Practice of Epidemiology

Validity of a Multipass, Web-based, 24-Hour Self-Administered Recall for Assessment of Total Energy Intake in Blacks and Whites

Lenore Arab*, Chi-Hong Tseng, Alfonso Ang, and Patricia Jardack

* Correspondence to Dr. Lenore Arab, David Geffen School of Medicine, University of California, Los Angeles, 10960 Wilshire Blvd., 12-262 Factor Building, Box 951736, Los Angeles, CA 90095-1736 (e-mail: larab@mednet.ucla.edu).

To date, Web-based 24-hour recalls have not been validated using objective biomarkers. From 2006 to 2009, the validity of 6 Web-based DietDay 24-hour recalls was tested among 115 black and 118 white healthy adults from Los Angeles, California, by using the doubly labeled water method, and the results were compared with the results of the Diet History Questionnaire, a food frequency questionnaire developed by the National Cancer Institute. The authors performed repeated measurements in a subset of 53 subjects approximately 6 months later to estimate the stability of the doubly labeled water measurement. The attenuation factors for the DietDay recall were 0.30 for blacks and 0.26 for whites. For the Diet History Questionnaire, the attenuation factors were 0.15 and 0.17 for blacks and whites, respectively. Adjusted correlations between true energy intake and the recalls were 0.50 and 0.47 for blacks and whites, respectively, for the DietDay recall. For the Diet History Questionnaire, they were 0.34 and 0.36 for blacks and whites, respectively. The rate of underreporting of more than 30% of calories was lower with the recalls than with the questionnaire (25% and 41% vs. 34% and 52% for blacks and whites, respectively). These findings suggest that Web-based DietDay dietary recalls offer an inexpensive and widely accessible dietary assessment alternative, the validity of which is equally strong among black and white adults. The validity of the Web-administered recall was superior to that of the paper food frequency questionnaire.

diet; energy metabolism; mental recall; questionnaires; validation studies

Abbreviations: BMI, body mass index; DHQ, Diet History Questionnaire; DLW, doubly labeled water; FFQ, food frequency questionnaire; OPEN, Observing Protein and Energy Nutrition; REI, reported energy intake; TEE, total energy expenditure; UCLA, University of California, Los Angeles.

Despite the increased attention being paid to the role of health disparities in nutrition-related diseases among different ethnic groups, there has been very little research on the validity of dietary assessment instruments across diverse ethnic subpopulations (1). Differential performance of dietary assessment tools in subgroups can bias findings, prevent comparisons, and hinder health disparity investigations of ethnic groups (2). Unless proven otherwise, differential responses need to be assumed. This concern has been substantiated by the comparison of responses to food frequency questionnaires (FFQs) in a multiethnic cohort in Hawaii and Los Angeles, which showed much better Pearson correlation coefficients with dietary recalls among white males and females (average correlations of 0.57 and 0.48, respectively)

than among black males and females (average correlations of 0.30 and 0.26, respectively) (2). Similar results have been found for women in the Women's Health Trial Feasibility Study in Minority Populations (3) and for young adults in the Coronary Artery Risk Development in Young Adults Study (4). Given the differences that are seen when FFQs and diet histories are used, the need for novel methods that are equally valid across ethnic groups has become pressing.

Dietary assessment methods used in nutritional epidemiology include prospective methods, such as diet records, and retrospective methods, such as multiple 24-hour recalls, FFQs, and diet history interviews. Although these methods all have individual strengths, each suffers from limitations. Twenty-four-hour recalls have recently become favored based on the

finding in a US population that a single 24-hour recall out-performed an FFQ (5). Attenuation factors were twice as high with the 24-hour recall, and correlations with biomarker-based estimates of true intakes were 0.34 as compared with 0.19 with the FFQ in a largely white population.

Recalls are open-ended, do not require literacy, and have a relatively low respondent burden (6). However, a major drawback is that a single day or a few days of collection is likely not representative of an individual's habitual diet because of considerable day-to-day variability in dietary intake (7). Multiple days of recall, unless automated, require repeated contacts with subjects, which results in high personnel costs (8–10) and drives interest in less-expensive alternatives. Self-conducted recalls over the Internet offer such an option, but their validity remains to be demonstrated.

Recovery biomarkers have become the gold standard for validation of dietary assessment methods in the population under study. The most robust of these biomarkers is the measurement of energy expenditure through doubly labeled water (DLW) isotopes in individuals who are healthy and in energy balance. Under these conditions, energy expenditure should equal energy intake, and valid reporting instruments should yield reported intakes that correlate with individual expenditures. Because of expense and prior limitations in supply, DLW has not been widely used to date. In only 3 studies with sample sizes greater than 150 subjects have investigators examined the energy validity of their dietary methods (8, 9, 11). All 3 of these have been tested among older (>40 years of age) and largely white US subjects.

The present report addresses racial differences in dietary assessment validity using Web-based 24-hour recalls in the Energetics Study, the first large DLW-based study to compare energy validity among blacks and whites and to include younger adults in substantial numbers. To our knowledge, it is also the first objective DLW-based validation of Web-based approaches to dietary assessment. Both multiple Web-based 24-hour recalls and the Diet History Questionnaire (DHQ), a food frequency questionnaire developed by the National Cancer Institute, were tested in a biracial population of Los Angelinos. In this article, we present the correlations between self-reported and DLW-based energy intake measurements; the attenuation factors derived from using this biomarker as a gold standard by race, age, and gender; the stability in the measurement of DLW within individuals at a 6-month interval (rho); and the correlations of the method with the estimated true intake after adjustment for rho. We also present the extent of underreporting and overreporting by race with both methods. The cost and benefit of each additional day of reporting is also presented for the first 6 reports by race, using the magnitude of the correlation with DLW as an indicator of improvement.

