we have used for many years in face-to-face interviews, from which the present questionnaire is adapted, was developed for and validated in multiethnic populations. 2) Although ethnic diets are distinct, many ethnic foods are eaten sufficiently often by all groups in our population to necessitate their inclusion in all of the questionnaires; thus, although frequency of consumption and quantity consumed would differ by ethnic group, the food items listed in the different questionnaires would have a great deal of overlap, thereby reducing the value of creating ethnic-specific instruments. 3) We did not know with certainty the ethnicity of many subjects when we mailed the questionnaire to them, and we would have risked having an individual respond to an

inappropriate questionnaire if the QFFQ was ethnic-specific.

The questionnaire includes eight frequency categories for foods and nine for beverages to permit adequate specificity in defining daily intakes. For food items, the highest frequency category is ≥ 2 times/day, whereas, for beverages, the highest category is ≥ 4 times/day.

The 3-day measured food records noted above revealed wide variation in amounts consumed within and between ethnic groups. These differences were utilized to identify three typical serving sizes for each category of food, so that the respondent could choose the size that most often typified his/her usual serving. Because it is difficult to visualize a particular serving size for some items (e.g., 3 oz (85 g) of beef), we printed photographs in the questionnaire illustrating three serving sizes of selected food groups. We also collected information on additions to breads, coffee, and tea, and on intakes of condiments. Modifications to nutrients for selected items were made based on the subject's use of various fats and oils in cooking and on his/her usual practice of eating the fat on meats and the skin on chicken.

Calibration substudy

It was considered likely that the QFFQ would assess dietary intake differently in each ethnic-sex group. The calibration study was therefore designed to include subjects from each of the ethnic-sex groups, and in essence a separate calibration study was performed for each of these subgroups.

For each ethnic-sex subgroup, a random sample of approximately 260 subjects was targeted for inclusion in the calibration study, based on lists of cohort members. An initial 24-hour dietary recall (of the previous day's diet) was performed after telephone contact had been established and then two more recalls were taken on randomly selected days of the week approximately one month apart. Participating subjects were not informed in advance when the 24-hour dietary recall was to occur. Interviewers were registered dietitians, two working at each location (Hawaii and Los Angeles) who were trained in eliciting 24-hour dietary information. The interviewers were instructed to give an assessment of the quality of each 24-hour recall at the time that it was performed, based on their impressions of how forthcoming the subject had been during the interview.

Calibration substudy second questionnaire

An additional QFFQ identical to the initial questionnaire was mailed to the calibration study subjects

approximately 4–6 weeks after the three 24-hour recalls were completed. This second questionnaire covers the reference period of the 24-hour recalls. In addition, the calibration substudy required several years to complete, making the time between administration of the initial questionnaire and the 24-hour recalls vary greatly across calibration substudy participants. Therefore, in this analysis, nutrients derived from the second QFFQ were compared with the corresponding 24-hour recall values, and subjects who did not return the second QFFQ were excluded.

Food composition table

The same basic food composition data were used to compute nutrients from the 24-hour recalls and the QFFQ. However, the food composition data were modified to account for the grouped foods in the QFFQ. We prepared a customized, and in part ethnic-specific, food composition table for the cohort QFFQ. To develop this database, we utilized the three 24-hour recalls from the first 1,362 participants in the calibration study. Individual foods from the recalls and their frequencies were classified according to the listing of the food items in the QFFQ. Several items, such as coleslaw and cauliflower, were listed as single foods on the QFFQ. However, other QFFQ items were composites, e.g., "pizza" included many different types. Final nutrient composition values for each composite food item were weighted averages of the nutrients for the component foods, where the weights were determined by the frequencies of consumption reported in the recalls.

The questionnaire food consumption database has separate records for meats trimmed of all visible fat and for chicken with skin removed. Based on a cohort member's response as to their usual practice in regard to fat/skin removal, the appropriate food composition value was used. Similarly, the respondent's usual preference for a specific type of oil (e.g., sunflower), butter, or margarine (e.g., whipped) was used to assign the appropriate nutrients to these items. A few composite food items in the questionnaire were assigned different weighting of the nutrients depending on ethnicity. For instance, the item "corn tortillas, corn muffins, or cornbread" was assigned a different nutrient composition for Latinos, where corn tortillas were given a higher weight (based on the 24-hour recall data). The food composition data were based primarily on US Department of Agriculture values (6, 7), with supplementation from laboratory analyses (unpublished data) and other research and commercial publications (8–10). In addition to the values of energy and macro- and micronutrients, we included values for carotenoids (11), tocopherols, isoflavonoids, and indi-

vidual fiber components (unpublished data). The food composition database includes over 1,500 foods and 700 recipes, which were developed from the food records and 24-hour recalls, as well as from established cookbooks.

Statistical methods

Correlation between QFFQ-derived intake and true intake. The standard assumption behind the use of 24-hour recalls to assess the correlations between the QFFQ-based intakes and true nutrient intakes, and to make predictions of true intakes based on the QFFQ intakes, is that each 24-hour recall provides an unbiased estimate of the true consumption of a nutrient (or of some aspect of diet, such as percent calories from fat). For a given nutrient, let x_i indicate the true long-term intake of subject i , and let z_{ij} be his/her intake estimated from the j th 24-hour recall ($j = 1, 2, 3$). We use the standard method for calculating the correlation, R_{xQ} , between x_i and the nutrient value from the QFFQ, Q_i (3–5). Letting \bar{z}_i be the average of the three 24-hour recall nutrient values, we first calculate the correlation, R_{zQ} , between \bar{z}_i and R_{xQ} . R_{xQ} is then estimated from R_{zQ} by adjusting for the fact that \bar{z}_i is based on only three recall values rather than a large number. The adjustment formula is

$$R_{xQ} = R_{zQ} \sqrt{1 + (\sigma_w^2/m)/\sigma_b^2},$$

where σ_w^2 is the within-person variance of the 24-hour recalls, σ_b^2 is the between-person variance of the 24-hour recalls, and m is the number of recalls per person, here 3. The quantities σ_b^2 and σ_w^2 are estimated from a simple variance components analysis of the within- and between-person variability of the 24-hour recalls. The method can be straightforwardly generalized to cover the situation where each QFFQ does not have the same number of associated 24-hour recalls. The R_{xQ} correlations are referred to as “corrected correlations”.

