

# Validation of nutrient intake using an FFQ and repeated 24 h recalls in black and white subjects of the Adventist Health Study-2 (AHS-2)

Karen Jaceldo-Siegl<sup>1,2,\*†</sup>, Synnove F Knutsen<sup>3</sup>, Joan Sabaté<sup>1,2</sup>, W Lawrence Beeson<sup>1</sup>, Jacqueline Chan<sup>1</sup>, R Patti Herring<sup>3</sup>, Terrence L Butler<sup>1</sup>, Ella Haddad<sup>2</sup>, Hannelore Bennett<sup>1</sup>, Susanne Montgomery<sup>3</sup>, Shalini S Sharma<sup>1</sup>, Keiji Oda<sup>1</sup> and Gary E Fraser<sup>1</sup>

<sup>1</sup>Department of Epidemiology & Biostatistics, School of Public Health, Loma Linda University, Loma Linda, CA, USA; <sup>2</sup>Department of Nutrition, School of Public Health, Loma Linda University, Loma Linda, CA, USA;

<sup>3</sup>Department of Health Promotion, School of Public Health, Loma Linda University, Loma Linda, CA, USA

Submitted 15 June 2009; Accepted 19 October 2009; First published online 8 December 2009

## Abstract

**Objective:** To validate a 204-item quantitative FFQ for measurement of nutrient intake in the Adventist Health Study-2 (AHS-2).

**Design:** Calibration study participants were randomly selected from the AHS-2 cohort by church, and then subject-within-church. Each participant provided two sets of three weighted 24 h dietary recalls and a 204-item FFQ. Race-specific correlation coefficients (*r*), corrected for attenuation from within-person variation in the recalls, were calculated for selected energy-adjusted macro- and micronutrients.

**Setting:** Adult members of the AHS-2 cohort geographically spread throughout the USA and Canada.

**Subjects:** Calibration study participants included 461 blacks of American and Caribbean origin and 550 whites.

**Results:** Calibration study subjects represented the total cohort very well with respect to demographic variables. Approximately 33% were males. Whites were older, had higher education and lower BMI compared with blacks. Across fifty-one variables, average deattenuated energy-adjusted validity correlations were 0.60 in whites and 0.52 in blacks. Individual components of protein had validity ranging from 0.40 to 0.68 in blacks and from 0.63 to 0.85 in whites; for total fat and fatty acids, validity ranged from 0.43 to 0.75 in blacks and from 0.46 to 0.77 in whites. Of the eighteen micronutrients assessed, sixteen in blacks and sixteen in whites had deattenuated energy-adjusted correlations  $\geq 0.4$ , averaging 0.60 and 0.53 in whites and blacks, respectively.

**Conclusions:** With few exceptions validity coefficients were moderate to high for macronutrients, fatty acids, vitamins, minerals and fibre. We expect to successfully use these data for measurement error correction in analyses of diet and disease risk.

**Keywords**  
Epidemiological methods  
Ethnic groups  
Questionnaires  
Validation studies

The self-administered FFQ, although a more practical tool for assessing individual diets in large cohorts, has associated measurement errors that usually attenuate relative risk estimates in studies of diet and disease risk<sup>(1–3)</sup>. One approach to improving effect estimates is through a calibration study where assessment of individual diets by questionnaire is compared with a more precise method, such as repeated 24 h recalls<sup>(4)</sup>. The goal of the present report is to assess the validity of nutrient intake as measured by the FFQ used in the Adventist Health Study-2

(AHS-2) by comparison with repeated 24 h dietary recalls in a sample of black and white Adventists.

The AHS-2, a prospective study of over 96 000 adult Adventists in the USA and Canada, has the goal of relating lifestyle choices, particularly diet, to health outcomes<sup>(5)</sup>. The AHS-2 cohort is relatively unique for its wide range of dietary patterns compared with the general Western population. Specifically, 8% are vegan, 28% are lacto-ovo-vegetarian, 16% are semi- or pesco-vegetarian, and 48% are non-vegetarian. There is also large variation in the consumption of plant foods, such as nuts, soya, other legumes and grains. The FFQ administered in the study was designed specifically to assess the wide range of

† Correspondence address: 24785 Stewart St, Loma Linda University, Loma Linda, CA 92350, USA.

\*Corresponding author: Email [kjaceldo@llu.edu](mailto:kjaceldo@llu.edu)

© The Authors 2009

dietary intake among a population where a large proportion is vegetarian. A calibration and validation study is an integral part of the AHS-2; the current paper describes that study and examines the validity of macro- and micronutrient intakes estimated from FFQ data.

## Experimental methods

### Calibration sub-study design

The AHS-2 cohort consists of adult members of Seventh-day Adventist churches geographically spread throughout the USA and Canada. Over 25 000 cohort members are blacks; the remaining subjects are of other races, mostly white. Recruitment and selection methods of the parent cohort have been described previously<sup>(5)</sup>. Participants in the calibration study were randomly selected from the parent cohort by church, and then subject-within-church. Because of the special focus on black Adventists as a minority group in the AHS-2, the calibration study was designed such that there would be approximately equal numbers of blacks and whites. No significant differences in the distribution of gender, age, education or vegetarian status between the calibration sub-study sample (*n* 1011) and the AHS-2 cohort (*n* 96 592) were observed. When

stratifying by ethnicity, only the distributions of mean age and age categories in blacks were nominally significantly different ( $P = 0.0018$  and  $0.0032$ , respectively) between the parent cohort and the calibration study sample (Table 1). This is easily consistent with chance given the multiple testing involved.

