American Journal of Epidemiology

Vol. 157, No. 8
Printed in U.S.A.
DOI: 10.1093/aje/kwg025

PRACTICE OF EPIDEMIOLOGY

Measuring Dietary Change in a Diet Intervention Trial: Comparing Food Frequency Questionnaire and Dietary Recalls

Cynthia A. Thomson[1], Anna Giuliano[1], Cheryl L. Rock[2], Cheryl K. Ritenbaugh[3], Shirley W. Flatt[2], Susan Faerber[2], Vicky Newman[2], Bette Caan[4], Ellen Graver[1], Vern Hartz[1], Robin Whitacre[1], Felicia Parker[1], John P. Pierce[2], and James R. Marshall[1]

[1] University of Arizona, Arizona Cancer Center, Tucson, AZ.
[2] University of California, San Diego, La Jolla, CA.
[3] Kaiser Center for Health Research, Portland, OR.
[4] Kaiser Permanente Northern California, Oakland, CA.

Received for publication April 3, 2002; accepted for publication October 17, 2002.

Measurement of dietary change was assessed in a systematic quota subsample (n = 397) of women recruited into the Women's Healthy Eating and Living Study between 1996 and 1998, a multicenter, randomized dietary intervention trial among breast cancer survivors. Women from the intervention and comparison arms completed the Arizona Food Frequency Questionnaire (AFFQ) and 24-hour dietary recalls at baseline (prerandomization) and at year 1 (postrandomization). Both dietary measurement methods demonstrated significant changes in intake of key intervention-associated nutrients at year 1 in the intervention group subjects compared with minimal or no change in the comparison group subjects. The reliability of the AFFQ and recalls was measured in the comparison group and showed correlations of 0.63 and 0.43, respectively. Both instruments captured differences in dietary intake associated with the diet intervention. These results demonstrate the utility of using a multimode, multimethod approach (AFFQ and 24-hour dietary recalls) to measure differences in self-reported dietary intake over time as shown in this dietary intervention trial being conducted among breast cancer survivors.

breast neoplasms; clinical trials; food habits; questionnaires; recall

Abbreviations: AFFQ, Arizona Food Frequency Questionnaire; CSFII, Continuing Survey of Food Intake by Individuals; FFQ, food frequency questionnaire; WHEL, Women's Healthy Eating and Living.

The ability of dietary assessment instruments to measure dietary change in the context of diet intervention trials being conducted in an effort to reduce risk for primary cancer or cancer recurrence is of central importance to cancer prevention research (1). Interpretation of study outcomes in relation to diet will be dependent on not only the reliability and validity of the instrument(s) used but also the capacity of the instrument to capture changes in intake over the time course of the study. Semiquantitative food frequency questionnaires (FFQs) are used to estimate dietary intake when a description of habitual diet is desired. When the dietary intervention trial involves a large number of subjects who will be followed longitudinally with multiple measurement time points, the FFQ provides a cost-effective method to examine nutrient-disease associations. However, concerns remain regarding measurement error associated with FFQs (2–4). Despite these concerns, the FFQ remains the most commonly used dietary measurement method for dietary intervention studies. Dietary recalls may also be used in dietary intervention trials, particularly when the goal is to measure more recent dietary intake rather than an estimate of usual intake over a longer time span. Recalls have the advan-

Reprint requests to Dr. Cynthia A. Thomson, Department of Nutritional Sciences, University of Arizona, Arizona Cancer Center, 1515 N. Campbell Avenue, Tucson, AZ 85724 (e-mail: cthomson@u.arizona.edu).

TABLE 1. Demographic and other selected characteristics of the study cohort and the remaining Women's Healthy Eating and Living Study participants (1996–1999)

Demographic characteristic*	Validation study cohort (n = 397)	Remaining WHEL cohort (n = 2,693)
Age (mean years)	52.7 (9.3)†	53.3 (8.9)
Ethnicity (%)		
White	84.9	85.4
Hispanic	4.8	5.4
African American	4.0	3.8
Asian/Pacific Islands	2.8	4.0
Other	3.5	1.3
Education (%)		
College/postcollege	53.7	54.4
Some college/post-high school	37.8	32.8
High school graduate	8.3	11.9
Less than high school	0.3	0.9
Breast cancer stage (%)		
I	40.3	38.2
II	56.2	56.6
IIIA	3.5	5.2
Body mass index (mean kg/m^2)	27.4 (6.3)	27.2 (6.1)

* No significant difference between validation cohort and remaining Women's Healthy Eating and Living (WHEL) cohort for any descriptive characteristic (chi-square analysis, $p > 0.05$).

