Validation of a food frequency questionnaire measurement of selected nutrients using biological markers in African-American men

European Journal of Clinical Nutrition (2007) 61, 1328–1336
 0954-3007/07 $30.00
www.nature.com/ejcn

ORIGINAL ARTICLE

Validation of a food frequency questionnaire measurement of selected nutrients using biological markers in African-American men

MD Holmes[1], IJ Powell[2], H Campos[3], MJ Stampfer[1,4], EL Giovannucci[1,4] and WC Willett[1,4]

[1]*Channing Laboratory, Department of Medicine, Harvard Medical School and Brigham and Women's Hospital, Boston, MA, USA;* [2]*Department of Urology, Wayne State University School of Medicine, Harper Professional Building, Detroit, MI, USA;* [3]*Department of Nutrition, Harvard School of Public Health, Boston, MA, USA and* [4]*Department of Epidemiology and Department of Nutrition, Harvard School of Public Health, Boston, MA, USA*

Objective: To validate selected nutrients assessed by the food frequency questionnaire (FFQ) used in the Harvard cohort studies in an African-American group.
Design: Blood aliquots were pooled for each decile of intake of two carotenoids and alpha tocopherol as measured by FFQ. These pooled samples were analyzed for nutrient content, and the resultant blood levels were plotted against the median for each decile of intake. In addition, adipose tissue samples taken from each man were analyzed for content of specific fatty acids. We calculated the Spearman correlations comparing intakes of specific fatty acids as percent of total fat intake, adjusted for energy intake, as measured by FFQ, with the percentage of the corresponding fatty acid in adipose tissue.
Subjects and settings: African-American men ($N=104$) with prostate cancer were recruited from a Detroit physician's practice and completed a detailed FFQ.
Results: Comparing decile 10 with decile 1 intake of nutrients as measured by FFQ, there was a 32% higher blood level of lycopene, a 288% higher blood level of beta carotene and a 100% higher blood level of alpha tocopherol. The Spearman correlation coefficients between intakes of linoleic acid, alpha linolenic acid, long-chain n-3 fatty acids and *trans* fatty acid measured by FFQ and the corresponding adipose tissue levels were between 0.10 and 0.47.
Conclusion: The FFQ was able to distinguish meaningful differences in biochemical measurements of selected nutrients and presumably corresponding differences in the extremes of intake in African-American men with prostate cancer who were likely to be motivated to report accurately. However, the results found are similar to those found in other populations.
Sponsorship: This work was supported by United States National Institutes of Health Grants CA42182 and CA72036.
European Journal of Clinical Nutrition (2007) **61,** 1328–1336; doi:10.1038/sj.ejcn.1602641; published online 7 February 2007

Keywords: African-American men; nutrition surveys; reproducibility of results; biological markers

Introduction

The food frequency questionnaire (FFQ) has many advantages for dietary assessment in epidemiological studies. However, performance of these questionnaires is often specific to demographic and cultural groups. Although the FFQ developed by Willett *et al*. (1983a) has been validated over a wide range of women and men (Ascherio *et al*., 1992; Longnecker *et al*., 1993; Willett, 1998), these studies have been conducted in populations that were mainly European American. We undertook this study to validate this questionnaire in a group of African-American men living in the USA.

Correspondence: Dr MD Holmes, Channing Laboratory, Department of Medicine, Harvard Medical School and Brigham and Women's Hospital, 181 Longwood Ave, Boston, MA, USA.
E-mail: michelle.holmes@channing.harvard.edu
Contributors: MH planned the study, supervised data collection, performed statistical analyses and contributed to writing of the paper. IP planned the study, supervised data collection, contributed to writing of the paper. HC advised on the methods of handling, transport and analysis of specimens, supervised the analysis of specimens, contributed to writing of the paper. MS sought funding and contributed to the writing of the paper. EG sought funding and contributed to the writing of the paper. WW planned the study and contributed to the writing of the paper.
Received 4 January 2006; revised 30 November 2006; accepted 30 November 2006; published online 7 February 2007

African-American men have a substantially elevated risk of prostate cancer incidence and mortality compared to other men (SEER and NCHS Mortality by race and age 1973–1995; Powell, 1997). Dietary hypotheses relating to prostate cancer incidence and survival are important to evaluate in this population. For this reason, we compared the intake of nutrients as assessed by FFQ with biological measures in blood and adipose tissue among a group of African-American men who were prostate cancer survivors. To reduce the costs associated with biochemical analyses, we used a previously described method to pool blood aliquots for categories of intake defined by the FFQ (Willett, 1998).

In these men with prostate cancer, we focused on several nutrients hypothesized to have an association with prostate cancer, including lycopene, alpha linolenic acid and alpha tocopherol (Giovannucci *et al*., 1993a; The Alpha tocopherol, Beta carotene Cancer Prevention Study Group, 1994; Giovannucci *et al*., 1995; Vogt *et al*., 2002). In addition, we studied beta carotene, *trans* unsaturated fats, linoleic acid and other n-3 fatty acids other than alpha linolenic.

Materials and methods

Refinement of the FFQ

Foods were selected by several criteria on our original FFQ; (1) by using step-wise regression, predicting total nutrient intakes, to select from a longer list of potential foods, (2) other foods were added if there were of interest for specific hypotheses and (3) periodic pilot studies indicated that use of certain foods had increased since our original studies, such as low fat versions of foods. Because of further modifications of the questionnaire used in this analysis, and the population selected was different, the study we describe was conducted.

