Public Health Nutrition: 7(8), 1097–1106 DOI: 10.1079/PHN2004642

Performance of a short tool to assess dietary intakes of fruits and vegetables, percentage energy from fat and fibre

Frances E Thompson[1,*], Douglas Midthune[1], Amy F Subar[1], Lisa L Kahle[2], Arthur Schatzkin[1] and Victor Kipnis[1]

[1]Risk Factor Surveillance and Methodology Branch, Applied Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, EPN 4016, 6130 Executive Blvd, MSC 7344, Bethesda, MD 20892-7344, USA: [2]Information Management Services, Inc., Silver Spring, MD, USA

Submitted 14 January 2004: Accepted 21 May 2004

Abstract

Objectives: We describe the methods used to develop and score a 17-item 'screener' designed to estimate intake of fruit and vegetables, percentage energy from fat and fibre. The ability of this screener and a food-frequency questionnaire (FFQ) to measure these exposures is evaluated.

Design: Using US national food consumption data, stepwise multiple regression was used to identify the foods to be included on the instrument; multiple regression analysis was used to develop scoring algorithms. The performance of the screener was evaluated in three different studies. Estimates of intakes measured by the screener and the FFQ were compared with true usual intake based on a measurement error model.

Setting: US adult population.

Subjects: For development of instrument, $n = 9323$ adults. For testing of instrument, adult men and women in three studies completing multiple 24-hour dietary recalls, FFQ and screeners, $n = 484$, 462 and 416, respectively.

Results: Median recalled intakes for examined exposures were generally estimated closely by the screener. In the various validation studies, the correlations between screener estimates and estimated true intake were 0.5–0.8. In general, the performances of the screener and the full FFQ were similar; estimates of attenuation were lower for screeners than for full FFQs.

Conclusions: When coupled with appropriate reference data, the screener approach described may yield useful estimates of intake, for both surveillance and epidemiological purposes.

Keywords
Diet
Epidemiological methods
Surveillance
Validation study
Data collection
Questionnaires

In 1998, the US National Cancer Institute (NCI) had the opportunity to develop a supplement to the National Health Interview Survey 2000, a survey of 40 000 adults aged 20 years and above. The supplement, called the Cancer Control Module (CCM), could average no more than 20 min. Areas of interest included cancer-related behaviours, such as participation in cancer screening, smoking, sunlight avoidance, physical activity and diet. The challenge in developing the dietary component was to provide meaningful information about diet in no more than 20 questions.

This paper describes our efforts to develop and evaluate a short dietary assessment tool, a 'screener' to estimate usual dietary intake of fruits and vegetables, percentage energy from fat and fibre. We describe the approach used to develop the instrument and analytical scoring procedures and how the new instrument was evaluated in community samples of adults, and report the instrument's ability to estimate mean intakes and its estimated correlation with true intake.

Methods

Development of instrument

Identification of foods

We were primarily interested in assessing intakes of servings of fruits and vegetables, percentage energy from fat and grams of fibre. To identify which foods best predicted these dietary factors, we analysed dietary intake data (two non-consecutive 24-hour dietary recalls) collected from a nationally representative US sample of the US Department of Agriculture's (USDA) 1994–96 Continuing Survey of Food Intakes by Individuals (CSFII 94–96). These data include estimated daily individual intakes of fat (g), energy (kcal), fibre (g) and servings of

*Corresponding author: Email ThompsoF@mail.nih.gov © The Authors 2004

fruits and vegetables, as defined by national dietary guidance[1]. For fruits, a serving is defined as a whole fruit, such as a medium apple; 1/2 cup of cut-up fruit; or 3/4 cup of fruit juice. For vegetables, a serving is defined as 1 cup of raw leafy vegetables, like lettuce; 1/2 cup of other vegetables; or 3/4 cup of vegetable juice. CSFII 94–96[2] provides the number of Food Guide Pyramid fruit and vegetables servings per 100 g of the food for each of over 5000 foods.

We categorised these CSFII foods into 193 mutually exclusive food groups, and calculated, for each food group and each individual in the survey, the average number of times per day the individual reported consuming a food in that group. The resulting food group variables represent possible frequency of use questions that could be asked on a short instrument in order to estimate dietary factors of interest.

We performed weighted least-squares regression on the adults (aged 18 years or older) in the CSFII 94–96 ($n = 9323$), using the CSFII sample weights, stratifying by gender. We used stepwise selection procedures to identify the food groups that best predicted each of the dietary factors of interest. The response variables were: (1) square-root of Pyramid servings of fruits and vegetables per day; (2) percentage energy from fat; and (3) cube-root of dietary fibre in g per day. The square-root and cube-root transformations were applied to better approximate a normal distribution. Before the analysis, we excluded extreme values of each dietary variable, defined as values more than two interquartile ranges below the first quartile or two interquartile ranges above the third quartile of that variable's distribution.

Because we were primarily interested in estimating fruit and vegetable intake, we first performed stepwise regression for fruits and vegetables and selected the six food groups that individually contributed at least 2% to the increase in the R^2 value, the proportion of variation in the response variable explained by the regression. We added lettuce salad, even though it did not meet the criterion, because it is included in existing short tools for fruits and vegetables, and we wanted to maintain comparability of instruments as much as possible. We then performed stepwise regression for percentage energy from fat, after including the seven previously selected food group variables in the model, and selected in the same way six more foods to predict percentage energy from fat. Finally, we performed stepwise regression for fibre, after including the 13 previously selected food group variables into the model, and selected three more foods to predict fibre.