† Numbers in parentheses, standard deviation.

tage of providing more specific information regarding intake than do FFQs, including the ability to more accurately quantify intake of food groups and to describe intake patterns. Depending on the nutrients of interest, the number of recalls and time frame for record collection can add to costs in terms of time, participant burden, and money. Response rates are also likely to decline when subjects must be contacted numerous times. Issues such as questionnaire format, participant motivation, perceived study burden, and repeated administration of the same instrument can each influence our ability to gather reliable and valid self-reported dietary intake data. Despite the numerous limitations, self-reported dietary intake remains the principal method for assessing diet. Although dietary measurement using this approach is not precise, the ability of dietary instruments to capture differences in intake over time is worthy of evaluation, given its importance to the interpretation of results from dietary intervention trials.

The purpose of this study was to compare reported change in dietary intake among breast cancer survivors involved in a diet intervention trial using two dietary measurement methods, the Arizona Food Frequency Questionnaire (AFFQ) and repeat 24-hour telephone recalls. The responsiveness and sensitivity of measurements of dietary change have been generally understudied (5, 6). The Women's Healthy Eating and Living (WHEL) diet intervention trial provides an opportunity to demonstrate changes in dietary intake as measured by repeated administration of the AFFQ and repeated dietary recalls in the context of an ongoing clinical trial (7).

MATERIALS AND METHODS

This comparison study was conducted within the context of the WHEL Study. The WHEL Study is a randomized, controlled trial of diet change, conducted among women treated for stage I, II, or IIIA breast cancer. The intervention goal consists of substantial dietary change, implementing a plant-based, reduced-fat diet. The object of the intervention is to test the effect of diet on the risk for recurrence of breast cancer (8). Dietary intake was measured using the AFFQ and by repeated dietary recalls. Our goal in this study was to assess the change in dietary intake among participants in a dietary intervention trial with two commonly used instruments, an FFQ and dietary recalls.

Study population

This substudy used a systematic quota sample of the first 400 breast cancer survivors who were randomized into the WHEL Study and who completed diet assessment using both instruments at baseline and 1 year. The WHEL Study design has been previously described (8). Within each study group (intervention and comparison), we selected the first 100 women with complete data who were not more than 50 years of age and the first 100 women who were 51 or more years

TABLE 2. Baseline and year 1 mean daily energy and nutrient intakes by dietary intervention group for Women's Healthy Eating and Living substudy subjects (1996–1999) estimated using the Arizona Food Frequency Questionnaire

Nutrient	Comparison (n = 200)		Intervention (n = 197)	
	Baseline	Year 1	Baseline	Year 1
Energy (kcal)	1,900 (799)*	1,787 (755)†	1,954 (827)	2,057 (870)†
Carbohydrate (g)	288 (137)	276 (127)†	297 (136)	344 (154)†
Protein (g)	70 (31)	70 (31)	73 (33)	77 (36)†
Energy from fat (%)	27 (6)	25 (6)†	27 (6)	22 (6)†
Fiber (g)	25 (14)	26 (16)	27 (16)	37 (20)†
Folate (μg)	375 (209)	444 (244)†	404 (245)	534 (322)†
Vitamin C (mg)	180 (127)	189 (112)†	206 (141)	275 (214)†
Vitamin A (retinol equivalent)	1,889 (2,740)	1,998 (2,279)	2,185 (2,891)	8,885 (6,674)†
α-Carotene (μg)	1,292 (2,287)	1,416 (1,836)	1,621 (2,289)	6,476 (4,912)†
β-Carotene (μg)	5,993 (6,822)	6,513 (6,526)	6,874 (7,495)	22,349 (15,470)†
Lutein (μg)	3,900 (2,934)	4,641 (4,527)	4,291 (3,871)	7,625 (6,014)†
Lycopene (μg)	9,255 (7,295)	8,615 (5,912)	9,866 (8,203)	12,246 (8,484)†
β-Cryptoxanthin (μg)	244 (261)	264 (255)†	291 (305)	360 (339)†

* Numbers in parentheses, standard deviation.

