New Carotenoid Values for Foods Improve Relationship of Food Frequency Questionnaire Intake Estimates to Plasma Values[1]

Cheryl Ritenbaugh,[2] Yei Mei Peng, Mikel Aickin, Ellen Graver, Michelle Branch, and David S. Alberts

Department of Family and Community Medicine and Graduate Program in Public Health, University of Arizona Health Sciences Center, Tucson, Arizona 85724 [C. R.]; Departments of Analytical Chemistry [Y. M. P.], Family and Community Medicine [M. A.], Biometry [M. A.], and Medicine [D. S. A.], Arizona Cancer Center, Tucson, Arizona 85724; and Nutrition Core, Arizona Cancer Center, Tucson, Arizona 85719 [E. G., M. B.]

Abstract

Carotenoid consumption is of great interest in disease prevention studies. Until recently, carotenoid food composition data have not been available from a single laboratory source with high validity/reliability characteristics. With the availability of a new carotenoid food composition data base, we examined the impact of the new data base on the intake estimates as measured by a food frequency questionnaire and on the relationship of those estimates to plasma values to ascertain what, if any, improvement is achieved through use of the new values. Plasma samples were available for 162 healthy adults participating in cancer prevention studies at the Arizona Cancer Center, including men and women, smokers and nonsmokers. A single laboratory analyzed plasma samples for β-carotene, α-carotene, lutein, and lycopene. All subjects had completed a modified version of the Block food frequency questionnaire, which calculates carotenoids using a literature-based algorithm. A new carotenoid composition data base using recently published data (A. R. Mangels *et al.*, J. Am. Diet. Assoc., *93:* 284–296, 1993) was then directly substituted for the Block data base. There were high correlations between intake estimates derived from the two data bases for all four carotenoids (range, $r = 0.76–0.96$). Average intake estimates based on the Mangels *et al.* data base were significantly higher for β-carotene and lycopene; however, correlations between intakes and plasma values were significantly different only for β-carotene ($r = 0.44$ for Mangels *versus* 0.32 for Block, $P = 0.015$).

Introduction

Numerous epidemiological studies have now demonstrated associations between fruit and vegetable consumption and a lower risk of cancer (1, 2) and heart disease (3). The unanswered questions revolve around the specific compounds responsible for this lowering of risk. The carotenoids have been implicated as potentially playing a role, and a recent international study on diet and heart disease has defined the ranges of plasma carotenoid values that are associated with minimum heart disease risk (3). The results of recent β-carotene supplementation trials (14, 15) confirmed that individuals with the highest baseline plasma β-carotene levels were at the lowest risk for lung cancer; however, supplementation at around 10 times a moderately high daily dose resulted in higher risk. High baseline plasma β-carotene in these studies could be a marker of lower tobacco exposure and of consumption of other important compounds from fruits and vegetables; alternatively, it may be that optimal plasma β-carotene levels are below those achieved in supplementation.

One impediment to discerning whether consumption of specific carotenoids can account for risk differentials has been the lack of an adequate food composition data base. Block and colleagues (4, 5) developed a carotenoid data base to accompany her HHHQ[3] based on available literature. The literature values, however, were developed using a variety of different laboratory protocols, including variation in sample collection and handling and in HPLC procedures. Block (5) recognized that the values were provisional.

Chug-Ahuja *et al.* (6) and Mangels *et al.* (7) have recently reported on the development of a carotenoid data base for fruits and vegetables that is based on standardized laboratory assessment. Forman *et al.* (8) used these new data to evaluate the relationship of the intake records and FFQ estimates to plasma carotenoid levels and have shown correlation values similar to those reported by others; however, no published studies have compared the two data bases to one another, neither in terms of carotenoid intake estimates nor for their correlations with plasma levels.

The University of Arizona has utilized a scannable modified version of the HHHQ in its cancer prevention studies since 1985. In several studies, data were available on plasma carotenoids and on intake as estimated by FFQs. Once the Mangels *et al.* (7) data base became available, the Nutrition Core staff of the Arizona Cancer Center developed a data base for the FFQ, which used the new values. We report here a direct comparison of carotenoid intake estimates from the two data bases and a comparison of each result with measured plasma concentrations.