In a study carried out among low-income black women in Atlanta Georgia, culturally specific foods were identified which were not measured by a standard FFQ but which contributed significantly to nutrients commonly measured in cancer epidemiologic research (Coates *et al*., 1991). In a similar manner, we had the Willett FFQ reviewed by two registered dieticians, supervised by a PhD-level researcher, all of whom were engaged in a trial to reduce dietary fat and energy intake among overweight pre-menopausal Detroit women for cancer prevention purposes (Djuric *et al*., 1999). The dieticians reviewed food records from the 49 African-American women in that trial (43% of the total) and identified eight new foods to add to the FFQ: honeydew melon, macaroni and cheese, scalloped potatoes, malt liquor, reduced fat cookies, olestra-containing chips, salt substitutes such as Mrs Dash and organ meats such as chitterlings, gizzards and hog maws. In addition, culturally appropriate prompts were added to six pre-existing questions (e.g. adding black-eyed peas to the question presently asking about beans or lentils).

The foods used as prompts were similar to the specific food from the database that was used for the calculation. The prompt was to help specifically identify a food eaten in this category (e.g. black-eyed peas are a lentil but are a specific lentil that is eaten in this population). As the prompt food was only an example of the type of food asked about, there was no need to change the nutrient database.

The final FFQ used had eight categories (vitamin supplements; dairy foods; fruits, vegetables; eggs, meat and fish; cereals, breads and starches; beverages; sweets, baked goods; and miscellaneous), 12 items on vitamins and other dietary supplements and 158 food items. The new foods identified by the dieticians were incorporated into the appropriate category. Nutrient data for the new foods not already contained within our own database (harvardsffq.120899) was obtained from the USDA (1996) (chitterlings and malt beer), Bowes and Church's Food Values of Portions Commonly Used (Pennington *et al*., 1998) (salt substitute) and the manufacturer (olestra-containing chips).

The database used for the nutrient analysis was a specifically designed program – harvardsffq.120899 (8 December 1999). The foundation of the database is the US Department of Agriculture Nutrient Database for Standard Reference, releases 12, with additional information from McCance and Widdowson' The Composition of Foods (4th and 5th editions), journals and manufacturers (Paul and Southgate, 1978; Holland *et al*., 1991; USDA, 1998).

Nutrient amounts were calculated using the FFQ frequencies multiplied by the nutrient data of the foods from the specifically designed database to determine a daily intake of the individual. Multivitamin and mineral supplements were included in the analysis.

Eligibility and recruitment

The study protocol and informed consent form was approved by the Human Subjects Committee of Brigham and Women's Hospital and Wayne State Medical School. African-American men with prostate cancer diagnosed within the past 5 years were identified from the patient records of a physician at Wayne State University, Detroit, MI, USA. Excluded were those with metastatic disease at diagnosis, those who could not speak English and those with severe cognitive impairment. A registered nurse research assistant recruited men into the study, administered the FFQ and took biological samples. Biological samples included 15 ml of blood collected by venipuncture. As intake of many fatty acids can be more reliably assessed from fat aspirates than from serum samples, men were also asked to give a fat aspirate sample. This procedure, which causes approximately the same amount of discomfort as a venipuncture (Beynen and Katan 1985), involves passing a 19-g needle into the lateral buttock or abdominal subcutaneous fat and aspirating into a syringe. Each participant was paid $50.00 for his time.

Sample size

Previous validation studies using 100–200 participants have provided stable correlation coefficients between 0.5. and 0.7 comparing the FFQ with a comparison method (Willett, 1998). In the present study we enrolled 104 participants over a 15-month period.

Collection and storage of biological samples

Men who were not fasting were recruited during a 15-month period between 1998 and 1999. Biological samples were stored at -80°C at Wayne State Medical Center until the end of the collection period, then express mailed to the Harvard School of Public Health and again stored at -80°C until analyzed. For carotenoids and alpha tocopherol, participants were divided into deciles of energy-adjusted intake according to the FFQ (Willett *et al*., 1985). Blood aliquots were pooled for each decile of intake, and mixed carefully. In addition, the 3rd and 4th decile were combined, the 5th and 6th decile were combined and the 7th and 8th decile were combined, leaving a total of seven categories of intake. A different set of pooled aliquots was prepared for each nutrient. Nutrient analysis was performed on each pooled blood sample for lycopene, beta carotene and alpha tocopherol by a high-performance liquid chromatography (HPLC) method in 2001; a replicate was performed in 2003. The results in the tables are based on the first measurement.

Fat samples cannot be pooled in the same manner as blood because different amounts of fat are collected from each individual, and there is no easy way to weigh each sample for pooling. Instead, individual samples were analyzed by gas–liquid chromatography (GLC). Four types of fatty acids were of interest: linoleic acid, alpha linolenic acid, long-chain omega-3 fatty acids from marine sources and *trans* fatty acids.

Laboratory methods

Concentrations of lycopene, alpha carotene, beta carotene and alpha tocopherol in blood samples were measured as previously described (El-Sohemy *et al*., 2002). Briefly, plasma samples (250 mcl) were mixed with ethanol, extracted with hexane, evaporated to dryness under nitrogen and reconstituted in ethanol, dioxane and acetonitrile. Samples were measured by HPLC (Restek, Corp., Bellefonte, PA, USA). The system manager software (D-7000, Version 3.0) was used for peak integration and data acquisition (Hitachi, San Jose, CA, USA). The minimum detection limits in plasma were (in mcg/l) 8.49 for lycopene, 7.31 for beta carotene and 285.3 for alpha tocopherol. Every run included two replicates each of a two-level plasma pool sample set. For external quality control, our laboratory participates in the standardization program for carotenoid analysis from the National Institute of Standards and Technology USA. Measurement of nutrient levels by HPLC was carried out two times as documented above; once in 2001 and once in 2003. The coefficients of variation were between 1.8 and 2.4%. The percent differences of the high and low mean vs the laboratory expected values were all within $\pm 10\%$, with the exception of beta carotene; the difference between the high mean and laboratory high expected value was -12.5%.

