

# Evaluation of 2 brief instruments and a food-frequency questionnaire to estimate daily number of servings of fruit and vegetables<sup>1-3</sup>

Frances E Thompson, Victor Kipnis, Amy F Subar, Susan M Krebs-Smith, Lisa L Kahle, Douglas Midthune, Nancy Potischman, and Arthur Schatzkin

## ABSTRACT

**Background:** Measurement of fruit and vegetable intake is important in the surveillance of populations and in epidemiologic studies that examine the relations between diet and disease. Some situations require the use of brief dietary assessment tools. **Objective:** Our objective was to evaluate the performance of 2 brief dietary assessment instruments, a 7-item standard screener and a new 16-item screener, and a complete food-frequency questionnaire (FFQ) in measuring total fruit and vegetable consumption.

**Design:** About 800 men and women from the National Institutes of Health–AARP Diet and Health Study completed an FFQ, 1 of the 2 screeners, and two 24-h dietary recalls. Fruit and vegetable intakes as measured by each screener and the FFQ were compared with estimated true usual intake by using a measurement-error model.

**Results:** Median daily servings of fruit and vegetables were underestimated by both screeners. The estimated agreement between true intake and the screener was higher for the new screener than for the standard screener and was higher for women than for men. The estimated agreement between true intake and the FFQ was higher than that for both screeners. Attenuation coefficients for the FFQ and screeners were comparable.

**Conclusions:** For estimating median intakes of fruit and vegetables and the prevalence of recommended intakes being met, the use of screeners without appropriate adjustment is suboptimal. For estimating relative risks in the relations between fruit and vegetable intake and disease, screeners and this FFQ are similar in performance. *Am J Clin Nutr* 2000;71:1503–10.

**KEY WORDS** Diet, epidemiologic methods, surveillance, validation study, data collection, questionnaires, screener, food-frequency questionnaire, epidemiology, nutrition, nutritional epidemiology, NIH-AARP Diet and Health Study

## INTRODUCTION

Fruit and vegetables provide essential vitamins, minerals, and other components such as fiber (1), glucosinolates, and indoles in our diets. In addition, high intakes of fruit and vegetables may displace other, less desirable components of the diet (2). Dietary intake of fruit and vegetables has been related to a variety of health outcomes, including cancer and cardiovascular disease

(3–5). One of the national health objectives for the year 2000 is to “increase complex carbohydrate- and fiber-containing foods in the diets of adults to 5 or more daily servings for vegetables (including legumes) and fruits”(6). Thus, measuring intakes of fruit and vegetables is important for surveillance and for studying relations between diet and disease.

These needs for data have stimulated great interest in how best to measure fruit and vegetable intakes (7). In the United States, 24-h recall data are used to monitor progress toward the Healthy People 2000 national fruit and vegetable intake objective (7). However, sometimes researchers have more limited purposes or resources, and they need brief dietary instruments that are easy to administer and process and that adequately satisfy a particular purpose.

After the 1991 5 a Day Baseline Survey (8), National Cancer Institute researchers and National 5 a Day Program grantees developed a brief dietary assessment tool, or screener, to measure total fruit and vegetable consumption. The results are used to track changes in fruit and vegetable intake and to measure program and intervention effectiveness (9). This standard screener includes 7 questions that ask how often fruit and vegetables were consumed in the past month (**Table 1**).

Our objectives were to test the performance of the standard screener, relative to criterion data from two 24-h dietary recalls, and to compare its performance with that of a new screener, composed of 16 questions, that we developed (**Table 2**). On the new screener, respondents are asked to divide the day into 3 periods and to report their usual consumption of fruit and vegetables separately for each period. A secondary objective was to compare the performance of both screeners with that of a food-frequency questionnaire (FFQ) that provides a more extensive assessment of the usual diet over the past year.

<sup>1</sup>From the National Cancer Institute, Bethesda, MD, and Information Management Services Inc, Silver Spring, MD.

<sup>2</sup>An earlier version of this article was presented at the Third International Conference on Dietary Assessment Methods in Arnhem, Netherlands, May 1998.

<sup>3</sup>Address reprint requests to FE Thompson, Risk Factor Surveillance and Methodology Branch, Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, EPN 4016, 6130 Executive Boulevard, MSC 7344, Bethesda, MD 20892-7344. E-mail: ft5h@nih.gov.

Received July 12, 1999.

Accepted for publication November 10, 1999.

**TABLE 1**

Questions asked on standard fruit and vegetable screener

---



---

In the past month, about how often did you drink (or eat):

- 100% orange juice or grapefruit juice?
- ... other 100% fruit juices, NOT COUNTING fruit drinks?
- ... green salad (with or without other vegetables)?
- ... French fries or fried potatoes?
- ... baked, boiled, or mashed potatoes?

In the past month, about how many servings of vegetables did you eat, NOT COUNTING potatoes and salad?

In the past month, about how many servings of fruit did you eat, NOT COUNTING juices?

Response categories are: never, 1–3 times/mo, 1–2 times/wk, 3–4 times/wk, 5–6 times/wk, 1 time/d, 2 times/d, 3 times/d, 4 times/d, and  $\geq 5$  times/d.