The six important predictors of fruits and vegetables, in order of selection, were: vegetables (other than potatoes, salads and dried beans); fruit juice; fruit; fried potatoes; other white potatoes; and cooked dried beans. Beyond these six and lettuce salad, the most important predictors of percentage energy from fat were: regular-fat bacon and sausage; potato chips; regular-fat salad dressing; nuts/seeds; skimmed milk; and regular-fat hot dogs. The additional three important predictors of fibre were: ready-to-eat cereals; whole-grain breads; and pasta. Taken together, frequency of these 16 food groups accounted for 45% of the variability in fruits and vegetables, 27% in percentage energy from fat and 41% in fibre among the CSFII 94–96 adult population.

In addition to the above quantitative results, we considered results from cognitive testing of how to enhance understanding of and response to food frequency-type questions[3,4] to develop the 17-question test instrument, called the Multifactor Screener (Appendix). The screener asks frequency questions for 16 food groups, including milk. A separate question about usual type of milk consumed is used to define four alternative types of milk (whole, 2%, 1% and skimmed), and the frequency is attributed to the specific type reported.

Development of analytical scoring algorithms

After deciding which questions to include in the Multifactor Screener, we developed scoring algorithms to translate an individual's frequency responses into estimated dietary intakes, using CSFII 94–96 recall data. We used the variables for average number of times per day the individual reported consuming a food in a given food group, for the 19 food groups consistent with the questions asked on the Multifactor Screener (15 non-milk food groups and four milk food groups). Although not originally planned, we decided to add portion size information from CSFII 94–96 in the scoring algorithms.

Portion size of many foods varies with gender and age[5]. For each of the 19 food groups, we computed from CSFII 94–96 the median portion size per mention of the food group for each of 14 population groups, defined by two gender and seven age groups (18–27 years, 28–37 years, …, 78 years and over). For analysis of fruits and vegetables, the portion size estimates were calculated in terms of servings of fruits and vegetables; for analyses of percentage energy from fat and fibre, the portion size estimates were calculated in terms of grams of food. These portion size estimates are found elsewhere[6].

For percentage energy from fat and fibre, the model took the form:

$$E\ (\text{dietary factor}) = \beta_0 + \beta_1 N_{\mathrm{FG}1} P_1 + \beta_2 N_{\mathrm{FG}2} P_2 + \ldots + \beta_{19} N_{\mathrm{FG}19} P_{19}. \quad (1)$$

Fibre was cube-root-transformed to approximate normality; no transformation was necessary for percentage energy from fat. $N_{\mathrm{FG}k}$ is the usual number of times per day an individual consumed food group k as reported on the screener; P_k is the median portion size of group k; and k indexes the 19 food groups. We calculated weighted least-squares estimates of the regression coefficients β_k, $k = 0, \ldots, 19$ on CSFII 94–96 adults aged 18 years

and above, stratifying by gender and excluding extreme exposure values. We first included all 19 food groups in the regression model. After examining the results, we dropped food groups that failed to attain statistical significance at the $\alpha = 0.25$ level to form more parsimonious final models. Because of the complex survey design, the analysis was performed using SUDAAN (RTI Inc., Research Triangle Park, NC, USA).

For servings of fruits and vegetables, the scoring algorithm was derived from the regression equation:

$$E\ (\text{dietary factor}^{1/2}) = \beta_0 + \beta_1(N_{\text{FG}1}P_1 + N_{\text{FG}2}P_2 + \ldots + N_{\text{FG}7}P_7)^{1/2}, \quad (2)$$

where the dietary factor is square-root-transformed to approximate normality; $N_{\text{FG}k}$ is the usual number of times per day an individual consumed food group k as reported on the screener; P_k is the median portion size of group k; and k indexes the seven fruit and vegetable food groups. We calculated weighted least-squares estimates of the regression coefficients β_0 and β_1 on the adults in the CSFII 94–96, stratifying by gender and excluding extreme exposure values. Comparisons in the validation datasets indicated that model (2) performs better than model (1) to estimate servings of fruits and vegetables.

For percentage energy from fat and fibre, the estimated parameters are shown in Table 1. For servings of fruits and vegetables, the estimated parameters are: $\beta_0 = 0.90679$ and $\beta_1 = 0.75856$ for men; $\beta_0 = 0.81956$ and $\beta_1 = 0.73086$ for women.

Testing of the instrument

Subjects and study design

We tested the performance of the Multifactor Screener in the NCI's Observing Protein and Energy (OPEN) Study, described in detail elsewhere[7]. OPEN enrolled 484 individuals, aged 40–69 years at baseline in 1999 and living in the Maryland suburban area near Washington, DC, to examine the structure of measurement error in self-reported dietary assessment tools. Data collected included two separate administrations of the NCI's new food-frequency questionnaire (FFQ) – the Diet History Questionnaire (DHQ) – and a 24-hour recall, administered approximately 3 months apart, and a single administration of the Multifactor Screener, 2 weeks from baseline.

Two other validation studies were also available to examine the performance of parts of the Multifactor Screener. The Eating at America's Table Study (EATS) included 462 individuals, aged 20–60 years at baseline in 1998, living throughout the USA[8]. Four non-consecutive 24-hour recalls (one per season) were administered by telephone, in addition to the DHQ and two versions of fruit and vegetable screeners that included portion size questions[4]. For purposes of the current analyses, only the frequency questions identical to those on the Multifactor Screener were used.