Fatty acids in adipose tissue were determined as described previously (Baylin *et al*., 2002). Briefly, fatty acids were extracted from a hexane:isopropanol mixture containing the sample and esterified with methanol and acetyl chloride. After esterification, the methanol and acetyl chloride were evaporated, and the fatty acid methyl esters were re-dissolved in iso-octane and the methyl esters were quantified by GLC. Samples were separated on a Supelco SP-2560 column ($100\,\text{m} \times 0.25\,\text{mm}$ ID with $0.2\,\mu$m film), housed within a Hewlett Packard 6890 plus gas chromatograph with flame ionization detector. Samples were injected using a Hewlett Packard 7683 autosampler/injector. Chromatograms were analyzed using Hewlett Packard ChemStation Software. Peak retention times and area percentages of total fatty acids were identified by injecting known standards (NuCheck Prep, Elysium, MN, USA).

Statistical methods

For the analysis of carotenoids and tocopherol, we used generalized estimating equations (GEEs) (Liang and Zeger, 1993) with a working independence correlation matrix and an identity link function. We regressed each individual's pooled decile on their measured FFQ intake value. GEEs account for the clustering by decile by estimating the true variance as $V_2 = V_1[1 + \rho(n\phi - 1]$, where V_1 is the variance estimated by conventional linear regression,

$$V_1 = \sigma^2 / \sum_i \sum_j (x_{ij} - \tilde{x})^2 = \sigma^2 / V_T$$

and

$$\phi = \sum_{i=1}^{k} n(\tilde{x}_i - \tilde{x})2 / V_T$$

is the fraction of the total variance among the x's that is caused by variation among cluster means $(\tilde{x}_i)$s, rather than variation in x's within clusters (Liang and Zeger, 1993). We fit the model:

blood_i = beta_0 + beta_1 × nutrient_f_i + b_i + e_{ij},

where blood_i is the value of the pooled mean of the nutrient in the decile to which participant i belongs (units), nutrient_f_i is participants i's average daily intake of the nutrient as measured by the FFQ, b_i is the random between deciles variation of blood levels after adjusting for individual intake, and e_{ij} is the random within-deciles variation of blood levels after adjusting for individual intake, beta_0 is the blood level for someone with 0 lycopene intake and beta_1 is the regression slope for the change in blood levels expected for a one unit increase in nutrient intake.

As all individuals within a decile of intake have the same blood level, we could not calculate cluster-adjusted *P*-values using the same approach. In order to calculate cluster-adjusted *P*-values, we used SAS PROC MIXED, with the pooled blood values as the independent variables and the individual intakes as measured by FFQ as the dependent variable.

We also calculated the percent difference in blood levels of nutrients comparing the highest to the lowest decile of intake as measured by FFQ.

For intake of each fatty acid, we calculated the percent of total fat intake, as measured by FFQ. We then calculated the Spearman correlations and linear regression coefficient comparing this with the percentage of fatty acids in adipose tissue, adjusted for total energy intake, by adding total energy intake in the model (SAS PROC CORR and PROC REG) (SAS Institute Inc., 1996).

We would expect an association between fatty acid intake and percent of adipose tissue made up by that fatty acid only if there was no endogenous synthesis. Fatty acids chosen for this analysis were those hypothesized to be associated with prostate cancer, and those for which there is no endogenous synthesis. In another analysis, the percent of fatty acids in the adipose tissue (which is the only reasonable way to express fatty acid value in such measurements) correlated best with intakes also expressed as a percentage of total fatty acid intake (Fawzi *et al*., 2004).

Results

One hundred and six men agreed to participate in the study. Only one man who was approached refused to participate. Two men did not complete the FFQ, leaving a total of 104 participants. The characteristics of the study subjects are shown in Table 1. Mean age was 63 years, mean time since prostate cancer diagnosis was 32 months and mean total energy intake was 8309 kJ/day (1985 kcal).

Figure 1' shows median intake of each category of lycopene, beta carotene and alpha tocopherol plotted against its corresponding blood level, and the predicted regression line generated using individual FFQ values. In addition, we show the percent increase in blood levels of nutrients across extreme deciles of intake as measured by FFQ. Comparing decile 10 with decile 1, there was a 32% increase in blood levels of lycopene, a 288% increase in blood levels of beta carotene and a 100% increase in blood levels of alpha tocopherol.