MATERIALS AND METHODS

Subjects and study design

The University of California, Los Angeles (UCLA) Energetics Study was conducted between July 2006 and June 2009. Participants were recruited through public postings and online advertisements (10). Eligible subjects were generally healthy,

nonsmoking black and white adults between the ages of 21 and 69 years who resided within 50 miles (80.5 km) of the UCLA campus. Inclusion requirements included having had a stable weight for the previous 6 months and being willing to maintain current dietary and physical activity habits for the duration of the study. Subjects were also required to be able to read and speak English, to have a working telephone, and to be available for all clinic visits and self-administered paper or Web-based questionnaires. Potential subjects were excluded if they suffered from any of the following conditions: diabetes mellitus, hemophilia, alcoholism, mental disorder, hypothyroidism, bipolar disorder, seizure disorders, congestive heart failure, renal failure, or other conditions that affected fluid balance. Subjects were also excluded if they were undergoing treatment with supplemental oxygen or antiretroviral, antineoplastic, antiulcer/antireflux, or central nervous system drugs.

During the consent process, subjects received a detailed explanation regarding the DLW biomarker procedures, as well as instruction on how to access and complete the DietDay Web-based 24-hour recall. Screening and collection of demographic and baseline information were done using Web-based questionnaires or electronic case report forms, as described previously (12). All subjects completed 2 study visits, 8 dietary recalls, and 1 DHQ, which estimated frequencies of consumption of 124 food items over the past year, as well as a DLW determination of total energy expenditure (TEE). To allow for determination of the intraclass correlation coefficient for the DLW biomarker, repeat values for 53 subjects were used. Both the original and repeat studies were approved by the UCLA institutional review board, and the procedures followed were in accordance with the institutional review board's ethical standards. Written informed consent was obtained from all subjects for both studies.

Dietary assessment

Twenty-four-hour dietary recalls were self-administered using the Web-based DietDay system (http://www.24hrrecall.com); details of the method have been described in a separate report (10). Briefly, DietDay applies multipasses similar to the US Department of Agriculture-designed multipass approach with automatic branching, complex skip routines, range checks, edit checks, and prompts to subjects as they report their dietary information (8). DietDay contains 9,349 foods and more than 7,000 food images; portion sizes are quantified using images of household measures. Data on food preparation methods, use of condiments, time of day of meals/snacks, and consumption of nutritional supplements were also collected. Nutrient values in the program were based on US Department of Agriculture values, with expansion to include mixed dishes and product-labeling information.

The first 6 DietDay 24-hour recalls were conducted over a 2-week period, with the first conducted at the consent visit, 1 conducted at each of the 2 study visits to the UCLA General Clinical Research Center, and the other 3 self-administered by subjects via the Internet between the 2 study visits. The final 2 DietDay recalls, which were not used in the present analysis, were completed by subjects approximately 1 and 2 months after the second study visit. Subjects were notified by automatic e-mail of the need for DietDay recall completion,

Table 1. Demographic Characteristics of the Sample Population in the Energetics Study, Los Angeles, California, 2006–2009

Population Characteristic	All Subjects			White Subjects			Black Subjects		
	No.	%	Median (IQR)	No.	%	Median (IQR)	No.	%	Median (IQR)
All subjects	233	100		118	100		115	100	
Gender									
Female	158	67.8		74	62.7		84	73.0	
Male	75	32.2		44	37.3		31	26.9	
Median age, years			33.3 (12.5)			29.2 (13.2)			38.3 (11.8)
Age, years									
21–<30	95	40.8		63	53.4		32	27.8	
30–50	87	37.3		29	24.6		58	50.4	
>50–69	51	21.9		26	22.0		25	21.7	
Median body mass index[a]			25.0 (6.1)			23.6 (4.8)			26.9 (6.7)
Body mass index[a]									
<25	112	48.1		77	65.3		35	30.4	
25–30	68	29.2		26	22.0		42	36.5	
>30	53	22.8		15	12.7		38	33.0	
Educational level									
Less than high school	1	0.4		0			1	0.9	
High school graduate	6	2.6		0			6	5.2	
Some college	86	36.9		29	24.6		57	49.6	
College graduate	105	45.1		62	52.5		43	37.4	
Postgraduate	35	15.0		27	22.9		8	6.9	

Abbreviation: IQR, interquartile range.

[a] Weight (kg)/height (m)^2.

and nonresponders were pursued through the use of personalized e-mails and telephone calls. Additionally, to allow pooling of findings with those of the Observing Protein and Energy Nutrition (OPEN) Study, subjects provided habitual food consumption information for the prior year using the 36-page paper-based DHQ, which allows recording of portion size and frequency of consumption for 124 individual foods, as well as additional queries regarding specifications (e.g., fat content and seasonal consumption) for selected foods and supplement usage (9, 13). Subjects self-administered the DHQ between the consent visit and the first study visit. Return of the completed DHQ was required for study entry. Nutrient values assigned to foods recorded in the DHQ were based on the National Cancer Institute's nutrient database. Subjects also completed a Web-based general questionnaire, a computer-assisted diet history, a computerized version of the International Physical Activity Questionnaire (14), and an exit questionnaire. The feasibility of this Web-based approach has been described elsewhere (10).