The screener validation part of the National Institutes of Health–American Association of Retired Persons (NIH–AARP) Diet and Health Study[9] included 416 individuals, aged 50–69 years at baseline in 1995, living throughout the USA. Two non-consecutive 24-hour recalls were administered by telephone, in addition to two FFQs and a screener to estimate percentage energy from fat[10]. The questions asked on the NIH–AARP Diet and Health Study percentage energy from fat screener differed somewhat from those asked on the Multifactor Screener. For purposes of the current analyses, only eight of the frequency questions – those that corresponded to foods asked on the Multifactor Screener – were analysed. Similar

Table 1 Estimated regression coefficients for foods as predictors of percentage energy from fat and fibre (g), by gender

	Men		Women	
Parameter	Percentage energy from fat	Fibre (cube root)	Percentage energy from fat	Fibre (cube root)
Intercept	31.93268	2.08423	31.36357	1.89847
Cold cereals	−0.02672	0.00209	−0.05797	0.00389
Whole milk	0.00653	0.00013	0.00842	0.00009
2% milk	0.00215	0.00013	0.00272	0.00011
1% milk	−0.00149	0.00022	−0.00196	0.00024
Skimmed milk	−0.00841	0.00028	−0.00867	0.00034
Bacon or sausage	0.13831	−0.00139	0.23128	−0.00201
Hot dogs	0.04078	0	0.10160	−0.00141
Whole-grain bread	0	0.00283	0	0.00337
100% fruit juice	−0.00533	0.00019	−0.01011	0.00025
Fruit	−0.00932	0.00103	−0.01201	0.00105
Salad dressing	0.15036	0	0.23974	0
Salad	0	0	0	0
Fried potatoes	0.02734	0.00160	0.04272	0.00156
Other white potatoes	0.00580	0.00071	0.00618	0.00066
Dried beans	−0.00526	0.00275	−0.00608	0.00380
Other vegetables	0	0.00084	0	0.00093
Pasta	−0.00504	0.00075	−0.00540	0.00082
Nuts	0.12454	0.00546	0.26018	0.00603
Chips	0.05376	0.00528	0.13144	0.00456

to procedures described earlier, a scoring algorithm was developed using CSFII data. Portion size-adjusted daily frequencies of each of the eight foods (cold cereals, skimmed milk, fruit, citrus fruit juice, mayonnaise/salad dressing, bacon, hot dogs, French fries) were regressed on percentage energy from fat to derive the appropriate regression coefficients. These CSFII-derived regression coefficients were used in conjunction with individual responses on the NIH–AARP Diet and Health Study percentage energy from fat screener to estimate percentage energy from fat.

Data collection and processing, and creation of variables

24-Hour dietary recalls. The 24-hour dietary recalls in OPEN were administered using a standardised five-pass methodology developed by the USDA for national dietary surveillance[11]; see Subar *et al.*[7] for a description of the procedures used. The 24-hour dietary recalls in EATS and the NIH–AARP Diet and Health Study were administered by telephone[8,9]. To derive Pyramid servings of fruits and vegetables, individual food codes reported were matched to corresponding CSFII 94–96 food codes which contained Pyramid serving information.

FFQ. The DHQ is a 36-page booklet that asks about frequency of intake of 124 individual food items (122 items include three-category portion size questions) in the past year. Its development and validation are reported elsewhere[8,12]. The DHQ was administered in both OPEN and EATS. The FFQ administered in the NIH–AARP Diet and Health Study was an earlier version of the DHQ and is described in reference 9.

Multifactor Screener. Each frequency category was assigned a value equal to the midpoint of that category and converted to times per day. Intake of fruits and vegetables, percentage energy from fat and fibre were estimated for each individual from the Multifactor Screener using the scoring algorithm described previously (called regression model estimates). For servings of fruits and vegetables, we calculated also an 'unadjusted screener' estimate:

$$\mathrm{F\&V_{freq}} = N_{\mathrm{FG1}} + N_{\mathrm{FG2}} + \ldots N_{\mathrm{FG7}},$$

where $N_{\mathrm{FG}k}$ is the usual number of times per day an individual consumed food group k, as reported on the screener.

The unadjusted fruit and vegetable screener estimate was computed and compared with the regression model estimate because earlier work[4] suggested that the information captured by this variable is sufficient to attain reasonable correlations.

Table 2 summarises the variables examined for each exposure. We also used the same algorithms with CSFII frequency variables to compute estimated intakes of the three variables in CSFII using only the screener-type information. We call this a pseudo-screener, since CSFII subjects were not administered a screener.

Measurement error model

True usual intake is not observable in free-living populations. However, we can estimate the distribution of true intake in the population by use of appropriate reference data and statistical methods. Our reference instrument is multiple non-consecutive 24-hour recalls. We applied a measurement error model described by Freedman *et al.*[13] to estimate relationships between true intake and the test instruments, assuming that the reference instrument (24-hour recalls) is unbiased at the individual level and contains only within-person error.