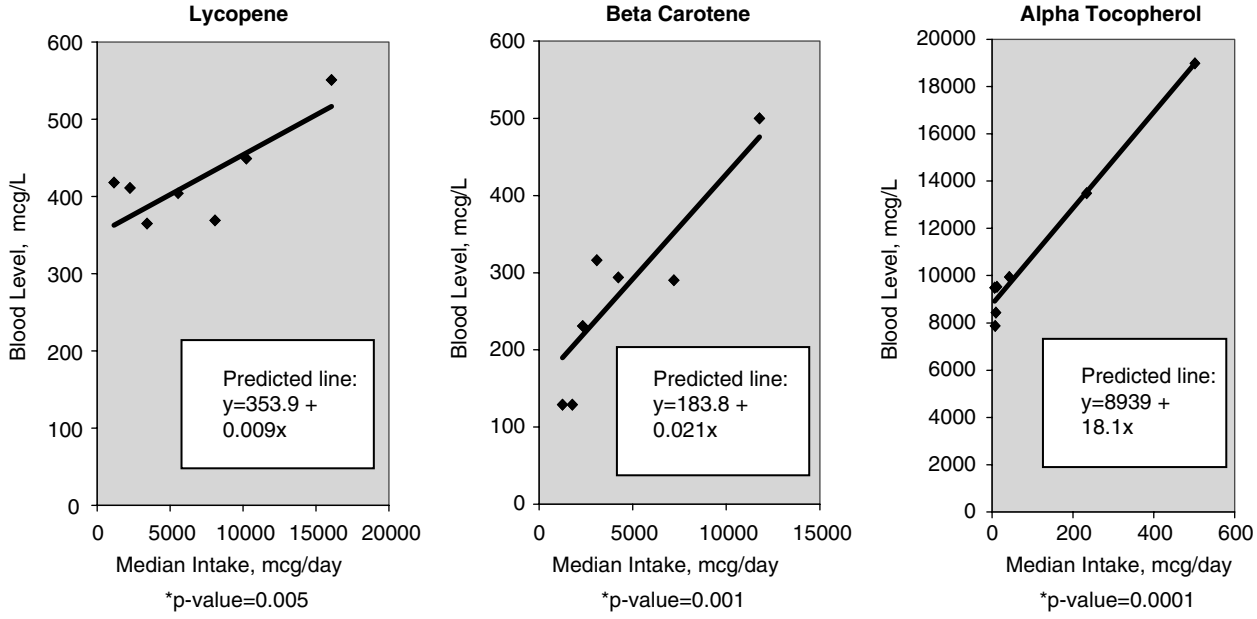
The use of pooled unfasted blood samples makes it difficult to control for multiple potential confounders such as age, smoking and blood lipid levels. Table 2 shows the distribution of age and current smoking across categories of reported intake of lycopene, beta carotene and alpha tocopherol. Age did not vary appreciably across categories of nutrient intake. However, those consuming the lowest amounts of carotenoids and alpha tocopherol were more likely to be current smokers.

Table 1 Characteristics of men with prostate cancer

Age in years (mean ± s.d.)	*63.3 ± 7.5*
Time since diagnosis in months (mean ± s.d.)	32 ± 19
Current smokers (%)	24
Family history of prostate cancer (%)	29
Extra prostatic extension of tumor at diagnosis (%)	18
Gleason stage at diagnosis (%)	
4–5–6	39
7–8–9	57
Unknown	4
Disease recurred (%)	14
Total energy intake (mean kJ/day ± s.d.)	8309 ± 3516
Total fat intake (mean % of energy intake/day ± s.d.)	32 ± 7
Polyunsaturated fat intake (mean % of energy intake/day ± s.d.)	6.3 ± 2.1
n-3 fatty acid intake (mean % of energy intake/day ± s.d.)	0.8 ± 0.3
Vitamin E intake (mean mg/day ± s.d.)	85 ± 156
Beta carotene intake (mean mcg/day ± s.d.)	4330 ± 3712
Lycopene intake (mean mcg/day ± s.d.)	6239 ± 4447

We performed several analyses to explore the effect of pooling. We simulated the effect of pooling in a dataset of 121 men from the Health Professionals' Follow-up Study (HPFS) for whom we had individual measures of blood lycopene and beta carotene levels, as well as nutrient intake as measured by FFQ (Michaud *et al*., 1998). We calculated the mean of the blood-level values for lycopene and beta carotene by category of FFQ intake. For each nutrient, we used this value as the dependent variable and individual nutrient intake as calculated by the FFQ as the independent variable, and computed the regression coefficient for the predicted line using GEEs (SAS PROC GENMOD). We then compared the regression coefficients obtained using all data points and those obtained using the simulated pools. In addition, we examined the effects of controlling and not controlling for confounders; the results are shown in Table 3. The parameter estimates obtained in the pooled, energy-adjusted analysis from the present study compared with the simulated pooled, energy-adjusted parameter estimates from the HPFS were very similar for both lycopene and beta carotene. Comparing the individual vs the simulated pooled parameter estimates in the HPFS, pooling minimally changed the estimates for lycopene and increased the estimates somewhat for beta carotene. There may have been a modest amount of confounding by smoking in the present study, although this was minimal in the HPFS. Comparing energy-adjusted vs multivariate-adjusted results in the HPFS, the additional control for covariates (age, smoking and blood lipid levels) caused the estimates for lycopene and beta carotene to change very little. These additional analyses provide some reassurance that the pooling method is valid.

Table 4 shows the comparison of fatty acid content of diet and adipose tissue. Three samples were not usable because of insufficient quantity leaving a total of 101 samples to analyze. We assessed intake of three of the most common



decile	median daily intake (mcg)	blood level (mcg/L)	% increase in blood levels from decile 1 to 10
1	1145	418	
2	2237	411	
3+4	3410	365	
5+6	5546	404	
7+8	8074	369	
9	10225	449	
10	16054	551	32%

decile	median daily intake (mcg)	blood level (mcg/L)	% increase in blood levels from decile 1 to 10
1	1256	129	
2	1773	129	
3+4	2321	231	
5+6	3084	316	
7+8	4239	294	
9	7204	290	
10	11777	500	288%

decile	median daily intake (mcg)	blood level (mcg/L)	% increase in blood levels from decile 1 to 10
1	6.5	9494	
2	7.8	7867	
3+4	9.8	8442	
5+6	12.4	9523	
7+8	43.1	9942	
9	233.9	13487	
10	501.6	18987	100%

5

* p-values are cluster adjusted using SAS PROC MIXED, where individual dietary intake was the dependent variable

Figure 1 Median intake of lycopene, beta carotene and alpha tocopherol, per decile, vs actual and predicted blood levels of the same nutrients.