---

## SUBJECTS AND METHODS

### Subjects

The screeners were offered to a random sample of members of the Calibration Study of the National Institutes of Health (NIH)-AARP (formerly the American Association of Retired Persons) Diet and Health Study. In the main study, an extensive FFQ was mailed in 1995–1996 to 3.5 million AARP members aged 50–69 y in 6 states (California, Florida, Louisiana, New Jersey, North Carolina, and Pennsylvania) and 2 cities (Atlanta and Detroit). This FFQ asked about the usual frequency of consumption and portion size for 120 food groups, as well as other dietary patterns. Those who responded by January 1996 and who were not excluded because of questionable dietary data or a medical condition were eligible for enrollment in the Calibration Study. Questionable dietary data were defined as  $\geq 2$  adjacent pages of the FFQ skipped;  $> 10$  items with multiple frequency responses marked;  $< 10$  items with any reported consumption; or energy intake too high ( $> 4200$  kcal or  $17573$  kJ for men or  $> 3500$  kcal or  $14644$  kJ for women) or too low ( $< 800$  kcal or  $3347$  kJ for men or  $< 600$  kcal or  $2510$  kJ for women). Medical conditions such as end-stage renal disease, self-reported poor health, or history of cancer were also used as exclusion criteria.

The baseline FFQ data were used to assess individuals' intakes of 4 dietary variables of interest: percentage of energy from fat, dietary fiber intake, servings of fruit and vegetables, and servings of red meat. For each of these variables, 5 strata were defined for each sex and individuals were classified into these strata. To increase the sample's variability for these 4 dietary measures, extreme intakes (strata 1 and 5) on any of the 4 variables were sampled at the highest rate, proportionately fewer individuals were sampled from strata 2 and 4, and even fewer were sampled from stratum 3. Two additional strata were defined as high alcohol use ( $\geq 90$  g alcohol/d on baseline FFQ) by sex. Within each stratum, individuals were randomly chosen to participate in the Calibration Study. Of the 2053 members of the Calibration Study, 1986 provided two 24-h dietary recalls (described below) and the rest provided one. All 2053 members of the Calibration Study were mailed a second FFQ in October 1996. A total of 1415 FFQs were returned, but 52 FFQs were later excluded because of questionable data quality on the basis of the criteria described above.

By June 1997, 80 of the original 2053 persons had withdrawn from further participation in the Calibration Study. The remaining 1973 individuals were randomly divided into 3 groups and were mailed the standard fruit and vegetable screener, the

new fruit and vegetable screener, or a new screener designed to assess percentage of energy from fat (not reported here).

For the standard screener, 2 of those mailed were undeliverable and 455 of the 657 mailed were returned (response rate of 69%). For the new screener, which was mailed to a different subgroup, 2 of those mailed were undeliverable and 454 of the 656 mailed were returned (response rate of 69%).

For these analyses, we excluded participants in the high-alcohol-use strata of the calibration sample ( $n = 11$  in the standard-screener group and  $n = 9$  in the new-screener group). We also excluded from further analyses participants with extreme fruit and vegetable values on either 24-h recall ( $\geq 30$  servings of fruit and vegetables on either day;  $n = 2$  in each screener group) because we did not want our results to be influenced unduly by extreme values. Finally, we excluded individuals with poor-quality data on the screeners, as described below (6 and 5 subjects in the standard- and new-screener groups, respectively).

Of the 436 participants in the standard-screener group, 47% were men, 34% were aged 50–59 y, and 66% were aged  $\geq 60$  y. Of the 438 participants in the new-screener group, 49% were men, 32% were aged 50–59 y, and 68% were aged  $\geq 60$  y.

Data collection for the NIH–AARP Diet and Health Study was approved by the National Cancer Institute Institutional Review Board and the Westat Inc Special Studies Institutional Review Board. Informed consent was obtained from all participants.

### Data collection and processing and variable creation

#### Collection and coding of 24-h dietary recalls

The 24-h dietary recalls were administered by trained interviewers over the telephone from February through August 1996. The second recall was obtained an average of 24 d (median of 20 d) after the first recall. A set of measuring guides was sent to participants before the interview; 94% of the participants reported

**TABLE 2**

Questions asked on new fruit and vegetable screener

---



---

In the last month, about how often did you drink (or eat):

- 100% orange juice or 100% grapefruit juice?
- ... other 100% fruit juices, such as apple or grape juice? (Do not count fruit drinks such as Hi-C.)
- ... French fries or fried potatoes?
- ... baked, boiled, or mashed potatoes?

Response categories are: never, 1–3 times last mo, 1–2 times/wk, 3–4 times/wk; 5–6 times/wk, 1 time/d, 2 times/d, and  $\geq 3$  times/d.

For morning, lunchtime and afternoon, and evening and nighttime separately:

On how many days did you eat fruit for your (timeframe) meals or snacks? (Do not count juices.)

Response categories are: none, 1–3 d last mo, 1–2 d/wk, 3–4 d/wk, 5–6 d/wk, and every day.

When you ate fruit in the (timeframe), how many total portions of fruit did you usually eat? (Count each piece or one-half cup you ate as one portion, whether it was one fruit or different fruits.)

Response categories are: 1 portion or less, 2 portions,  $\geq 3$  portions.

On how many days did you eat vegetables for your (timeframe) meal and snacks? (Do not count potatoes.)

Response categories are: none, 1–3 d last mo, 1–2 d/wk, 3–4 d/wk, 5–6 d/wk, and every day.

When you ate vegetables in the (timeframe), how many total portions of vegetables did you usually eat? (Count each one-half cup you ate as one portion, whether it was one vegetable or different vegetables.)

Response categories are:  $\leq 1$  portion, 2 portions,  $\geq 3$  portions.