We fitted the measurement error model, by gender, to the FFQ, screener and 24-hour recall data, and estimated measurement error characteristics of the FFQ and screener. For each instrument, we estimated the slope (β_1) and intercept (β_0) in the regression of reported intake on true intake (μ_{T}), the correlation coefficient between reported and true intake (R), and the attenuation coefficient ($\lambda = R^2/\beta_1$). The parameters β_1, β_0 and R reflect validity. If the test instrument is unbiased at the individual level, the regression of questionnaire-reported intake on true intake should have a slope of 1 and an intercept of 0. The squared correlation coefficient (R^2) measures the proportion of variation in true intake explained by reported intake. In studies relating diet and disease, $1 - R^2$ represents the loss of power to detect diet–disease associations due to measurement error in the dietary assessment instrument. R^2 is inversely proportional to the sample size needed to obtain the desired power to detect a relationship, while λ indicates the extent to which the observed log relative risk is biased

Table 2 Variables analysed for each dietary factor by study

Study and dietary factor	Reference instrument	Screener	FFQ
OPEN			
Fruits and vegetables (servings)	Two 24-hour recalls	Unadjusted screener; regression model estimate	DHQ
Percentage energy from fat	Two 24-hour recalls	Regression model estimate	DHQ
Fibre (g)	Two 24-hour recalls	Regression model estimate	DHQ
EATS			
Fruits and vegetables (servings)	Four 24-hour recalls	Unadjusted screener; regression model estimate	DHQ
NIH–AARP Diet and Health Study			
Percentage energy from fat	Two 24-hour recalls	Regression model estimate	FFQ

FFQ – food-frequency questionnaire; OPEN – Observing Protein and Energy study; EATS – Eating at America's Table Study; NIH–AARP – National Institutes of Health–American Association of Retired Persons; DHQ – Diet History Questionnaire.

towards zero by measurement error. An attenuation coefficient close to 1 indicates a small amount of attenuation, whereas λ close to 0 indicates severe attenuation. The more severe the attenuation, the more important it is to be able to correct for attenuation via a calibration sub-study.

We calculated maximum-likelihood estimates of the parameters β_0, β_1, R, λ and μ_T and their standard errors under the assumption of normality. We excluded outliers, as previously defined, in order to avoid their undue influence.

Results

Table 3 presents gender-specific estimates of dietary intake for the various instruments in the OPEN, EATS and NIH–AARP Diet and Health studies. Table 4 shows the estimated slopes in the regressions of the test instruments on true intake, correlations between the instruments and true intake, and the attenuation factors for the test instruments for intake of each dietary factor estimated for the screener and the DHQ/FFQ in the studies.

Servings of fruits and vegetables

For both men and women in both OPEN and EATS, the unadjusted screener significantly underestimated median intake (Table 3). The screener regression model medians were closer to the estimated true median intake. In OPEN, the screener regression model medians were somewhat lower than estimated true median intake (1 and 0.7 servings per day less for men and women, respectively); in EATS, they were not statistically significantly different from estimated true median intake. For both men and women in both studies, the DHQ medians were higher than the screener medians. In comparison with recall medians, the DHQ was similar for men, but significantly higher for women. Differences between men and women in mean intakes apparent from recall estimates were reflected well by screener estimates and less so by DHQ estimates.

The parameters for the unadjusted screener model and the regression model were not significantly different from each other: for example, for men, correlations were about 0.6–0.7 and for women, 0.5 in EATS and 0.6–0.7 in OPEN (Table 4). This confirms earlier findings[4] that the unadjusted screener is sufficient to obtain reasonable

Table 3 Estimated median daily intake of dietary factor (and 95% confidence interval) by dietary factor, study* and instrument, according to gender

Dietary factor, study and instrument	Gender	
	Men	Women
Servings of fruits and vegetables		
OPEN		
Recalls	6.3 (5.95–6.75)	5.4 (5.01–5.72)
Screener:		
unadjusted screener	2.9 (2.71–3.09)†	3.3 (3.03–3.49)†
regression model estimate	5.3 (5.13–5.56)†	4.7 (4.46–4.87)†
DHQ	6.2 (5.80–6.57)	6.2 (5.75–6.57)†
EATS		
Recalls	5.8 (5.39–6.22)	4.2 (3.96–4.49)
Screener:		
unadjusted screener	2.9 (2.66–3.09)†	3.1 (2.85–3.29)†
regression model estimate	5.5 (5.22–5.73)	4.5 (4.32–4.72)
DHQ	6.6 (6.06–7.26)†	6.2 (5.81–6.68)†
Percentage energy from fat		
OPEN		
Recalls	31.8 (30.9–32.7)	32.0 (30.9–33.0)
Screener	32.0 (31.6–32.4)	30.5 (30.0–31.0)†
DHQ	30.6 (29.7–31.5)†	30.0 (29.0–31.0)†
NIH–AARP Diet and Health Study		
Recalls	30.7 (30.1–31.2)	30.1 (29.6–30.7)
Screener	30.2 (29.5–30.8)	28.6 (27.9–29.4)†
FFQ	30.5 (30.4–30.5)	30.0 (30.0–30.1)
Fibre (g)		
OPEN		
Recalls	21.2 (20.2–22.4)	16.5 (15.5–17.5)
Screener	18.3 (17.7–18.9)†	14.1 (13.6–14.7)†
DHQ	18.2 (17.2–19.3)†	15.7 (14.8–16.6)

OPEN – Observing Protein and Energy study; DHQ – Diet History Questionnaire; EATS – Eating at America's Table Study; NIH–AARP – National Institutes of Health–American Association of Retired Persons; FFQ – food-frequency questionnaire.

*Number of subjects per study: OPEN – $n = 261$ men, $n = 223$ women; EATS – $n = 202$ men, $n = 260$ women; NIH–AARP Diet and Health Study – $n = 211$ men, $n = 205$ women.

† Significantly different from recall, $P < 0.05$.