Received 2/2/96; revised 7/29/96; accepted 8/1/96.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] Supported in part by Grant NIH-P01-CA27502.
[2] To whom requests for reprints should be addressed, at Graduate Program in Public Health, University of Arizona, Arizona Health Sciences Center, 1501 North Campbell Avenue, Tucson, AZ 85724. Phone: (520) 626-3200; Fax: (520) 626-3206.

[3] The abbreviations used are: HHHQ Health Habits and History Questionnaire, definition; HPLC, high-performance liquid chromatography; FFQ, food frequency questionnaire; BMI, body mass index.

Table 1 Summary characteristics of the study sample

	Females (n = 69)	Males (n = 93)
Age (yr)	57.3 ± 11.7[a]	57.5 ± 11.0
BMI (kg/m^2)	24.9 ± 4.8	25.7 ± 3.5
No. of smokers (%)	12 (17.3)	25 (26.9)

[a] Mean ± SD.

Materials and Methods

Population. One hundred sixty-two healthy adult men and women participating in four studies at the Arizona Cancer Center provided FFQs and blood samples for carotenoid determinations. Three of the studies examined the plasma pharmacokinetics of supplementation with carotenoids and vitamin E (9). Subjects in these studies were healthy men and women, smokers and nonsmokers, between the ages of 50 and 65 years, who did not consume vitamin supplements. FFQs and plasma levels used in this article are all from the baseline data collection. The fourth study was an evaluation of the relationship of diet, plasma, and tissue levels among certain patients and controls (10). The data used here are from the control group, a population-based sample identified through random-digit dialing. Plasma β-carotene was measured in all subjects; other carotenoids were measured in subsamples. Smokers and nonsmokers were present in all study populations.

FFQs. The questionnaire in use was the Arizona FFQ, a scannable modification of Block's HHHQ. Modifications made in the form, beyond those adjustments required to make it scannable, were made to elicit more complete information on high-fiber foods. No foods were added to the standard form that were not available in the Block data base.

Data Base Development. For this article, carotenoid intake was initially assessed using Block's carotenoid file output, which includes β-carotene, α-carotene, lutein, and lycopene based on algorithms developed from the literature (5). A revised data base was then developed using the laboratory determinations (7). The same FFQs were reanalyzed using Block's program; however, in the second analysis the final four nutrients of the standard food composition data base were replaced with the carotenoids. Comparisons presented here are from the same FFQs analyzed by the two different data bases. Because Block's values are imputed by algorithm, we made no direct comparison of carotenoid values on an item-by-item or line-by-line basis.

Plasma Carotenoid Methods. Approximately 5 ml of blood were drawn by venipuncture into a foil-wrapped, green top tube and immediately placed on ice. The blood samples were centrifuged at 1200 × g for 10 min at 4°C. Plasma was collected and aliquoted 0.25 ml/tube. The samples were sorted at −70°C until analysis.

Extractions were carried out under red dim light using our published procedures with some modifications (11, 12). Briefly, 250 μl of 1% SDS in ethanol containing 0.1% butylated hydroxytoluene (w/v/w) were added to 0.25 ml of plasma. After mixing, the micronutrients were extracted with 500 μl of hexane containing 0.1% butylated hydroxytoluene (w/v) twice. The dried extract was stored at −20°C until HPLC analysis within 3 days.

HPLC analysis was performed as described using a Waters 600E multisolvent delivery system, a Waters 715 ULTRA WISP autoinjector, and a Hewlett-Packard 1040M photodiode array detector with a chem station (11). Two Novapak C_{18} columns (4 μm, 300 × 3.9 mm; Waters Associates, Milford, MA) connected in a series were used to separate the micronutrients. To simultaneously separate the carotenoids, a gradient mobile phase at a flow rate of 1.3 ml/min was used. Mobile phases A and B consisted of acetonitrile, tetrahydrofuran, methanol, and 1% ammonium acetate in different proportions. Additions of 0.1% triethylamine (v/v) to both mobile phases improved the consistency of carotenoid values with time. The total run time, including reequilibration, was 47 min. The HPLC effluent was monitored at 452 nm for the carotenoids. The analytical standards and the validity of the extraction and HPLC procedure have been reported previously (11).