† Significant ($p < 0.05$) difference within dietary intervention group between baseline and year 1 (based on paired t test of log-transformed data).

of age. Women who had changed their tamoxifen usage over this period were excluded.

Dietary intervention

Subjects in the WHEL Study were randomly assigned to one of two dietary intervention groups, stratified by stage of disease, age, and clinical site. The comparison group was advised to consume a daily diet currently recommended for cancer prevention (five servings of vegetables and fruit per day, 20 g of fiber, and not more than 30 percent of energy from fat). The intensive diet intervention group was provided daily dietary goals of five servings of vegetables, three servings of fruit, 16 fluid ounces (0.473 liter) of vegetable juice, 30 g of fiber, and from 15 percent to 20 percent of energy from fat. A telephone-based diet-counseling protocol, 12 group meetings/cooking classes, and printed materials were used to achieve intervention group study goals. The comparison group attended four cooking classes unrelated to the dietary intervention and was also provided the standard self-help material available from governmental agency sources.

Study design and data collection

The 400 women enrolled in this study were mailed an AFFQ prior to their initial clinic visit along with written instructions on how to complete the form. The women then attended a baseline clinic visit during which they returned their completed AFFQs, which were reviewed for completeness and reasonableness of response. Subjects were then given instructions on the 24-hour recall process including general guidelines on serving size estimation (including pictures of commonly consumed portions), the need to avoid written record keeping, and the telephone approach to data collection. Over the course of the next 2–3 weeks the women completed four 24-hour telephone recalls conducted by trained dietary assessors at the University of California, San Diego. During their next clinic visit, women were randomized to either the intervention or the comparison group. After 1 year, the AFFQ and recalls were repeated, replicating the time sequence at baseline.

Arizona Food Frequency Questionnaire. The AFFQ is a 153-item, semiquantitative, scannable questionnaire that elicits information regarding the usual foods consumed and the frequency of consumption, using age- and gender-specific estimates of portions estimated as small, medium, or large. The original AFFQ was a modification of the Block-National Cancer Institute Health Habits and History Questionnaire (9). Earlier versions of the AFFQ have been calibrated against diet records (10). The version of the AFFQ used in this clinical trial was revised in 1995 upon completion of the WHEL Feasibility Study to ensure that the food items and dietary practices associated with the prescribed dietary intervention were adequately represented in the instrument, including an expanded list of fruits and vegetables, inclusion of vegetable juice and modification of food items, and preparation for fat content. Participants were instructed to describe their dietary intake patterns during the previous 3 months. The form required an estimated 40 minutes to complete in this study population.

The AFFQ analysis program is a four-module system of programs written in Stata software (Stata Corporation, College Station, Texas) that reduces data from scanned questionnaires to individual nutrients per day. The database used to quantify nutrient intake from the AFFQ is derived from

TABLE 3. Baseline and year 1 mean daily energy and nutrient intakes by dietary intervention group for Women's Healthy Eating and Living substudy subjects (1996–1999) estimated using dietary recalls

Nutrient	Comparison (n = 200)		Intervention (n = 197)	
	Baseline	Year 1	Baseline	Year 1
Energy (kcal)	1,761 (422)*	1,636 (384)†	1,729 (364)	1,663 (346)†
Carbohydrate (g)	245 (62)	231 (60)†	244 (59)	270 (66)†
Protein (g)	69 (19)	68 (19)	67 (16)	67 (17)
Energy from fat (%)	28 (7)	30 (7)†	28 (7)	21 (7)†
Fiber (g)	22 (9)	22 (9)	21 (8)	30 (10)†
Folate (μg)	322 (129)	319 (128)	311 (119)	450 (156)†
Vitamin C (mg)	149 (82)	147 (72)	150 (81)	235 (103)†
Vitamin A (retinol equivalent)	1,471 (919)	1,353 (900)†	1,428 (1,114)	4,578 (2,697)†
α-Carotene (μg)	1,522 (2,290)	1,467 (2,230)	1,539 (2,404)	11,794 (9,586)†
β-Carotene (μg)	5,364 (5,127)	4,967 (4,845)	5,136 (5,678)	26,618 (18,526)†
Lutein (μg)	3,210 (2,226)	3,208 (2,359)	3,016 (1,849)	5,702 (3,704)†
Lycopene (μg)	6,164 (6,016)	5,655 (5,484)	5,913 (4,794)	16,329 (14,251)†
β-Cryptoxanthin (μg)	253 (595)	233 (258)†	202 (187)	382 (582)†

* Numbers in parentheses, standard deviation.