Table 4 Estimates of slope in the regression of reported dietary factor on true intake*, correlations between reported and true intakes, and attenuation coefficients resulting from measurement error in reported intake for each dietary factor, study and test instrument by gender

Dietary factor, study and test Instrument	Men			Women		
	Slope (SEE)	Correlation (SEE)	Attenuation (SEE)	Slope (SEE)	Correlation (SEE)	Attenuation (SEE)
Servings of fruits and vegetables						
OPEN						
Screener:						
unadjusted screener	0.54 (0.085)	0.59 (0.065)	0.64 (0.080)	0.95 (0.176)	0.78 (0.079)	0.64 (0.068)
regression model estimate	0.43 (0.069)	0.57 (0.066)	0.75 (0.097)	0.70 (0.146)	0.76 (0.087)	0.81 (0.094)
DHQ	0.83 (0.115)	0.64 (0.058)	0.49 (0.051)	1.01 (0.222)	0.62 (0.083)	0.38 (0.052)
EATS						
Screener:						
unadjusted screener	0.62 (0.073)	0.69 (0.055)	0.77 (0.079)	0.58 (0.082)	0.50 (0.059)	0.43 (0.058)
regression model estimate	0.50 (0.062)	0.66 (0.058)	0.87 (0.095)	0.42 (0.061)	0.48 (0.060)	0.56 (0.078)
DHQ	1.07 (0.135)	0.67 (0.060)	0.42 (0.047)	0.86 (0.114)	0.54 (0.059)	0.34 (0.043)
Percentage energy from fat						
OPEN						
Screener: regression model estimate	0.31 (0.052)	0.55 (0.068)	0.96 (0.130)	0.42 (0.089)	0.60 (0.085)	0.88 (0.130)
DHQ	0.95 (0.116)	0.76 (0.051)	0.60 (0.049)	1.04 (0.187)	0.76 (0.073)	0.56 (0.056)
NIH–AARP Diet and Health Study						
Screener: regression model estimate	0.55 (0.081)	0.55 (0.069)	0.56 (0.072)	0.56 (0.093)	0.51 (0.073)	0.46 (0.068)
FFQ	0.93 (0.067)	0.72 (0.029)	0.55 (0.024)	0.85 (0.067)	0.65 (0.032)	0.49 (0.027)
Fibre (g)						
OPEN						
Screener: regression model estimate	0.41 (0.056)	0.54 (0.059)	0.70 (0.088)	0.43 (0.072)	0.55 (0.069)	0.69 (0.098)
DHQ	0.75 (0.089)	0.59 (0.052)	0.46 (0.048)	0.56 (0.106)	0.45 (0.070)	0.36 (0.060)

SEE – standard error of the estimate; OPEN – Observing Protein and Energy study; DHQ – Diet History Questionnaire; EATS – Eating at America's Table Study; NIH–AARP – National Institutes of Health–American Association of Retired Persons; FFQ – food-frequency questionnaire.
* True intake estimated from multiple non-consecutive 24-hour recalls in a measurement error model.

correlations. In comparison with the DHQ, estimated correlations between the screener and fruit and vegetable intake were similar. Estimated attenuation coefficients were larger for the screener than for the DHQ.

Percentage energy from fat

For men, the screener means were not statistically significantly different from the recall estimates (Table 3) in both OPEN and the NIH–AARP Diet and Health Study. For women, however, the screener means were significantly lower than the recall means in both studies. The screener estimates were closer to estimated true intake than the DHQ/FFQ estimates among men in OPEN, and vice versa among women in the NIH–AARP Diet and Health Study.

For the screener, the estimated correlations with true intake were 0.5–0.6 for both genders and studies (Table 4). In comparison, for the DHQ/FFQ, estimated correlations with true intake were 0.7–0.8. Again, there was less attenuation for the screener than for the FFQ, but only in OPEN.

Fibre

In the OPEN study, recall medians were 21.2 for men and 16.5 for women (Table 3). The screener medians of 18.3 for men and 14.1 for women were statistically significantly lower than the recall medians. The DHQ medians were also lower than recall medians. For both genders, the estimated correlations between the screener and true intake were similar to estimated correlations between the DHQ and true intake, about 0.5 (Table 4), although there was less attenuation for the screener than for the DHQ.

Discussion

Short dietary assessment instruments have been developed for a variety of dietary factors and purposes. The method of identifying from national data the main sources of a nutrient was used by Block *et al.* to develop a 13-item fat screening tool to be used to screen populations at higher average risk[14]. Kristal *et al.* pioneered the development of questionnaires asking about specific types of behaviour, e.g. trimming fat from chicken, and have related scores to intake estimated by more detailed methods[15]. Others have modified or developed additional questions about behaviours to be used in a clinical setting to encourage specific dietary behaviours[16]. Similar to Block *et al.*, we used national data, but unlike them, our aim was not to identify the major food sources of particular nutrients, but to identify the foods contributing most to explaining the variability in the intake of particular nutrients. Our purpose was to develop a short tool useful to estimate population intakes for surveillance and epidemiology. Our results suggest that the Multifactor Screener with its scoring algorithms shows promise for measuring mean or median intakes of servings of fruits

and vegetables, percentage energy from fat and grams of fibre.

