

Using a correction factor to correct for overreporting in a food-frequency questionnaire does not improve biomarker-assessed validity of estimates for fruit and vegetable consumption

Citation for published version (APA):

Bogers, R. P., Dagnelie, P. C., Westerterp, K. R., Kester, A. D. M., van Klaveren, J. D., Bast, A., & van den Brandt, P. A. (2003). Using a correction factor to correct for overreporting in a food-frequency questionnaire does not improve biomarker-assessed validity of estimates for fruit and vegetable consumption. *Journal of Nutrition*, 133(4), 1213-1219. <https://doi.org/10.1093/jn/133.4.1213>

Document status and date:

Published: 01/01/2003

DOI:

[10.1093/jn/133.4.1213](https://doi.org/10.1093/jn/133.4.1213)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 17 Apr. 2024

Using a Correction Factor to Correct for Overreporting in a Food-Frequency Questionnaire Does Not Improve Biomarker-Assessed Validity of Estimates for Fruit and Vegetable Consumption¹

Rik P. Bogers,² Pieter C. Dagnelie, Klaas R. Westerterp,* Arnold D. M. Kester,[†] Jacob D. van Klaveren,** Aalt Bast[‡] and Piet A. van den Brandt

*Department of Epidemiology, Maastricht University, Maastricht, The Netherlands; *Department of Human Biology, Maastricht University, Maastricht, The Netherlands; †Department of Methodology & Statistics, Maastricht University, Maastricht, The Netherlands; **State Institute for Quality Control of Agricultural Products, Wageningen, The Netherlands; and ‡Department of Pharmacology & Toxicology, Maastricht University, Maastricht, The Netherlands*

ABSTRACT To correct for overreporting of fruit and vegetable (FV) consumption in a food-frequency questionnaire, summary questions about consumption of main FV groups are often used to calculate correction factors. This study compared the ability to rank people according to their FV intake of those summary questions and the sum of questions on individual FV items within categories, and of corrected or uncorrected estimates of specific sorts of FV. Healthy middle-age women ($n = 161$) completed a food-frequency questionnaire about FV consumption during the previous month and gave a single fasting blood sample. Correction factors were calculated as the reported frequency on a summary question divided by the summed frequencies of all items in a category. Plasma carotenoids and vitamin C served as biomarkers of FV consumption. Significant correlations between FV consumption and biomarkers were observed (e.g., Spearman's correlation coefficient r with total carotenoids/vitamin C: 0.32/0.34 for vegetables, 0.30/0.25 for fruits). Summary estimates of cooked, raw and total vegetable consumption correlated higher with biomarkers than sum estimates. For fruits no differences in correlations between sum and summary estimates were observed. Applying a correction factor on the consumption of carrots and total cabbage resulted in lower correlations with relevant biomarkers. For broccoli/cauliflower, Brussels sprouts and citrus fruits, correlations with biomarkers did not change after correction. We conclude that summary questions may suffice to rank individuals according to their intake of main FV categories, and that correction for overreporting of individual FV items is probably not advisable when ranking individuals according to intake of these items. *J. Nutr.* 133: 1213–1219, 2003.

KEY WORDS: • food-frequency questionnaire • validity • biomarkers • fruits • vegetables

Many human observational studies suggest that consumption of fruit and vegetables (FV)³ or associated micronutrients is beneficial in the prevention of cancer (1–4) and cardiovascular disease (5–7).

For epidemiological studies aimed at further establishing associations between FV consumption and disease, a food-frequency questionnaire (FFQ) that adequately classifies persons according to their usual FV intake is a useful tool. Such a questionnaire not only must be suitable for estimating total FV consumption, but also must cover consumption of all relevant sorts of FV to calculate intake of specific (groups of) FV and of essential nutrients and other bioactive compounds.

However, estimates of consumption derived from a FFQ are never free of errors and result in biased estimates of reported

consumption frequency and portion size. Most subjects have difficulties in estimating the frequency of consumption of specific food items (in this case specific fruits or vegetables). The sum of all items within a certain food category usually results in an estimate that is higher than the true overall frequency of consumption (8,9). On the other hand, estimates derived from the sum of items might better represent true intake, because in that case a more appropriate weight belonging to each specific item instead of a mean weight for all the items can be used in the calculations (e.g., an apple weighing more than a tangerine).

In studies where the intake of specific items or nutrients is of interest, a common practice to correct for this overreporting is to include a summary question (how often do you eat vegetables?) preceding or following questions on intake of specific items within a food category. In data analysis, frequencies of intake of specific items are then calibrated so that the sum of all items equals the overall frequency (10). However, the question arises whether in the case of vegetables this is a

¹ Supported by the Dutch Product Board for Horticulture.

² To whom correspondence should be addressed.

E-mail: rik.bogers@epid.unimaas.nl.

³ Abbreviations used: FFQ, food-frequency questionnaire; FV, fruits and vegetables.

valid method, because more than one type of vegetable may be served at the same meal. Also for fruits, different sorts may be eaten during a single day. Amanatidis et al. (9) found that applying a correction factor did not substantially affect the ranking of subjects according to intake of various micronutrients. However, they did not have an independent reference method against which validity of corrected and uncorrected estimates could be determined.

The concept of validity consists of several aspects, each of which can be addressed depending on the purpose and context of the method used (11,12). The type of validity we refer to in the present study concerns the relative ranking of subjects by two different methods. A correlation coefficient is the typical measure used to evaluate this aspect of validity (12).