† Significant (p < 0.05) difference within dietary intervention group between baseline and year 1 (based on paired t test of log-transformed data).

the US Department of Agriculture Food Composition Database and the Nutrient Database for Standard Reference (versions 11–13) (11) and from the Continuing Survey of Food Intake by Individuals (1994–1996, 1998; CSFII) (12). Briefly, a standardized, systematic approach for assigning nutrient values to the AFFQ food lines has been instituted. This method is based on the determination of national consumption patterns for each AFFQ food line, in which all possible foods from the CSFII (1994–1996) that could be included on each AFFQ food line are identified. Based on the relative consumption of foods derived from the CSFII national consumption data, an appropriate nutrient mix is then calculated and assigned for that AFFQ food line. Many nutrients not included in the CSFII database were added from the Nutrient Database for Standard Reference, resulting in a final data set that includes 82 nutrients and 21 food groups. Carotenoids were updated from the US Department of Agriculture-Nutrition Coordinating Center Carotenoid Database (1998) (11). Separate databases are maintained for questionnaires completed before and after January 1, 1998 (the year that the Food and Drug Administration-mandated folic acid fortification of cereal-grain products was enforced).

Adjustments have been made to the nutrient values on the basis of methodological research from Subar et al. (13). All AFFQs were double scanned to ensure the accuracy of data entry, followed by computerized checks for outliers and reasonableness of data. Any FFQs missing more than 10 items were not included in the final analysis. For this study, there were three FFQs that did not meet this criterion, all of which were collected from intervention group subjects. Log-transformation was used to normalize skewed data; no cases were removed because of extreme values.

Repeated 24-hour dietary recalls. Each study participant provided four 24-hour dietary recalls at baseline and at 1 year, including two week-day recalls and two weekend recalls over a 3-week period. Data were collected by trained dietary assessors, using a scheduled telephone interview, with data collection centralized at the WHEL Study Coordinating Center at the University of California, San Diego. Dietary assessment staff members were trained regarding the 24-hour dietary recall data collection procedure by a registered dietitian experienced in recall data collection and analysis. Assessors were blinded to the intervention or comparison group assignment of the subjects. Nutrient calculations were performed using Nutrition Data System for Research software (version 4.03), developed by the Nutrition Coordinating Center, University of Minnesota, Minneapolis, Minnesota, for the Food and Nutrient Database 31, released in November 2000 (14).

Statistical methods

Descriptive analyses were used to describe the study sample, and chi-square and/or t-test analysis was used to compare the sociodemographic characteristics of the study population in this substudy and the remaining study subjects enrolled in the WHEL Study. Means and standard deviations were calculated for crude (unadjusted) nutrients from the AFFQ and dietary recalls. The four recalls were averaged for the analysis. The distribution of response variables was examined for normality prior to modeling, resulting in the log-transformation of all dietary variables. Paired t-test analysis was used to examine differences in group mean intake at baseline versus year 1; this analysis was performed for dietary intake reported by the AFFQ and separately for

TABLE 4. Reliability of the Arizona Food Frequency Questionnaire and 24-hour recall instruments among comparison group Women's Healthy Eating and Living substudy participants (1996–1999): Spearman's rank correlation coefficients for the baseline and year 1 energy and nutrient intakes (*n* = 200)

Nutrient	AFFQ*	24-hour recalls
Energy (kcal)	0.668	0.497
Carbohydrate (g)	0.646	0.534
Protein (g)	0.656	0.438
Energy from fat (%)	0.469	0.471
Fiber (g)	0.708	0.599
Folate (μg)	0.678	0.454
Vitamin C (μg)	0.633	0.426
Vitamin A (RE*)	0.650	0.453
α-Carotene (μg)	0.582	0.345
β-Carotene (μg)	0.639	0.396
Lutein (μg)	0.621	0.464
Lycopene (μg)	0.568	0.234
β-Cryptoxanthin (μg)	0.628	0.330
Mean	0.627	0.434

* AFFQ, Arizona Food Frequency Questionnaire; RE, retinol equivalent.

dietary intake reported on repeated recalls. Spearman's rank correlation coefficient was used to describe the reproducibility of the data for the comparison group at baseline and year 1 using either the AFFQ or repeated recalls. In most cases, the mean estimates of dietary intake variables did not significantly change in the comparison group between the two time points. Therefore, the comparison group was used to estimate instrument reliability. Finally, responsiveness was used as a measure of an evaluation of the sensitivity of the instrument to change, as previously described in the literature (15, 16). Responsiveness, as defined by Kristal et al. (16), is the observed intervention effect divided by its standard deviation, where the intervention effect is calculated as the mean change in the intervention group minus the mean change in the comparison group. Here, the responsiveness of dietary assessment measures for the key dietary intervention target variables included the percentage of energy intake from fat, dietary fiber, α- and β-carotene, and folate. Statistical significance for all analyses was set at $p < 0.05$.