Energy biomarker

TEE was measured using the DLW method, essentially as described by Subar et al. (9), using isotope measurement methods described by Schoeller et al. (15) and Cole and Coward (16). Subjects were instructed to refrain from consuming any foods or beverages for 4 hours before dosing. DLW was administered in a volume of one-fourth to one-half

cup (approximately 59–118 mL) at a dose of approximately 2 g of 10 atom percent ^{18}O-labeled water and 0.12 g of 99.9 atom percent deuterium-labeled water per kilogram of measured body weight; subjects also consumed a 50-mL water rinse from the DLW bottle. A 250-mL glass of water was provided 1 hour after dosing, and an additional 250-mL glass of water was provided between hours 2 and 4 after dosing. All volumes and times of consumption were recorded. Spot urine samples for determination of isotope enrichment were collected before dosing and at 1, 2, and 3–4 hours after dosing, as well as twice on the 15th day after dosing. Subjects 60 years of age or older also provided a blood sample to allow for determination of isotope enrichment in blood in the case of incomplete bladder emptying as indicated by incomplete isotope equilibration. Deuterium and ^{18}O levels in urine samples were quantified by using mass spectroscopy, and the values were used for calculation of TEE according to the plateau method (15). All isotopic analyses were conducted at the University of Wisconsin (Madison, Wisconsin). The DLW determination of TEE was repeated for a subset of 53 subjects 6 months after the first assessment to allow calculation of the intraclass correlation coefficient, rho.

Statistical methods

The main purpose of the present article is to compare by race the validity of DietDay Web-based 24-hour recall for assessment of energy intake with that of the DHQ. Summary
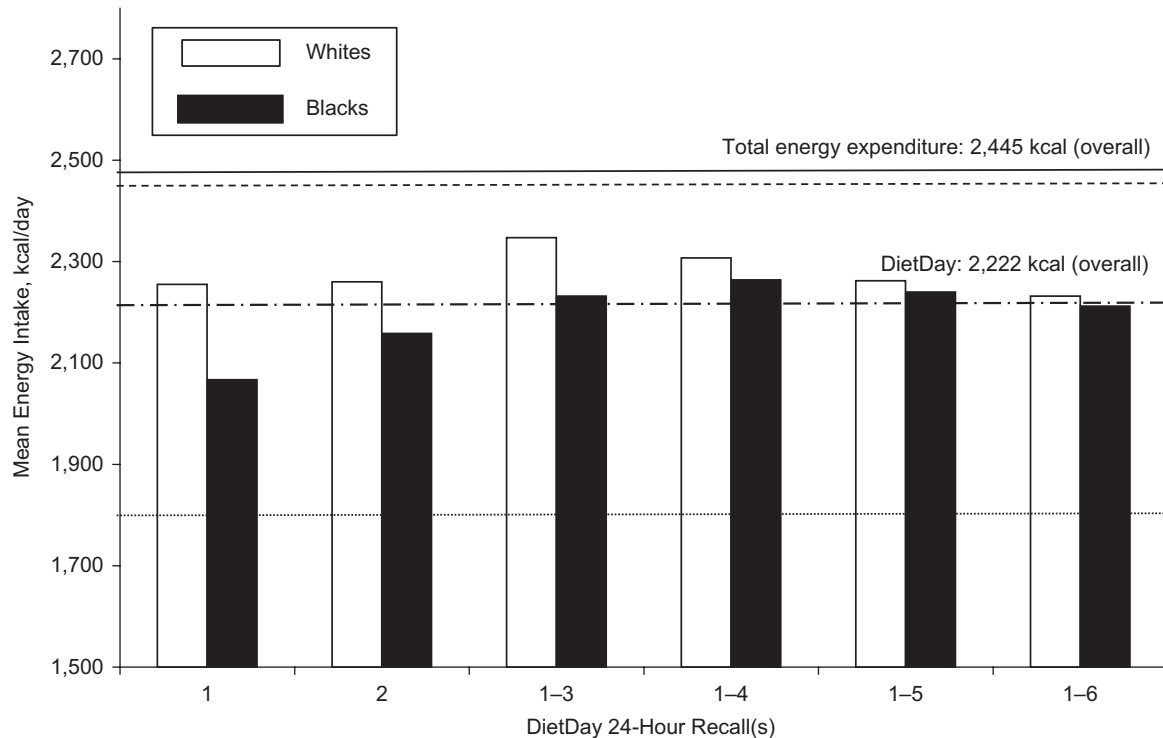
Figure 1. Mean energy intake based on DietDay 24-hour recall by race/ethnicity in the University of California, Los Angeles Energetics Study, 2006–2009. The dashed line indicates the mean total energy expenditure (TEE) of 2,442 kcal/day for whites; the solid line indicates the mean TEE of 2,448 kcal/day for blacks; and the dotted line indicates the mean overall National Cancer Institute Diet History Questionnaire expenditure of 1,783 kcal/day for whites and blacks. The numbers on the x-axis refer to the DietDay 24-hour recalls used (first, second, first through third, etc.).

statistics (mean, standard deviation, median, interquartile range, and frequency distributions) were generated to characterize the study population. We applied log-transformations to energy estimated by using DietDay recalls, the DHQ, and DLW. The DietDay energy intake was estimated using the mean of the first through sixth DietDay recalls. To allow for comparability of our findings with those of the OPEN Study (17), we excluded extreme outlying values that fell outside the interval delimited by the 25th percentile minus twice the interquartile range and the 75th percentile plus twice the interquartile range. We also excluded participants with fewer than 6 Web-based dietary assessments. For this reason, 19 blacks and 10 whites were excluded from the analyses.