In our correlational analyses, we estimate separate correlations for each of the eight ethnic-sex subgroups. These analyses were carried out for each nutrient by considering the nutrients in absolute terms, as nutrient densities (nutrient/total calories – alcohol), and as nutrient values adjusted for calories based on the method of residuals (3).

Prediction of true consumption. We have adopted the standard approach to predicting true nutrient intake, x_i , from Q_i (3–5). Writing the mean of the x_i 's for a specific Q_i as $(Q_i|Q_i)$, we assume that

$$E(x_i|Q_i) = a + bQ_i,$$

and estimate a and b by ordinary regression of \bar{z}_i on Q_i . In this situation, the slope term, b , is key to measurement error correction of the relations between disease outcome and QFFQ estimates of nutrient intake in the entire cohort. In its simplest form, the error correction of a log relative risk estimate derived from the main cohort analysis of disease and nutrient intake from the QFFQ is accomplished by dividing the estimate by b (4, 5).

We estimated separate intercept, a , and slope, b , parameters for each of the ethnic-sex subgroups. Only the slope parameters are needed to combine estimates of nutrient effect on disease risk across subgroups, but the intercept parameters are required in addition for the comparison of true intakes between subgroups. We performed additional analyses to investigate whether the slope, b , depends on other factors beyond sex and ethnic group, such as age, educational level, and body mass index (BMI). Such analyses of the calibration study data may be important in later work with the cohort, in order to assess the comparability of risk estimates across these variables as well.

Performance characteristics of the calibration study

Consider the contribution to the variance of an error-corrected log relative risk estimate due to sampling error in our estimate of slope in the calibration study. In its simplest form, the corrected log relative risk estimate may be written as $\hat{\beta}/\hat{b}$, where $\hat{\beta}$ is the log of the uncorrected relative risk estimate and \hat{b} is the estimate of the calibration slope b . For a calibration study with a very large number of subjects, there is no error in the estimation of b so that the variance of the corrected estimate is then

$$\text{Var}(\hat{\beta})/b^2. \quad (1)$$

The ratio of the variance of the corrected estimate based on a finite sized (denoted n) calibration study to this value is

$$\frac{\text{Var}(\hat{\beta}/\hat{b}_n)}{\text{Var}(\hat{\beta})/b^2}. \quad (2)$$

Under independence of $\hat{\beta}$ and \hat{b}_n , equation 1 can be written, after a little algebra, as

$$b^2 E(1/\hat{b}_n)^2 \left(1 + \left\{ 1 + \frac{E(\hat{\beta})^2}{\text{Var}(\hat{\beta})} \right\} \frac{\text{Var}(1/\hat{b}_n)}{E(1/\hat{b}_n)^2} \right). \quad (3)$$

If we approximate $E(1/\hat{b}_n)$ as $1/b$ and $\text{Var}(1/\hat{b}_n)$ as $\text{Var}(\hat{b}_n)/b^4$, then we see that equation 2 is approximately equal to

$$1 + c \times [\text{SE}(\hat{b}_n)/b]^2 \quad (4)$$

where $c = 1 + [E(\hat{\beta})/\text{SE}(\hat{\beta})]^2$. The square root of equation 4 approximates the inflation in the size of confidence intervals for the true log relative risk which is due to correction using only a finite number of subjects in the calibration study. Note that the ratio of expectation to standard error, i.e., $E(\hat{\beta})/\text{SE}(\hat{\beta})$, determines the power of the main study to detect a nonzero log relative risk, using the uncorrected estimate. Based on standard sample size considerations, for a main study that is designed to have well specified type I and II error properties,

$$E(\hat{\beta})/\text{SE}(\hat{\beta}) = (Z_{1-\alpha/2} + Z_{1-\beta})^2.$$

Thus, for a study with 80 percent power to reject the null hypothesis at the 95 percent level of significance, c must be set to $1 + (1.96 + 0.84)^2 = 8.84$.

Note also that the estimate of $b/\text{SE}(\hat{b}_n)$ (the inverse of which appears in equation 4) is the usual t statistic for testing for a nonzero calibration slope estimate. If

$t > 6.5$ in the calibration study, then for a reasonably sized main study (one with ~80 percent power to detect an effect), from equation 4 the inflation in the confidence interval for true log relative risk due to errors in the calibration study estimate will be less than 10 percent. If $t = 9.3$, then the inflation factor is 5 percent, and if $t = 4.5$, the inflation is 20 percent. We report below the range of values of t observed in the calibration study for the eight different sex-ethnic group combinations, and use this statistic as a performance criteria for judging the adequacy of the sample size in the calibration.

RESULTS

Calibration study members completed their initial QFFQs over the period from May 1993 to January 1996. The 24-hour recalls were collected from March 1994 to June 1997, and the second questionnaires were collected from May 1994 to June 1997. Table 1 gives participation rates in the calibration substudy. Because the 24-hour recalls were done by phone, subjects had to be reachable by this mechanism. Our main losses to participation were caused by cohort members whose phone numbers were unavailable. Of those we were able to contact, the majority (range 76–90 percent) agreed to participate and completed all three recalls. However, many of the subjects failed to return the sec-

TABLE 1. Participation in the calibration study: multiethnic cohort in Hawaii and Los Angeles, 1993–1996

	African-Americans		Japanese-Americans		Latinos		Whites	
	Male	Female	Male	Female	Male	Female	Male	Female
No. selected	378	397	395	374	333	316	423	427
Not able to be contacted								
Deceased	16	10	7	1	8	5	5	4
No phone number available*	53	45	27	36	61	53	34	44
Relocated and unable to locate	15	2	3	1	11	14	12	11
Total	84	57	37	38	80	72	51	59
Able to be contacted								
Refused or too ill	32	37	64	49	25	30	31	28
Other reasons†	15	10	15	7	5	4	12	2
Dropped out after 1 or 2 recalls	11	11	7	11	9	10	5	6
Completed all 3 recalls	236	282	272	269	214	200	324	332
Total	294	340	358	336	253	244	372	368
Returned second QFFQ‡								
No	153	199	228	225	140	127	267	267
(%§)	(65)	(70)	(84)	(84)	(65)	(52)	(72)	(80)
No. in analysis¶	151	186	224	222	136	123	264	264

* Phone number was not supplied by respondent and could not be found in other sources.