For servings of fruits and vegetables, estimates of intake based only on frequency were lower than recall or record estimates, consistent with other studies[17–21]. The earlier EATS validation study, using similarly worded frequency questions but with portion size questions, produced estimates within 1.2 serving per day of median true intakes[4]. When the scoring algorithms are applied to the screener frequency responses, screener estimates of medians differed by no more than 1.0 serving per day from estimated true intakes. Thus, in this study, the portion size-adjusted regression coefficients applied to the frequency questions appeared to adequately substitute for portion size questions. The Multifactor Screener fruits and vegetables component was strongly correlated with estimated true intake, at about 0.6–0.7 among men and 0.5–0.8 among women. In the EATS study, correlations using respondent-reported portion size information[4] were similar to correlations using the portion size-adjusted regression coefficients. Consistent with that study, correlations for the Multifactor Screener fruit and vegetable component were higher than those found in other studies of adults using records or recalls and the Behavioural Risk Factor Surveillance Survey (correlations ranged from 0.29 to 0.56)[17,20] or 5 a Day instruments (correlations ranged from 0.24 to 0.52)[21,22], indicating that the enhancements in the wording of the fruit and vegetable questions are associated with improved performance.

For percentage energy from fat, Multifactor Screener estimates of mean intake were similar to recall estimates for men and somewhat lower for women, in both OPEN and the NIH–AARP Diet and Health Study. This is consistent with the previous NIH–AARP validation study of a 13-item percentage energy from fat screener[10]. The development of that full screener was similar to the current work. The favourable results in both studies indicate the effectiveness of the development process. While other instruments have been developed to briefly assess grams of fat intake[14,23] or ranking using a fat index[24], we know of no other that quantitatively estimates percentage energy from fat. Absolute fat intake is less meaningful without a measure of total energy intake, as requirements for fat increase with increased energy requirements. We know of no short instrument that assesses total energy intake, nor do we think that task is feasible with a short instrument.

While the Multifactor Screener underestimated fibre intake, its estimated correlation with true intake was about 0.54. The Block seven-item instrument plus beans[23] correlated 0.62 with grams of fibre estimated by a 100-item Block FFQ. Kristal and co-workers' 33-item fat- and fibre-related behaviour questionnaire correlated 0.50 with fibre g/1000 kcal estimated by an FFQ[25]. However, in both studies, the correlation was estimated across men and women, and an FFQ was used as the reference measure; thus, due to FFQ measurement error, these correlations are likely to substantially overestimate the within-gender correlations between the screener or questionnaire and true intake.

Table 5 Observed relative risk (RR) at different levels of true RR and attenuation

If true RR is:	and attenuation is:	Observed RR at level of attenuation is:
5.0	0.4	1.90
	0.6	2.63
	0.8	3.62
	0.9	4.26
3.0	0.4	1.55
	0.6	1.93
	0.8	2.41
	0.9	2.69
1.8	0.4	1.27
	0.6	1.42
	0.8	1.60
	0.9	1.70

For servings of fruits and vegetables and fibre, estimated correlations between the screeners and true intake and a complete FFQ and true intake were very similar. These results indicate that for the adult population studied, the variation in intakes appears to be captured as well with a short targeted list of foods without portion size information as with a longer list with portion size information. Moreover, relative risk estimates based on data from the screener may be less attenuated than those based on data from an FFQ. For example, if the true relative risk between a disease and lower fruit and vegetable intake is 3.0, the relative risks observed among men would be $3^{0.75} = 2.3$ using the screener and $3^{0.49} = 1.7$ using the FFQ. The effect of attenuation on observed relative risk is illustrated in Table 5. Less attenuation for the screener means that the need to de-attenuate observed relative risks, and thus the need for a calibration sub-study, is less critical with the screener than the FFQ.

We have taken an empirical approach to model building, using data from a recent, large, representative survey of the diets of the US population. Ideally, we would have liked to estimate the regression coefficients for the screener from a large study in which subjects had been administered both the screener and 24-hour recalls. Since such a study was not available, we created a pseudo-screener in the CSFII data, which may not act precisely like the actual screener. For this reason it was especially important to assess the performance of the regression model estimates within the validation studies.

An additional concern is that the relationships between responses on the screener and the exposure variable may differ substantially by race/ethnicity or some other factor. This might occur, for example, if, among different subgroups, there were substantially different measurement errors in reporting on the screener; if typical portion

size consumed of the individual food groups differed; or if the food patterns differed greatly. There would not be a problem if the diets differed, but only if relationships between the foods asked on the screener and the dietary factor differed. We explored this possibility by comparing gender-specific models with gender/race/ethnicity-specific models in our validation datasets. We found little difference in the level of validity, as compared with the 24-hour recalls. The largest difference was found for the other race/ethnicity category. However, our reference dataset (CSFII 94–96) included relatively few individuals in that category and we think the category is too non-specific to be useful. Therefore, without further reference data, we prefer the more parsimonious gender-specific models.

A final concern with the use of the pseudo-screener in CSFII is that cognitive factors affecting individuals' actual response to the specific screener questions are not reflected in these variables. To take into account this dimension, we performed further analyses of the EATS and NIH–AARP Diet and Health Study. We estimated from those studies regression coefficients for screener questions similar to the Multifactor Screener, and applied those coefficients to the Multifactor Screener items for the OPEN sample. We found no improvement in ability to predict the reference estimates of true intake. Thus, we think that the CSFII's advantages of size and representativeness outweigh its limitations.