To evaluate the reliability of reported energy intake (REI), we first examined dietary intake underreporting. We calculated the ratio of reported intakes to TEE measurements to indicate the presence of reporting bias and underreporters and overreporters. We evaluated the percentage classified as misreporters using the cutoffs of REI/DLW-measured energy expenditure ratios, which were multiplied by 100 to generate percentages. We then defined participants who reported less than 70% of DLW-quantified energy intake as severe underreporters and those who reported less than 80% as moderate underreporters. Participants who reported greater than 120% of DLW-quantified energy intake were labeled overreporters. Bootstrap analysis was used to compare the percentages of

underreporting and overreporting in the DietDay recalls with those in the DHQ.

In the second reliability analysis, we used the unadjusted correlation to assess the agreement between the log-transformed reported intakes (DietDay and DHQ) and the DLW biomarker (TEE) for all subjects and by baseline characteristics. We used the bootstrap method with replacement using 1,000 replications to construct the 95% confidence interval around the correlations comparing the performance of DietDay and DHQ.

To examine the structure of the measurement errors in the dietary instruments, the following classical measurement error model is assumed:

$$Z_{1i} = a_1 + b_1 X_i + r_{1i} + \varepsilon_{z1i},$$
$$Z_{2i} = a_2 + b_2 X_i + r_{2i} + \varepsilon_{z2i},$$

and

$$W_i = \mu + X_i + \varepsilon_{wi},$$

where Z_{1i} is the reported energy intake from DietDay, Z_{2i} is the reported energy intake from the DHQ, X_i is the long-term true energy intake, and W_i is the energy intake measured by using TEE. a_1 and a_2 represent the overall average biases of the reported energy intake. b_1 and b_2 are the intake-related

Table 2. Pearson Correlation Coefficients for Reported Energy Intake and Total Energy Expenditure Based on Doubly Labeled Water Measurements, the Energetics Study, Los Angeles, California, 2006–2009[a]

	All Subjects (n = 233)		White Subjects (n = 118)		Black Subjects (n = 115)	
	r	95% CI	r	95% CI	r	95% CI
DietDay 24-hour recall						
First	0.22	0.10, 0.34	0.26	0.08, 0.42	0.20	0.02, 0.37
First through second	0.40	0.28, 0.49	0.31	0.13, 0.46	0.45	0.30, 0.59
First through third	0.41	0.30, 0.51	0.35	0.18, 0.50	0.46	0.30, 0.59
First through fourth	0.41	0.30, 0.51	0.37	0.21, 0.52	0.44	0.28, 0.58
First through fifth	0.45	0.35, 0.55	0.44	0.28, 0.57	0.47	0.31, 0.60
Mean of first through sixth	0.45	0.35, 0.55	0.44	0.28, 0.58	0.47	0.31, 0.60
National Cancer Institute Diet History Questionnaire	0.33	0.21, 0.44	0.34	0.17, 0.49	0.32	0.14, 0.47

Abbreviation: CI, confidence interval.

[a] Reported energy intakes were from either the DietDay 24-hour Web-based recall or the Diet History Questionnaire; total energy expenditure was based on the doubly labeled water biomarker. Analysis included subjects who completed the study and had evaluable doubly labeled water biomarkers, dietary intake for at least six 24-hour recalls, and an evaluable Diet History Questionnaire.

biases. ε_{z1i}, ε_{z2i}, and ε_{wi} are independent zero mean random measurement errors that are independent of X_i, and r_{1i} and r_{2i} are zero mean correlated subject-specific biases that are independent of X_i. This model assumes that the energy intake

measured by TEE is unbiased, whereas the estimates from dietary assessment might be biased.

In the present study, TEE and dietary assessments were replicated in a subset of subjects approximately 6 months

Table 3. Estimated Attenuation Factor (λ) for Energy Intake Reported in the DietDay 24-Hour Recall or the National Cancer Institute Diet History Questionnaire, by Age, Gender, Race/Ethnicity, and Body Mass Index, the Energetics Study, Los Angeles, California, 2006–2009

Subgroup	No. of Subjects	DietDay 24-Hour Recall[a]		National Cancer Institute Diet History Questionnaire[a]	
		λ	95% CI	λ	95% CI
Total[a]	233	0.28	0.21, 0.35	0.16	0.10, 0.21
Race					
White*	118	0.26	0.18, 0.34	0.17	0.08, 0.25
Black*	115	0.30	0.20, 0.40	0.15	0.07, 0.22
Gender					
Female*	158	0.21	0.11, 0.30	0.08	0.02, 0.13
Male	75	0.12	0.04, 0.18	0.07	−0.00, 0.13
Age, years					
<30*	95	0.38	0.29, 0.47	0.18	0.10, 0.27
30–50	87	0.28	0.16, 0.39	0.16	0.05, 0.26
>50	51	0.08	−0.06, 0.23	0.07	−0.04, 0.18
Body mass index[b]					
<25*	113	0.32	0.22, 0.41	0.22	0.13, 0.31
25–30	67	0.17	0.08, 0.26	0.13	0.05, 0.21
>30*	53	0.30	0.10, 0.47	−0.01	−0.15, 0.11

Abbreviation: CI, confidence interval.