Table 2 Mean age and percent of current smokers across categories of reported lycopene, beta carotene and alpha tocopherol intake

Decile	*Mean age in years*[a] *(± s.d.)*			*Percent of current smokers*[b] *(count)*		
	Lycopene	*Beta carotene*	*Alpha tocopherol*	*Lycopene*	*Beta carotene*	*Alpha tocopherol*
1 + 2	62.3 (8.8)	61.1 (7.3)	63.4 (7.4)	45% (9/20)	30% (6/20)	53% (10/19)
3 + 4	62.9 (8.3)	63.4 (7.9)	63.0 (8.2)	14% (3/21)	24% (5/21)	27% (6/22)
5 + 6	60.8 (7.1)	63.2 (7.2)	63.7 (6.3)	24% (5/21)	43% (9/21)	14% (3/21)
7 + 8	66.0 (6.6)	65.2 (6.5)	63.8 (7.2)	24% (5/21)	14% (3/21)	19% (4/21)
9 + 10	64.3 (5.9)	63.2 (8.4)	62.5 (8.6)	14% (3/21)	10% (2/21)	10% (2/21)

[a] *P*-value for *t*-test comparing mean age of deciles 1 + 2, 3 + 4, 5 + 6, 7 + 8 and deciles 9 + 10 against each other, within nutrient category: for lycopene, decile 5 + 6 vs decile 7 + 8, $P = 0.02$. All other comparisons $P > 0.05$.
[b] *P*-value for χ^2 for lycopene, $P = 0.14$; for beta carotene, $P = 0.09$; and for alpha tocopherol, $P = 0.01$.

types of *trans* fatty acids in the diet: *trans* 18:1, *trans* 18:2 and total *trans*. The highest correlation we found was between total *trans* measured by FFQ and total *trans* measured in adipose tissue; the correlation coefficient was 0.26. For n-6 fatty acids, we assessed intake of linoleic acid (18:2 – *cis* only) with the FFQ, and the percent of linoleic acid in adipose tissue. The correlation coefficient was 0.33.

We assessed several n-3 fatty acids: eicosapentanoic acid (EPA, 20:5), docosahexanoic acid (DHA, 22:6), alpha linolenic acid (18:3), the sum of EPA + DHA and total intake of n-3 fatty acids (EPA + DHA + alpha linolenic). Correlation coefficients ranged from a low of 0.19 (EPA) to a high of 0.47 (DHA).

All results were adjusted for total energy intake by adding total energy intake to the model. Although not strictly necessary, as fat intake was expressed as a percent of energy, in any analysis of disease end points we would typically adjust for total energy, and thus we did so here. Analyses carried out without adjustment for total energy gave very similar results.

In addition, we calculated the predicted regression coefficient (β) using individual values and simulating the pooling done with the blood samples of lycopene and beta carotene. We obtained similar results regardless of the method (Table 4). Adjustment for age and deleting current smokers (25 men) did not substantially change the results.

Table 3 Parameter estimates of the regression slopes for the prediction of blood nutrient levels by food frequency questionnaire intake using linear regression: comparison of pooled, energy-adjusted analysis in Detroit men ($N = 104$) with analyses in the HPFS ($N = 121$)

PopulationAnalysis		*Parameter estimate*	
		Lycopene	*Beta carotene*
1. Detroit	Pooled , energy adjusted	0.009	0.021
2. HPFS	Individual, energy adjusted	0.011	0.016
3. HPFS	Individual, multivariate adjusted[a]	0.012	0.014
4. HPFS	Simulated pooled, energy adjusted	0.011	0.022
5. HPFS	Simulated pooled, multivariate adjusted[a]	0.011	0.020

Abbreviation: HPFS, health professionals follow-up study.
[a]Adjusted for energy, age, smoking and blood lipids.

Discussion

Several studies have examined the correlation between intake of lycopene, beta carotene and alpha tocopherol as measured by FFQ, and blood levels measured in *individuals*. Correlations in several studies have been close to zero for lycopene (Coates *et al*., 1991; Ascherio *et al*., 1992; Resnicow *et al*., 2000). One reason for these low correlations may be because the bioavailability of dietary lycopene is a critical issue. For instance, taking into account the cooking of tomato products may improve the correlation with lycopene (Giovannucci *et al*., 2002). However, one study which examined diet records in 98 non-smoking pre-menopausal women found a correlation of 0.50 (Yong *et al*., 1994), and the correlation among non-smoking men in the HPFS was 0.47 (Michaud *et al*., 1998). Other reasons include the distribution of lycopene in the foodstuffs and because of the difficulties of the questionnaire to evaluate lycopene intake using the food composition tables.

Correlations for beta carotene have been in the range from 0.24 to 0.44 (Stryker *et al*., 1988; Coates *et al*., 1991; Ascherio *et al*., 1992; Campbell *et al*., 1994; Goodman *et al*., 1996; Resnicow *et al*., 2000); although in general higher in non-smokers (Stryker *et al*., 1988). Likewise, correlations for alpha tocopherol have been in the range of 0.28–0.53 (Willett *et al*., 1983a; Stryker *et al*., 1988; Coates *et al*., 1991; Ascherio *et al*., 1992; Jacques *et al*., 1993), with much lower correlations found ($r = 0.12$) when unadjusted for serum lipids (Willett *et al*., 1983a). In a randomized trial, 16 weeks of 30 mg of a beta carotene supplement daily caused a tripling of blood carotenoid levels, and 800 IU of an alpha tocopherol supplement daily caused a doubling of blood levels (Willett *et al*., 1983b).