One way of validating FFQ is by using biochemical markers of dietary intake. Although there are several drawbacks in the use of biomarkers (13), they provide an objective measure of intake of which the errors are largely independent of the errors associated with FFQ (14). Frequently used biomarkers for FV intake are plasma carotenoids and vitamin C, which have been shown to be responsive to intake of FV (15–18). Significant (although moderate) correlations between plasma or serum concentrations of carotenoids and carotenoid intake (19–24) or FV consumption (24–28) have been observed, and plasma/serum vitamin C was found to be significantly correlated with FV intake (26,28) or vitamin C intake (29). This study focuses on two questions. First, are summary questions on FV consumption better able to rank subjects according to their usual FV consumption than the sum of specific items within each category? This question was studied for the categories total vegetables (excluding potatoes), cooked vegetables (excluding potatoes), raw vegetables, fruits and fruit juice. Second, do calibrated estimates of three (groups of) FV (cabbage, carrots and citrus fruits) lead to a better ranking of subjects than uncalibrated estimates? Plasma carotenoids (α -carotene, β -carotene, lutein, lycopene, β -cryptoxanthin) and vitamin C were used to serve as independent estimates of FV consumption as criterion variables against which validity was assessed. At the same time, this study can be regarded as a validation study of the FFQ we developed for estimating FV consumption.

SUBJECTS AND METHODS

Study design. The study was part of a larger project aimed at increasing FV consumption in children aged 7–10 y and their mothers, for which we received approval by the medical ethical committee of the University Hospital in Maastricht. Participants completed an FFQ about FV consumption in the previous month, after which blood was taken from the mothers for the analysis of biomarkers. FV consumption was calculated both with summary questions about FV categories (cooked vegetables, raw vegetables, fruits, fruit juice) and with extensive lists of single items within each category. Consumption of specific fruits and vegetables was calculated with and without correction for overreporting. All data presented in this publication were collected in March 2001.

Participants. From the population registry of the municipality of Maastricht, a random sample was obtained containing 2000 addresses of mothers with children aged 7–10 y. A letter was sent inviting the mother and the selected child in each family to participate. An additional 1100 letters were distributed in elementary schools in Maastricht. The women had to be apparently healthy nonsmokers who agreed not to use vitamin supplements from 1 mo before the first blood collection to the end of the study period. We explained to the participants that the purposes of the study were to develop questionnaires to assess both FV intake and determinants of FV consumption (the latter, in another part of the project, are not described here). A total of 207 volunteers were recruited (6.7% response rate), of whom

163 women (79%) fully completed the study, making an overall response rate of 5.3%. Two subjects having FFQ with >20% missing values were excluded from further analyses, leaving a total number of 161 subjects.

Food-frequency questionnaire. The FFQ was designed to measure FV intake during the past month and was completed in March 2001. We distributed the questionnaires by mail and asked the participants to fill out the FFQ on the day before blood collection.

The FFQ we used was a modified version of the VEG-FFQ (from “vegetable”) developed at the Dutch State Institute for Quality Control of Agricultural Products (30). This VEG-FFQ was a 172-item semiquantitative food-frequency questionnaire that was developed and first used as a part of the Third Dutch National Food Consumption Survey in 1998 (31). Questions were asked separately for summer and winter. In the design of the VEG-FFQ, existing Dutch FFQ (32–37) served as guidelines. Specific features of the VEG-FFQ were as follows: the use of household measures and natural units in questions on amounts eaten; ordering of the questions according to the Dutch meal pattern; questions at the level of meals; addition of overall questions on each product category (e.g., fruits) for correction for overreporting; grouping products within the same food category and ordering foods within a category according to frequency of consumption of products; taking into account seasonal differences in FV consumption. The food categories included in the VEG-FFQ were soup, warm meals, potatoes, mashed vegetables, salads, cooked vegetables, fruits (including apple sauce), drinks and snacks. Questions on foods were noted as consumption frequency and portion size, except for sweet peppers, mushrooms, tomatoes and onions, which were estimated by asking the amount eaten per month or week. The seven frequency categories used were: never or less than 1 d/mo, 1 d/mo, 2–3 d/mo, and 1, 2–3, 4–5 and 6–7 d/wk.

Before the present study the VEG-FFQ had been used in 1592 out of 5958 participants of the Third Dutch National Food Consumption Survey in which subjects recorded their 48-h food intake on two consecutive days. These 1592 participants completed the VEG-FFQ 4–8 wk after the 48-h records. It was found that mean FV intake calculated using the VEG-FFQ corresponded well with mean intake calculated from the 48-h records. After correction for overreporting using summary questions on food categories, mean vegetable intake was 105 g/d according to the VEG-FFQ and 123 (95) g/d according to the records; for fruit intake, this was 114 and 105 (114) g/d, respectively (30). Further validation was not performed.

For the present study, we slightly modified the VEG-FFQ to fit into the study design. Separate questions on consumption of FV in summer and winter were combined into questions on FV consumption during the past month. Open-ended questions on average portion size were changed to multiple-choice questions so that the questionnaires could be read by an optical scanner. The result was a 106-item semiquantitative FFQ. Calculated FV categories were fruits (including fresh orange and grapefruit juice), cooked vegetables (not including potatoes), raw vegetables, total vegetables (the sum of cooked and raw vegetables) and fruit juice (from a carton or bottle). For each of these categories, a summary question was asked, followed by a list of items belonging to a category that we used for calculating the sum estimates. The numbers of items were 17 (fruits), 21 (cooked vegetables), 14 (raw vegetables) and 5 (fruit juice).

Assessment of biomarkers in blood. Blood was taken shortly after the participants completed the FFQ. After an overnight fast, 14 mL venous blood was collected between 0700 and 1100 h in two K₂EDTA-coated plastic tubes (Vacutainer Systems; Becton Dickinson, Plymouth, UK). Both tubes were immediately placed on ice in the dark.

Total ascorbic acid was determined according to the method of Speek et al. (38). Briefly, blood was collected in a 4-mL tube in which 0.08 mL EGTA-glutathione solution had been injected before blood sampling. A 250- μ L aliquot of blood was stabilized in 1 mL trichloroacetic acid and centrifuged after 20 to 60 min (15,000 \times g, 10 min, 4°C). A pilot study showed that ascorbic acid in trichloroacetic acid placed on ice is stable for at least 60 min of incubation. Of the supernatant, 350 μ L was transferred in amber Eppendorf vials (Eppendorf-Netherler-Hinz-GmbH, Hamburg, Germany) and stored at –80°C. The samples were analyzed within 2 wk with HPLC with

fluorometric detection. For this purpose, L-ascorbic acid was enzymatically oxidized to dehydro-L-ascorbic acid and the latter was condensed with *o*-phenylenediamine to its fluorescent quinoxaline derivative. The HPLC mobile phase consisted of 0.08 mol/L potassium dihydrogen phosphate and methanol (4:1 v/v).