---

that they used the guides for the interview. Participants were not told in advance when they would be interviewed. Because diets differ by day of the week (10, 11), participant interviews were scheduled to reflect the distribution of weekdays and weekend days. Of the 1986 participants who reported 2 d of intake, 55% reported about 2 weekdays, 37% reported about 1 weekday and 1 weekend day, and 8% reported about 2 weekend days. Participants were asked to report all foods and beverages consumed on the day before the interview, for the entire 24 h from midnight to midnight. Interviewers used a Food Probe List containing standardized probes specific to the various foods in each of >100 food categories. The probes addressed such information as preparation method, type of fat used, brand, additions to the food, dilution, and quantity. Data were coded by using the FOOD INTAKE ANALYSIS SYSTEM, version 2.3 (12). This system uses food codes comparable with those used by the US Department of Agriculture (USDA) and includes extensive error-checking procedures for identifying potentially erroneous data. In addition, data checks were performed on records with extremely high values for fat, total energy, and total fruit and vegetable intakes and corrections were made to the data when warranted.

#### *Conversion of intake data to servings of fruit and vegetables*

National dietary guidance (13–15) was used to define which foods count as fruit and vegetables and to quantify a serving for all the instruments. For fruit, a serving is defined as a whole fruit (eg, medium apple), three-fourths cup (178 mL) fruit juice, or one-half cup (120 mL) cut-up fruit. For vegetables, a serving is defined as 1 cup (240 mL) raw leafy vegetables (eg, lettuce), one-half cup other vegetables, or three-fourths cup vegetable juice. Data released from the USDA Continuing Survey of Food Intakes by Individuals (CSFII) for 1994–1996 (16) made it possible to translate food intake data coded with USDA codes into servings from the food guide pyramid groups. The data files provide the number of servings from each food guide pyramid group per 100 g of the food code, for each of >5000 food codes. With this system, the various components of food mixtures can be assigned to their appropriate food guide pyramid group.

Most of the food codes used to code the NIH–AARP 24-h dietary recalls corresponded to codes in the 1994–1996 CSFII; this was the case for 3271 of 3759 food codes reported. When there was no matching 1994–1996 CSFII code, a similar database developed for the USDA 1989–1991 CSFII (17) was used to assess servings of fruit and vegetables (140 food codes). A small number of food codes (348) did not match codes in either USDA database. For these, a food code from the 1994–1996 CSFII database that best matched the food with regard to relative composition of fruit, vegetables, meat, and starch was chosen. After all the food codes reported on the NIH–AARP 24-h dietary recalls had been matched to CSFII food codes, the total number of servings of fruit and vegetables from each food item was computed and these were summed across all foods to give an estimate of the total fruit and vegetable servings consumed by each individual. This total reflected all sources of fruit and vegetables, including fried vegetables such as French fries, miscellaneous sources such as potato chips, and small amounts such as those found in yogurt or desserts.

For the screener instruments, various assumptions were made to estimate the total servings of fruit and vegetables. On the standard screener, the amount of food consumed each time is not asked;

instead, we assumed one serving as defined by the food guide pyramid. For the new screener, in the sections asking about intake by time of day (Table 2), individuals were asked to report both the frequency and the number of portions per occurrence. They were instructed that a portion is one-half cup fruit, one whole fruit, or one-half cup vegetables. Servings were assigned to the 3 portion categories as follows:  $\leq 1$  portion = 1 serving, 2 portions = 2 servings,  $\geq 3$  portions = 3 servings. Responses to these questions were also used to estimate the total number of servings for orange and grapefruit juice, other 100% fruit juices, French fries or fried potatoes, and baked, boiled, or mashed potatoes, for which the number of portions was not asked.

For the FFQ, 1994–1996 CSFII survey data were used to construct a database that included, for each food asked about on the FFQ, a mean number of servings of fruit and vegetables for each of 6 different strata, defined by sex and portion size (small, medium, or large) (18). For each individual, the total daily number of servings of fruit and vegetables was computed by multiplying that individual's reported frequency by the number of fruit and vegetable servings assigned for his or her stratum (defined by sex and his or her reported portion size) for each food on the FFQ, summing across all foods, and standardizing to reflect daily average intake.

Missing responses on both of the screeners were examined and the following decisions were made. For the standard screener, the case was excluded from analysis if there was no frequency response on either the "other vegetables" question or the "fruit" question. When the frequency response for any other question was missing, it was assumed that the respondent never eats that food, and a 0 was assigned. For the new screener, a 0 was assigned when both the frequency and the corresponding portion question were not answered. When the frequency question was answered and the portion question was not, the portion most frequently reported by that respondent for that same type of food (fruit or vegetable) was used as the default value. For both screeners, if any question had multiple responses or entirely missing data, the case was excluded from analysis.

#### **Analytic methods**

The variable of interest for all analyses was true usual intake of fruit and vegetables expressed in daily servings. We used the measurement-error model developed by Freedman et al (19), which assumes that measurement error in a test instrument may include systematic bias as well as within-person random error, but that error in the criterion instrument (such as a 24-h recall) includes only within-person variation. The measurement-error model is written as follows:

$$Q_i = \beta_0 + \beta_1 T_i + e_i \quad (1)$$

$$F_{ij} = T_i + u_{ij} \quad (2)$$

where  $Q_i$  is the questionnaire value (ie, from standard screener, new screener, or FFQ) for the  $i$ th individual,

$T_i$  is the true usual intake of fruit and vegetables for the  $i$ th individual,

$\beta_0$  is the intercept and  $\beta_1$  is the slope of the linear regression of  $Q_i$  on  $T_i$ ,

$e_i$  is random within-person error,

$F_{ij}$  is the  $j$ th repeat criterion measurement (24-h recalls) for person  $i$ ,

$u_{ij}$  is within-person random error for person  $i$ , repeat measurement  $j$ , assumed to be independent of  $T_i$  and  $e_i$ . Because the repeat

**TABLE 3**

Median servings of fruit and vegetables on the basis of estimated true intake, the standard screener, the new screener, and a food-frequency questionnaire (FFQ) among men and women

	Estimated true intake <sup>1</sup>	Screener	FFQ
	servings/d		
<b>Men</b>			
Standard-screener sample (n = 205)	6.6	3.7	6.6
New-screener sample (n = 216)	6.5	4.4	6.3
<b>Women</b>			
Standard-screener sample (n = 231)	5.5	4.2	6.2
New-screener sample (n = 222)	5.7	5.0	6.0

<sup>1</sup>From 24-h recall data, adjusted for within-person random measurement error.