Data Analysis. Files from the four studies were merged into a single file and screened for outliers. Retained values of carotenoids were compared between the two data bases using two-sided paired t tests. For each data base, partial correlation with the plasma value was computed by correlating the standardized residuals from regressions on the BMI and energy intake within sex and smoking status groups. Significance of differences between partial correlations were determined by testing the slope of the regression of the difference between the standardized residual intake values on the standardized residual plasma

Table 2 Summary statistics of carotenoid intake estimates using Block's and Mangels' data bases by sex and smoking status

	Female nonsmokers (n = 56)	Female smokers (n = 12)	Male nonsmokers (n = 67)	Male smokers (n = 25)	Total (n = 160)
β-carotene (mg/day)					
Mangels'	3.42 ± 2.06[a]	2.90 ± 2.11	3.64 ± 2.23	3.27 ± 1.60	3.45 ± 2.07[b]
Block's	2.61 ± 1.56	2.30 ± 2.39	3.26 ± 2.82	2.70 ± 1.56	2.87 ± 2.24[b]
α-carotene (mg/day)					
Mangels'	0.53 ± 0.35	0.51 ± 0.43	0.61 ± 0.58	0.46 ± 0.29	0.55 ± 0.45[c]
Block's	0.52 ± 0.37	0.42 ± 0.36	0.77 ± 0.89	0.43 ± 0.29	0.60 ± 0.64[c]
Lycopene (mg/day)					
Mangels'	0.39 ± 0.27	0.38 ± 0.30	0.45 ± 0.30	0.46 ± 0.31	0.43 ± 0.29[d]
Block's	0.40 ± 0.29	0.38 ± 0.32	0.39 ± 0.32	0.40 ± 0.33	0.40 ± 0.31[d]
Lutein (mg/day)					
Mangels'	1.63 ± 0.84	1.93 ± 2.26	1.97 ± 1.39	2.12 ± 1.79	1.89 ± 1.47[b]
Block's	1.93 ± 1.13	2.39 ± 3.28	2.11 ± 1.74	2.31 ± 2.39	2.13 ± 1.97[b]

[a] Mean ± SD.
[b] Two-sided P values by paired t test, P = 0.0001.
[c] Two-sided P values by paired t test, 0.1339.
[d] Two-sided P values by paired t test, 0.0063.

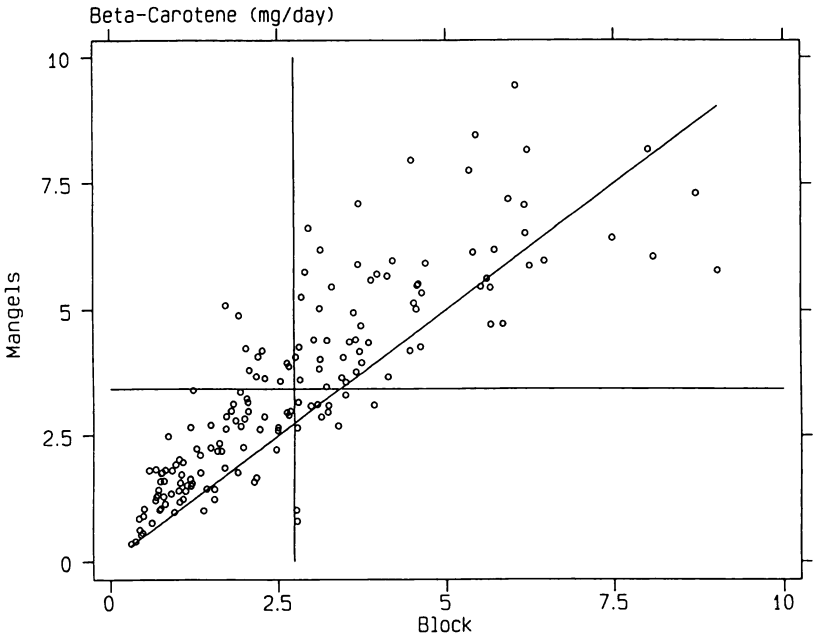

Fig. 1. Comparison of dietary β-carotene intakes as estimated by FFQ using two different food composition data bases, Block's (5) and Mangels' *et al.* (7). *Horizontal* and *vertical lines*, respective population means. *Diagonal line*, line of identity. $r = 0.82$ for both males and females.