The duration of the calibration study was 9–12 months for any one subject (see Fig. 1). We obtained a set of three variably timed 24 h dietary recalls (one Saturday, one Sunday and one weekday intake) during the first two months, which then was duplicated approximately six months later. The aim was to obtain two sets of recalls (a total of six 24 h recalls) from each subject. During the 6-month interval, subjects completed a self-administered food FFQ that was identical to the baseline FFQ and attended a local study clinic. The study was approved by the institutional review board of Loma Linda University, and all subjects gave written informed consent.

### Dietary assessment

The 24 h dietary recalls were unannounced and obtained by telephone. A two-dimensional food portion visual (2D Food Portion Visual; Nutrition Consulting Enterprises, Framingham, MA, USA) was sent to each participant before the first recall to assist with portion size estimates.

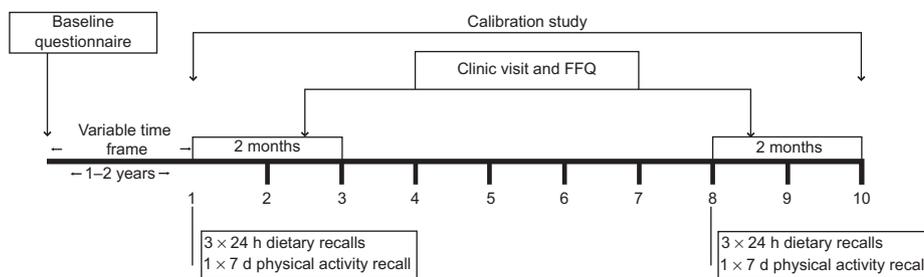
**Table 1** Comparison of selected characteristics between the Adventist Health Study-2 (AHS-2) cohort and calibration study participants according to race, 2003–8

	AHS-2 cohort† ( <i>n</i> 96 592)				Calibration study‡ ( <i>n</i> 1011)			
	Whites ( <i>n</i> 69 740)		Blacks ( <i>n</i> 25 286)		Whites ( <i>n</i> 550)		Blacks ( <i>n</i> 461)	
	% or Mean	SD	% or Mean		% or Mean	SD	% or Mean	SD
Gender								
Males (%)	37.0		29.8		35.5		30.7	
Females (%)	63.0		70.2		64.5		69.3	
Age (years)§*								
<50 (%)	27.4		43.0		25.9		35.3	
50–59 (%)	23.2		25.7		23.6		29.4	
60–69 (%)	21.6		17.3		22.5		20.8	
70–79 (%)	18.1		10.4		19.3		12.3	
80 or more (%)	9.7		3.6		8.7		2.2	
Mean age (years)§**	60.3	14.4	53.8	13.7	60.6	13.9	55.6	12.7
Education								
High school or less (%)	21.6		24.5		20.5		23.7	
Some college (%)	39.0		40.9		38.2		42.4	
College graduate or more (%)	39.4		34.6		41.3		33.9	
Diet patterns								
Vegan (%)	9.8		7.1		11.5		8.7	
Lacto-ovo-vegetarian (%)	34.3		13.7		35.6		12.9	
Pesco-vegetarian (%)	8.4		14.4		9.8		15.0	
Semi-vegetarian (%)	6.0		3.8		5.2		1.9	
Non-vegetarian (%)	41.4		60.9		37.9		61.5	
BMI (kg/m <sup>2</sup> )								
<25 (%)	44.8		29.7		46.9		29.6	
25–29.9 (%)	33.6		36.9		33.5		36.1	
>30 (%)	21.6		33.4		19.6		34.3	
Mean BMI (kg/m <sup>2</sup> )	26.6	5.9	28.6	6.4	26.4	5.7	28.8	6.4

†Proportion of missing data: ethnicity, 1.4%; gender, 1.4%; age, 1.7%; education, 2.8%; dietary patterns, 4.6%; BMI, 6.2%.

‡Proportion of missing data: education, 1.8%; dietary patterns, 3.2%; BMI, 5.3%.

§Differences in the age distributions between the AHS-2 cohort and the calibration study participants were significant only in blacks (\* $P = 0.0032$ , \*\* $P = 0.0018$ ); otherwise, no significant differences in the distribution of gender, age, education or vegetarian status between the AHS-2 cohort and calibration study participants according to race were observed.



**Fig. 1** Adventist Health Study-2 study design

Trained research dietitians used standard probes and a multiple-pass approach methodology to collect detailed information on all foods, beverages and supplements consumed by each subject during the previous 24 h. Each recall interview was entered using Nutrition Data System for Research (NDS-R) version 4.06 or 5.0 (The Nutrition Coordinating Center, Minneapolis, MN, USA) and the conversation digitally recorded for subsequent quality check. An experienced research dietitian later evaluated randomly selected recall interviews (~5%) and compared them with the recording, as a quality control measure.

The AHS-2 FFQ is a quantitative and comprehensive 22-page instrument consisting of 204 foods, fifty-four questions about food preparation and forty-six fields for open-ended questions. Frequency categories vary with food type to allow respondents to define their daily intake with greater specificity. Thus, the lowest category for most foods and beverages is never or rarely, and for cold cereals and vegetarian protein products, 1–3 per month. The highest frequency category for vegetables, soups, cereals, pasta, dressing, meats, fish, vegetarian protein products and soya milk is 2 or more per day; for nuts and seeds, 4 or more per day; for breads, eggs, dairy products, snacks and beverages, 6 or more per day. Portion sizes include three levels: standard,  $\frac{1}{2}$  or less, and  $1\frac{1}{2}$  or more. Standard portions are based on serving sizes using familiar household units such as cup, tablespoon, slice, patty and others. Pictures of common foods or beverages typically served together were included with the questionnaire to assist subjects in estimating portion sizes. For example, to represent a standard portion, spaghetti, broccoli and steak are arranged on an 11-inch dinner plate, each food measured in a standard portion. Representations of  $\frac{1}{2}$  of standard and  $1\frac{1}{2}$  times the standard portion size are provided as well, using the same foods and dinner plate.