† No reply to repeated phone calls and other reasons.

‡ QFFQ, quantitative food frequency questionnaire.

§ % of those completing 3 recalls.

¶ Subjects were excluded if total energy intake from the second QFFQ fell outside range of 500–8,000 Kcal or if more than one 24-hour recall was rated as unreliable.

ond QFFQ, with response rates ranging from 52 percent to 84 percent. Approximately 95 percent of 24-hour recalls were assessed as being reliable. No unreliable recalls were used in analysis. Additionally, persons were excluded if they had fewer than two reliable recalls or if their calorie intakes from the second QFFQ were outside the range of 500–8,000 kcal per day.

Although a substantial proportion of the originally selected subjects were not included in the calibration analyses, the final sample was reasonably representative of the entire cohort. Table 2 gives background information (age, body weight, BMI, education, and smoking status) on the subjects retained in the analysis compared with all cohort members, by sex and ethnicity. The calibration subjects who completed the second

QFFQ and who were retained in the analysis well mirror the full cohort for age, weight, BMI, and smoking status. However, they are somewhat better educated than the full cohort. This implies that it may be especially important, here and in future work, to investigate whether the performance of the questionnaire in assessing diet varies by educational attainment.

The average 24-hour recall estimates of total energy intake were 1,811 kcal/day for males and 1,428 kcal/day for females averaged over all the ethnic groups. While these estimates are low even in comparison with what may be expected in sedentary older populations (12), the impact of underreporting on the 24-hour recalls on calibration slopes and correlations is presently unknown (see Discussion).

TABLE 2. Background characteristics of calibration subjects completing the second quantitative food frequency questionnaire (QFFQ) and multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

	African-Americans		Japanese-Americans		Latinos		Whites	
	Male	Female	Male	Female	Male	Female	Male	Female
No. in calibration	151	186	224	222	136	123	264	264
No. in cohort	12,851	22,256	26,964	29,957	22,818	24,620	22,857	26,502
Age (years) at time of first QFFQ, mean (SD*)								
Calibration	63.1 (8.3)	61.3 (8.9)	61.8 (9.1)	60.0 (8.7)	62.6 (7.4)	61.3 (7.8)	59.8 (8.8)	59.2 (9.1)
Cohort	62.0 (8.5)	60.8 (8.8)	60.7 (9.1)	60.4 (8.9)	59.9 (7.8)	58.9 (7.7)	58.8 (9.1)	58.6 (9.0)
Weight (kg), mean (SD)								
Calibration	85.5 (16.4)	78.8 (18.2)	70.9 (10.8)	57.0 (10.5)	80.9 (15.4)	71.0 (16.5)	83.1 (13.5)	70.1 (16.8)
Cohort	85.7 (15.6)	77.9 (17.7)	71.0 (11.2)	56.9 (10.2)	80.6 (18.1)	70.2 (14.4)	83.7 (14.6)	68.6 (15.5)
BMI (kg/m ²), mean (SD)								
Calibration	27.3 (4.2)	29.4 (6.4)	25.0 (3.2)	23.6 (4.0)	26.5 (4.0)	28.1 (6.2)	26.1 (3.8)	26.0 (6.1)
Cohort	27.2 (4.5)	29.1 (6.4)	25.2 (3.4)	23.6 (4.0)	27.4 (4.3)	28.0 (5.7)	26.6 (4.2)	25.8 (5.6)
Education by frequency (%)								
Calibration								
<11 years	11	11	7	6	22	33	4	4
11–12 years	25	26	27	26	24	29	14	24
Some college†	36	43	31	27	36	26	30	32
College graduates	28	20	35	40	18	11	52	40
Cohort								
<11 years	17	14	9	11	44	48	7	7
11–12 years	26	28	29	32	21	24	16	24
Some college†	35	36	30	29	23	19	29	33
College graduates	22	22	32	28	12	8	48	36
Smoking status by frequency (%)								
Calibration								
Never	21	45	36	67	36	70	35	46
Previous smoker	54	36	51	23	48	21	50	42
Current smoker	25	19	13	10	16	9	16	12
Cohort								
Nonsmoker	25	46	30	69	33	66	32	45
Previous smoker	48	34	56	22	49	24	51	38
Current smoker	27	20	15	9	18	10	17	17

* SD, standard deviation.

† Including vocational school.

Corrected correlations and regression slopes of mean 24-hour recalls on QFFQ

Tables 3 and 4 give corrected correlations for the regression of mean 24-hour intakes on the QFFQ intakes by ethnic-sex group for 10 selected dietary

TABLE 3. Corrected correlations between 24-hour recalls and the second quantitative food frequency questionnaire (QFFQ) for selected nutrients* among male multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