* $P < 0.05$ (significant difference in the correlations between DietDay and the National Cancer Institute Diet History Questionnaire).

[a] Data are based on mean intake for the first 6 DietDay 24-hour recalls or the National Cancer Institute Diet History Questionnaire.

[b] Weight (kg)/height (m)^2.

Table 4. Pearson Correlation Coefficients for Reported Dietary Intake and Estimated True Intake, by Race, the Energetics Study, Los Angeles, California, 2006–2009[a,b]

Dietary Assessment	All Subjects (n = 233)		White Subjects (n = 118)		Black Subjects (n = 115)	
	r	95% CI	r	95% CI	r	95% CI
DietDay 24-hour recall						
First	0.24	0.10, 0.36	0.28	0.09, 0.47	0.21	0.03, 0.35
First through second	0.42	0.28, 0.56	0.33	0.14, 0.49	0.48	0.27, .065
First through third	0.44	0.30, 0.57	0.38	0.20, 0.53	0.49	0.28, 0.66
First through fourth	0.44	0.31, 0.56	0.40	0.23, 0.55	0.47	0.27, 0.63
First through fifth	0.48	0.34, 0.59	0.47	0.32, 0.62	0.50	0.32, 0.65
Mean of first through sixth	0.48	0.35, 0.59	0.47	0.30, 0.62	0.50	0.32, 0.65
National Cancer Institute Diet History Questionnaire	0.35	0.21, 0.47	0.36	0.16, 0.53	0.34	0.16, 0.51

Abbreviation: CI, confidence interval.

[a] Reported dietary intakes were from either the DietDay 24-hour recalls or the National Cancer Institute Diet History Questionnaire, and estimations of true intake were based on the doubly labeled water biomarkers. Analysis included subjects who completed the study and had evaluable doubly labeled water measurements, dietary intake for at least six 24-hour recalls, and an evaluable National Cancer Institute Diet History Questionnaire.

[b] Adjustment factor for estimating estimated true intake was based on correlations for doubly labeled water measurements performed approximately 6 months apart in the current study, in which $r = 0.88$.

after their baseline visits. The intraclass correlation of the repeated TEE measurement was estimated to be 0.88 based on the linear mixed-effects model. We used the correlation between Z and X and the attenuation factor associated with the dietary assessment to demonstrate and compare the validity of REIs. The correlations between Z and X and the attenuation factor associated with the dietary assessment were estimated on the basis of the moment estimators (18) according to the above model. The attenuation factor (λ) was estimated by using the regression coefficient from the regression model of W on Z, and the correlation between Z and X was estimated as the covariance of W and Z divided by the square root of the product of the variance of W, the variance of Z, and the intraclass correlation coefficient of W (rho). The 95% confidence intervals were constructed using the bootstrap method.

RESULTS

Subject characteristics and disposition

As detailed elsewhere, 262 of the 326 subjects who responded to advertisements and passed the eligibility criteria consented to participate, enrolled in the Energetics Study, and completed all clinic visits. The demographic profile of the enrolled subjects was similar to that of the group who screened positive (10). Validation analyses of the DietDay 24-hour recall method were performed using the subset of 233 subjects for whom the following data were available: at least 6 of the 8 planned DietDay 24-hour recalls, the DHQ, and assessment of TEE using the DLW method. Demographic and baseline characteristics for this subset (Table 1) were similar to those observed in the overall sample population (10). The majority of subjects were white (51%), and the rest were black (49%). The median age of all subjects was 33.3 years. Although 78% of all participants were 50 years of

age or younger, the distribution differed by race: The majority (50%) of black participants were 30–50 years of age, and the majority (53%) of white participants were 21–30 years of age. Notably, 52% of all subjects had body mass indexes (BMIs, defined as weight in kilograms divided by height in meters squared) greater than 25 (indicating higher than normal weight for height); 12.7% of white subjects and 33% of black subjects had a BMI of greater than 30, which qualified them as obese. The educational levels ranged widely, but most subjects had some college education or were college graduates.

Mean energy intakes and TEEs

Subject-reported mean energy intake among the entire cohort was on average 279–402 kcal higher when assessed using the DietDay 24-hour recall (mean of the first recall, mean of the second recall, and the averages of multiple days) than when assessed using the DHQ. As shown in Figure 1, whites reported higher DietDay energy intakes than did blacks. Reported mean intakes declined with increasing numbers of recall days among whites but not among blacks, suggesting racial bias in decay in reporting over the duration of the study. Figure 1 also demonstrates the gap between mean TEE determined by the DLW biomarker for the entire cohort (2,445 kcal/day) and the mean intake reported using either DietDay or the DHQ. The difference in the means (TEE − REI) was 223 kcal/day for DietDay recalls, as seen in Figure 1 (the mean of the first 6 DietDay recalls), and 662 kcal/day for the DHQ.