The AHS-2 FFQ was designed to include foods commonly eaten by US Adventists and later modified to accommodate foods specific to black Adventists of US and Caribbean origin<sup>(6)</sup>. The questionnaire was sent to each subject, completed at home, and then mailed back to AHS-2. Respondents were asked to report on their intake over the previous one year. Upon receipt of the questionnaire, study personnel reviewed the

questionnaire for completeness and as necessary followed-up by telephone to clarify any ambiguous or incomplete information.

Completed questionnaires were optically scanned using the NCS 5000i Image Scanner with ScanTools Plus software (Pearson NCS, Bloomington, MN, USA). Standardized processing of open-ended questions was done using the Food Write-In Processing software (Adventist Health Study-2, Loma Linda, CA, USA), a network-based application created in Microsoft<sup>®</sup> Access (Microsoft Corporation, Redmond, WA, USA).

Nutrient composition of foods reported from 24 h recalls and FFQ were based on the NDS-R 5.0\_35 database (The Nutrition Coordinating Center), an analytic database of over 20 000 foods updated annually while maintaining nutrient profiles true to the version used for data collection<sup>(7)</sup>. Nutrient profiles of foods and supplements not found in the NDS database were obtained from the US Department of Agriculture, manufacturers, and the Caribbean Food and Nutrition Institute. Considerable attention was given to creating recipes for home-cooked vegetarian dishes ( $n > 500$ ), home-made and commercial soya and nut milks ( $n > 180$ ) and commercial meat analogues ( $n 309$ ) frequently consumed among our study population. For the latter we contacted manufacturers or worked with a senior food technologist with experience in this industry, to create recipes.

### Statistical analysis

Within each of the two sets of 24 h dietary recalls, each day was weighted appropriately to produce a synthetic week (Saturday intake + Sunday intake +  $5 \times$  weekday intake) and then divided by 7 to obtain mean daily nutrient estimate. Dietary habits among Adventists are often rather different on Saturdays, Sundays and typical weekdays. Thus, for those who completed six 24 h recalls ( $n 950$ ), mean daily intake was calculated by dividing the total of the two synthetic weeks by 14. For those sixty-one subjects who completed only three, four or five recalls, only the first three recalls (first synthetic week) were used in these analyses. Nutrient estimates from FFQ data were calculated using the product-sum method<sup>(8)</sup>. Thus, nutrient intake = sum [(weighted frequency of use of a food)  $\times$  (weighted portion size consumed of that

food)  $\times$  (amount of that nutrient in a standard serving size of that food)]. FFQ with estimated energy intake  $<2093$  kJ ( $<500$  kcal) or  $>18\,833$  kJ ( $>4500$  kcal) were excluded from analyses. Arithmetic means and standard deviations were calculated for nutrient intakes assessed by 24 h recalls (R) and FFQ (Q). To determine differences, Student's *t* tests were used for continuous variables and  $\chi^2$  tests for categorical variables.

Nutrients were energy-adjusted using the residual method<sup>(9)</sup>. However, for vitamins, minerals and fatty acids also commonly used as supplements, unadjusted supplemental intake was added to energy-adjusted dietary values, resulting in energy-adjusted total intake. Nutrient validity correlations,  $\text{corr}(R,Q)$ , were calculated between averages of the two synthetic weeks of recalls and intake measured from FFQ. Where the linear regression implied by the correlation coefficient had residuals with skew  $\geq 2.0$ , both dietary quantities were log-transformed.

After energy adjustment, we determined correlations for  $\text{corr}(Q,R)$  and then corrected the correlation coefficients for attenuation due to within-person variation in the recalls<sup>(3)</sup>. The deattenuated energy-adjusted correlation coefficients,  $r_c$ , that we report are estimates of  $\text{cov}(Q,T)/\sqrt{\text{var}(Q) \times \text{var}(T)}$ , where  $\text{var}(T)$  is estimated using the repeated measurements of R. This allowed us to also use the 6% of subjects who did not complete the second week of recalls, assuming that their within-person variances did not systematically differ from the others.

A bias factor was calculated as the proportional bias,  $(\text{crude} - \text{true})/\text{true}$ , that would be observed in a regression coefficient if the uncorrected FFQ estimate were used as the independent variable in a regression where in fact the mean of a large number of recall estimates is the truth.

We calculated 95% confidence intervals for all validity coefficients using bootstrap resampling and the BCa method<sup>(10)</sup>. Analyses were performed using the SAS statistical software package version 9.1 (SAS Institute Inc., Cary, NC, USA).

## Results

The 1011 subjects in the calibration study represented the total AHS-2 cohort very well, both with respect to demographic variables and dietary practices (Table 1). There were more females than males regardless of race. Whites were older, had higher education and lower BMI compared with blacks. A greater proportion of blacks than whites were semi- and non-vegetarians.

Nutrient intake estimates from the FFQ and the 24 h recalls by race are shown in Table 2. Intakes of nutrients estimated from 24 h recalls were in general higher for whites than blacks, except for *trans* and very-long-chain fatty acids, protein, animal protein and vitamin D. Nutrient intake estimates from the FFQ were also generally higher in whites; however, blacks had higher intakes of

very-long-chain fatty acids, total energy, protein, animal protein, carbohydrate, non-fibre carbohydrate, Fe, K, and vitamins D and C.