Nutrient by unit	African-Americans	Japanese-Americans	Latinos	Whites
Total calories (minus alcohol) (Kcal)	0.16	0.34	0.33	0.48
Protein				
g†	0.17	0.31	0.27	0.51
% calories‡	0.44	0.46	0.35	0.43
Kcal-adjusted g§	0.43	0.46	0.32	0.44
Fat				
g†	0.29	0.41	0.33	0.57
% calories‡	0.60	0.68	0.70	0.69
Kcal-adjusted g§	0.59	0.70	0.75	0.69
Saturated fat				
g†	0.35	0.49	0.33	0.61
% calories‡	0.71	0.71	0.68	0.72
Kcal-adjusted g§	0.71	0.71	0.75	0.71
Carbohydrate				
g†	0.20	0.43	0.43	0.49
% calories‡	0.47	0.54	0.57	0.68
Kcal-adjusted g§	0.47	0.56	0.59	0.68
Calcium				
mg†	0.25	0.47	0.34	0.52
mg/1,000 Kcal¶	0.65	0.72	0.46	0.65
Kcal-adjusted mg§	0.64	0.71	0.48	0.64
Vitamin A				
IU†	0.30	0.45	0.62	0.59
IU/1,000 Kcal¶	0.61	0.62	0.79	0.66
Kcal-adjusted IU§	0.58	0.61	0.80	0.64
Vitamin C				
mg†	0.41	0.67	0.48	0.63
mg/1,000 Kcal¶	0.64	0.74	0.65	0.72
Kcal-adjusted mg§	0.66	0.76	0.65	0.71
Beta-carotene				
mcg†	0.40	0.50	0.60	0.64
mcg/1,000 Kcal¶	0.67	0.63	0.68	0.70
Kcal-adjusted mcg§	0.65	0.63	0.67	0.68
Dietary fiber				
g†	0.36	0.58	0.47	0.59
g/1,000 Kcal¶	0.69	0.79	0.68	0.71
Kcal-adjusted g§	0.69	0.78	0.68	0.71
Average correlations over all nutrients				
Amount	0.30	0.48	0.43	0.57
Density	0.61	0.66	0.62	0.66
Kcal-adjusted	0.60	0.66	0.63	0.66

* All correlations were based on log transformed nutrients and were corrected to account for the fact that there were only three recall values rather than a large number.

† Correlation for absolute nutrient intake.

‡ Correlation for % calories from the nutrient.

§ Correlation for calorie-adjusted nutrient values.

¶ Correlation for density of nutrient per 1,000 Kcal of calorie intake.

components. Nutrient densities generally have better correlations than do absolute values of nutrients. Nutrient densities and calorie-adjusted nutrient values give very similar results, as would be expected.

The subgroups with the highest average correlations for absolute intakes in tables 3 and 4 are White

TABLE 4. Corrected correlations between 24-hour recalls and the second quantitative food frequency questionnaire (QFFQ) for selected nutrients* among female multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

Nutrient by unit	African-Americans	Japanese-Americans	Latinos	Whites
Total calories (minus alcohol) (Kcal)	0.17	0.19	0.40	0.28
Protein				
g†	0.22	0.25	0.35	0.38
% calories‡	0.61	0.39	0.36	0.65
Kcal-adjusted g§	0.61	0.39	0.40	0.67
Fat				
g†	0.24	0.32	0.57	0.39
% calories‡	0.62	0.62	0.52	0.76
Kcal-adjusted g§	0.64	0.60	0.46	0.77
Saturated fat				
g†	0.24	0.48	0.54	0.45
% calories‡	0.54	0.78	0.65	0.75
Kcal-adjusted g§	0.56	0.76	0.63	0.76
Carbohydrate				
g†	0.26	0.30	0.29	0.37
% calories‡	0.63	0.54	0.34	0.80
Kcal-adjusted g§	0.64	0.54	0.31	0.80
Calcium				
mg†	0.33	0.59	0.34	0.49
mg/1,000 Kcal¶	0.76	0.68	0.55	0.69
Kcal-adjusted mg§	0.76	0.66	0.54	0.69
Vitamin A				
IU†	0.22	0.49	0.52	0.58
IU/1,000 Kcal¶	0.52	0.54	0.75	0.79
Kcal-adjusted IU§	0.51	0.53	0.73	0.77
Vitamin C				
mg†	0.38	0.55	0.30	0.66
mg/1,000 Kcal¶	0.50	0.64	0.63	0.72
Kcal-adjusted mg§	0.50	0.67	0.55	0.73
Beta-carotene				
mcg†	0.16	0.48	0.49	0.58
mcg/1,000 Kcal¶	0.49	0.60	0.62	0.74
Kcal-adjusted mcg§	0.45	0.61	0.61	0.71
Dietary fiber				
g†	0.34	0.56	0.40	0.46
g/1,000 Kcal¶	0.70	0.72	0.72	0.75
Kcal-adjusted g§	0.68	0.69	0.70	0.73
Average correlations over all nutrients				
Amount	0.26	0.45	0.42	0.48
Density	0.60	0.61	0.57	0.74
Kcal-adjusted	0.59	0.61	0.55	0.74

* All correlations were based on log transformed nutrients and were corrected to account for the fact that there were only three recall values rather than a large number.

† Correlation for absolute nutrient intake.

‡ Correlation for % calories from the nutrient.

§ Correlation for calorie-adjusted nutrient values.

¶ Correlation for density of nutrient per 1,000 Kcal of calorie intake.

males and females (respective R 's = 0.57 and 0.48); the subgroups with the lowest average correlations are African-American males and females (respective R 's = 0.30 and 0.26). For nutrient densities, all groups have relatively high average correlations, varying from a low of 0.57 in Latino females to a high of 0.74 in White females.

Tables 5 and 6 show the slope estimates from the calibration equations for males and females, respectively. The extent to which these slopes differ from 1.0 reflects the amount of adjustment needed to "correct" the observed QFFQ nutrient intakes for measurement

TABLE 5. Regression slopes between average of 24-hour recalls and the second quantitative food frequency questionnaire (QFFQ) for selected nutrients* among male multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