Plasma carotenoids were analyzed according to Hess et al. (39), with modifications described by Oostenbrug et al. (40). Briefly, after blood was drawn in a 10-mL tube it was centrifuged (1000 × *g*, 10 min, 4°C), and the plasma was transferred into amber Eppendorf vials and stored at -80°C. The plasma was analyzed within 5 mo using HPLC with UV detection. First, 500 μL double-distilled water was added to 500 μL plasma and this mixture was denatured with 1 mL of a mixture of absolute ethanol and methanol (1:1 v/v). Carotenoids were twice extracted into 3 mL hexane. The pooled supernatant was evaporated to dryness under nitrogen at 37°C and resuspended in 200 μL ethanol-dioxan (1:1 v/v), after which 300 μL acetonitrile was added. The HPLC mobile phase consisted of acetonitrile, tetrahydrofuran, methanol and double-distilled water (684:220:68:28 v/v/v/v). Ethyl-β-apo-8'-carotenoate was used as an internal standard.

Calculations of FV intake and statistical analysis. Statistical analyses were performed by use of the SPSS program (version 10.0; SPSS, Chicago, IL). Standard amounts as indicated in a Dutch table for measures and weights (41,42) were used for converting household measures to portion sizes. Missing values for portion sizes on a particular item were substituted with mean portion sizes for the item concerned. FV intake (in g/d) was calculated as the product of frequency of intake and portion size. Mean values were used to substitute missing values for FV intake.

Besides being included under the heading "raw vegetables," consumption of sweet peppers and tomatoes was recorded as the absolute amount eaten both raw and cooked per week or month. After subtraction of raw sweet peppers and tomatoes we added these amounts to cooked vegetables. We supposed that onions and mushrooms, which were also included as raw and cooked amounts eaten per month, were eaten predominantly cooked, and added these two items to cooked vegetables. The correction factor for calibrating items was calculated as the reported frequency on the summary question divided by the sum of frequencies of all items belonging to a category. Corrected intake of items was calculated by multiplying intake of an item by the correction factor. Corrected intake of a category was the sum of all corrected items in a category. The relative validity of corrected and uncorrected intakes was determined for carrots, citrus fruits and cabbage (total cabbage, Brussels sprouts and broccoli/cauliflower), because of their known high contents of α-carotene, β-cryptoxanthin and lutein, respectively.

Agreement between the summary question and the sum of individual items within a category was determined by cross-classifying quartiles of FV intake according to the sum of items in a category with quartiles according to the summary question. Validity was calculated as correlation coefficients between FV intake and biomarkers. We used Spearman's correlation coefficients because of the skewed distribution of most food consumption data. To compare the validity of two estimates of FV consumption, equality of correlations was tested using a formula stated by Olkin and Siotani (43,44).

RESULTS

Table 1 provides a summary of some of the participants' characteristics. **Table 2** presents a comparison between mean FV consumption derived from one summary question or from summing all items in a category. The greatest difference in intake between these two methods was for cooked vegetables: 110 g/d using the summary question (summary amount) vs. 204 g/d using the sum of items (sum amount). For cooked vegetables the correlation between frequencies ($r = 0.35$) and intake ($r = 0.49$) measured with either method was lowest; however, the mean correction factor of 0.89 was closer to 1.00, and the mean difference between sum and summary frequencies was lower than for raw vegetables and fruit. For fruit juice the sum amount approached the summary amount most closely (67 vs. 79 g/d), and for fruits this was 156 vs. 195 g/d. The

TABLE 1

Characteristics of the study population¹

Characteristic	
Age, y	41 ± 4.0
Body mass index, kg/m ²	24.1 ± 3.5
Educational level, %	
Low	6
Intermediate	52
High	42

¹ Values are means ± SD, $n = 161$ women.

mean correction factor for fruit juice was almost 1.00 because most people drank only orange juice, so that the sum frequency consisted of only one item.

Table 3 illustrates the agreement in results when the summary question or the sum of items in a category was used to calculate FV intake. The best agreement was found for fruits and fruit juice, where 95 and 93% of the subjects were classified in the same or adjacent quartile, respectively. Again, the lowest agreement was observed for cooked vegetables (83% in the same or adjacent quartile).

Table 4 details the descriptive statistics for the biomarkers. The carotenoid most abundant in plasma was β-cryptoxanthin, and the least abundant was α-carotene. The median values corresponded well with mean values. **Table 5** presents the results of the validation of the two methods for calculating FV consumption against plasma carotenoids and vitamin C. Values of correlations with biomarkers were consistently higher for summary amounts than for sum amounts. Some of these differences were statistically significant, especially for total vegetable consumption. In contrast, for fruit consumption the correlations with biomarkers were comparable for the summary and the sum amounts. For cooked vegetables, the highest correlation was with vitamin C (Spearman's correlation coefficient $r = 0.35$). Consumption of raw vegetables correlated highest with α-carotene and total carotenoids ($r = 0.31$ for both), and fruit consumption with β-cryptoxanthin ($r = 0.43$). Consumption of fruit juice was not significantly correlated with any of the biomarkers, regardless of whether it was expressed as summary or sum amount.

Table 6 shows the relative validity of the estimated consumption of cabbage, carrots and citrus fruits when a correction factor was used or not. For carrots, correlations with α- and β-carotene were significantly higher when consumption was not corrected for overreporting; this was also the case for α-carotene and consumption of total cabbage. However, there were no differences in correlations with biomarkers between corrected and uncorrected consumption of broccoli/cauliflower and Brussels sprouts, two specific sorts of cabbage. The correlation between consumption of citrus fruits and its most important marker, β-cryptoxanthin, did not change when consumption was corrected for overreporting.