All uncorrected and deattenuated correlation coefficients reported for selected macro- and micronutrients were energy-adjusted (Table 3 and Fig. 2). Correlations improved after correction for attenuation in the recalls on average by 28% and 34% in white and black subjects, respectively. Average deattenuated validity correlations were higher in whites compared with blacks ( $r=0.60$  and  $0.52$ , respectively). The lowest validity correlations in both races were for energy:  $r=0.09$  (95% CI  $-0.01, 0.18$ ) in blacks and  $r=0.24$  (95% CI  $0.15, 0.32$ ) in whites. Nutrients with the highest deattenuated correlations were docosapentaenoic acid (22:5n-3) in blacks ( $r=0.75$ , 95% CI  $0.70, 0.79$ ) and animal protein in whites ( $r=0.85$ , 95% CI  $0.83, 0.87$ ). Individual components of protein (animal, vegetable, soya and dairy protein) had validity correlations ranging from  $0.40$  to  $0.68$  in blacks and from  $0.63$  to  $0.85$  in whites, but total protein was much lower in both races ( $r=0.39$  and  $0.18$  in whites and blacks, respectively; see further comment in the discussion section). Deattenuated correlations for total fat and fatty acids ranged from  $0.43$  to  $0.75$  in blacks and from  $0.46$  to  $0.77$  in whites. Regardless of race, fibre and its components had higher validity compared with total carbohydrate. Of the eighteen micronutrients assessed, sixteen in blacks and sixteen in whites had deattenuated correlations  $\geq 0.4$ , averaging  $0.60$  and  $0.53$  in whites and blacks, respectively.

The average bias factor was greater in blacks compared with whites ( $-0.70$  and  $-0.58$ , respectively), but in both ethnic groups it is clear that use of uncorrected FFQ data in a regression would result in severely attenuated effect estimates.

## Discussion

Participants of this relatively large ( $n=1011$ ) calibration study were representative of the parent cohort with respect to gender, age, education, BMI and dietary patterns. Because of our initial selection by church, subjects also were geographically spread throughout the USA and Canada. As a reference measure the present study obtained six 24 h recalls, whereas other cohorts with large calibration samples typically used fewer recalls<sup>(4,11,12)</sup> or a 4 d food record<sup>(13)</sup> that did not systematically combine weekdays and weekends.

Intake estimates of most nutrients from both FFQ and recall data generally were higher in whites compared with blacks. However, blacks reported higher intakes of protein, animal protein, very-long-chain fatty acids and vitamin D. This is not surprising, given the greater proportion of non- and semi-vegetarian eating patterns followed by this minority population in the study. As expected, nutrient intake

**Table 2** Comparison of nutrient intake as measured by the FFQ and 24 h recall† in the Adventist Health Study-2 calibration study participants, 2003–8