24-h recalls were reasonably separated in time in this case, it is further assumed that random error  $u_{ij}$  and  $u_{ij'}$  corresponding to 2 repeat measurements for the same person are independent.

There are many parameters of interest when evaluating how a test instrument performs relative to true dietary intake. For each instrument comparison, we estimated the mean and median, slope ( $\beta_1$ ), intercept ( $\beta_0$ ), correlation coefficient between test instrument and true intake ( $R$ ), and attenuation coefficient ( $\lambda$ ). The parameters  $\beta_1$ ,  $\beta_0$ , and  $R$  reflect validity, which is whether the test instrument measures what it is supposed to measure. If the test instrument is unbiased, ie, if  $E(Q_i|T_i) = T_i$ , the regression of measured on true intake would show a straight line at a 45° angle ( $\beta_1 = 1.0$ ), intercepting at 0 ( $\beta_0 = 0$ ). The squared correlation coefficient ( $R^2$ ) for the correlation between reported and true intake indicates the proportion of variation in intake measured by the screener that is explained by true intake. The attenuation coefficient ( $\lambda = R^2/\beta_1$ ) is the amount by which the relative risk between the exposure (ie, fruit and vegetable intake) and some disease would be distorted because of measurement error in the test instrument. Although, in general,  $\lambda$  can take on any value, it typically ranges from 0 to 1 for dietary variables. A  $\lambda$  close to 1 indicates minimum attenuation, whereas a  $\lambda$  close to 0 indicates maximum attenuation.

The variable “servings of fruit and vegetables” had a nonnormal distribution for all instruments, whereas the square root of that variable was approximately normally distributed. We therefore used a square-root transformation of the data, redefining  $Q_i$ ,  $T_i$ , and  $F_{ij}$  in the measurement-error model to refer to their square-root-transformed values. We then calculated maximum-likelihood estimates of the parameters  $\beta_0$ ,  $\beta_1$ ,  $\sigma_e^2$ ,  $\sigma_u^2$ ,  $u_T$ , and  $\sigma_T^2$  under the assumption of normality. From these parameter estimates, we calculated estimates of the percentiles of the distributions of  $Q$  and  $T$  and squared them to get estimates of the percentiles of the distribution of the untransformed variables.

## RESULTS

### Comparison of screeners with the estimated truth

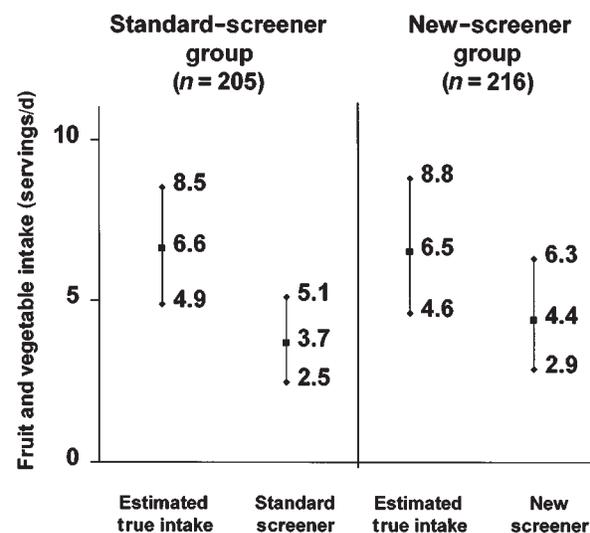
The estimated median values for total fruit and vegetable intake, from each instrument for men and women separately, are shown in **Table 3**. In **Figures 1** and **2**, the estimated median, 25th percentile, and 75th percentile values are shown. Among men (Figure 1), in the standard-screener group, the estimated true

median number of servings per day was significantly higher than the median servings estimated by the screener (6.6 and 3.7, respectively;  $P < 0.0001$ ). Among men in the new-screener group, the estimated true median number of servings per day was also significantly higher than the median servings estimated by the screener (6.5 and 4.4, respectively;  $P < 0.0001$ ). The new screener was significantly ( $P = 0.003$ ) better than was the standard screener in estimating true fruit and vegetable intakes.

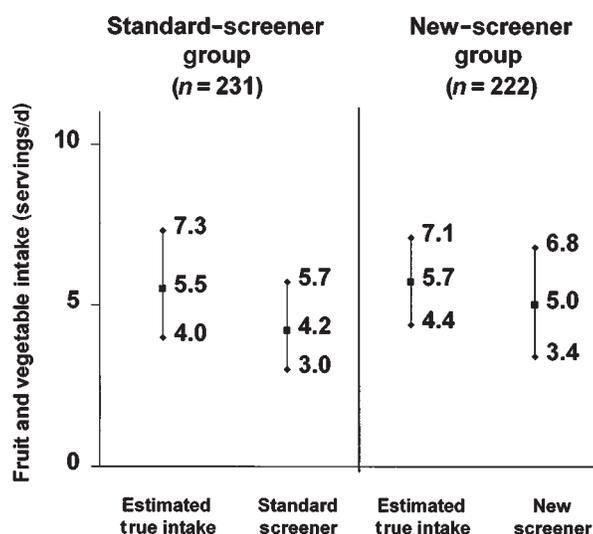
Among women (Figure 2), the findings were similar, with both screeners significantly underestimating the median servings per day relative to the estimated true intake. However, the discrepancies between the estimates of true intake and the screener results were not nearly as great among women as they were among men. Among women in the standard-screener group, the estimated true median number of servings per day was 5.5, compared with 4.2 from the screener ( $P < 0.0001$ ). In the new-screener group, the estimated true median number of servings per day was 5.7, compared with 5.0 from the screener ( $P < 0.0001$ ). The new screener was significantly ( $P = 0.007$ ) better than the standard screener in estimating true fruit and vegetable intakes.