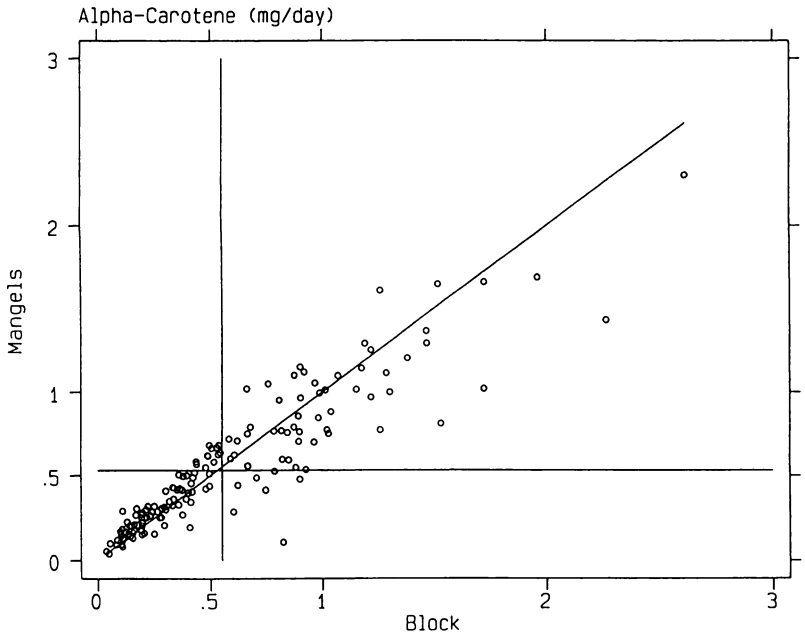

Fig. 2. Comparison of dietary α-carotene intakes as estimated by FFQ using two different food composition data bases, Block's (5) and Mangels' *et al.* (7). *Horizontal* and *vertical lines*, respective population means. *Diagonal line*, line of identity. $r = 0.85$ and 0.90 for males and females, respectively.

value. All analyses were carried out using Stata (Stata Corp., College Station, TX).

Results

Table 1 shows the characteristics of the individuals in the study. These are normal healthy older adults consistent with the populations seen in cancer prevention studies.

Table 2 compares carotenoid intakes estimated using the two data bases, designated as Block and Mangels for convenience. Because of intake differences between males and females and smokers and nonsmokers and because the diet-plasma relationships may vary by these same categories, we have presented the data separately and combined. There is no single consistent pattern of differences across all carotenoids and all groups between the Block and Mangels carotenoid estimates. However, the values for β-carotene and lycopene from the Mangels data base are higher across all groups but one, and the overall differences are significant at $P = 0.0001$ and 0.0063, respectively. The Block data base gives higher values for lutein than the Mangels data base, significant at $P = 0.0001$. The α-carotene values are not significantly different overall ($P = 0.1339$).

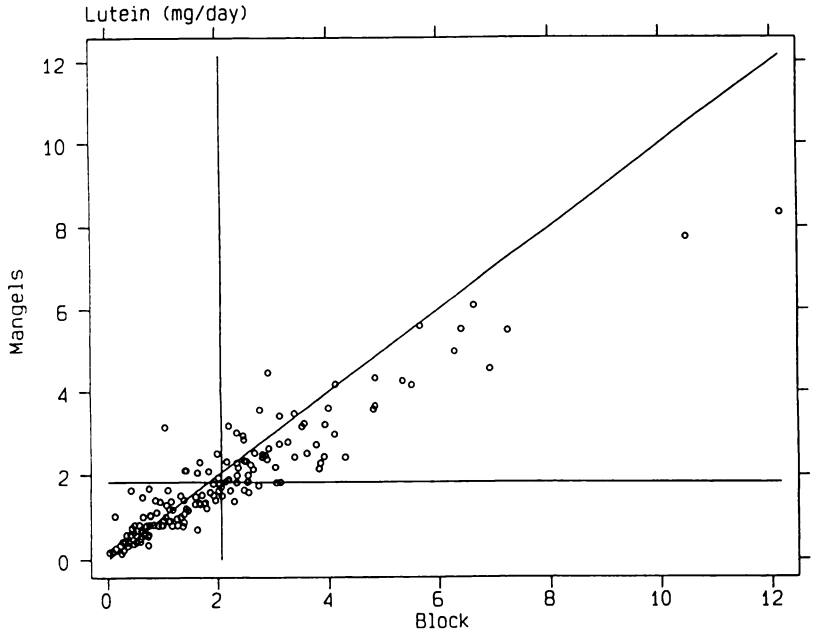

Fig. 3. Comparison of dietary lutein intakes as estimated by FFQ using two different food composition data bases, Block's (5) and Mangels' *et al.* (7). *Horizontal* and *vertical lines*, respective population means. *Diagonal line*, line of identity. r = 0.94 and 0.96 for males and females, respectively.