One study has similarly compared pooled blood levels of nutrients with intake estimated by FFQ. In a study of 72 African American and 132 European American pregnant

Table 4 Spearman rank correlations (*P*-values), predicted regression coefficient (β) and predicted regression coefficient simulating pooled data between energy-adjusted dietary fatty acids (from FFQ) and adipose tissue fatty acids

Fatty acids		*Correlation*	P-*value*	*Predicted* β	*Predicted* β *simulating pooled data*
Trans	Total *Trans*	0.26	0.009	0.10	0.08
	Trans 18:1	0.11	0.26	0.001	0.007
	Trans 18:2	0.20	0.04	0.83	0.85
N-6	Linoleic (18:2)	0.33	0.0009	0.20	0.19
N-3	EPA (20:5)	0.19	0.07	0.04	0.04
	DHA (22:6)	0.47	0.0001	0.09	0.10
	Alpha Linolenic (18:3)	0.22	0.03	0.11	0.10
	EPA + DHA	0.43	0.0001	0.07	0.07
	Total n-3	0.29	0.003	0.14	0.15

DHA, docosahexanoic acid; DPA, docosapentanoic acid; EPA, eicosapentanoic acid; FFQ, food frequency questionnaire.

Table 5 Comparison of correlation coefficients between fatty acid intake assessed by food frequency questionnaire and adipose tissue in various studies

Author	*Linoleic (18:2n-6)*	*Alpha Linolenic (18:3n-3)*	*EPA (20:5n-3) and/or DHA (22:6n-3)*	*Trans*	*Comments*
Holmes (2007)	0.31	0.22	0.43 (EPA + DHA)	0.26	101 Detroit African-American men with prostate cancer
Baylin (2002)	0.58	0.34	0.15 (EPA) 0.18 (DHA)	0.54	Costa Rican men (367) and women (136)
Garland (1997)	0.37	0.34			140 women
Godley (1996)			0.38 (EPA only)		29 African-American men and 98 white men, 89 of these men had prostate cancer
Ascherio (1995)			0.49 (EPA only)		127 Boston men
Hunter (1992)	0.37–0.48			0.29–0.34	118 Boston men
London (1991)	0.35	0.12	0.48 (EPA + DHA)	0.51	115 postmenopausal Boston women, described as 'mostly white'

Abbreviations: EPA, eicosapentanoic acid; DHA, docosahexanoic acid.

women, Fawzi *et al*. (2004) reported the percent difference in blood nutrient levels comparing the highest to the lowest deciles of intake. They reported a 21% increase for African-American women and 64% increase for European American women for lycopene, compared with the 32% increase we found (Figure 1). We found a 288% increase in blood levels between extreme deciles of intake of beta carotene. In comparison, Fawzi *et al*. did not study beta carotene; however, they did study alpha carotene, and found a similar increase of 275% for African-American women and 152% for European American women in blood levels of alpha carotene comparing extreme deciles of intake. In addition, where we studied alpha tocopherol and found a 100% increase in blood levels between extreme deciles of intake, Fawzi *et al*. studied gamma tocopherol. They found a 42% increase in blood levels for African-American women and a 12% increase in blood levels for European American women, comparing extreme deciles of gamma tocopherol intake.

In calculating the *P*-value for the predicted lines, the pooled biomarker levels are indeed order statistics of repeated nutrient intake and thus will be to some extent correlated. However, we examined empirically the correlations between all 10 712 pair-wise combinations of median of the blood-level decile except between themselves, the dependent variable in our regression models. We found that the correlation was very small (about −0.008). Although the problem of induced correlation between validation study participants owing to using the decile cutoffs is theoretically possible, it was not evident to any measurable degree in our study.

The use of pooled samples has not been used much previously in FFQ validation studies; thus the validity of this novel technique discussed by Forman *et al*. (1990) and Willett (1998) should be explored. Wahrendorf *et al*. measured both pooled and individual blood vitamin levels in 50 samples (10 pools) from an intervention trial in China. They found good agreement between the measurement of pooled blood retinol and beta carotene levels and the mean of individually measured levels within pooling groups (Wahrendorf *et al*., 1986). Weinberg and Umbach (1999) have proposed pooling of randomly grouped sets of biological samples from cases and controls in case–control studies, and show that a valid estimated exposure odds ratio can be calculated.

Several other studies have compared fatty acid intake as measured by FFQ with the levels of those fatty acids in adipose tissue in various populations (London *et al*., 1991; Hunter *et al*., 1992; Ascherio *et al*., 1995; Godley *et al*., 1996; Garland *et al*., 1998; Baylin *et al*., 2002). Some of the other studies, as well as the present study, are summarized in Table 5. In general, correlations were between 0.3 and 0.5, and there was good agreement between our results and the results published by others, regardless of the population.

The FFQ we used has been useful to predict diseases in a predominantly European American population: heart disease (Ascherio *et al*., 1994, 1996), colon cancer (Giovannucci *et al*., 1993b, 1994) and prostate cancer (Giovannucci *et al*., 1998) among others. Given this previous history, we feel the use of pooled samples provides an attractive cost-saving option to validate the FFQ in another population.