DISCUSSION

The present study indicates that, based on biomarkers as a reference for validation of a FFQ, estimates of consumption of vegetables based on one summary question ("summary estimate") are better able to rank persons according to their FV intake than estimates obtained by summing all separate items across each category ("sum estimate"). Our data suggest that this applies to both raw and, even more so, cooked vegetables.

TABLE 2

Comparison between mean fruit and vegetable consumption derived from one summary question or from summing all items in a category, and Spearman's correlation coefficients between these two methods^{1,2}

	Summary ³	Sum ³	Spearman's <i>r</i> (<i>P</i>) ⁴	Summary frequency ³	Sum frequency ³	Spearman's <i>r</i> (<i>P</i>) ⁴	Correction factor ^{3,5}
	<i>g/d</i>	<i>g/d</i>		<i>times/d</i>	<i>times/d</i>		
Total vegetables	151 ± 69	264 ± 123	0.59 (0.00)	1.09 ± 0.45	1.83 ± 1.03	0.61 (0.00)	0.69 ± 0.30
Cooked	110 ± 51	204 ± 94	0.49 (0.00)	0.65 ± 0.22	0.83 ± 0.39	0.35 (0.00)	0.89 ± 0.42
Raw	41 ± 38	60 ± 66	0.80 (0.00)	0.44 ± 0.35	1.01 ± 0.82	0.80 (0.00)	0.51 ± 0.27
Fruit	156 ± 116	195 ± 128	0.82 (0.00)	0.69 ± 0.28	1.34 ± 0.78	0.71 (0.00)	0.59 ± 0.27
Fruit juice	67 ± 84	79 ± 95	0.81 (0.00)	0.32 ± 0.33 ⁵	0.37 ± 0.35 ⁶	0.81 (0.00)	0.97 ± 0.60

¹ Assessed with a food-frequency questionnaire in 161 women.

² Values are means ± SD.

³ All differences (paired *t* test and Wilcoxon signed rank test) between summary and sum amounts and frequencies are significant at *P* < 0.001.

⁴ *P*-values are in parentheses.

⁵ Calculated as (frequency of consumption indicated in summary question)/(sum of frequencies of consumption of all items in a category).

⁶ Sum and summary significantly different at *P* = 0.002.

For fruits, the correlations with biomarkers for summary and sum estimates are comparable. Furthermore, when intake of individual fruits and vegetables is estimated in a FFQ, correction for overreporting seems unnecessary or may even lower correlations with biomarkers.

Several possible explanations can be given as to why the summary estimate is as valid as the sum estimate for fruits, whereas this is not the case for vegetables. First, for fruits portion size is easier to indicate because, contrary to vegetables, fruits are mostly served in natural units. Second, it may be easier for subjects to estimate consumption frequencies of specific fruits because the number of different sorts of fruits eaten is usually smaller than the number of different vegetables (31). However, the mean correction factor (based on frequencies) for fruits was close to 0.5, suggesting a high degree of overestimating the frequency of consumption of individual fruits relative to the summary frequency. Rather than overestimating the frequency of consumption of individual sorts of fruits, this could be attributed to the fact that people eat more than one sort of fruit a day. The finding that the sum amount was only 25% higher than the summary amount supports this idea. In contrast, the mean correction factor was closest to unity for cooked vegetables, although consumption of cooked

vegetables (expressed in *g/d*) showed the largest difference between summary and sum estimates. This was the result of the great between-subject variation in the correction factor for cooked vegetables, where, not contrary to fruits, almost every subject reported a higher sum than summary frequency.

The observation that correction for overreporting may not improve or even lower correlations with biomarkers may be attributable to the fact that the correction factor includes estimate errors of all separate items. The sum of these errors may likely be higher than the estimate error of a single item, which would reduce the validity of an estimate of a specific fruit or vegetable when correction for overreporting is applied. As already mentioned, the consumption frequency of fruits may be easier to estimate, and as a result the error of the correction factor for fruits would probably be less than that of vegetables.

We used biomarkers as a reference method because the errors in biomarkers (e.g., biological variation between persons in uptake or metabolism of vitamins) are to a high degree statistically independent of the errors in intake estimates based on questionnaires. Although the correlations we observed are moderate, the use of biomarkers for our purpose can be justified. Because the variation in our criterion measure (the biomarkers) remains the same when we calculate correlations with either of the two dietary measures, differences between the two correlation coefficients (biomarkers vs. summary/sum

TABLE 3

Cross-classification of quartiles derived from calculating fruit and vegetable intake as the sum of all items in a category by quartiles of intake using a summary question¹

	Same quartile using summary	Adjacent quartile using summary	Two quartiles' difference using summary	Extreme opposites using summary
	%			
Total vegetables	47.2	39.2	10.6	3.1
Cooked	42.2	41.0	17.9	3.1
Raw	56.5	36.6	6.7	0.0
Fruits	59.7	35.5	4.9	0.0
Fruit juice	67.1	25.9	4.9	1.8

¹ Assessed with a food-frequency questionnaire in 161 women.

TABLE 4

Percentiles of plasma concentrations of vitamin C and carotenoids¹

	P10	P25	P50	P75	P90
	$\mu\text{mol/L}$				
Vitamin C	39.93	47.89	56.24	65.64	72.65
Total carotenoids	1.24	1.49	1.90	2.20	2.94
Lutein	0.24	0.29	0.38	0.48	0.61
β -Cryptoxanthin	0.30	0.39	0.51	0.71	1.04
Lycopene	0.14	0.22	0.26	0.36	0.43
α -Carotene	0.06	0.09	0.13	0.19	0.24
β -Carotene	0.21	0.29	0.45	0.65	1.02

¹ P10 to P90 are the 10th to the 90th percentiles, *n* = 161 women.