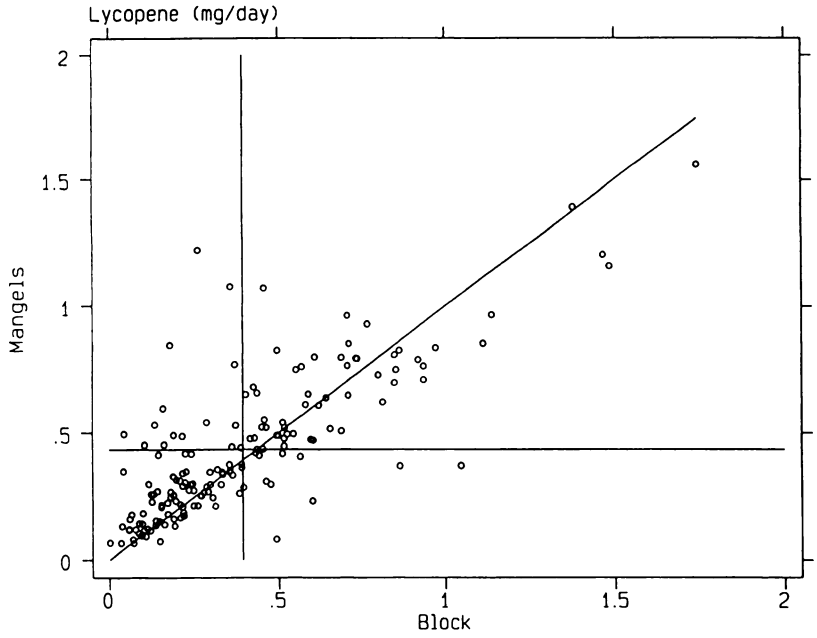

Fig. 4. Comparison of dietary lycopene intakes as estimated by FFQ using two different food composition data bases, Block's (5) and Mangels' *et al.* (7). *Horizontal* and *vertical lines*, respective population means. *Diagonal line*, line of identity. r = 0.84 and 0.78 for males and females, respectively.

Figs. 1-4 relate the Block values to the Mangels values for each carotenoid. The *horizontal* and *vertical lines* indicate the population means, and the *diagonal line* is the line of identity. Correlations between the two estimates are high, as would be expected. For males and females, respectively, the values are as follows: β-carotene: 0.82, 0.82; α-carotene: 0.85, 0.90; lutein: 0.94, 0.96; and lycopene: 0.84, 0.78. Note that two β-carotene extreme outliers are omitted from Fig. 1 and all calculations.

Table 3 provides summary information for plasma carotenoid concentrations across sex and smoking groups. Plasma β-carotene and α-carotene concentrations are lower in smokers than in nonsmokers. Plasma carotenoid concentrations were also negatively related to BMI, perhaps due to storage in adipose tissue. Subsequent analyses using plasma levels were adjusted for BMI.

Table 4 shows the correlations between plasma values and intake estimates using the two data bases overall and by sex and smoking status. All correlations were adjusted for BMI and energy. The total correlations are also adjusted for sex and smoking status. Addition of other adjustment factors [*e.g.*, alcohol intake (8)] resulted in minor and inconsistent changes (data not shown) and were considered to add no additional information. The correlation coefficients shown here could not be deattenuated because multiple measures of intake and

Table 3 Summary statistics of plasma carotenoid concentrations by sex and smoking status

	Female nonsmokers	Female smokers	Male nonsmokers	Male smokers	Total
β-carotene (μmol/liter)	0.56 ± 0.44[a]	0.26 ± 0.12	0.41 ± 0.34	0.23 ± 0.12	0.43 ± 0.36
n	56	12	63	25	157
α-carotene (μmol/liter)	0.15 ± 0.13	0.09 ± 0.05	0.11 ± 0.12	0.08 ± 0.05	0.12 ± 0.11
n	56	12	55	21	144
Lycopene (μmol/liter)	0.64 ± 0.26	0.60 ± 0.36	0.66 ± 0.27	0.58 ± 0.31	0.63 ± 0.28
n	56	12	55	21	144
Lutein (μmol/liter)	0.18 ± 0.08	0.18 ± 0.07	0.20 ± 0.10	0.14 ± 0.05	0.18 ± 0.09
n	31	12	28	21	92

[a] Mean ± SD.