Nutrient	24 h recalls‡				FFQ‡			
	Whites		Blacks		Whites		Blacks	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<b>Energy and fatty acids</b>								
Energy (kJ)	6886**	1719	6004	1776	7839	2868	7898	3499
Fat (g)	55.4	7.4	54.6	8.9	71.4**	15.8	66.5	18.4
Saturated fat (g)	14.8**	3.8	14.4	3.9	16.2	5.5	14.7	5.6
Monounsaturated fat (g)	21.3	4.6	21.1	3.7	29.3*	9.6	27.3	10.5
Polyunsaturated fat (g)	15.0	3.9	14.7	3.1	20.7*	5.3	19.4	6.1
<i>Trans</i> fatty acids								
18:1 <i>n</i> -9 (g)	2.7*	1.6	3.0	1.2	3.2*	2.0	2.9	2.1
18:2 <i>n</i> -6 (g)	19.7*	4.4	19.3	3.7	26.4*	9.4	24.4	10.3
18:3 <i>n</i> -3 (g)	13.2	3.6	12.9	2.8	18.4**	4.9	17.0	5.7
20:4 <i>n</i> -6 (g)	1.7*	0.79	1.5	0.62	2.1*	0.82	1.9	0.67
20:5 <i>n</i> -3 (g)	0.046**	0.21	0.086	0.058	0.036**	0.046	0.064	0.063
20:5 <i>n</i> -3 dietary (g)	0.017**	0.039	0.059	0.075	0.018**	0.042	0.047	0.077
20:5 <i>n</i> -3 total (g)	0.029**	0.056	0.070	0.11	0.089	0.21	0.12	0.25
22:5 <i>n</i> -3 (g)	0.009	0.011	0.023	0.021	0.023	0.048	0.040	0.060
22:6 <i>n</i> -3 dietary (g)	0.036**	0.056	0.11	0.13	0.051	0.12	0.11	0.17
22:6 <i>n</i> -3 total (g)	0.046	0.06	0.13	0.18	0.10	0.20	0.18	0.27
<i>n</i> -3 total (g)	1.7	0.80	1.6	0.70	2.2	0.87	2.2	0.85
Long-chain fatty acids	0.12	0.14	0.29	0.34	0.26	0.48	0.40	0.61
Dairy fat (g)	5.9*	2.4	4.7	2.9	4.8*	5.1	3.7	4.4
<b>Proteins</b>								
Protein (g)	54.8**	7.4	59.1	9.0	71.8	13.6	72.3	18.0
Animal protein (g)	19.5**	4.4	25.7	15.8	16.4*	13.4	18.9	15.8
Vegetable protein (g)	35.2*	5.9	33.2	11.0	54.5	16.4	52.5	19.8
Soya protein (g)	6.2	5.0	6.4	5.4	6.1	6.3	6.5	7.5
Dairy protein (g)	7.6**	2.7	5.3	3.6	8.3**	8.4	5.8	6.6
<b>Carbohydrate and fibre</b>								
Carbohydrate (g)	223*	14.9	218	26.8	253*	38.9	263	50.1
Total fibre (g)	24.0**	4.9	21.4	6.7	34.2	8.7	33.3	10.3
Soluble fibre (g)	7.1**	2.7	6.5	1.8	9.5	2.6	9.6	3.5
Insoluble fibre (g)	16.1**	4.0	14.1	4.8	22.5*	6.4	21.2	7.6
Non-fibre carbohydrate (g)	199	14.1	196	23.3	219**	35.6	230	45.6
<b>Minerals</b>								
Ca dietary (mg)	692**	26.3	614	105	833*	277	783	262
Ca total (mg)	993**	31.5	767	279	1252*	557	1041	504
P (mg)	1014*	31.8	981	127	1272*	222	1230	259
Mg (mg)	384**	19.6	325	96.0	511**	200	463	174
Fe dietary (mg)	14.8*	3.8	13.7	2.7	18.3	3.9	18.0	5.0
Fe total (mg)	21.0	20.5	19.3	18.0	27.2	32.1	30.5	41.7
Zn (mg)	16.2**	11.7	12.7	6.5	17.8	11.0	15.4	9.8
Cu (mg)	2.2**	1.5	1.8	0.70	2.9*	2.6	2.6	1.3
Na (mg)	2435*	49.3	2342	376	3599*	1567	3218	1828
K (mg)	2536**	50.4	2381	443	3412	723	3524	952
<b>Vitamins</b>								
Vitamin D dietary (µg)	2.9*	1.5	3.3	1.9	2.9	2.6	2.9	2.3
Vitamin D total (µg)	8.5	7.9	14.8	13.8	8.8	8.0	8.1	8.0
Vitamin E dietary (µg)	8.8**	3.0	7.7	2.0	11.3**	4.8	9.9	3.8
Vitamin E total (µg)	54.2**	7.4	26	41.1	87.6*	123	72.2	118
Vitamin C dietary (mg)	116	10.8	109	38.4	175**	83.3	203	123
Vitamin C total (mg)	304	367	268	322	467	479	495	503
Thiamin (mg)	8.5	17.0	8.1	9.4	10.3	17.8	10.0	15.9
Riboflavin (mg)	7.5	14.6	4.8	7.5	8.3	12.1	7.1	15.2
Niacin (mg)	36.9	40.7	31.1	31.1	45.9	41.6	40.4	36.3
Pantothenic acid (mg)	14.4*	18.9	10.6	10.4	16.9	27.3	14.3	20.9
Vitamin B <sub>6</sub> (mg)	10.5**	23.5	6.7	10.1	11.9	21.8	11.3	21.5
Folate (µg)	634	25.2	534	207	799*	382	746	373
β-Carotene (µg)	4550**	6920	4817	2553	9604**	12 579	11 326	13 338

†Average of two sets of three weighted 24 h recalls.

‡Intakes of these nutrients were significantly different in blacks compared with whites: \* $P < 0.05$ , \*\* $P < 0.0001$ .

estimates derived from FFQ tended to be higher than those from 24 h recalls. This probably results when people are asked to recall the frequency of intake of a large number of foods in an FFQ and thus tend to overestimate their actual

intake<sup>(9)</sup>. In a study that assessed bias of food frequency-based measures of fruit and vegetable intakes, Kristal and colleagues compared intake estimates between recalls (or food records) *v.* a brief questionnaire with only seven items

**Table 3** Energy-adjusted validity correlation† between the FFQ and 24h dietary recalls in Adventist Health Study-2 calibration study participants, 2003–8