A limitation of the study is the assumption that 24-hour recalls provide unbiased data. It is recognised that 24-hour recalls underestimate overall energy intake[7,26–29], but it is not known whether other dietary components are proportionately underestimated. It seems probable that some foods are underreported more than others[30], given that underreporting is a sociological/psychological phenomenon. In the OPEN study, the finding that underreporting of energy (10–14%) in recalls was about double the level of underreporting of protein (6–7%) supports this possibility[31]. The presence and extent of underreporting for fruits and vegetables and for fibre-containing foods are unknown at this time. Reporting biases for percentage of energy from fat are also unknown. However, there is some evidence that measurement error for density variables may be less than for absolute variables[7,31].

We believe that the approach we used of empirical data analysis to identify important foods and their scoring algorithms shows merit for continued evolution of short instruments. However, the instrument developed is highly dependent on the suitability of the data, both for the population of interest and the time period. For example, the national US data we used are less useful for groups of individuals whose eating patterns differ substantially. As the food supply of the USA changes and as the population changes its eating patterns, newer reference data will be needed.

Acknowledgements

The authors wish to acknowledge and thank the following individuals for their invaluable contributions to this project. We thank the following staff from Westat, Inc.: Paul Hurwitz, for overall study management in the NIH–AARP Diet and Health Study; Susie McNutt, for management of the 24-hour dietary recall collection and effort; and Anna McIntosh, for screener questionnaire formatting and management.

Data collection for the OPEN study, the EATS study and the NIH–AARP Diet and Health Study were approved by the NCI Institutional Review Board and the Westat, Inc. Special Studies Institutional Review Board. Informed consent was obtained from all participants.

References

1 US Department of Agriculture, Human Nutrition Information Service. USDA's food guide pyramid. *Home and Garden Bulletin* 1992: 252.
2 US Department of Agriculture (USDA). *What we eat in America 1994–96, Continuing Survey of Food Intakes by Individuals (CSFII) 1994–1996* [public use CD-ROM]. Washington, DC: USDA, 1998.
3 Subar AF, Thompson FE, Smith AF, Jobe JB, Ziegler RG, Potischman N, *et al.* Improving food frequency questionnaires: a qualitative approach using cognitive interviewing. *Journal of the American Dietetic Association* 1995; **95**: 781–8.
4 Thompson FE, Subar AF, Smith AF, Midthune D, Radimer KL, Kahle LL, *et al.* Fruit and vegetable assessment: performance of two new short instruments and a food frequency questionnaire. *Journal of the American Dietetic Association* 2002; **102**: 1764–72.
5 Krebs-Smith SM, Guenther PM, Cook A, Thompson FE, Cucinelli J, Udler J. *Foods Commonly Eaten in the United States: Quantities Consumed per Eating Occasion and in a Day, 1989–91*. NFS Report No. 91–3. Washington, DC: US Department of Agriculture, 1997.
6 US National Cancer Institute, Division of Cancer Control and Population Sciences, Applied Research Program. http://appliedresearch.cancer.gov/surveys/nhis/multifactor/scoring.html
7 Subar AF, Kipnis V, Troiano R, Midthune D, Schoeller DA, Bingham S, *et al.* Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the Observing Protein and Energy Nutrition (OPEN) Study. *American Journal of Epidemiology* 2003; **158**: 1–13.
8 Subar AF, Thompson FE, Kipnis V, Hurwitz P, McNutt S, McIntosh A, *et al.* Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: The Eating at America's Table Study (EATS). *American Journal of Epidemiology* 2001; **154**: 1089–99.
9 Schatzkin A, Subar AF, Thompson FE, Harlan LC, Tangrea J, Hollenbeck AR, *et al.* Design and serendipity in establishing a large cohort with wide dietary intake distributions: The National Institutes of Health–American Association of Retired Persons Diet and Health Study. *American Journal of Epidemiology* 2001; **154**: 1119–25.
10 Thompson FE, Kipnis V, Subar AF, Schatzkin A, Potischman N, Kahle L *et al.* Performance of a short instrument to estimate usual dietary intake of percent calories from fat. *European Journal of Clinical Nutrition* 1998; **52**(Suppl. 2): S63.