Table 4 Partial correlations between plasma concentrations and dietary intake by FFQ: sex and smoker specific correlations adjusting for BMI and energy (kcal)

	Female nonsmokers	Female smokers	Male nonsmokers	Male smokers	Total
β-carotene					
Mangels'	0.577	0.506	0.319	0.556	0.441[a]
Block's	0.507	−0.037	0.254	0.397	0.317[a]
n	56	12	67	25	160
α-carotene					
Mangels'	0.369	0.171	0.507	0.671	0.431[b]
Block's	0.405	0.089	0.490	0.595	0.432[b]
n	56	12	58	21	147
Lycopene					
Mangels'	0.374	0.319	0.208	0.150	0.265[c]
Block's	0.374	0.349	0.226	0.301	0.295[c]
n	56	12	58	21	147
Lutein					
Mangels'	0.449	−0.014	0.358	0.405	0.301[d]
Block's	0.390	−0.012	0.325	0.364	0.255[d]
n	31	12	29	21	93

[a] Correlations differ at $P = 0.005$.
[b] Correlations differ at $P = 0.963$.
[c] Correlations differ at $P = 0.562$.
[d] Correlations differ at $P = 0.176$.

plasma levels, which are needed to develop the deattenuation procedure, were not available. Thus, the actual level of the relationships between intake measures and plasma levels is stronger than shown in Table 4. The only carotenoid for which the correlations differ significantly between the intake estimates for the two data bases is β-carotene, with the Mangels' data base correlated more strongly with the plasma values than the Block data base (0.441 *versus* 0.317, respectively, $P = 0.015$). There were no consistent sex differences in the patterns of correlations. Similarly, there were no consistent differences between the correlations for smokers and nonsmokers.

Discussion

It can be seen from the correlations between the two methods that the carotenoid estimates used by Block (5) based on previous literature were extremely close to those subsequently measured in the laboratory and reported by Chug-Ahuja *et al.* (6) Mangels *et al.* (7).

In this study, neither smoking status nor gender seemed to play a consistent role in the relationship of intake estimates to plasma values. The correlations between diet intake estimates and measured plasma values of carotenoids tended to be higher for women than for men for β-carotene and lycopene and higher for men than for women for α-carotene and mixed for lutein, indicating that the women were not necessarily better reporters than men. The correlations for female nonsmokers were higher than for female smokers on all four carotenoids; this may be a mixture of biological and small sample size effects. Among males, there was a tendency for higher correlations among the smokers, contrary to expectations.

For research on diet and cancer prevention, we are largely dependent on self-reports, whether through food records or FFQs. It has been shown that the relationships between intake estimates measured by FFQs and biological markers are generally less strong than those between food records and these markers (8). This has been attributed to human factors involved in the responses to the questionnaires, such as knowledge of diet and willingness to be honest. The human factors are difficult to change, and thus there has been a reluctant acceptance of the notion that we may not be able to improve the level of accuracy provided by FFQs.

It is encouraging to note that this is not the case and that some improvements in the relationship of intake estimates to biological parameters can still be made through improvements on the questionnaire side, in this case by improving the food composition data base. The change in correlation from 0.32 to 0.44 doubles the percentage of variance explained (0.10–0.19); from another perspective, this suggests that there might be an 18% reduction in the bias of a regressive coefficient for dietary β-carotene. Although these are not huge changes, they are good

news, and the new values are a welcome addition. Other data have now been published which may similarly assist in cancer prevention studies (13).

The improvement in correlation, although small, also may encourage us to look more closely at the appropriateness of the FFQ structure and nutrient and portion size data bases for each population under study, as was recommended by the First International Conference on Dietary Assessment Methods in Minneapolis in 1992. Validity and reliability estimates are not a measure of the questionnaire *per se*, but rather are a function of the interaction of a specific questionnaire and a particular target population. There are likely to be ethnic and regional differences in typical portion sizes, recipes, and in the most commonly consumed foods. Researchers will need to identify such differences as new populations are studied. We are hopeful that attention to such detail will lead to improvements on the same order of magnitude or greater than those shown here for β-carotene.