Many factors could serve to diminish the correlations we found between intake as measured by FFQ and the biochemical measures. First of all, both methods contain measurement errors. However, an advantage of this study is that these methods are sufficiently different such that the errors should not be correlated with each other. In addition, we were unable to control in the regression analysis for factors such as body mass index or serum lipids. In the pooled samples, we were unable to control for even those factors which we did have information on, for example, age and smoking status. In future studies using pooled analysis, it may be preferable to adjust dietary intakes for these variables before creating categories to be used for pooling. The regression slopes we estimated do not provide the same information as do correlation coefficients, which would require individual-level measurements. The correlation coefficient is an important measure of validity, but the regression coefficient also provides useful information, and is the basis of error–correction procedures. Also, when comparing assessments among various populations, differences in

correlation coefficients can be due to differences in distributions of nutrient intakes and biological measures rather than the functioning of the questionnaire itself, but the regression slope will not be influenced in this way.

Men with prostate cancer are likely to be highly motivated to report their diet accurately. Therefore, the choice of prostate cancer patients may impede the extension of the validation of this questionnaire to normal subjects who might be less attentive than patients in answering the questionnaire. However, the results found in this study were similar to the results in other populations, with and without disease.

In conclusion, we found that the FFQ was able to distinguish meaningful differences in biochemical measurements of nutrients and presumably corresponding differences in the extremes of intake. The associations in this population of African-American men with prostate cancer were similar to the associations seen in healthy predominantly European American populations of men and women.

References

The Alpha tocopherol, Beta carotene Cancer Prevention Study Group (1994). The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers **330**, 1029–1035.

Ascherio A, Hennekens CH, Buring JE, Master C, Stampfer MJ, Willett WC (1994). Trans fatty acids intake and risk of myocardial infarction. *Circulation* **89**, 94–101.

Ascherio A, Rimm EB, Giovannucci EL, Spiegelman D, Stampfer MJ, Willett WC (1996). Dietary fat and risk of coronary heart disease in men: cohort follow up study in the United States. *Br Med J* **313**, 84–90.

Ascherio A, Rimm EB, Stampfer MJ, Giovannucci E, Willett WC (1995). Dietary intake of marine n-3 fatty acids, fish intake and the risk of coronary disease among men. *N Engl J Med* **332**, 977–982.

Ascherio A, Stampfer MJ, Colditz GA, Rimm EB, Litin L, Willett WC (1992). Correlations of vitamin A and E intakes with the plasma concentrations of carotenoids and tocopherols among American men and women. *J Nutr* **122**, 1792–1801.

Baylin A, Kabagambe EK, Siles X, Campos H (2002). Adipose tissue biomarkers of fatty acid intake. *Am J Clin Nutr* **76**, 750–757.

Beynen AC, Katan MB (1985). Rapid sampling and long-term storage of subcutaneous adipose-tissue biopsies for determination of fatty acid composition. *Am J Clin Nutr* **42**, 317–322.

Campbell DR, Gross MD, Martini MC, Grandits GA, Slavin JL, Potter JD (1994). Plasma carotenoids as biomarkers of vegetable and fruit intake. *Cancer Epidemiol Biomarkers Prev* **3**, 493–500.

Coates RJ, Eley JW, Block G, Gunter EW, Sowell AL, Grossman C *et al.* (1991). An evaluation of a food frequency questionnaire for assessing dietary intake of specific carotenoids and vitamin E among low-income black women. *Am J Epidemiol* **134**, 658–671.

Djuric Z, Uhley VE, Depper JB, Brooks KM, Lababidi S, Heilbrun LK (1999). A clinical trial to selectively change dietary fat and/or energy intake in women: the women's diet study. *Nutr Cancer* **34**, 27–35.

El-Sohemy A, Baylin A, Kabagambe E, Ascherio A, Spiegelman D, Campos (2002). Individual carotenoid concentrations in adipose tissue and plasma as biomarkers of dietary intake. *Am J Clin Nutr* **76**, 172–179.

Fawzi WW, Rifas-Shiman SL, Rich-Edwards JW, Willett WC, Gillman MW (2004). Calibration of a semi-quantitative food frequency questionnaire in early pregnancy. *Ann Epidemiol* **14**, 754–762.

Forman D, Sitas F, Newell DG, Stacey AR, Boreham J, Peto R *et al.* (1990). Geographic association of helicobacter pylori antibody prevalence and gastric cancer mortality in rural China. *Int J Cancer* **46**, 608–611.

Garland M, Sacks FM, Colditz GA, Rimm EB, Sampson LA, Willett WC *et al.* (1998). The relation between dietary intake and adipose tissue composition of selected fatty acids in US women. *Am J Clin Nutr* **67**, 25–30.

Giovannucci E, Ascherio A, Rimm EB, Stampfer MJ, Colditz GA, Willett WC (1995). Intake of carotenoids and retinol in relation to risk of prostate cancer. *J Natl Cancer Inst* **87**, 1767–1776.

Giovannucci E, Rimm EB, Colditz GA, Stampfer MJ, Ascherio A, Chute CC *et al.* (1993a). A prospective study of dietary fat and risk of prostate cancer. *J Natl Cancer Inst* **85**, 1571–1579.

Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Willett WC (1994). Intake of fat, meat, and fiber in relation to risk of colon cancer in men. *Cancer Res* **54**, 2390–2397.

Giovannucci E, Rimm EB, Wolk A, Ascherio A, Stampfer MJ, Colditz GA *et al.* (1998). Calcium and fructose intake in relation to risk of prostate cancer. *Cancer Res* **58**, 442–447.

Giovannucci E, Rimm EB, Liu Y, Stampfer MJ, Willett WC (2002). A prospective study of tomato products, lycopene, and prostate cancer risk. *J Natl Cancer Inst* **94**, 391–398.

Giovannucci E, Stampfer MJ, Colditz GA, Rimm EB, Trichopolous D, Rosner BA *et al.* (1993b). Folate, methionine, and alcohol intake and risk of colorectal adenoma. *J Natl Cancer Inst* **85**, 875–884.