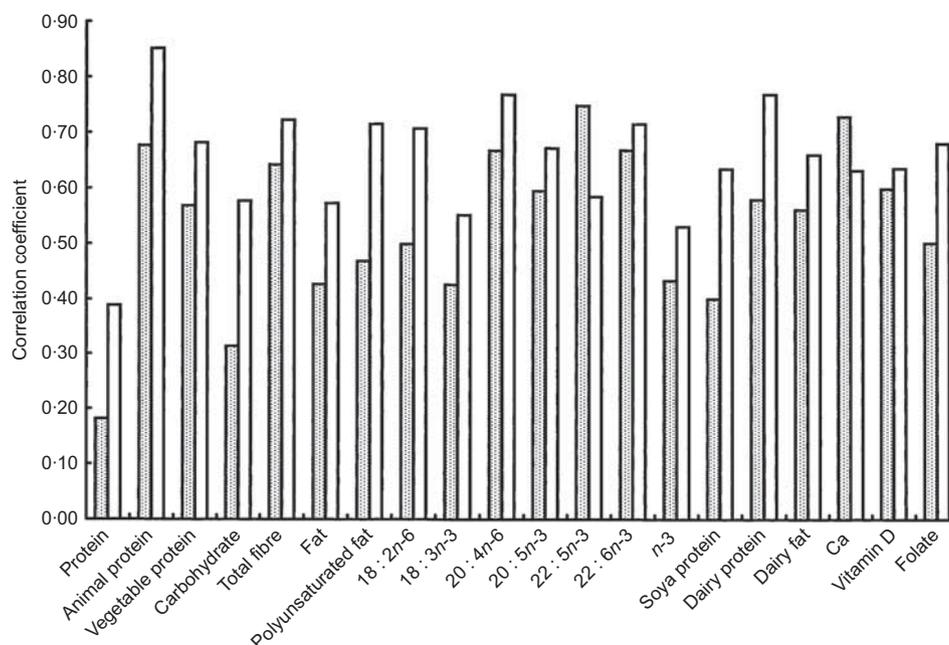
Nutrient	Whites				Blacks			
	Uncorrected correlation	Deattenuated‡ correlation	95% CI	Bias factor§	Uncorrected correlation	Deattenuated‡ correlation	95% CI	Bias factor§
<b>Energy and fatty acids</b>								
Energy	0.20	0.24	0.15, 0.32	-0.85	0.07	0.09	-0.01, 0.18	-0.96
Total fat	0.42	0.57	0.51, 0.63	-0.69	0.32	0.43	0.34, 0.50	-0.79
Saturated fat	0.59	0.71	0.67, 0.75	-0.41	0.36	0.45	0.37, 0.52	-0.69
Monounsaturated fat	0.34	0.48	0.41, 0.54	-0.79	0.30	0.45	0.37, 0.53	-0.84
Polyunsaturated fat	0.47	0.72	0.67, 0.75	-0.58	0.32	0.47	0.39, 0.54	-0.76
Trans fatty acids	0.35	0.51	0.45, 0.57	-0.71	0.35	0.51	0.44, 0.58	-0.72
18:1n-9	0.32	0.46	0.38, 0.52	-0.81	0.31	0.45	0.37, 0.53	-0.84
18:2n-6	0.44	0.71	0.66, 0.75	-0.62	0.34	0.5	0.42, 0.57	-0.76
18:3n-3	0.40	0.55	0.49, 0.61	-0.38	0.27	0.43	0.34, 0.50	-0.67
20:4n-6	0.56	0.77	0.73, 0.80	-0.38	0.52	0.67	0.61, 0.72	-0.39
20:5n-3 dietary	0.40	0.58	0.52, 0.64	-0.67	0.41	0.61	0.54, 0.67	-0.61
20:5n-3 total	0.54	0.67	0.62, 0.72	-0.65	0.43	0.60	0.53, 0.65	-0.67
22:5n-3	0.42	0.58	0.53, 0.64	-0.82	0.51	0.75	0.70, 0.79	-0.73
22:6n-3 dietary	0.40	0.57	0.51, 0.63	-0.77	0.45	0.68	0.62, 0.73	-0.68
22:6n-3 total	0.56	0.72	0.67, 0.75	-0.66	0.47	0.67	0.61, 0.72	-0.62
n-3 total	0.39	0.53	0.47, 0.59	-0.54	0.30	0.43	0.35, 0.51	-0.68
Long-chain fatty acids	0.47	0.62	0.56, 0.67	-0.76	0.50	0.68	0.62, 0.73	-0.68
Dairy fat	0.53	0.66	0.61, 0.71	-0.38	0.38	0.56	0.49, 0.62	-0.63
<b>Proteins</b>								
Total protein	0.29	0.39	0.31, 0.46	-0.74	0.12	0.18	0.09, 0.27	-0.91
Animal protein	0.76	0.85	0.83, 0.87	-0.06	0.59	0.68	0.62, 0.73	-0.33
Vegetable protein	0.57	0.68	0.63, 0.73	-0.57	0.46	0.57	0.50, 0.63	-0.68
Soya protein	0.46	0.63	0.58, 0.68	-0.57	0.31	0.40	0.32, 0.48	-0.82
Dairy protein	0.64	0.77	0.73, 0.80	-0.46	0.45	0.58	0.51, 0.64	-0.68
<b>Carbohydrate and fibre</b>								
Total carbohydrate	0.45	0.58	0.52, 0.63	-0.62	0.26	0.31	0.22, 0.40	-0.83
Total fibre	0.63	0.72	0.68, 0.76	-0.40	0.56	0.64	0.58, 0.70	-0.58
Soluble fibre	0.57	0.70	0.65, 0.74	-0.49	0.40	0.50	0.43, 0.57	-0.74
Insoluble fibre	0.59	0.70	0.65, 0.74	-0.43	0.55	0.64	0.58, 0.70	-0.59
Non-fibre carbohydrate	0.40	0.54	0.48, 0.60	-0.67	0.21	0.27	0.18, 0.36	-0.86
<b>Minerals</b>								
Ca dietary	0.39	0.53	0.47, 0.59	-0.68	0.31	0.54	0.46, 0.60	-0.78
Ca total	0.54	0.63	0.58, 0.68	-0.55	0.61	0.73	0.68, 0.77	-0.60
P	0.45	0.59	0.53, 0.64	-0.62	0.31	0.44	0.36, 0.52	-0.78
Mg	0.49	0.56	0.50, 0.62	-0.61	0.52	0.61	0.55, 0.67	-0.66
Fe dietary	0.37	0.54	0.48, 0.60	-0.56	0.33	0.48	0.40, 0.55	-0.74
Fe total	0.50	0.6	0.54, 0.65	-0.50	0.43	0.54	0.46, 0.60	-0.67
Zn	0.58	0.65	0.60, 0.70	-0.38	0.46	0.57	0.50, 0.63	-0.60
Cu	0.29	0.33	0.25, 0.41	-0.88	0.48	0.59	0.52, 0.65	-0.69
Na	0.18	0.29	0.21, 0.37	-0.93	0.21	0.33	0.24, 0.42	-0.93
K	0.46	0.58	0.52, 0.63	-0.63	0.42	0.54	0.46, 0.60	-0.75
<b>Vitamins</b>								
Vitamin D dietary	0.50	0.65	0.59, 0.69	-0.58	0.33	0.51	0.43, 0.57	-0.72
Vitamin D total	0.56	0.64	0.58, 0.68	-0.47	0.50	0.60	0.53, 0.66	-0.55
Vitamin E dietary	0.29	0.39	0.31, 0.46	-0.75	0.31	0.48	0.40, 0.55	-0.74
Vitamin E total	0.67	0.76	0.72, 0.79	-0.57	0.40	0.45	0.37, 0.52	-0.84
Vitamin C dietary	0.34	0.46	0.39, 0.53	-0.76	0.37	0.52	0.45, 0.59	-0.84
Vitamin C total	0.58	0.66	0.61, 0.71	-0.50	0.50	0.58	0.51, 0.64	-0.55
Thiamin	0.58	0.63	0.57, 0.68	-0.40	0.46	0.57	0.50, 0.63	-0.61
Riboflavin	0.61	0.66	0.61, 0.71	-0.40	0.51	0.61	0.55, 0.67	-0.60
Niacin	0.57	0.65	0.59, 0.69	-0.53	0.36	0.45	0.36, 0.52	-0.79
Pantotheic acid	0.62	0.69	0.64, 0.73	-0.42	0.50	0.60	0.54, 0.66	-0.60
Vitamin B <sub>6</sub> total	0.63	0.68	0.64, 0.73	-0.44	0.48	0.56	0.49, 0.63	-0.67
Folate total	0.55	0.68	0.63, 0.72	-0.58	0.41	0.50	0.42, 0.57	-0.72
β-Carotene	0.37	0.51	0.44, 0.57	-0.78	0.23	0.35	0.26, 0.43	-0.87