The interquartile ranges for both screeners were smaller than were those for true intakes. For example, for men in the standard-screener group, the interquartile range for true intake was estimated as 3.6, compared with 2.6 for the screener. In the new-screener group, the interquartile range for true intake was estimated as 4.2, compared with 3.4 for the screener.

The percentages of individuals who met the Healthy People 2000 objective of eating  $\geq 5$  servings of fruit and vegetables daily were estimated by using the same model (**Table 4**). In the standard-screener group, the estimated true prevalence among men was 73%, compared with 24% on the basis of the screener results. In the new-screener group, the estimated true prevalence among men was 70%, compared with 42% on the basis of the screener results. Among women, the estimated true prevalence and prevalence from the standard screener were 59% and 36%, respectively; the estimated true prevalence and prevalence from the new screener were 64% and 50%, respectively. Both screeners



**FIGURE 1.** Median, 25th percentile, and 75th percentile values for fruit and vegetable intakes: estimated true intake versus the standard screener and new screener in men.



**FIGURE 2.** Median, 25th percentile, and 75th percentile values for fruit and vegetable intakes: estimated true intake versus the standard screener and new screener in women.

substantially underestimated the prevalence of the Healthy People 2000 fruit and vegetable objective being met, but the discrepancy was smaller for the new screener than for the standard screener.

In **Table 5**, we show the estimated slope ( $\beta_1$ ), correlation with true intake ( $R$ ), and attenuation coefficient as a result of measurement error in the reported intake for each screener and the FFQ by sex. In **Figure 3** (men) and **Figure 4** (women), the regression lines created from the estimated slopes and intercepts ( $\beta_0$ ) are shown graphically. Among men, the estimated slopes were well below 1.0 for both screeners, showing the flattened slope syndrome. This meant that individuals who ate a lot underreported their intake on the screeners and those who ate little overreported their intake on the screeners. There was substantial bias in both screeners, both from the intercept estimates of 0.67 (standard) and 0.75 (new), which were  $>0$ , and from the slope estimates of 0.49 (standard) and 0.53 (new), which were  $<1.0$ . Estimated correlation coefficients between true values and the standard and new-screener results were nearly identical (0.52 and 0.54, respectively), as were the attenuation coefficients (0.56 and 0.55, respectively).

Among women, the estimated slopes for both screeners were  $<1.0$ , again showing the flattened slope syndrome. However, the slope of 0.82 for the new screener was significantly ( $P = 0.02$ ) better than the slope of 0.49 for the standard screener. The correlations between estimated true intake and the new-screener and standard-screener results were 0.59 and 0.50, respectively (NS). The attenuation coefficients between estimated true intake and the new-screener and standard-screener results were 0.43 and 0.52, respectively (NS).

#### Comparison of FFQ and screeners with estimated true intakes

We examined the performance of a complete FFQ relative to the screeners in estimating usual intake of fruit and vegetables (the relation between the FFQ and 24-h recalls for the entire calibration sample will be examined in a future publication).

Among men in the standard-screener group, the median estimated true intake was 6.6 servings, compared with the FFQ esti-

mate of 6.6 servings and the screener estimate of 3.7 servings (**Table 3**). For men in the new-screener group, the median estimated true intake was 6.5 servings, compared with the FFQ estimate of 6.3 servings and the screener estimate of 4.4 servings. Among women, the median estimated true intakes were intermediate between the FFQ and screener estimates in both screener groups. Among both men and women, median intakes of fruit and vegetables were estimated most closely by the complete FFQ and were underestimated by both screeners, although much less so for women than for men. Median intakes from the new screener were closer to true intake than were those from the standard screener. The FFQ was better than either screener at estimating the prevalence of the Healthy People 2000 fruit and vegetable objective being met (**Table 4**).

Measures reflective of validity (ie, the slope and the correlation coefficient) and the attenuation coefficients were generally similar for the FFQ and screener for each comparison by screener type and sex (**Table 5**). Only one comparison between the FFQ and a screener showed a significant difference: among men, the slope of 0.81 for the FFQ was significantly better than was the slope of 0.49 for the standard screener ( $P = 0.0009$ ). The estimated attenuation coefficients indicate that there would be substantial attenuation in calculating the relative risk of a disease on the basis of dietary exposure data from any of these instruments.

## DISCUSSION

The methods used in this validation study are novel in 2 ways. First, the criterion data used here provide estimates of fruit and vegetable intake quantified into servings, as defined by the food guide pyramid, for each food reported. Previously published validity studies (20, 21) did not have the 1994–1996 CSFII food guide pyramid servings database available and therefore used other, less precise means of quantifying fruit and vegetable intake in relation to recommended servings. In those studies, certain uses of fruit and vegetables (eg, condiments, sauces, soups, and fried potatoes), small amounts ( $<30$  g of solid fruit or vegetable), and particular vegetables and fruit (eg, dried beans) were excluded from consideration. Because our purpose here was to consider all fruit and vegetables to reflect total exposure, we included all uses, however small, of fruit and vegetables reported

**TABLE 4**

Percentages of study samples meeting the Healthy People 2000 fruit and vegetable objective on the basis of estimated true intake, the standard screener, the new screener, and a food-frequency questionnaire (FFQ) among men and women<sup>1</sup>

	Estimated true intake <sup>2</sup>	Screener	FFQ
	%		
Men			
Standard-screener sample (n = 205)	73	24	69
New-screener sample (n = 216)	70	42	65
Women			
Standard-screener sample (n = 231)	59	36	65
New-screener sample (n = 222)	64	50	64

<sup>1</sup>Healthy People 2000 fruit and vegetable objective: eat  $\geq 5$  servings of fruit and vegetables/d.