Nutrient by unit	African-Americans	Japanese-Americans	Latinos	Whites
Total calories (minus alcohol) (Kcal)	0.09	0.23	0.19	0.32
Protein				
g†	0.10	0.18	0.15	0.31
% calories‡	0.40	0.36	0.28	0.36
Kcal-adjusted g§	0.39	0.34	0.25	0.34
Fat				
g†	0.17	0.29	0.24	0.39
% calories‡	0.58	0.62	0.57	0.53
Kcal-adjusted g§	0.57	0.65	0.58	0.53
Saturated fat				
g†	0.24	0.37	0.25	0.46
% calories‡	0.69	0.64	0.56	0.64
Kcal-adjusted g§	0.69	0.65	0.57	0.64
Carbohydrate				
g†	0.12	0.34	0.24	0.37
% calories‡	0.42	0.51	0.46	0.54
Kcal-adjusted g§	0.41	0.53	0.46	0.54
Calcium				
mg†	0.19	0.40	0.27	0.43
mg/1,000 Kcal¶	0.67	0.77	0.50	0.69
Kcal-adjusted mg§	0.66	0.76	0.52	0.68
Vitamin A				
IU†	0.22	0.29	0.46	0.51
IU/1,000 Kcal¶	0.55	0.41	0.64	0.63
Kcal-adjusted IU§	0.52	0.40	0.65	0.61
Vitamin C				
mg†	0.32	0.69	0.43	0.64
mg/1,000 Kcal¶	0.60	0.78	0.70	0.84
Kcal-adjusted mg§	0.61	0.80	0.70	0.82
Beta-carotene				
mcg†	0.36	0.33	0.48	0.59
mcg/1,000 Kcal¶	0.75	0.46	0.66	0.70
Kcal-adjusted mcg§	0.72	0.45	0.66	0.68
Dietary fiber				
g†	0.25	0.47	0.32	0.55
g/1,000 Kcal¶	0.70	0.73	0.68	0.80
Kcal-adjusted g§	0.70	0.72	0.68	0.80

* All nutrient amounts were log transformed.

† Correlation for absolute nutrient intake.

‡ Correlation for % calories from the nutrient.

§ Correlation for calorie-adjusted nutrient values.

¶ Correlation for density of nutrient per 1,000 Kcal of calorie intake.

error. For example, based on the results in table 6, African-American and Latino women will require the greatest adjustment to their QFFQ fat density values.

Table 7 gives the overall between-ethnic group significance test results for differences in the slope estimates for nutrient densities. Significant differences ($p < 0.05$) in the regression slopes of the calibration equations are found between ethnic groups for total calories and vitamin A in males, where the slopes are lower for African-Americans and Japanese-Americans, respectively. For females, significant differences were found for protein, fat, saturated fat, and

TABLE 6. Regression slopes between average of 24-hour recalls and the second quantitative food frequency questionnaire (QFFQ) for selected nutrients* among female multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

Nutrient by unit	African-Americans	Japanese-Americans	Latinos	Whites
Total calories (minus alcohol) (Kcal)	0.10	0.10	0.24	0.18
Protein				
g†	0.13	0.14	0.19	0.23
% calories‡	0.45	0.27	0.27	0.56
Kcal-adjusted g§	0.45	0.27	0.28	0.56
Fat				
g†	0.14	0.23	0.36	0.29
% calories‡	0.35	0.55	0.39	0.69
Kcal-adjusted g§	0.37	0.53	0.34	0.68
Saturated fat				
g†	0.15	0.37	0.36	0.39
% calories‡	0.40	0.69	0.51	0.78
Kcal-adjusted g§	0.41	0.65	0.49	0.76
Carbohydrate				
g†	0.16	0.18	0.17	0.28
% calories‡	0.42	0.54	0.22	0.71
Kcal-adjusted g§	0.43	0.54	0.20	0.73
Calcium				
mg†	0.21	0.44	0.29	0.46
mg/1,000 Kcal¶	0.52	0.66	0.64	0.69
Kcal-adjusted mg§	0.52	0.64	0.63	0.68
Vitamin A				
IU†	0.21	0.35	0.31	0.41
IU/1,000 Kcal¶	0.53	0.41	0.43	0.53
Kcal-adjusted IU§	0.52	0.41	0.42	0.52
Vitamin C				
mg†	0.43	0.61	0.27	0.58
mg/1,000 Kcal¶	0.63	0.75	0.61	0.72
Kcal-adjusted mg§	0.63	0.76	0.54	0.72
Beta-carotene				
mcg†	0.18	0.32	0.39	0.44
mcg/1,000 Kcal¶	0.60	0.42	0.50	0.57
Kcal-adjusted mcg§	0.56	0.42	0.50	0.55
Dietary fiber				
g†	0.26	0.36	0.28	0.37
g/1,000 Kcal¶	0.71	0.63	0.65	0.75
Kcal-adjusted g§	0.70	0.60	0.63	0.73

* All nutrient amounts were log transformed.

† Correlation for absolute nutrient intake.

‡ Correlation for % calories from the nutrient.

§ Correlation for calorie-adjusted nutrient values.

¶ Correlation for density of nutrient per 1,000 Kcal of calorie intake.

TABLE 7. Tests of significance of differences in slope by ethnic group* among multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

Nutrient	Males	Females
Total calories (minus alcohol)	0.007	0.19
Protein (% calories)	0.72	0.02
Fat (% calories)	0.83	<0.001
Saturated fat (% calories)	0.63	<0.001
Carbohydrate (% calories)	0.72	<0.001
Calcium (mg/1,000 Kcal)	0.16	0.24
Vitamin A (mg/1,000 Kcal)	0.03	0.70
Vitamin C (mg/1,000 Kcal)	0.32	0.68
Beta carotene (mg/1,000 Kcal)	0.05	0.34
Dietary fiber (g/1,000 Kcal)	0.65	0.69

* p values from tests of homogeneity of the slope estimates in table 4.

carbohydrate, generally resulting from the higher slopes in Whites.

Performance characteristics of the study

The values for $t = b/SE(\hat{b})$, ranged from 1.6 to 12.6, for the nutrients and sex-ethnic group combina-

tions considered in the previous tables, with a median value of 5.7. The median value corresponds to an inflation in confidence intervals for the true log relative risk, attributable to variability in the calibration slope estimates, of 13 percent. The estimates of the performance criteria, t , showed patterns similar to the results for the corrected correlations. In general, the t values were smaller for total calories (median 3.8) and absolute nutrients (median 4.4) than for nutrient densities (median 6.8).