†Uncorrected and deattenuated correlations were energy-adjusted.

‡Corrected for attenuation due to within-person error in the recalls.

§Bias factor is the proportional bias in a univariate regression if the uncorrected FFQ variable is the independent variable.

||Quantities from FFQ and recalls were log-transformed where the linear regression implied by the correlation coefficient had residuals with skew  $\geq 2.0$ .



**Fig. 2** Comparison of selected deattenuated energy-adjusted correlation coefficients between the FFQ and 24 h dietary recalls in black subjects (▨) and white subjects (□) of the Adventist Health Study-2, 2003–8

or an FFQ with thirty or more items<sup>(14)</sup>. They found that estimated intakes were lower from the brief questionnaire, but higher from the longer questionnaire when compared with recalls or food records.

Despite differences between the recalls and FFQ in the current study, deattenuated correlation coefficients were  $\geq 0.40$  for forty-three of fifty-one nutrients examined in both ethnicities combined. Estimates from the FFQ of individual fatty acids, total long-chain fatty acids and fibre were moderately to highly correlated, and the micronutrients (except Na, Cu and dietary vitamin E) were also generally highly correlated with recall data. This shows that the AHS-2 FFQ has the ability to provide relatively good estimates of these nutrients.

Deattenuated correlation coefficients with values less than 0.40 included energy, total protein, total carbohydrate (in blacks), non-fibre carbohydrate (in blacks), Na, dietary vitamin E (in whites) and  $\beta$ -carotene (in blacks). Energy intakes in both races were lower compared with others<sup>(4,12,15–18)</sup>. For carbohydrate, our results in blacks and whites were lower or within the range of those observed by other investigators<sup>(12,15,17,19)</sup>.

Our deattenuated validity correlations for total protein were within the range of  $r^{(4,20)}$  or lower<sup>(15,17,21)</sup> than those reported in other validation studies. It was particularly noticeable in our study that the individual protein components (animal and vegetable protein) that contribute to this composite nutrient had greatly higher coefficients than total protein. In examining the results from the FFQ and recall data, it appears that the questionnaire overestimates the amount of vegetable protein intake, but underestimates the amount of animal protein compared

with the recalls. The most likely explanation is the differential scaling of animal and plant sources of protein in our FFQ, which includes a long list of plant proteins (in part due to the attention we give to meat analogues) but a comparatively shorter list of animal protein foods. In calculating the deattenuated correlation coefficient for total protein, we noted that the variance of animal protein (APRO) from recalls (315.9) was much greater than the variance of animal protein from the FFQ (211.34), yet the variance of vegetable protein (VPRO) from recalls (165.24) was much smaller than the variance of vegetable protein from the FFQ (325.13).

We resolved this problem by putting the animal and vegetable protein FFQ values on the same scale as each other and the recall values. In a separate analysis, we first calculated for both APRO and VPRO  $E(\text{recalls} | \text{FFQ})$ , summed the two estimates and then correlated this with recall values for total protein. Our correlations for total protein subsequently improved from 0.28 to 0.54 in all subjects combined. This interesting observation suggests that the product-sum method of calculating FFQ intakes may cause distortion at times, depending on the form of the FFQ, and we plan to investigate this issue more fully for other variables (e.g. total energy).

In the current validation study energy-adjusted uncorrected correlations improved after correction for attenuation in the recalls, as others have found. We also observed improved correlations in general when supplements were added, which emphasizes not only the importance of assessing both dietary and supplemental sources of micronutrients, but also that supplements contribute importantly to total nutrient intake in this population.

The  $\lambda$  values reported by Johansson *et al.*<sup>(16)</sup> address the same issue as the bias factor ( $\lambda - 1$ ) results in the current study. They illustrate quite dramatically the severe proportional biases that result even when validity correlations are well within values commonly held to be acceptable in FFQ work. For example, a univariate regression where uncorrected energy-adjusted PUFA from the FFQ in white subjects (deattenuated validity correlation of 0.72) was used as the independent variable would produce a  $\beta$  coefficient that was biased downwards by 58%.