Correlations between DietDay 24-hour recall and energy intake by race, age, and BMI

Correlational analyses were used to compare the 2 methods of dietary assessment (the Web-based, self-administered
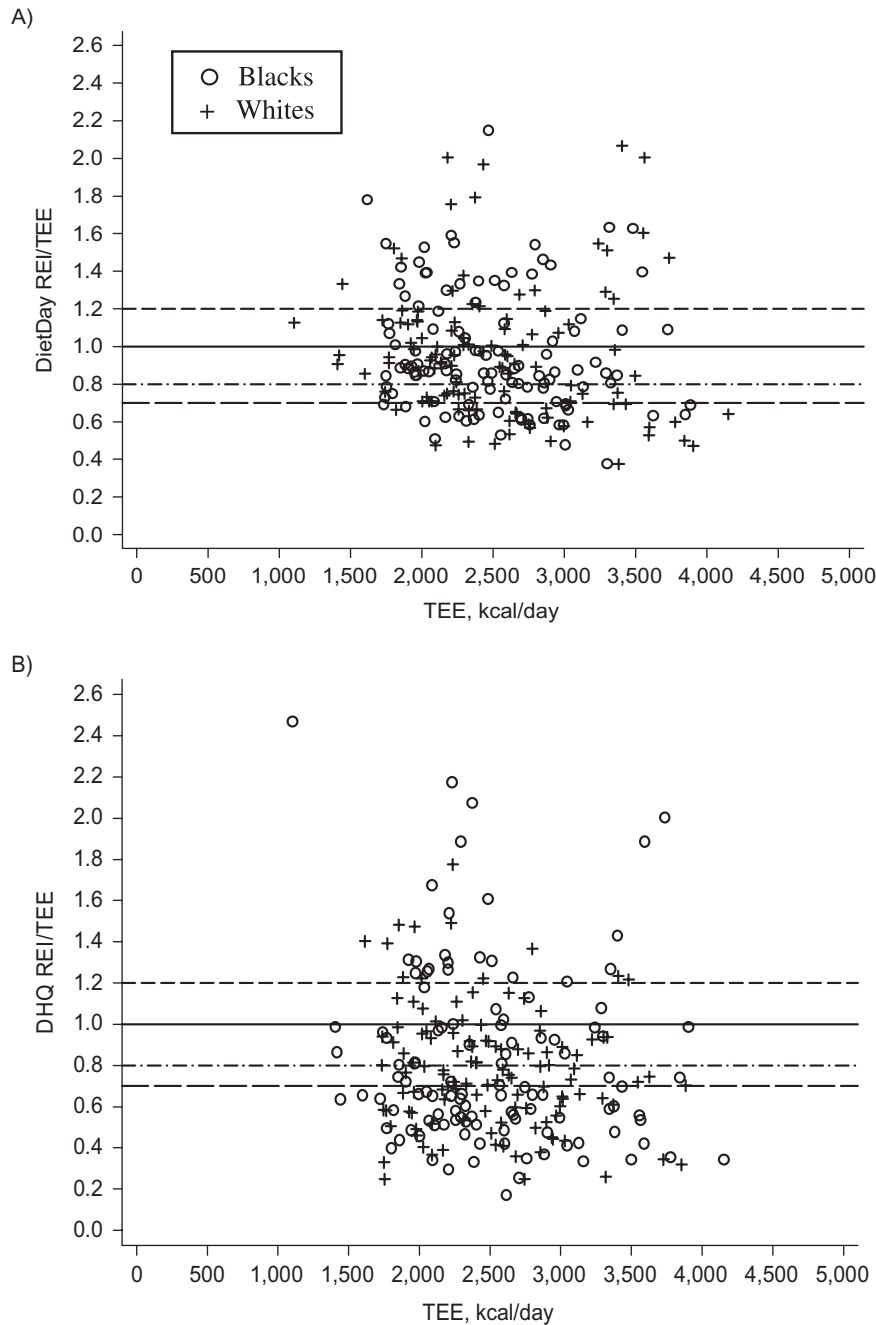
A)

B)

Figure 2. A) Reported energy intake (REI) contrasted with the degree of misreporting for A) DietDay 24-hour recalls and B) the National Cancer Institute Diet History Questionnaire, University of California, Los Angeles Energetics Study, 2006–2009. The dashed lines represent underreporting at 70% and 80% and overreporting at 120%. TEE, total energy expenditure.

DietDay system and the DHQ) in terms of their validity for reporting energy intake using TEE determined by the DLW method. Crude correlations of TEE determined by DLW and dietary intake calculated using the means of the first 6 Diet-Day recalls were 0.45 for the overall cohort, 0.47 for blacks, and 0.44 for whites (Table 2). Crude correlations of TEE and dietary intake determined from the DHQ were 0.33 for the

overall cohort, 0.32 for blacks, and 0.34 for whites. As shown in Table 2, correlations with TEE were higher for DietDay-reported intake than for the DHQ-reported intake, and correlations for TEE with DietDay-reported intake also improved as the number of recalls increased. Notably, except for the first DietDay recall, correlations between DietDay intake and TEE were higher for blacks than for whites and stabilized in

Table 5. Percentage of Participants Classified as Underreporters and Overreporters of Mean Energy Intake by Age, Gender, Race, and Body Mass Index, the Energetics Study, Los Angeles, California, 2006–2009