<sup>2</sup>From 24-h recall data, adjusted for within-person random measurement error.

**TABLE 5**

Slope of the regression of reported intake on true intake, correlation between reported and true intakes, and attenuation coefficient resulting from measurement error in reported intake for the food-frequency questionnaire (FFQ) and each screener by sex<sup>1</sup>

	Slope		Correlation coefficient		Attenuation coefficient	
	FFQ	Screener	FFQ	Screener	FFQ	Screener
<b>Men</b>						
Standard-screener sample ( <i>n</i> = 205)	0.81 ± 0.129	0.49 ± 0.093	0.62 ± 0.067	0.52 ± 0.076	0.47 ± 0.058	0.56 ± 0.089
New-screener sample ( <i>n</i> = 216)	0.63 ± 0.091	0.53 ± 0.082	0.55 ± 0.061	0.54 ± 0.066	0.48 ± 0.062	0.55 ± 0.076
<b>Women</b>						
Standard-screener sample ( <i>n</i> = 231)	0.65 ± 0.102	0.49 ± 0.081	0.50 ± 0.063	0.50 ± 0.067	0.39 ± 0.055	0.52 ± 0.078
New-screener sample ( <i>n</i> = 222)	0.81 ± 0.140	0.82 ± 0.137	0.53 ± 0.068	0.59 ± 0.070	0.35 ± 0.050	0.43 ± 0.058

<sup>1</sup>± SE.

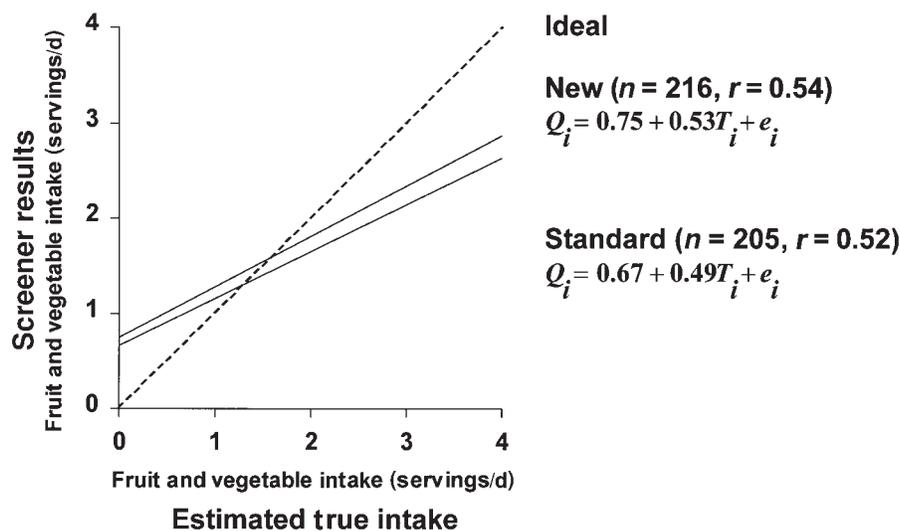
on the 24-h recalls. For example, potato chips, fruit used in condiments such as preserves, and other miscellaneous uses were counted in this study. Our estimates of true intake in our sample of adults aged 50–69 y were generally higher than such estimates for adults in other published research that used different methodologies (8, 22, 23). However, even when we used the same methodology, we found that our sample estimates were somewhat higher than those in the 1994–1996 CSFII (24) sample of persons aged 50–69 y (median for men: 5.2; median for women: 4.4), indicating a relatively higher fruit and vegetable intake among respondents in the NIH-AARP study than in the US population.

A second novel feature of the present study is the use of an advanced measurement-error model to estimate a variety of parameters useful in understanding the performance of instruments that measure fruit and vegetable intake. The usual comparisons of means and correlations between reported and true intakes were supplemented with estimates of slopes in the regressions of reported intakes on true intakes and attenuation coefficients resulting from measurement error, allowing a fuller understanding of each instrument's utility in epidemiologic research.

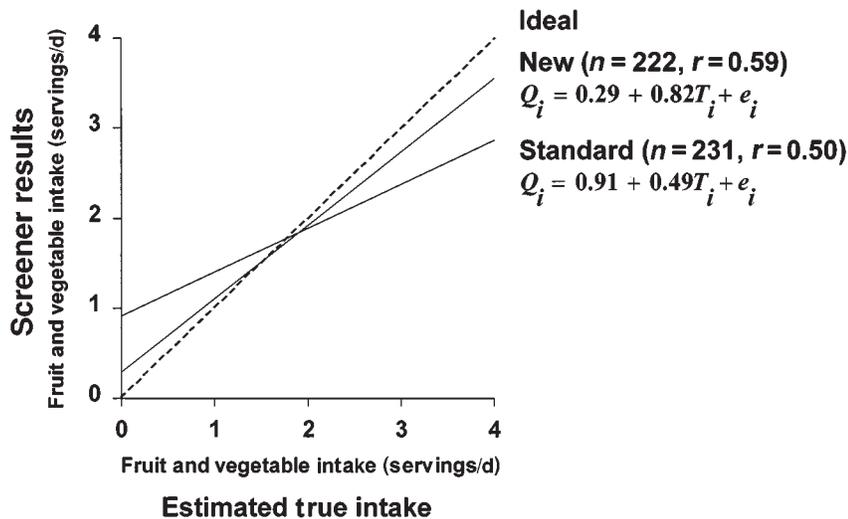
The measurement-error-model approach used here uses 24-h recalls (2 for 95% of the samples and 1 for 5% of the samples) as criterion data. There is no requirement that the two 24-h recalls accurately estimate the true usual intake of the individual. Rather, the criterion instrument may involve error of its