11 Moshfegh AJ, Raper N, Ingwersen L, Cleveland L, Anand J, Goldman J, *et al*. An improved approach to 24-hour dietary recall methodology. *Annals of Nutrition & Metabolism* 2001; **45**(Suppl. 1): 156.
12 Subar AF, Midthune D, Kulldorff M, Brown CC, Thompson FE, Kipnis V, *et al*. Evaluation of alternative approaches to assign nutrient values to food groups in food frequency questionnaires. *American Journal of Epidemiology* 2000; **152**: 279–86.
13 Freedman LS, Carroll RJ, Wax Y. Estimating the relation between dietary intake obtained from a food frequency questionnaire and true average intake. *American Journal of Epidemiology* 1991; **134**: 310–20.
14 Block G, Clifford C, Naughton MD, Henderson M, McAdams M. A brief dietary screen for high fat intake. *Journal of Nutrition Education* 1989; **21**: 199–207.
15 Kristal AR, Shattuck AL, Henry HG. Patterns of dietary behavior associated with selecting diets low in fat: reliability and validity of a behavioral approach to dietary assessment. *Journal of the American Dietetic Association* 1990; **90**: 214–20.
16 Smith-Warner SA, Elmer PJ, Tharp TM, Fosdick L, Randall B, Gross M, *et al*. Increasing vegetable and fruit intake: randomized intervention and monitoring in an at-risk population. *Cancer Epidemiology, Biomarkers & Prevention* 2000; **9**: 307–17.
17 Serdula M, Coates R, Byers T, Mokdad A, Jewell S, Chavez N, *et al*. Evaluation of a brief telephone questionnaire to estimate fruit and vegetable consumption in diverse study populations. *Epidemiology* 1993; **4**: 455–63.
18 Field AE, Colditz GA, Fox MK, Byers T, Serdula M, Bosch RJ, *et al*. Comparison of 4 questionnaires for assessment of fruit and vegetable intake. *American Journal of Public Health* 1998; **88**: 1216–8.
19 Campbell MK, Demark-Wahnefried W, Symons M, Kalsbeek WD, Dodds J, Cowan A, *et al*. Fruit and vegetable consumption and prevention of cancer: The Black Churches United for Better Health Project. *American Journal of Public Health* 1999; **89**: 1390–6.
20 Smith-Warner SA, Elmer PJ, Fosdick L, Tharp TM, Randall B. Reliability and comparability of three dietary assessment methods for estimating fruit and vegetable intakes. *Epidemiology* 1997; **8**: 196–201.
21 Thompson FE, Kipnis V, Subar AF, Krebs-Smith SM, Kahle LL, Midthune D, *et al*. Evaluation of 2 brief instruments and a food-frequency questionnaire to estimate daily number of servings of fruit and vegetables. *American Journal of Clinical Nutrition* 2000; **71**: 1503–10.
22 Plesko M, Cotugna N, Aljadir L. Usefulness of a brief fruit and vegetable FFQ in a college population. *American Journal of Health Behavior* 2000; **24**: 201–8.
23 Block G, Gillespie C, Rosenbaum EH, Jenson C. A rapid food screener to assess fat and fruit and vegetable intake. *American Journal of Preventive Medicine* 2000; **18**: 284–8.
24 Yaroch AL, Resnicow K, Khan LK. Validity and reliability of qualitative dietary fat index questionnaires: a review. *Journal of the American Dietetic Association* 2000; **100**: 240–4.
25 Shannon J, Kristal AR, Curry SJ, Beresford SA. Application of a behavioral approach to measuring dietary change: the fat- and fiber-related diet behavior questionnaire. *Cancer Epidemiology, Biomarkers & Prevention* 1997; **6**: 355–61.
26 Johnson RK, Soultanakis RP, Matthews DE. Literacy and body fatness are associated with underreporting of energy intake in US low-income women using the multiple-pass 24-h recall: a doubly labeled water study. *Journal of the American Dietetic Association* 1998; **98**: 1136–40.
27 Kroke A, Klipstein-Grobusch K, Voss S, Mosender J, Thielecke F, Noack R, *et al*. Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *American Journal of Clinical Nutrition* 1999; **70**: 439–47.
28 Bathalon GP, Tucker KL, Hays NP, Vinken AG, Greenberg AS, McCrory MA, *et al*. Psychological measures of eating behavior and the accuracy of 3 common dietary assessment methods in healthy postmenopausal women. *American Journal of Clinical Nutrition* 2000; **71**: 739–45.
29 Tran KM, Johnson RK, Soultanakis RP, Matthews DE. In-person vs telephone-administered multiple-pass 24-hour recalls in women: validation with doubly labeled water. *Journal of the American Dietetic Association* 2000; **100**: 777–80, 783.
30 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. *European Journal of Clinical Nutrition* 2000; **54**: 281–7.
31 Freedman LS, Midthune D, Carroll RJ, Krebs-Smith S, Subar AF, Troiano RP, *et al*. Adjustments to improve the estimation of usual dietary intake distributions in the population. *Journal of Nutrition* 2004; **134**: 1836–43.

Appendix – Questions asked on Multifactor Screener

How many times per day, week, or month did you usually eat (or drink):

…cold cereals?

…milk, either to drink or on cereal?

What kind of milk did you usually use?

Response categories are: whole milk, 2% fat, 1% fat, 0.5% fat, non-fat or skimmed, did not drink milk in past month.

…bacon or sausage, not including low-fat, light or turkey varieties?

…hot dogs made of beef or pork?

…whole-grain bread including toast, rolls and in sandwiches? Whole-grain breads include whole wheat, rye, oatmeal and pumpernickel.

…100% fruit juice such as orange, grapefruit, apple and grape juices? Do not count fruit drinks such as Kool-Aid, lemonade, cranberry juice cocktail, Hi-C and Tang.

…fruit? Count fresh, frozen or canned fruit. Do not count juices.

…regular-fat salad dressing or mayonnaise, including on salad and sandwiches? Do not include low-fat, light or diet dressings.

…lettuce or green leafy salad, with or without other vegetables?

…French fries, home fries or hash brown potatoes?

…other white potatoes? Count baked potatoes, boiled potatoes, mashed potatoes and potato salad. Do not include yams or sweet potatoes.

…cooked dried beans, such as refried beans, baked beans, bean soup and pork and beans?

…other vegetables? Count any form of vegetable: raw, cooked, canned or frozen. Do not count: lettuce salads, white potatoes, cooked dried beans, rice.

…any kind of pasta? Count spaghetti, noodles, macaroni and cheese, pasta salad, rice noodles, soba and any other kind of pasta.

…peanuts, walnuts, seeds or other nuts? Do not include peanut butter.

…regular-fat potato chips, tortilla chips or corn chips? Do not include low-fat chips.

Frequency response categories are: never, 1–3 times last month, 1–2 times per week, 3–4 times per week, 5–6 times per week, 1 time per day, 2 times per day, 3 times per day, 4 or more times per day.