When comparing our validation results by race, we found that the deattenuated energy-adjusted validity correlations of several nutrients in our study were smaller in blacks than whites. This is consistent with the reports of others<sup>(4,13,18,20,22)</sup>. Presumably these may be explained by somewhat lesser educational attainment<sup>(13,18)</sup> on average among blacks or less familiarity with research studies of this sort. In an older biracial sample, investigators also reported lower validity in the very old (aged 79 years and older) compared with those aged 68–78 years<sup>(20)</sup>. In our calibration study, a greater proportion of whites achieved higher education, but they were somewhat older compared with blacks.

We provide in this report the validation of an unusually comprehensive list of micronutrients and fatty acids, and with few exceptions find moderate to high validity. Thus we expect to be able to use most of these results to good effect in both traditional analyses of disease risk and analyses incorporating recall- or biomarker-guided measurement error correction.

## Acknowledgements

*Sources of funding:* This work was supported by a grant from the National Cancer Institute RO1 CA94594. *Conflict of interest declaration:* None of the authors have conflicts of interest. *Authors' contributions:* study design, G.E.F.; coordination of data collection, K.J.-S., S.F.K., J.S., W.L.B., J.C., R.P.H., T.L.B., E.H., H.B., S.M.; data analyses, K.O., G.E.F., K.J.-S.; interpretation of data, K.J.-S., S.F.K., J.S., W.L.B.; writing of the manuscript, K.J.-S., G.E.F., S.S.S.; editing of the manuscript, S.F.K., J.S., W.L.B., J.C., R.P.H.

## References

1. Fraser GE (2003) A search for truth in dietary epidemiology. *Am J Clin Nutr* **78**, 3 Suppl., 521S–525S.
2. Thomas D, Stram D & Dwyer J (1993) Exposure measurement error: influence on exposure–disease. Relationships and methods of correction. *Annu Rev Public Health* **14**, 69–93.
3. Willett W (1998) Correction for the effects of measurement error. In *Nutritional Epidemiology*, 2nd ed., pp. 302–320 [W Willett, editor]. New York: Oxford University Press.
4. Stram DO, Hankin JH, Wilkens LR *et al.* (2000) Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. *Am J Epidemiol* **151**, 358–370.

5. Butler TL, Fraser GE, Beeson WL *et al.* (2008) Cohort profile: the Adventist Health Study-2 (AHS-2). *Int J Epidemiol* **37**, 260–265.
6. Akbar JA, Jaceldo-Siegl K, Fraser G *et al.* (2007) The contribution of soul and Caribbean foods to nutrient intake in a sample of blacks of US and Caribbean descent in the Adventist Health Study-2: a pilot study. *Ethn Dis* **17**, 244–249.
7. Schakel SF, Sievert YA & Buzzard IM (1988) Sources of data for developing and maintaining a nutrient database. *J Am Diet Assoc* **88**, 1268–1271.
8. Willett W (1998) Food frequency methods. In *Nutritional Epidemiology*, 2nd ed., pp. 74–100 [W Willett, editor]. New York: Oxford University Press.
9. Willett W & Stampfer M (1998) Implications of total energy intake for epidemiologic analysis. In *Nutritional Epidemiology*, 2nd ed., pp. 273–301 [W Willett, editor]. New York: Oxford University Press.
10. Efron B & Tibshirani RJ (1993) Better bootstrap intervals. In *An Introduction to the Bootstrap*, pp. 184–188. New York: Chapman and Hall.
11. Slimani N, Bingham S, Runswick S *et al.* (2003) Group level validation of protein intakes estimated by 24-hour diet recall and dietary questionnaires against 24-hour urinary nitrogen in the European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study. *Cancer Epidemiol Biomarkers Prev* **12**, 784–795.
12. Subar AF, Thompson FE, Kipnis V *et al.* (2001) Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. *Am J Epidemiol* **154**, 1089–1099.
13. Kristal AR, Feng Z, Coates RJ *et al.* (1997) 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* **146**, 856–869.
14. Kristal AR, Vizenor NC, Patterson RE *et al.* (2000) Precision and bias of food frequency-based measures of fruit and vegetable intakes. *Cancer Epidemiol Biomarkers Prev* **9**, 939–944.
15. Jackson M, Walker S, Cade J *et al.* (2001) Reproducibility and validity of a quantitative food-frequency questionnaire among Jamaicans of African origin. *Public Health Nutr* **4**, 971–980.
16. Johansson I, Hallmans G, Wikman A *et al.* (2002) Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. *Public Health Nutr* **5**, 487–496.
17. Kumanyika SK, Mauger D, Mitchell DC *et al.* (2003) Relative validity of food frequency questionnaire nutrient estimates in the Black Women's Health Study. *Ann Epidemiol* **13**, 111–118.
18. Mayer-Davis EJ, Vitolins MZ, Carmichael SL *et al.* (1999) Validity and reproducibility of a food frequency interview in a multi-cultural epidemiology study. *Ann Epidemiol* **9**, 314–324.
19. Jain MG, Rohan TE, Soskolne CL *et al.* (2003) Calibration of the dietary questionnaire for the Canadian Study of Diet, Lifestyle and Health cohort. *Public Health Nutr* **6**, 79–86.
20. Morris MC, Tangney CC, Bienias JL *et al.* (2003) Validity and reproducibility of a food frequency questionnaire by cognition in an older biracial sample. *Am J Epidemiol* **158**, 1213–1217.
21. Marks GC, Hughes MC & van der Pols JC (2006) The effect of personal characteristics on the validity of nutrient intake estimates using a food-frequency questionnaire. *Public Health Nutr* **9**, 394–402.
22. Coates RJ & Monteilh CP (1997) Assessments of food-frequency questionnaires in minority populations. *Am J Clin Nutr* **65**, 4 Suppl., 1108S–1115S.