RESULTS

Complete data for dietary intake were available for 397 women, including 200 from the comparison group and 197 from the intervention group. Table 1 compares the demographic and other selected characteristics for this subgroup with those of the remainder of subjects enrolled in the WHEL Study. Participants were predominantly non-Hispanic White, well-educated women, with ages ranging from 28 to 74 years. They had a mean body mass index of 27.4 kg/m^2. Chi-square analysis indicated that subjects in this substudy were not significantly different from the remainder of the WHEL Study cohort with regard to any of the descriptive characteristics. In addition, comparison and intervention participants were equivalent with respect to sociodemographic characteristics (data not shown).

Means and standard deviations for unadjusted nutrient intake, at baseline and year 1 by study group, are presented in tables 2 (AFFQ) and 3 (24-hour recalls). No significant differences in baseline intake by dietary group assignment were observed regardless of the dietary methodology used to assess intake. Differences in self-reported nutrient intake between baseline and year 1 for intervention group subjects were observed for energy, carbohydrate, total fat, fiber, folate, vitamins A and C, α-carotene, β-carotene, lutein, lycopene, and β-cryptoxanthin using either dietary measurement method. However, only the AFFQ suggested a difference in protein intake. Significant differences in energy, carbohydrate, percentage of energy from fat, and β-cryptoxanthin were also seen in the comparison group using either diet assessment method, while only the AFFQ showed significant differences in folate and vitamin C intake in this group and only recalls showed a decline in vitamin A. Differences may be attributed to differences in the nutrient/carotenoid databases used for the AFFQ and recalls.

Table 4 presents the reproducibility or reliability of the AFFQ and repeated recalls using baseline and year 1 AFFQ dietary data from only the comparison group. The average correlation across nutrients for the AFFQ was 0.63, while for the repeated recalls it was 0.43. Correlation coefficients for nutrients and foods associated with the dietary intervention were 0.47 for the percentage of energy from fat for both the AFFQ and recalls and 0.71 and 0.60 for dietary fiber, 0.68 and 0.45 for folate, 0.58 and 0.35 for α-carotene, and 0.64 and 0.40 for β-carotene for the AFFQ and recalls, respectively. Thus, the repeated recall data had an overall lower reproducibility than did the AFFQ.

Finally, table 5 represents the responsiveness of the two dietary assessment measures for key dietary intervention target variables—the percentage of energy from fat, fiber, and α- and β-carotene and folate (as indicators of fruit and vegetable intake). Both dietary instruments demonstrated minimal change in diet among participants in the comparison group for the percentage of energy from fat and fiber. However, for α- and β-carotene and folate intakes, the AFFQ supported increased intake over 12 months in both intervention and comparison group subjects, while the recalls demonstrated an increase for intervention group subjects and a decrease for comparison group subjects. The point estimate for the intervention effect was somewhat greater for the recalls on each of the key dietary measures.

DISCUSSION

This study compared two diet assessment instruments, the AFFQ and repeated dietary recalls, in the context of a large diet intervention trial focused on promoting a major diet change among women previously treated for breast cancer. Historically, the accuracy of FFQ-based dietary data compared with dietary recalls or records has been evaluated principally in observational studies (17–21), and FFQs have

TABLE 5. Responsiveness* of the Arizona Food Frequency Questionnaire and recall in measuring intake of key dietary intervention target variables among Women's Healthy Eating and Living substudy subjects (1996–1999)