TABLE 5

Spearman's rank correlations of plasma carotenoid and vitamin C concentrations with fruit and vegetable intake calculated either with a summary question or as the sum of the items in a category^{1,2}

	Vitamin C	Total carotenoids	Lutein	β-Cryptoxanthin	Lycopene	α-Carotene	β-Carotene
Total vegetables							
Summary ³	0.34 (0.00)	0.32 (0.00)*	0.30 (0.00)	0.19 (0.02)*	0.00 (0.96)	0.37 (0.00)*	0.26 (0.00)*
Sum ³	0.23 (0.00)	0.22 (0.01)*	0.25 (0.00)	0.09 (0.25)*	0.03 (0.67)	0.20 (0.01)*	0.14 (0.08)*
Cooked vegetables							
Summary	0.35 (0.00)*	0.23 (0.00)	0.24 (0.00)	0.14 (0.07)	0.04 (0.62)	0.24 (0.00)	0.15 (0.05)
Sum	0.16 (0.04)*	0.19 (0.02)	0.21 (0.01)	0.08 (0.34)	0.06 (0.48)	0.14 (0.08)	0.11 (0.15)
Raw vegetables							
Summary	0.19 (0.02)	0.31 (0.00)	0.22 (0.01)	0.23 (0.00)	0.02 (0.84)	0.31 (0.00)	0.25 (0.00)
Sum	0.14 (0.08)	0.21 (0.01)	0.20 (0.01)	0.11 (0.18)	0.01 (0.88)	0.20 (0.01)	0.17 (0.03)
Fruit							
Summary	0.25 (0.00)	0.30 (0.00)	0.10 (0.21)	0.43 (0.00)	0.03 (0.69)	0.15 (0.07)	0.18 (0.02)
Sum	0.24 (0.00)	0.34 (0.00)	0.15 (0.06)	0.42 (0.00)	0.01 (0.89)	0.17 (0.04)	0.22 (0.01)
Fruit juice							
Summary	0.15 (0.06)	0.09 (0.28)	-0.02 (0.77)	0.08 (0.34)	0.02 (0.79)	0.02 (0.78)*	0.05 (0.52)
Sum	0.14 (0.08)	0.05 (0.56)	0.01 (0.93)	0.08 (0.31)	0.05 (0.50)	-0.10 (0.19)*	-0.02 (0.78)

¹ Assessed with a food-frequency questionnaire in 161 women.

² P-values are in parentheses.

³ Summary and sum refer to fruit and vegetable consumption (g/d) calculated with a summary question or as the sum of items in a category.

* Summary and sum correlation coefficients significantly different from each other at P < 0.05.

amounts or biomarkers vs. corrected/uncorrected estimates of FV) are only a result of the ability of the summary or sum questions to discriminate between persons. However, it should be noted that the biomarkers we used only allow testing whether a questionnaire is suitable for ranking persons according to their FV consumption because these biomarkers are not a measure of absolute intake. For the latter purpose, comparison of our results with food records or diaries may be helpful. Our finding of lower summary than sum amounts is in line with food consumption studies in which FV intake measured by short frequency-based screening modules was compared with intake as determined by 3-d diet records (45,46), repeated 24-h recalls (46,47) or a more extensive FFQ (48). All these

studies showed that the short instruments underestimated FV consumption relative to the reference methods used.

In contrast, screening modules gave slightly better (45) or similar (47) correlations with dietary records or recalls in comparison with an extensive FFQ, indicating that they were able to rank people according to their FV intake. However, because of correlated errors in the screener and reference method these correlations could be overestimated (49). We are aware of only one study in which summary and sum measures of FV consumption were compared with plasma carotenoids. Contrary to our findings, no substantial differences in correlations between the two measures of consumption were found (46). The different results may be attributed

TABLE 6

Spearman's rank correlations of plasma carotenoid and vitamin C concentrations with consumption of cabbage, carrots and citrus fruits with and without correction for under- or overreporting^{1,2}

	Vitamin C	Total carotenoids	Lutein	β-Cryptoxanthin	Lycopene	α-Carotene	β-Carotene
Carrots							
Uncorrected ³	0.19 (0.02)	0.29 (0.00)*	0.17 (0.03)	0.12 (0.12)	0.04 (0.58)	0.39 (0.00)*	0.23 (0.00)*
Corrected ³	0.24 (0.00)	0.21 (0.01)*	0.13 (0.11)	0.09 (0.26)	0.05 (0.56)	0.28 (0.00)*	0.15 (0.06)*
Cabbage							
Uncorrected	0.16 (0.04)	0.20 (0.01)	0.19 (0.01)	0.11 (0.17)	-0.10 (0.20)	0.21 (0.01)*	0.16 (0.04)
Corrected	0.13 (0.10)	0.11 (0.15)	0.15 (0.06)	0.08 (0.31)	-0.07 (0.40)	0.10 (0.23)*	0.10 (0.21)
Brussels sprouts							
Uncorrected	0.05 (0.52)	0.09 (0.24)	0.22 (0.01)	0.04 (0.61)	-0.08 (0.35)	0.10 (0.20)	0.05 (0.50)
Corrected	0.06 (0.43)	0.07 (0.40)	0.21 (0.01)	0.05 (0.55)	-0.08 (0.30)	0.08 (0.31)	0.04 (0.58)
Broccoli/cauliflower							
Uncorrected	0.18 (0.02)	0.14 (0.07)	0.20 (0.01)	0.04 (0.63)	-0.14 (0.09)	0.12 (0.13)	0.11 (0.19)
Corrected	0.18 (0.02)	0.10 (0.22)	0.19 (0.02)	0.02 (0.85)	-0.10 (0.21)	0.05 (0.56)	0.05 (0.51)
Citrus fruits							
Uncorrected	0.21 (0.01)	0.24 (0.00)	0.08 (0.35)	0.54 (0.00)	-0.02 (0.85)	-0.02 (0.82)	0.02 (0.76)
Corrected	0.24 (0.00)	0.21 (0.01)	0.06 (0.46)	0.54 (0.00)	-0.06 (0.48)	-0.03 (0.70)	-0.00 (0.96)

¹ Assessed with a food-frequency questionnaire in 161 women.