Additional analyses

As an illustration of further analyses that may be required for a full discussion of questionnaire performance and between-ethnic group comparisons, tables 8 and 9 expand the calibration model for two nutrients: percent of calories from fat, and density of vitamin C. A significant relation between educational level and the slope of the calibration equation for fat density was seen in males, with slopes generally increasing with years of education. However, the relation was not monotonic; for example, slopes for White men in the

TABLE 8. Expanded calibration model* for male multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

	Fat (% calories)			Vitamin C (mg/1,000 Kcal)		
	Estimate	SE†	p value‡	Estimate	SE	p value‡
Baseline slope	0.183			0.630		
Education						
<11 years	-0.191	0.107		0.091	0.150	
11–12 years	-0.048	0.084		-0.076	0.104	
Some college§	0.324	0.123		0.120	0.153	
College graduates¶	0	–	0.007	0	–	0.58
Ethnicity						
African-American	0.106	0.101		-0.226	0.116	
Japanese-American	0.067	0.080		-0.060	0.104	
Latino	0.115	0.104		-0.036	0.132	
White¶	0	–	0.61	0	–	0.26
Body mass index (kg/m ²)						
≤23	0.300	0.094		0.041	0.122	
23.01–26	0.193	0.092		0.281	0.114	
26.01–29	0.241	0.102		0.295	0.117	
>29¶	0	–	0.013	0	–	0.012
Age (years)						
<55	0.130	0.093		-0.008	0.118	
55–62	0.046	0.091		-0.103	0.127	
63–68	0.221	0.094		-0.126	0.120	
≥69¶	0	–	0.076	0	–	0.60

* Values shown are parameter estimates with all covariates in the model. Corrected values of R^2 were 0.77 for % fat and 0.72 for vitamin C density.

† SE, standard error.

‡ p values from tests of homogeneity of the slope estimates.

§ Including vocational school.

¶ Reference group.

TABLE 9. Expanded calibration model* for female multiethnic cohort subjects, Hawaii and Los Angeles, 1993–1996

	Fat (% calories)			Vitamin C (mg/1,000 Kcal)		
	Estimate	SE†	<i>p</i> value‡	Estimate	SE	<i>p</i> value‡
Baseline slope	0.719			0.411		
Education						
<11 years	-0.021	0.104		-0.237	0.136	
11–12 years	-0.050	0.076		-0.008	0.090	
Some college§	-0.108	0.132		-0.090	0.162	
College graduates¶	0	–	0.82	0	–	0.33
Ethnicity						
African-American	-0.350	0.087		-0.044	0.105	
Japanese-American	-0.181	0.087		0.043	0.104	
Latino	-0.328	0.106		-0.027	0.122	
White¶	0	–	<0.001	0	–	0.89
Body mass index (kg/m ²)						
≤23	0.040	0.092		0.054	0.107	
23.01–26	-0.005	0.094		-0.074	0.117	
26.01–29	0.024	0.107		0.073	0.123	
>29¶	0	–	0.94	0	–	0.58
Age (years)						
<55	-0.169	0.095		0.309	0.124	
55–62	-0.022	0.096		0.191	0.122	
63–68	0.006	0.100		0.263	0.130	
≥69¶	0	–	0.14	0	–	0.08

* Values shown are parameter estimates with all covariates in the model. Corrected values of R^2 were 0.73 for % fat and 0.55 for vitamin C density.

† SE, standard error.

‡ *p* values from tests of homogeneity of the slope estimates.

§ Including vocational school.

¶ Reference group.

highest age and BMI groups with <11 years of education, 11–12 years, some college, and college degrees were, respectively, -0.008 (= 0.183 + - 0.191), 0.135, 0.507, and 0.183. The difference between ethnic groups in the slopes for fat density for women remained significant after adjustment for the other variables.

The calibration slopes for both nutrients in both sexes were larger among subjects with BMI ≤ 29 (kg/m²) than among more obese subjects, and the difference was significant in males. This finding indicates that any future analyses of the joint relation between nutrients and BMI and risk of specific cancers may need to control carefully for differences in the performance of the questionnaire by BMI group.

DISCUSSION

Our correlation measures compare favorably, and those for nutrient densities compare very favorably, with those of other investigators and with our previous work. In the Nurses' Health Study (13), values ranged

from 0.26 for vitamin A without supplements to 0.73 for total vitamin C (unadjusted for energy), based on 28 days of food records for each subject. We report (table 5) for White females correlations that ranged from 0.38 for protein to 0.66 for vitamin C. For energy-adjusted nutrient intake among the nurses, Willett (3) reported correlations from 0.47 to 0.59 for protein, fat, saturated fat, and carbohydrate, compared with our range among White females of 0.65 to 0.59. The correlations for protein and fat were 0.33 and 0.39. Pietinen et al. (14) reported the following correlations: total fat, 0.60; vitamin A, 0.51; and vitamin C, 0.59. We report a wider range of correlations here; however this may be expected given that we are reporting values for a total of eight different ethnic-sex groups. In previous work in Hawaii, a preliminary version of this questionnaire was validated against 28 days of food records (15), in White, Native Hawaiian, Japanese, Chinese, and Filipino ethnic groups. Correlations for total fat, vitamin A, and vitamin C ranged from 0.21 to 0.76 for these groups, a range which is quite similar to our current results for Whites

and Japanese-Americans. The large multi-centered European Prospective Investigation into Cancer and Nutrition (EPIC) (16) has recently reported pilot-phase corrected correlation coefficients for energy-adjusted nutrient residuals, which may be compared with the results presented here. These correlations appeared to vary over a roughly similar range as seen in our results. For example, the correlations for energy-adjusted fat intake ranged from 0.09 to 0.87 in the EPIC study, compared with the range of 0.52 to 0.76 reported here. Our poor correlations for calories are in concordance with the poor correlations that have been observed (28) between questionnaires, daily food records, and physical or biologic measurements of total energy expenditure (e.g., doubly labeled water).

Based on between-group differences in the slope estimates in calibration equations, there were significant differences by ethnicity in the performance of the questionnaire in characterizing intake (table 7). These differences highlight the importance of performing separate calibration studies for each of the different ethnic-sex groups. The calibration slopes shown in tables 5 and 6 are needed as a part of the evaluation of the consistency of effects of diet on cancer risk between the ethnic groups, as described previously by Kaaks et al. (17).