Subgroup	No. of Subjects	REI:TEE					
		<70%		<80%		>120%	
		DD 24HR	DHQ	DD 24HR	DHQ	DD 24HR	DHQ
Total	233	30*	46*	44*	56*	14	15
Gender							
Female	258	28	47	41*	56*	13	14
Male	75	33	44	51	56	16	17
Age, years							
<30	95	25*	39*	44	55	14	14
30–50	87	38	51	44	54	15	18
>50	51	24*	53*	43	63	14	12
Body mass index[a]							
<25	113	23*	39*	39	50	19	12
25–30	67	37*	55*	48*	64*	12	15
>30	53	34	51	49	60	8*	21*
Race							
Black	115	25*,**	41*,**	44*	55*	9*,**	20*,**
White	118	34*,**	52*,**	43	57	19*,**	10*,**

Abbreviations: DD 24HR, DietDay 24-hour recall; DHQ, Diet History Questionnaire; REI, reported energy intake from either the DietDay 24-hour recall or the Diet History Questionnaire; TEE, total energy expenditure.

* $P < 0.05$ (significant difference between DietDay and the National Cancer Institute DHQ).

** $P < 0.05$ (significant difference between blacks and whites).

[a] Weight (kg)/height (m)^2.

fewer days among the blacks. For DHQ assessments, whites had higher correlations with TEE compared with blacks.

Comparison of attenuation factors by race, age, and BMI

The attenuation factor, λ, represents the degree to which correlations between dietary intake and true intake or regressions describing effects of dietary intake on disease risk are underestimated or overestimated because of random error in reporting intake using one of the usual dietary assessment methods (24-hour recall, multiday diet record, and annual FFQ). Attenuation factors for the DietDay and DHQ methods as applied to the current study are shown in Table 3. For the entire cohort, the attenuation factor was higher when reported energy intake was based on DietDay (λ = 0.28) than when reported intake was based on the DHQ (λ = 0.16), indicating a higher degree of measurement error with the DHQ. Similarly, attenuation factors were higher across all subsets examined when reported energy intake was based on DietDay than when reported intake was based on the DHQ. Blacks had better attenuation factors (value closer to 1) than did whites with the DietDay. The attenuation factors were significantly lower with the DHQ.

Correlations of reported energy intake and true intake by dietary assessment and method, race, age, and BMI

In separate analyses in which we used the rho to estimate the correlations of dietary intakes with estimated true energy

intakes, results were similar to, but slightly stronger than, the correlations with TEE. In the overall cohort, the correlations of energy intake reported via DietDay to true intake were higher than those reported using the DHQ for any number of days greater than 1 (Table 4). Correlations of DietDay (mean of first through sixth recalls) with TEI (Table 4) were slightly higher for blacks (r = 0.50) than for whites (r = 0.47). Correlations of estimated true intake with DHQ-reported intake were slightly higher for whites (r = 0.36) than for blacks (r = 0.34).

Underreporting and overreporting by dietary assessment method

Underreporting and overreporting of energy intake as measured by the ratio of REI to TEE multiplied by 100 are displayed graphically by race for DietDay versus TEE (Figure 2A) and DHQ versus TEE (Figure 2B). There was less underreporting and a narrower range of underreporters with DietDay than with the DHQ. On the basis of the dotted lines that show the cutoff for those who overreported by more than 20% in Figure 2B, it can be seen that blacks appeared much more likely to overreport greatly with the DHQ. The percentages of underreporters and overreporters in each category by gender, race, age, and BMI are presented in Table 5. In all cases, the DietDay 24-hour recall had fewer underreporters (30% reported <70% of intake, 44% reported <80% of intake, and 14% reported >120% of intake) than did the DHQ diet assessment (46%, 56%,

and 15%, respectively). Whites appeared more likely to underreport by 30% or more than did blacks. More whites overreported with DietDay and more blacks overreported with the DHQ. Tests of significance on the differences in the degree of underreporting between DietDay and DHQ assessments indicated that ethnicity played a role in differences in both underreporting and overreporting, with more substantial underreporting and overreporting among whites, regardless of the method of assessment.

DISCUSSION

The surprising finding of the OPEN Study that even in a culture with great intraindividual dietary variability, a few days of recall outperform an FFQ was confirmed in this study. We have demonstrated that a low-cost dietary assessment approach (Web-based recalls) can provide valid dietary intake reports in this volunteer population of black and white adults. Given that these findings have external validity, this fact opens the door for large-scale studies of gene-nutrient interactions with repeated measures of diet at frequent intervals in these groups.

The underreporting and misreporting of dietary intakes is nontrivial and has become the focus of statistical approaches to improve the precision of estimates of dietary effect. When the magnitudes of the correlations with objective recovery biomarkers are within an acceptable range and adequate for determination of attenuation coefficients, they can be applied to more precisely relate dietary intakes to health outcomes of interest. This has been demonstrated by the work on attenuation factors and their applications (17) and the development of regression calibration approaches to improving the estimates of effect in epidemiologic studies (19). As Prentice et al. (19) note, although the use of attenuation factors allows better approximation of the magnitude of effect, simple deattenuation does not enhance the power to detect associations. However, adjustment for the distorting effects of systematic biases can dramatically change epidemiologic findings (20). The present study identified areas of systematic bias by age, gender, and BMI that would benefit from biomarker-based calibration.