² P-values are in parentheses.

³ Uncorrected and corrected refer to uncorrected consumption and consumption after multiplication with a correction factor based on frequencies of consumption. * Uncorrected and corrected correlation coefficients significantly different from each other at P < 0.05.

either to the lower mean FV consumption in their population (50) or to the smaller number of items these authors used for calculating sum amounts, which could lead to a sum estimate closer to the summary estimate (51).

The correlations we found between FV consumption and biomarkers were moderate, but comparable to those from several other studies in which FV intake was related to plasma carotenoids and/or vitamin C. Correlations observed in these studies were approximately between 0.2 and 0.5, with the exception of lycopene, for which much lower correlations were found (24–28), as we did. A strong point of our study is that the results are not biased by the effect of smoking, nor by the use of vitamin supplements on concentrations of carotenoids and vitamin C in blood, given that all participants were nonsmokers who agreed not to use vitamin supplements. We are aware of the fact that plasma concentrations of carotenoids and vitamin C vary over time (20,27), and that a single measurement may not represent the 1-mo average. Repeated blood measurements could have increased the correlations we found between FV intake and biomarkers, although we expected this increase would be minor because of the short study period of 1 mo.

Although in literature correlation coefficients are reported with (25,26,52) or without (52–54) adjustment for blood cholesterol levels, we decided not to include this adjustment because these studies showed that correlations between plasma or serum carotenoids and FV or carotenoid intake changed only marginally after adjusting for cholesterol (25,26,52–54). It is also possible that adjustment for energy intake would yield different correlation coefficients. Because we wanted to assess FV intake elaborately, we made the concession not to determine energy intake because this would have resulted in a far more extended questionnaire, increasing the risk that participants would complete it less accurately.

Before generalizing the results of the present study, attention should be paid to the high degree of selective participation by subjects who were probably interested in healthy eating. How exactly this has affected our findings is difficult to say. On the one hand, our study population could have completed the questionnaire more carefully than would have been the case for the general population. On the other hand, it is also possible that our study population, being more aware of the healthy properties of FV than the average Dutch population, would be biased toward overreporting their FV intake. In addition, studying a more representative population could also have increased the observed correlation coefficients as a result of the greater range in FV intake we would have measured.

In conclusion, for the purpose of ranking people according to their intake of main groups of FV, it would appear that a small number of brief, general questions on FV intake suffice. This has important implications in the design of FFQ for use in studies related to health education, where total FV consumption is of interest. For epidemiological studies on the relation between single food items and disease or prognosis, correcting individual items with a correction factor based on frequencies does not seem to be advisable, given that this is likely to decrease rather than enhance the ability of a FFQ to rank individuals according to their intake of these items. In view of our specific study population, these findings need confirmation in future studies using more representative populations.

ACKNOWLEDGMENTS

The authors thank Willemieke Kroeze and Linda Joosten for practical assistance in the study, including laboratory analyses of

vitamin C; and Joan Senden for assistance in the HPLC analyses of carotenoids.