own, but the assumption is that this error has no systematic bias. Biomarker studies that used doubly labeled water and urinary nitrogen indicated systematic underreporting, at least at the population level, of energy and protein intakes when data were collected with dietary records and dietary recalls (25, 26). One study indicated that intakes of fruit and vegetables may also be underreported (27). If this were also true in the present study, the mean true intake would be underestimated and the differences between the screeners and the estimated true intake would be greater. The ranking of the instruments that we evaluated would remain the same, however. If the amount of underreporting on the 24-h recalls were similar across individuals, the other parameters of interest (slope, correlation coefficient, and attenuation coefficient) would be unaffected.

For surveillance and monitoring purposes, we are most interested in estimates of the mean or median true usual intake and the prevalence of consumption of recommended amounts of fruit and vegetables. In this study, both screeners underestimated median intake and the percentage of individuals consuming recommended amounts, but results from the new screener were closer to true intakes than were those from the standard screener. The median intake from a screener as a percentage of estimated true intake ranged from 56% (standard screener for men) to 88% (new screener for women). This finding is consistent with those of other, recent studies that found lower estimates of fruit and vegetable intakes



**FIGURE 3.** Ideal regression line compared with regressions of new-screener results on estimated true intake and standard-screener results on estimated true intake for men.



**FIGURE 4.** Ideal regression line compared with regressions of new-screener results on estimated true intake and standard-screener results on estimated true intake for women.

from screeners relative to multiple 24-h recalls in adolescents (28) and 3-d food records in adult, African American church members in rural North Carolina (29). These studies taken together indicate that the utility of the standard fruit and vegetable screeners used in the National 5 a Day Program and in the Behavioral Risk Factor Surveillance System should be re-evaluated. For estimating median total fruit and vegetable consumption from all sources, these standard screeners may be suboptimal. However, it is possible that estimates from these screeners could be improved by analytic adjustments developed in calibration studies relating screener data to data from food records or recalls, ideally in the population of interest. Our data do not allow for evaluation of the use of either screener for tracking intakes over time or for targeting populations for program intervention.

The new screener was designed to divide a complex reporting task into smaller, more manageable tasks. Thus, foods that are broadly defined and frequently consumed in different ways and at different times of the day (ie, the food groups *fruit* and *other vegetables*) were asked about in 3 different time periods. Recently, we conducted cognitive testing that confirmed that many individuals, when asked to report their usual intakes of fruit and other vegetables, do so by recalling their typical day, from morning to bedtime. The improved performance of the new screener compared with the standard screener in this study supports our theory and cognitive findings.

Further analyses were conducted to explore the sources of the discrepancies between the screener results and true intakes. It is unlikely that certain miscellaneous uses of fruit and vegetables (eg, potato chips and condiments) would be considered by respondents when reporting on either screener. These miscellaneous uses contributed <2% to the total fruit and vegetables reported on the 24-h recalls. The dried beans group is not clearly specified in either screener, but it too accounted for little (2%) of the total consumption of fruit and vegetables reported on the 24-h recalls. Respondents may not have fully considered the use of fruit and vegetables in mixtures, such as soups, casseroles, and pasta dishes, when answering the screener questions; these uses accounted for >13% of total fruit and vegetables reported on the

24-h recalls. In addition, intake of other vegetables was underreported on the standard screener relative to the number of servings of these vegetables reported on the 24-h recalls. These findings indicate opportunities to refine the wording of questions on screeners to be developed in the future.

The FFQ used in this study performed better than did both screeners in estimating median intakes. Its better performance is most likely a result of 2 factors: the number of questions asked (frequency and portion size of 65 individual foods that contributed to the total FFQ estimate of fruit and vegetable intake) and the database used to quantify all responses into recommended servings of fruit and vegetables.

For epidemiologic studies examining the relation between fruit and vegetable intake and some outcome measure, both screeners and the complete FFQ are similar in performance, on the basis of these analyses in this particular study. Relative risk estimates made on the basis of either screener or the FFQ would be attenuated. For example, for men, a true relative risk of 4.0 would be observed as 2.2 when data from the standard screener were used, 2.1 when data from the new screener were used, and 1.9 when data from a complete FFQ were used. For women, a true relative risk of 4.0 would be observed as 2.1 when using data from the standard screener, 1.8 when using data from the new screener, and 1.6–1.7 when using data from a complete FFQ.

We conclude that for surveillance purposes of estimating median intakes of fruit and vegetables, both screeners substantially underestimated true intakes and thus the prevalence of meeting recommended intakes. If screeners are to be used for these purposes, appropriate adjustment procedures need to be developed and tested. For examination of relative risk estimates between total fruit and vegetable intake and some disease, both screeners performed nearly as well as a much longer FFQ in this sample. However, if individual components of usual fruit and vegetable intake are to be examined, a complete FFQ that provides more specific intake information would be needed.

More cognitive research is needed to understand how individuals conceptualize their intake of fruit and vegetables and how questions can be worded so that they correspond to these

concepts. New approaches or, alternatively, more precise questions and refined language to make existing tools easier to complete, are needed. The resultant instruments should be tested in different populations, using accepted criterion data, and in comparison with another screener instrument. 