Additional contributions from this study include the encouraging finding that, unlike the experience with FFQs, which tend to have mixed results in different ethnic groups (with dishearteningly lower validity among minorities (2–4, 21)), the 24-hour recall performed equally well, if not better, among the blacks studied here than among the whites. We also report the finding that, for energy estimation, 2 or 3 days of recall might adequately characterize individuals of both races. This open question of the number of days needed to adequately characterize an individual is a fundamental part of decisions regarding the resource investment in capturing and managing dietary data. The number of days is also a major determinant of the subject burden of the diet assessment and affects the quality of the reports in the population at large. Although the number of days needed to assess impact depends on the nutrient under study and is driven by the intraindividual to interindividual variability in intake of that nutrient, in the case of energy intake, this short reference period in a population that adheres to a greatly varied diet is encouraging.

There remain a number of limitations to these findings. First of all, the population, although diverse in age, educational level, and BMI, was a select group of adult volunteers. The applicability of these findings to population-based studies as well as to other age, geographic, and ethnic groups will require validation. Second, measurement error is unavoidable in human studies. This has led to deattenuation approaches to derive more accurate risk estimates (17). However, bias in reporting is more noxious and can lead to inappropriate conclusions. Attention to bias is a prerequisite, and approaches such as regression calibration are being developed to take biases into account (11, 20). BMI in particular appears to remain correlated with biased reporting of intake, the degree of which depends upon the dietary assessment method, as found in this analysis. In our populations, the bias association with overweight and obesity is less extreme with a 24-hour recall than with an FFQ.

Another important point is that, despite the relatively large size of the present study, the derivation of precise attenuation estimates by race, age, educational level, BMI, and gender requires a much larger data pool. With this in mind, a pooling project led by the US National Cancer Institute is under way. This project will combine all of the large DLW-based validation studies to better understand measurement error using different dietary assessment tools and the incremental value of each additional day of recall.

Although this will add to our knowledge base, even this endeavor will not support extrapolation to other ethnic groups, other age groups, and other birth cohorts. The need for valid and therefore well-validated dietary assessment in other populations, both healthy and ill, will continue, as will granularity on the questions of which days of the week and which times of the year provide the best information if habitual intake is of interest. Preliminary results suggest that these might differ by age and race.

REFERENCES

1. Coates RJ, Monteilh CP. Assessments of food-frequency questionnaires in minority populations. *Am J Clin Nutr*. 1997; 65(suppl 4):1108S–1115S.
2. Stram DO, Hankin JH, Wilkens LR, et al. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. *Am J Epidemiol*. 2000;151(4):358–370.
3. 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(10):856–869.
4. Liu K, Slattery M, Jacobs D Jr, et al. A study of the reliability and comparative validity of the CARDIA dietary history. *Ethn Dis*. 1994;4(1):15–27.
5. Schatzkin A, Kipnis V, Carroll RJ, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) Study. *Int J Epidemiol*. 2003;32(6):1054–1062.
6. Willett WC. *Nutritional Epidemiology*. 2nd ed. New York, NY: Oxford University Press; 1998.
7. Basiotis PP, Welsh SO, Cronin FJ, et al. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. *J Nutr*. 1987;117(9): 1638–1641.
8. Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *Am J Clin Nutr*. 2008; 88(2):324–332.
9. Subar AF, Kipnis V, Troiano RP, et al. Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN Study. *Am J Epidemiol*. 2003;158(1):1–13.
10. Arab L, Wesseling-Perry K, Jardack P, et al. Eight self-administered 24-hour dietary recalls using the Internet are feasible in African Americans and whites: the Energetics Study. *J Am Diet Assoc*. 2010;110(6):857–864.
11. Neuhouser ML, Tinker L, Shaw PA, et al. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol*. 2008;167(10): 1247–1259.
12. Arab L, Hahn H, Henry J, et al. Using the web for recruitment, screen, tracking, data management, and quality control in a dietary assessment clinical validation trial. *Contemp Clin Trials*. 2010;31(2):138–146.
13. Smith A. Effects of caffeine on human behavior. *Food Chem Toxicol*. 2002;40(9):1243–1255.
14. Ainsworth BE, Macera CA, Jones DA, et al. Comparison of the 2001 BRFSS and the IPAQ Physical Activity Questionnaires. *Med Sci Sports Exerc*. 2006;38(9):1584–1592.
15. Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. *J Nutr*. 1988; 118(11):1278–1289.
16. Cole TJ, Coward WA. Precision and accuracy of doubly labeled water energy expenditure by multipoint and two-point methods. *Am J Physiol*. 1992;263(5):E965–E973.
17. Kipnis V, Subar AF, Midthune D, et al. Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol*. 2003;158(1):14–21.
18. Spiegelman D, Zhao B, Kim J. Correlated errors in biased surrogates: study designs and methods for measurement error correction. *Stat Med*. 2005;24(11):1657–1682.
19. Prentice RL, Huang Y, Tinker LF, et al. Statistical aspects of the use of biomarkers in nutritional epidemiology research. *Stat Biosci*. 2009;1(1):112–123.
20. Prentice RL, Shaw PA, Bingham SA, et al. Biomarker-calibrated energy and protein consumption and increased cancer risk among postmenopausal women. *Am J Epidemiol*. 2009;169(8):977–989.
21. Mayer-Davis EJ, Vitolins MZ, Carmichael SL, et al. Validity and reproducibility of a food frequency interview in a multi-cultural epidemiology study. *Ann Epidemiol*. 1999;9(5):314–324.