LITERATURE CITED

- Steinmetz, K. A. & Potter, J. D. (1996) Vegetables, fruit, and cancer prevention: a review. *J. Am. Diet. Assoc.* 96: 1027–1039.
- Gandini, S., Merzenich, H., Robertson, C. & Boyle, P. (2000) Meta-analysis of studies on breast cancer risk and diet: the role of fruit and vegetable consumption and the intake of associated micronutrients. *Eur. J. Cancer* 36: 636–646.
- Voorrips, L. E., Goldbohm, R. A., Verhoeven, D. T., van Poppel, G. A., Sturmans, F., Hermus, R. J. & van den Brandt, P. A. (2000) Vegetable and fruit consumption and lung cancer risk in the Netherlands Cohort Study on diet and cancer. *Cancer Causes Control* 11: 101–115.
- Slattery, M. L., Benson, J., Curtin, K., Ma, K. N., Schaeffer, D. & Potter, J. D. (2000) Carotenoids and colon cancer. *Am. J. Clin. Nutr.* 71: 575–582.
- Ness, A. R. & Powles, J. W. (1997) Fruit and vegetables, and cardiovascular disease: a review. *Int. J. Epidemiol.* 26: 1–13.
- Law, M. R. & Morris, J. K. (1998) By how much does fruit and vegetable consumption reduce the risk of ischaemic heart disease? *Eur. J. Clin. Nutr.* 52: 549–556.
- Liu, S., Manson, J. E., Lee, I. M., Cole, S. R., Hennekens, C. H., Willett, W. C. & Buring, J. E. (2000) Fruit and vegetable intake and risk of cardiovascular disease: the Women's Health Study. *Am. J. Clin. Nutr.* 72: 922–928.
- Haraldsdottir, J. (1993) Minimizing error in the field: quality control in dietary surveys. *Eur. J. Clin. Nutr.* 47(suppl.): S19–S24.
- Amanatidis, S., Mackerras, D. & Simpson, J. M. (2001) Comparison of two frequency questionnaires for quantifying fruit and vegetable intake. *Public Health Nutr.* 4: 233–239.
- Tjonneland, A., Overvad, K., Haraldsdottir, J., Bang, S., Ewertz, M. & Jensen, O. M. (1991) Validation of a semiquantitative food frequency questionnaire developed in Denmark. *Int. J. Epidemiol.* 20: 906–912.
- Frongillo, E. A., Jr. (1999) Validation of measures of food insecurity and hunger. *J. Nutr.* 129: 506S–509S.
- Willett, W. & Lenart, E. (1998) Reproducibility and validity of food-frequency questionnaires. In: *Nutritional Epidemiology*, 2nd ed. (Willett, W., ed.), pp. 101–147. Oxford University Press, New York, NY.
- Hunter, D. (1998) Biochemical indicators of dietary intake. In: *Nutritional Epidemiology*, 2nd ed. (Willett, W., ed.), pp. 174–243. Oxford University Press, New York, NY.
- van 't Veer, P., Kardinaal, A. F., Bausch Goldbohm, R. A. & Kok, F. J. (1993) Biomarkers for validation. *Eur. J. Clin. Nutr.* 47(suppl. 2): S58–S63.
- Le Marchand, L., Hankin, J. H., Carter, F. S., Essling, C., Luffey, D., Franke, A. A., Wilkens, L. R., Cooney, R. V. & Kolonel, L. N. (1994) A pilot study on the use of plasma carotenoids and ascorbic acid as markers of compliance to a high fruit and vegetable dietary intervention. *Cancer Epidemiol. Biomarkers Prev.* 3: 245–251.
- McEligot, A. J., Rock, C. L., Flatt, S. W., Newman, V., Faerber, S. & Pierce, J. P. (1999) Plasma carotenoids are biomarkers of long-term high vegetable intake in women with breast cancer. *J. Nutr.* 129: 2258–2263.
- Maskarinec, G., Chan, C. L. Y., Meng, L. X., Franke, A. A. & Cooney, R. V. (1999) Exploring the feasibility and effects of a high-fruit and -vegetable diet in healthy women. *Cancer Epidemiol. Biomarkers Prev.* 8: 919–924.
- Broekmans, W. M., Klopping Ketelaars, I. A., Schuurman, C. R., Verhagen, H., van den Berg, H., Kok, F. J. & van Poppel, G. (2000) Fruits and vegetables increase plasma carotenoids and vitamins and decrease homocysteine in humans. *J. Nutr.* 130: 1578–1583.
- Yong, L. C., Forman, M. R., Beecher, G. R., Graubard, B. I., Campbell, W. S., Reichman, M. E., Taylor, P. R., Lanza, E., Holden, J. M. & Judd, J. T. (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.
- Cooney, R. V., Franke, A. A., Hankin, J. H., Custer, L. J., Wilkens, L. R., Harwood, P. J. & Le Marchand, L. (1995) Seasonal variations in plasma micronutrients and antioxidants. *Cancer Epidemiol. Biomarkers Prev.* 4: 207–215.
- Scott, K. J., Thurnham, D. I., Hart, D. J., Bingham, S. A. & Day, K. (1996) The correlation between the intake of lutein, lycopene and beta-carotene from vegetables and fruits, and blood plasma concentrations in a group of women aged 50–65 years in the UK. *Br. J. Nutr.* 75: 409–418.
- Ritenbaugh, C., Peng, Y. M., Aickin, M., Graver, E., Branch, M. & Alberts, D. S. (1996) New carotenoid values for foods improve relationship of food frequency questionnaire intake estimates to plasma values. *Cancer Epidemiol. Biomarkers Prev.* 5: 907–912.
- Michaud, D. S., Giovannucci, E. L., Ascherio, A., Rimm, E. B., Forman, M. R., Sampson, L. & Willett, W. C. (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 Epidemiol. Biomarkers Prev.* 7: 283–290.
- Tucker, K. L., Chen, H., Vogel, S., Wilson, P. W., Schaefer, E. J. & Lammi Keefe, C. J. (1999) Carotenoid intakes, assessed by dietary questionnaire, are associated with plasma carotenoid concentrations in an elderly population. *J. Nutr.* 129: 438–445.
- Campbell, D. R., Gross, M. D., Martini, M. C., Grandits, G. A., Slavina, J. L.