We acknowledge and thank the following staff members of Westat Inc for their invaluable contributions to this project: Paul Hurwitz, for overall study management; Susie McNutt, for management of the 24-h dietary recall collection and effort; and Anna MacIntosh, for screener questionnaire formatting and management.

## REFERENCES

- National Research Council. Recommended dietary allowances. 10th ed. Washington, DC: National Academy Press, 1989.
- Committee on Diet and Health, National Research Council. Diet and health: implications for reducing chronic disease risk. Washington, DC: National Academy Press, 1989.
- Block G, Patterson B, Subar A. Fruits, vegetables, and cancer prevention: a review of the epidemiological evidence. *Nutr Cancer* 1992;18:1-29.
- Steinmetz KA, Potter JD. Vegetables, fruit, and cancer prevention: a review. *J Am Diet Assoc* 1996;96:1027-39.
- Ness AR, Powles JW. Fruit and vegetables, and cardiovascular disease: a review. *Int J Epidemiol* 1997;26:1-13.
- Public Health Service. Healthy people 2000: national health promotion and disease prevention objectives. Washington, DC: US Department of Health and Human Services, 1990. (US DHHS publication PHS 91-50212.)
- Krebs-Smith SM, Cook DA, Subar AF, Cleveland L, Friday J. US adults' fruit and vegetable intakes, 1989 to 1991: a revised baseline for the healthy people 2000 objective. *Am J Public Health* 1995; 85:1623-9.
- Subar A, Heimendinger J, Krebs-Smith SM, Patterson BH, Kessler R, Pivonka E. Five a day for better health: a baseline study of Americans' fruit and vegetable consumption. Bethesda, MD: National Cancer Institute, 1995.
- Havas S, Heimendinger J, Damron D, et al. 5 a day for better health—nine community research projects to increase fruit and vegetable consumption. *Public Health Rep* 1995;110:68-79.
- Thompson FE, Larkin FA, Brown MB. Weekend-weekday differences in reported dietary intake: the Nationwide Food Consumption Survey, 1977-78. *Nutr Res* 1986;6:647-62.
- Baranowski T, Smith MNM, Hearn MD, et al. Adult consumption of fruit and vegetables and fat related practices by meal and day. *Am J Health Promot* 1998;12:162-5.
- Douglass D. Food intake analysis system (FIAS), version 2.3. Houston: The University of Texas, 1993.
- US Department of Health and Human Services, US Department of Agriculture. Dietary guidelines for Americans. 3rd ed. Washington, DC: US Government Printing Office, 1989.
- US Department of Agriculture. The food guide pyramid. Hyattsville, MD: Human Nutrition Information Service, 1992. (Publication HG252.)
- US Department of Agriculture, Human Nutrition Information Service. Choose a diet with plenty of vegetables, fruits, and grain products. *Home and Garden Bulletin* 1993;July:253-5.
- US Department of Agriculture. What we eat in America 1994-96, Continuing Survey of Food Intakes by Individuals (CSFII), 1994-1996. Beltsville, MD: Agricultural Research Service, 1998. (Public use CD-ROM.)
- Cleveland LE, Cook DA, Krebs-Smith SM, Friday J. Method for assessing food intakes in terms of servings based on food guidance. *Am J Clin Nutr* 1997;65(suppl):1254S-63S.
- Subar AF, Midthune D, Kulldorff M, et al. An evaluation of alternative approaches to assigning nutrient values to food groups in food frequency questionnaires. *Am J Epidemiol* (in press).
- Freedman LS, Carroll RJ, Wax Y. Estimating the relation between dietary intake obtained from a food frequency questionnaire and true average intake. *Am J Epidemiol* 1991;134:310-20.
- Serdula M, Coates R, Byers T, et al. Evaluation of a brief telephone questionnaire to estimate fruit and vegetable consumption in diverse study populations. *Epidemiology* 1993;4:455-63.
- Baranowski T, Smith M, Baranowski J, et al. Low validity of a seven-item fruit and vegetable food frequency questionnaire among third-grade students. *J Am Diet Assoc* 1997;97:66-8.
- Patterson BH, Block G, Rosenberger WF, Pee D, Kahle LL. Fruit and vegetables in the American diet: data from the NHANES II survey. *Am J Public Health* 1990;80:1443-9.
- Serdula M, Coates RJ, Byers T, Simoes EJ, Mokdad A, Subar AF. Fruit and vegetable intake among adults in 16 states: results of a brief telephone survey. *Am J Public Health* 1995;85:236-9.
- Tippett KS, Cypel YS, eds. Design and operation: the Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, 1994-96. Beltsville, MD: US Department of Agriculture, Agricultural Research Service, 1998. (Nationwide Food Surveys report no. 96-1.)
- Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology. 2. Evaluating the results of published surveys. *Eur J Clin Nutr* 1991;45:583-99.
- Johnson RK, Goran MI, Poehlman ET. Correlates of over- and underreporting of energy intake in healthy older men and women. *Am J Clin Nutr* 1994;59:1286-90.
- Krebs-Smith SM, Graubard BI, Kahle LL, Subar AF, Cleveland LE, Ballard-Barbash R. Low energy reporters vs others: a comparison of reported food intakes. *Eur J Clin Nutr* 2000;54:281-7.
- Field AE, Colditz GA, Fox MK, et al. Comparison of 4 questionnaires for assessment of fruit and vegetable intake. *Am J Public Health* 1998;88:1216-8.
- Campbell MK, Demark-Wahnefried W, Symons M, et al. Fruit and vegetable consumption and prevention of cancer: the Black Churches United for Better Health Project. *Am J Public Health* 1999;89:1390-6.