In our data, the QFFQ appeared to characterize nutrient intakes somewhat better in Whites, and especially in White females, than in the other ethnic groups. This is particularly true for absolute values, but also true for energy-adjusted values. The average corrected correlations for nutrient densities in males and females, respectively, were 0.61 and 0.60 for African-Americans, 0.66 and 0.61 for Japanese-Americans, 0.62 and 0.57 for Latinos, and 0.66 and 0.74 for Whites (tables 3 and 4). These correlations are more than satisfactory for all ethnic-sex groups.

Comparisons of diet questionnaires with daily diet records have been performed in relatively few US minority populations. Coates et al. (18) found that food frequency questionnaires appeared to work less well in minority populations than among Whites generally, although this was not true in every study they reviewed. Recently, Kristal et al. (19) found that ethnicity was related to the effectiveness of the diet questionnaire used in the Women's Health Trial for a number of nutrients, with correlations among African-American females significantly lower than among Whites and Latinos. Liu et al. (20) also found that the questionnaire used in the Coronary Artery Risk Development in Young Adults (CARDIA) Study performed less well in African-Americans than in Whites, when compared with seven 24-hour recalls. However, unlike in the present study, they found that correlations were not neces-

sarily better for nutrient densities than for absolute nutrients; their average corrected correlations for absolute nutrients were 0.77 for White males, 0.56 for White females, 0.54 for African-American males and 0.37 for African-American females, with the corresponding means for nutrient densities being 0.72, 0.73, 0.44, and 0.34. (Our corresponding means were 0.57, 0.48, 0.30, and 0.26 for absolute nutrients, and 0.66, 0.66, 0.74, and 0.61 for nutrient densities.)

It is notable that for each nutrient and for each ethnic-sex group, the calibration slopes are always less than 1.0 and often less than 0.5. This appears to be a quite general finding in nutritional epidemiology. The meaning of the slope estimate of 0.69 for the group of White males, for example (table 3), is that an observed difference in percent of calories from fat, on the questionnaire, of 20 percent between subjects corresponds to an average true difference of 14 percent [$0.14 = 0.69 \times (0.40 - 0.20)$]. This contraction of differences in mean nutrient intakes between individuals has obvious implications for the interpretation of differences in disease risk from this (or any other) cohort study using diet questionnaires. For example, Hunter et al. (21) found that measurement error correction more than tripled the relative risk estimate per 25 g of fat in a pooled analysis of data from eight cohort studies.

Assumptions of the calibration method

Underlying the use of a calibration study for assessing the performance of the questionnaire are several assumptions, with the most important being that the 24-hour recalls provide unbiased estimates of true diet, and that errors in the 24-hour recalls are uncorrelated with errors in the QFFQ. Other issues, such as whether the 24-hour recalls are more accurate (i.e., less variable around true intake) than the questionnaire or whether the relation between the 24-hour recalls and the questionnaire values is linear (22), are less problematic. As long as a food record or recall is an unbiased estimate of daily food intake, even a single day of data can be used as the comparison instrument in a calibration study. One day of data per subject may give much less information about long-term intake of a particular nutrient than does the questionnaire; however, it is possible to compensate for this by increasing the number of subjects used in the calibration study (23, 24). If examination of a plot of the 24-hour recalls and questionnaire nutrient values indicates that a non-linear relation holds, then this may be modeled by the inclusion of polynomial or other terms in the regression calibration equation (25).

It is increasingly well recognized that self-reports of 24-hour food intake generally underreport total intake (26), and our study is no exception to this rule. Daily

energy intake in a sedentary population is expected to be 1.55 times the basal metabolism rate (12). Briefel et al. (27) reported for the National Health and Nutrition Examination Survey (NHANES) that 24-hour recalls yield energy intake estimates that averaged 1.47 times estimated basal metabolism rate for men and 1.26 times for women. Further evidence of underreporting of total energy intake has been found when comparisons of self-assessments of diet have been made with energy expenditure measurements using the doubly labeled water method (28).

Effect of underreporting in 24-hour recalls

At present, the significance in calibration and correlational analysis of the evident underreporting of total energy intake in 24-hour recalls is not fully understood either for this study or for others using the calibration study approach for measurement error correction. If, for example, the 24-hour recalls simply underestimate intake by some relatively fixed amount, or by a relatively constant fraction of intake, then calibration slope estimates on the arithmetic scale (in the first case), or the log scale (in the second) would be unaffected by the underreporting. Serious problems arise when errors in reporting on the two instruments are correlated, conditional on true intake, because this will result in positively biased correlations and calibration slope estimates. The type of experiments required to carefully address these issues of conditional correlation have been performed only on small numbers of subjects. One such study compared protein excretion in 24-hour urine samples with several diet assessment techniques and came to the conclusion that correlation between errors in the self-reports was substantial (29). However, the extent to which this effect is at play in calibration studies such as ours has not yet been carefully examined.

In our study, we found that men (but not women) with BMI >29 appeared to underreport on the questionnaire their percent of calories from fat and density of vitamin C intake. There have been reports (30, 31) of underreporting of food intake by obese subjects on both questionnaire and daily food records, when reported intake was compared with doubly labeled water measurements of energy expenditure. This observation while suggestive of the kind of correlation between errors, conditional on true intake, that would bias the results of a calibration study, does not constitute *prima facie* proof of the existence of such an effect. Obese people may be only underreporting more than nonobese on an absolute scale and not as a percent of true intake. If this is so, then multiplicative errors in the two instruments would not be correlated with the obesity of the subject, and on the log scale the

corrected correlations and calibration slopes should remain unbiased. Indeed, Lissner et al. (32), in an investigation involving intensive study of 63 adult women found that obese subjects underreported no more than nonobese subjects, on food diaries or 24-hour recalls, when they were compared at the same intake level. If, on the other hand, obese individuals are observed to underreport more on both instruments than do normal subjects *at the same true energy level*, then this would have impact on the calibration slope and correlation estimates. That is, this amounts to a violation of the assumption that errors in the two reports are independent conditional on true intake. The issue of underreporting in the 24-hour recalls is of serious concern. Further knowledge concerning the extent to which errors in the 24-hour recalls are correlated with the errors in the QFFQ will be needed in order to determine whether corrected correlations or calibration slopes have been seriously distorted in this or similar studies. In the absence of large, carefully designed and analyzed studies, the degree to which our study or other similar calibration studies may be subject to correlation between errors in the two dietary assessments is unknown.