- & Potter, J. D. (1994) Plasma carotenoids as biomarkers of vegetable and fruit intake. *Cancer Epidemiol. Biomarkers Prev.* 3: 493-500.
26. Drewnowski, A., Rock, C. L., Henderson, S. A., Shore, A. B., Fischler, C., Galan, P., Preziosi, P. & Hercberg, S. (1997) Serum beta-carotene and vitamin C as biomarkers of vegetable and fruit intakes in a community-based sample of French adults. *Am. J. Clin. Nutr.* 65: 1796-1802.
27. van Kappel, A. L., Steghens, J. P., Zeleniuch Jacquotte, A., Chajes, V., Toniolo, P. & Riboli, E. (2001) Serum carotenoids as biomarkers of fruit and vegetable consumption in the New York Women's Health Study. *Public Health Nutr.* 4: 829-835.
28. Block, G., Norkus, E., Hudes, M., Mandel, S. & Helzlsouer, K. (2001) Which plasma antioxidants are most related to fruit and vegetable consumption? *Am. J. Epidemiol.* 154: 1113-1118.
29. McKeown, N. M., Day, N. E., Welch, A. A., Runswick, S. A., Luben, R. N., Mulligan, A. A., McTaggart, A. & Bingham, S. A. (2001) Use of biological markers to validate self-reported dietary intake in a random sample of the European Prospective Investigation into Cancer United Kingdom Norfolk cohort. *Am. J. Clin. Nutr.* 74: 188-196.
30. van Dooren-Flipsen, M.M.H. & van Klaveren, J. D. (1999) Evaluation of Food Consumption Data Gathered with the ANI- and VEG-Food Frequency Questionnaire (in Dutch), p. 18. DLO-State Institute for Quality Control of Agricultural Products (RIKILT-DLO), Wageningen, The Netherlands.
31. Hulshof, K.F.A.M. (1998) The Consumption of Food Groups by Dutch Population Groups: Food Consumption Survey 1997-1998 (in Dutch). TNO Nutrition and Food Research, Zeist, The Netherlands.
32. Kardinaal, A.F.M., van 't Veer, P., Brants, H.A.M., Kruijssen, H.A.C.M., Hermus, R.J.J. & Kok, F. J. (1995) Comparison of biomarkers and dietary intake of antioxidants in relation to myocardial infarction. *Nutr. Metab. Cardiovasc. Dis.* 5: 225-230.
33. Brants, H.A.M., Brug, H., van Erp-Baart, M. & Kistemaker, C. (1991) Development of a Written Questionnaire on the Intake of Vitamin B6 by Dutch Adults (Food Survey 1990/92) (in Dutch). TNO Nutrition and Food Research, Zeist, The Netherlands.
34. Goldbohm, R. A., van den Brandt, P. A., Brants, H. A., van't Veer, P., Al, M., Sturmans, F. & Hermus, R. J. (1994) Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur. J. Clin. Nutr.* 48: 253-265.
35. Goldbohm, R. A., van 't Veer, P., van den Brandt, P. A., van 't Hof, M. A., Brants, H. A., Sturmans, F. & Hermus, R. J. (1995) Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. *Eur. J. Clin. Nutr.* 49: 420-429.
36. Ocke, M. C., Bueno de Mesquita, H. B., Goddijn, H. E., Jansen, A., Pols, M. A., van Staveren, W. A. & Kromhout, D. (1997) The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int. J. Epidemiol.* 26(suppl. 1): S37-S48.
37. Ocke, M. C., Bueno de Mesquita, H. B., Pols, M. A., Smit, H. A., van Staveren, W. A. & Kromhout, D. (1997) The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int. J. Epidemiol.* 26(suppl. 1): S49-S58.
38. Speek, A. J., Schrijver, J. & Schreurs, W. H. (1984) Fluorometric determination of total vitamin C in whole blood by high-performance liquid chromatography with pre-column derivatization. *J. Chromatogr.* 305: 53-60.
39. Hess, D., Keller, H. E., Oberlin, B., Bonfanti, R. & Schuep, W. (1991) Simultaneous determination of retinol, tocopherols, carotenes and lycopene in plasma by means of high-performance liquid chromatography on reversed phase. *Int. J. Vitam. Nutr. Res.* 61: 232-238.
40. Oostenbrug, G. S., Mensink, R. P., Hardeman, M. R., De Vries, T., Brouns, F. & Hornstra, G. (1997) Exercise performance, red blood cell deformability, and lipid peroxidation: effects of fish oil and vitamin E. *J. Appl. Physiol.* 83: 746-752.
41. Hulshof, K.F.A.M., Van der Heijden, L.J.M. & Donders-Engelen, M. (1992) Measures, Weights and Code Numbers 1992 (in Dutch). TNO Nutrition and Food Research, Zeist, The Netherlands.
42. Donders-Engelen, M. R., Van der Heijden, L. & Hulshof, K.F.A.M. (1997) Measures, Weights and Code numbers 1997 (in Dutch). Department of Human Nutrition, Wageningen Agricultural University, Wageningen & TNO Nutrition and Food Research, Zeist, The Netherlands.
43. Olkin, I. & Siotani, M. (1964) Asymptotic Distribution Functions of a Correlation Matrix. Stanford University Laboratory for Quantitative Research in Education, Stanford, CA.
44. Olkin, I. (1967) Correlation revisited. In: *Improving Experimental Design and Statistical Analysis* (Stanley, J. C., ed.). Rand McNally, Chicago, IL.
45. Smith Warner, S. A., Elmer, P. J., Fosdick, L., Tharp, T. M. & Randall, B. (1997) Reliability and comparability of three dietary assessment methods for estimating fruit and vegetable intakes. *Epidemiology* 8: 196-201.
46. Kristal, A. R., Vizenor, N. C., Patterson, R. E., Neuhauser, M. L., Shattuck, A. L. & McLerran, D. (2000) Precision and bias of food frequency-based measures of fruit and vegetable intakes. *Cancer Epidemiol. Biomarkers Prev.* 9: 939-944.
47. Field, A. E., Colditz, G. A., Fox, M. K., Byers, T., Serdula, M., Bosch, R. J. & Peterson, K. E. (1998) Comparison of 4 questionnaires for assessment of fruit and vegetable intake. *Am. J. Public Health* 88: 1216-1218.
48. Thompson, F. E., Kipnis, V., Subar, A. F., Krebs Smith, S. M., Kahle, L. L., Midthune, D., Potischman, N. & Schatzkin, A. (2000) Evaluation of 2 brief instruments and a food-frequency questionnaire to estimate daily number of servings of fruit and vegetables. *Am. J. Clin. Nutr.* 71: 1503-1510.
49. Kipnis, V., Midthune, D., Freedman, L. S., Bingham, S., Schatzkin, A., Subar, A. & Carroll, R. J. (2001) Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am. J. Epidemiol.* 153: 394-403.
50. Kristal, A. R., Patterson, R. E., Neuhauser, M. L., Thornquist, M., Neumark Sztainer, D., Rock, C. L., Berlin, M. C., Cheskin, L. & Schreiner, P. J. (1998) Olestra Postmarketing Surveillance Study: design and baseline results from the sentinel site. *J. Am. Diet. Assoc.* 98: 1290-1296.
51. Krebs Smith, S. M., Heimendinger, J., Subar, A. F., Patterson, B. H. & Pivonka, E. (1995) Using food frequency questionnaires to estimate fruit and vegetable intake: association between the number of questions and total intakes. *J. Nutr. Educ.* 27: 80-85.
52. Forman, M. R., Lanza, E., Yong, L. C., Holden, J. M., Graubard, B. I., Beecher, G. R., Meltz, M., Brown, E. D. & Smith, J. C. (1993) 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.
53. Martini, M. C., Campbell, D. R., Gross, M. D., Grandits, G. A., Potter, J. D. & Slavin, J. L. (1995) Plasma carotenoids as biomarkers of vegetable intake: the University of Minnesota Cancer Prevention Research Unit Feeding Studies. *Cancer Epidemiol. Biomarkers Prev.* 4: 491-496.
54. Kardinaal, A. F., van 't Veer, P., Brants, H. A., van den Berg, H., van Schoonhoven, J. & Hermus, R. J. (1995) Relations between antioxidant vitamins in adipose tissue, plasma, and diet. *Am. J. Epidemiol.* 141: 440-450.