Godley PA, Campbell MK, Miller C, Gallagher P, Martinson FE, Mohler JL *et al.* (1996). Correlation between biomarkers of omega-3 fatty acid consumption and questionnaire data in African American and Caucasian United States males with and without prostatic carcinoma. *Cancer Epidemiol Biomarkers Prev* **5**, 115–119.

Goodman EG, Thornquist M, Kestin M, Metch B, Anderson G, Omenn GS (1996). The association between participant characteristics and serum concentrations of b-carotene, retinol, retinyl palmitate, and α-tocopherol among participants in the Carotene and Retinol Efficacy Trial (CARET) for prevention of lung cancer. *Cancer Epidemiol Biomarkers Prev* **5**, 815–821.

Holland G, Welch AA, Unwin ID, Buss DH, Paul AA, Southgate DAT (1991). McCance and Widdowson's the composition of foods, 5th edn. The Royal Society of Chemists: UK.

Hunter DJ, Rimm EB, Sacks FM, Stampfer MJ, Colditz GA, Litin LB *et al.* (1992). Comparison of measures of fatty acid intake by subcutaneous fat aspirate, food frequency questionnaire, and diet records in a free-living population of US men. *Am J Epidemiol* **135**, 418–427.

Jacques PF, Sulsky SI, Sadowski JA, Phillips JC, Rush D, Willett WC (1993). Comparison of micronutrient intake measured by a dietary questionnaire and biochemical indicators of micronutrient status. *Am J Clin Nutr* **57**, 182–189.

Liang KY, Zeger SL (1993). Regression analysis for correlated data. *Ann Rev Publ Health* **14**, 43–68.

London SJ, Sacks FM, Caesar J, Stampfer MJ, Siguel E, Willett WC (1991). Fatty acid composition of subcutaneous adipose tissue and diet in post-menopausal US women. *Am J Clin Nutr* **54**, 340–345.

Longnecker MP, Lissner L, Holden JM, Flack VF, Taylor PR, Stampfer MJ *et al.* (1993). The reproduciblity and validity of a self-administered semi-quantitative food frequency questionnaire in subjects from South Dakota and Wyoming. *Epidemiology* **4**, 356–365.

Michaud DS, Giovannucci EL, Ascherio A, Rimm EB, Forman MR, Sampson L *et al.* (1998). Associations of plasma carotenoid concentrations and dietary intake of specific carotenoids in samples of two prospective cohort studies using a new carotenoid database. *Cancer Epidem Biomar Prevent* **7**, 283–290.

Paul AA, Southgate DA (1978). *McCance and Widdowson's the composition of foods*. Her Majesty's Stationery Office: London, England.

Pennington JAT, Bowes AD, Church HN (1998). *Bowes and Church's food values of portions commonly used*. Lippincott-Raven: Philadelphia.

Powell I (1997). Prostate cancer and African American men. *Oncology* **11**, 599–605.

Resnicow K, Odom E, Wang T, Dudley WN, Mitchell D, Vaughan R *et al.* (2000). Validation of three food frequency questionnaires and 24-hour recalls with serum carotenoid levels in a sample of African-American adults. *Am J Epidemiol* **152**, 1072–1080.

SAS Institute Inc. (1996). Version 6.12, SAS Institute, Inc.: Cary, NC.

SEER and NCHS Mortality by race and age 1973–1995.

Stryker WS, Kaplan LA, Stein EA, Stampfer MJ, Sober A, Willett WC (1988). The relation of diet, cigarette smoking, and alcohol consumption to plasma beta-carotene and alpha-tocopherol levels. *Am J Epidemiol* **127**, 283–296.

USDA (1996). U.S. Department of Agriculture Nutrient Database for Standard Reference, Release 11. Nutrient Data Laboratory Home Page: http://www.nal.usda.gov/fnic/foodcomp.

USDA (1998). U.S. Department of Agriculture Nutrient Database for Standard Reference, Release 12. Nutrient Data Laboratory Home Page: http://www.nal.usda.gov/fnic/foodcomp.

Vogt TM, Mayne ST, Graubard BI, Swanson CA, Sowell AL, Schoenberg JB *et al.* (2002). Serum lycopene, other serum carotenoids, and risk of prostate cancer in US Blacks and Whites. *Am J Epidemiol* **155**, 1023–1032.

Wahrendorf J, Hanck AB, Munoz N, Vuilleumier JP, Walker AM (1986). Vitamin measurements in pooled blood samples. *Am J Epidemiol* **123**, 544–550.

Weinberg CR, Umbach DM (1999). Using pooled exposure assessment to improve efficiency in case–control studies. *Biometrics* **55**, 718–726.

Willett W (1998). *Nutritional epidemiology*. Oxford University Press: New York.

Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J *et al.* (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* **122**, 51–65.

Willett WC, Stampfer MJ, Underwood BA, Speizer FE, Rosner B, Hennekens CH (1983a). Validation of a dietary questionnaire with plasma carotenoid and alpha-tocopherol levels. *Am J Clin Nutr* **38**, 631–639.

Willett WC, Stampfer MJ, Underwood BA, Taylor JO, Hennekens CH (1983b). Vitamins A, E, and carotene: effects of supplementation on their plasma levels. *Am J Clin Nutr* **38**, 559–566.

Yong L-C, Forman MR, Beecher GR, Graubard BI, Campbell WS, Reichman ME *et al.* (1994). Relationship between dietary intake and plasma concentrations of carotenoids in premenopausal women: application of the USDA–NCI carotenoid food-composition database. *Am J Clin Nutr* **60**, 223–230.