References

1. Steinmetz, K. A., and Potter, J. D. Vegetables, fruit, and cancer. I. Epidemiology. Cancer Causes & Control, *2:* 325–357, 1992.

2. Block, G., Patterson, B., and Subar, A. Fruits, vegetables, and cancer prevention: a review of the epidemiological evidence. Nutr. Cancer, *18:* 1–29, 1992.

3. Gey, K. F., Moser, U. K., Jordan, P., Stahelin, H. B., Eichholzer, M., and Ludin, E. Increased risk of cardiovascular disease at suboptimal plasma concentrations of essential antioxidants: an epidemiological update with special attention to carotene and vitamin C. Am. J. Clin. Nutr., *57* (Suppl)*:* 787s–797s, 1993.

4. Block, G., Hartmann, A. M., Dresser, C. M., Carroll, M. D., Gannon, J., and Gardner, L. A data based approach to diet questionnaire design and testing. Am. J. Epidemiol., *124:* 453–469, 1986.

5. Block, G. Health, Habits, and History Questionnaire—Diet History and Other Risk Factors, Personal Computer System Packet. Bethesda, MD: National Cancer Institute, 1989.

6. Chug-Ahuja, J. K., Holden, J. M., Forman, M. R., Mangels, A. R., Beecher, G. R., and Lanza, E. The development and application of a carotenoid database for fruits, vegetables, and selected multicomponent foods. J. Am. Diet. Assoc., *93:* 318–323, 1993.

7. Mangels, A. R., Holden, J. M., Beecher, G. R., Forman, M. R., and Lanza, E. The carotenoid content of fruits and vegetables: an evaluation of analytical data. J. Am. Diet. Assoc., *93:* 284–296, 1993.

8. Forman, M. R., Lanza, E., Yong, L-C., Holden, J. M., Graubard, B. I., Beecher, G. R., Melitz, M., Brown, E. D., and Smith, J. C. The correlation between two dietary assessments of carotenoid intake and plasma carotenoid concentrations: application of a carotenoid food-composition database. Am. J. Clin. Nutr., *58:* 519–524, 1993.

9. Xu, M. J., Plezia, P. M., Alberts, D. S., Emerson, S. S., Peng, Y. M., Sayers, S. M., Liu, Y., Ritenbaugh, C., and Gensler, H. L. Reduction in plasma or skin α-tocopherol concentration with long-term oral administration of β-carotene in humans and mice. J. Natl. Cancer Inst., *84:* 1559–1564, 1993.

10. Peng, Y. M., Peng, Y. S., Lin, T., Moon, T., Roe, D., Aickin, M., and Ritenbaugh, C. Concentrations and plasma-tissue-diet relationships of carotenoids, retinoids and tocopherols in humans. Nutr. Cancer, *23:* 233–246, 1995.

11. Peng, Y. S., and Peng, Y. M. Simultaneous liquid chromatographic determination of carotenoids, retinoids and tocopherols in human buccal mucosal cells. Cancer Epidemiol. Biomarkers & Prev., *1:* 375–382, 1992.

12. Peng, Y. M., Peng, Y. S., and Lin, Y. A nonsaponification method for the determination of carotenoids, retinoids, and tocopherols in solid human tissues. Cancer Epidemiol Biomarkers & Prev., *2:* 139–144, 1993.

13. Marlett, J. A. Content and composition of dietary fiber in 117 frequently consumed foods. J. Am. Diet. Assoc., *92:* 175–186, 1992.

14. The α-Tocopherol, Betacarotene Cancer Prevention Study Group. The effect of vitamin E and betacarotene on the incidence of lung cancer and other cancers in male smokers. N. Engl. J. Med., *330:* 1029–1035, 1994.

15. Omenn, G. S., Goodman, G. E., Thornquist, M. D., Balmes, J., Cullen, M. R., Glass, A., Keogh, J. P., Meyskens, F. L., Jr., Valanis, B., Williams, J. H., Barnhart, S., and Hammar, S. Effects of a combination of betacarotene and vitamin A on lung cancer and cardiovascular disease. N. Engl. J. Med., *334:* 1150–1155, 1996.

New carotenoid values for foods improve relationship of food frequency questionnaire intake estimates to plasma values.

C Ritenbaugh, Y M Peng, M Aickin, et al.

Cancer Epidemiol Biomarkers Prev 1996;5:907-912.