In summary, we found for all sex-ethnic groups that the questionnaire used in this multiethnic cohort gave calibration and correlation results for nutrient densities and energy corrected nutrients that are comparable with those reported by other groups, and highly satisfactory for energy-adjusted nutrients for all subgroups. The performance of the questionnaire varied by ethnic group, with higher correlations in Whites, and also varied by BMI and, to a lesser degree, education level. These findings indicate the need to control for the differences in performance of the questionnaire across subgroups.

ACKNOWLEDGMENTS

This work was supported in part by Public Health Service grant no. R01 CA54281 from the National Cancer Institute.

The authors thank the registered dietitians who collected the 24-hour recalls by telephone: Kapuanani Rothfus, Deborah Ishiyama, and Sabrina Umphress in Hawaii, and Gretchen Perea, Dawn Narvaez, and Kerrie Weaver in Los Angeles.

REFERENCES

1. Keys A. Dietary epidemiology. *Am J Clin Nutr* 1967;20: 1151-7.

2. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14:32-8.
3. Willett W. *Nutritional epidemiology*. New York, NY: Oxford University Press, 1990.
4. 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.
5. Rosner B, Spiegelman D, Willett WC. Correction of logistic relative risk estimates and confidence intervals for random within-person measurement error. *Am J Epidemiol* 1992;136:1400-9.
6. US Department of Agriculture. USDA nutrient database for standard reference: Release 10. Riverdale, MD: Nutrient Data Laboratory, Agricultural Research Service, 1993.
7. US Department of Agriculture. Nutrient database for individual food intake surveys and data sets used to create it: Release 7. Report from the National Technical Information Service. Springfield, VA: National Technical Information Service, 1994.
8. Chavez MM, Hernandez M, Roldan JA. In: *Comision nacional de alimentacion tablas de uso pactico del: Valor nutritivo de los alimentos de mayor consumo en Mexico*. Mexico City: Instituto Nacional de la Nutricion Salvador Zubiran, 1992.
9. Kagawa R. *Standard tables of food composition in Japan*. Tokyo: Women's University, 1995.
10. Pennington JAT. *Bowes & Church's food values of portions commonly used*. 17th ed. Philadelphia, PA: Lippincott, 1998.
11. Mangels AR, Holden JM, Beecher GR, et al. The carotenoid content of fruits and vegetables: an evaluation of analytic data. *J Am Diet Assoc* 1993;93:284-96.
12. National Research Council. *Recommended dietary allowances*. 10th ed. Washington, DC: National Academy Press, 1989.
13. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 1985;122:51-65.
14. 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.
15. Hankin J, Wilkens L, Kolonel L, et al. Validation of a quantitative diet history method. *Am J Epidemiol* 1991;133:616-28.
16. Kaaks R, Slamani N, Riboli E. Pilot phase studies on the accuracy of dietary intake measurements in the EPIC project: overall evaluation of results. *Int J Epidemiol* 1997;26:S26-36.
17. Kaaks R, Riboli E. Validation and calibration of dietary intake measurements in the EPIC project: methodological considerations. *Int J Epidemiol* 1997;26:S15-S25.
18. Coates R, Monteilh C. Assessments of food-frequency questionnaires in minority populations. *Am J Clin Nutr* 1997;65:S1108-15.
19. Kristal AR, Feng Z, Coates RJ, et al. Associations of race/ethnicity, education, and dietary intervention with the validity and reliability of a food frequency questionnaire. The Women's Health Trial Feasibility Study in Minority Populations. *Am J Epidemiol* 1997;146:856-69.
20. Liu KL, Slattery M, Jacobs D, et al. A study of the reliability and comparative validity of the CARDIA dietary history. *Ethn Dis* 1994;4:15-27.
21. Hunter DJ, Spiegelman D, Adami HO, et al. Cohort studies of fat intake and the risk of breast cancer—a pooled analysis. *N Engl J Med* 1996;334:356-61.
22. Lee J. Use of regression adjustments for ascertaining nutrient intakes. *J Am Diet Assoc* 1981;78:245-9.
23. Wacholder S, Armstrong B, Hartge P. Validation studies using an alloyed gold standard. *Am J Epidemiol* 1993;137:1251-8.
24. Stram DO, Longnecker MP, Shames L, et al. Cost-efficient design of a diet validation study. *Am Epidemiol* 1995;142:353-62.
25. Thomas D, Stram D, Dwyer J. Exposure measurement error: influence on exposure-disease relationships and methods of correction. *Ann Public Health* 1993;14:69-93.
26. Black AE, Prentice AM, Goldberg G, et al. Measurements of total energy expenditure provide insights into the validity of dietary measurements of energy intake. *J Am Diet Assoc* 1993;93:572-9.
27. Briefel R, Sempos C, McDowell M, et al. Dietary methods research in the third National Health and Nutrition Examination Survey: underreporting of energy intake. *Am J Clin Nutr* 1997;65(suppl):1203S-9S.
28. Black AE. Physical activity levels from a meta-analysis of doubly labeled water studies for validating energy intake as measured by dietary assessment. *Nutr Rev* 1996;54:170-4.
29. Plummer M, Clayton D. Measurement error in dietary assessment: an investigation using covariance structure models. Parts I and II. *Stat Med* 1993;12:925-48.
30. Sawaya A, Tucker K, Tsay R, et al. Evaluation of four methods for determining energy intake in young and older women: comparison with doubly labeled water measurements of total energy expenditure. *Am J Clin Nutr* 1996;63:491-9.
31. Bandini L, Schoeller D, Cyr H, et al. Validity of reported energy intake in obese and nonobese adolescents. *Am J Clin Nutr* 1990;52:421-5.
32. Lissner L, Habicht J, Strupp B, et al. Body composition and energy intake: do overweight women overeat and underreport? *Am J Clin Nutr* 1989;49:320-5.