Variable, arm, and time	Instrument	
	AFFQ†	24-hour recall
Energy from fat (%)		
Intervention		
Baseline	26.5 (6.5)‡	28.0 (6.7)
Year 1	21.9 (5.7)	21.2 (6.8)
Change	–4.6 (6.7)	–6.8 (7.0)
Comparison		
Baseline	26.7 (6.4)	28.0 (6.8)
Year 1	25.3 (6.2)	27.0 (7.0)
Change	–1.4 (6.4)	–1.0 (7.1)
Intervention effect	–3.2 (9.3)	–5.8 (10.0)
Responsiveness	0.34	0.58
Fiber (g)		
Intervention		
Baseline	26.7 (15.9)	21.0 (8.3)
Year 1	37.5 (20.3)	29.5 (10.0)
Change	10.7 (18.0)	8.5 (9.4)
Comparison		
Baseline	25.2 (14.2)	21.8 (8.6)
Year 1	26.5 (16.3)	21.9 (8.8)
Change	1.3 (10.6)	0.1 (7.6)
Intervention effect	9.4 (20.9)	8.4 (12.1)
Responsiveness	0.45	0.69
α-Carotene (μg)		
Intervention		
Baseline	1,621 (2,289)	1,539 (2,404)
Year 1	6,476 (4,911)	11,794 (9,586)
Change	4,855 (5,112)	10,255 (9,571)
Comparison		
Baseline	1,292 (2,287)	1,522 (2,290)
Year 1	1,416 (1,836)	1,467 (2,230)
Change	124 (1,728)	–54 (2,680)
Intervention effect	4,731 (5,396)	10,309 (9,939)
Responsiveness	0.88	1.05
β-Carotene (μg)		
Intervention		
Baseline	6,874 (7,495)	5,136 (5,678)
Year 1	22,349 (15,470)	26,618 (18,526)
Change	15,599 (15,463)	21,482 (18,517)
Comparison		
Baseline	5,993 (6,822)	5,364 (5,127)
Year 1	6,513 (6,526)	4,967 (4,845)
Change	521 (5,330)	–402 (5,652)
Intervention effect	15,078 (16,356)	21,884 (19,360)
Responsiveness	0.92	1.13
Folate (μg)		
Intervention		
Baseline	404 (245)	311 (119)
Year 1	534 (322)	450 (156)
Change	151 (286)	139 (152)
Comparison		
Baseline	375 (209)	322 (129)
Year 1	444 (244)	319 (128)
Change	69 (183)	–2 (130)
Intervention effect	82 (340)	141 (200)
Responsiveness	0.24	0.71

Table continues

TABLE 5. Continued

* Responsiveness is defined as the observed intervention effect divided by its standard deviation (Kristal et al. (16)). The observed intervention effect is defined as the mean change in the tabulated variable for the intervention group minus the mean change in the comparison group. The standard deviation of this difference is the square root of the sum of the variances for change in the intervention and comparison arms.

† AFFQ, Arizona Food Frequency Questionnaire.

‡ Numbers in parentheses, standard deviation.

been validated in healthy populations. The results from this study, conducted in a subpopulation of women participating in the WHEL Study, demonstrate that the AFFQ is a useful approach for measuring change in dietary intake in the context of a randomized, controlled dietary intervention trial among breast cancer survivors. Both of the dietary methodologies were shown to be responsive to the intervention efforts, with a similar degree of responsiveness for key target nutrients.

In this study, both diet assessment instruments reflected changes in dietary intake for intervention group subjects over time that corresponded with the dietary intervention. These included changes in the self-reported intakes of fat, fiber, folate, vitamin A, vitamin C, α-carotene, β-carotene, lutein, lycopene, and β-cryptoxanthin. Both methods were able to detect significant changes over time (year 1 compared with baseline intakes) in energy, carbohydrate, fat, fiber, folate, vitamins A and C, and all five carotenoids for the intervention group. The reliability of responses to the AFFQ (table 4) was likely reduced because even the comparison group subjects had made some changes in their usual dietary intakes prior to study entry (22). Other factors that may have altered the reliability of the AFFQ instrument include the uniqueness of the study sample, the use of National Cancer Institute guidelines as the comparison group diet, and the time interval of 1 year between measurements. Breast cancer survivors consenting to their participation in a long-term dietary study are motivated by the opportunity to develop what they perceive to be healthier dietary habits. Thus, assignment to the comparison group did not preclude them from attempting to improve their diet. Change in dietary intake demonstrated by comparison group subjects included significant reductions in fat intake and increased intakes of vitamin A, vitamin C, and β-cryptoxanthin (table 2). In a study of the reliability of the American Cancer Society dietary assessment instrument, median correlations of 0.68 for vitamins and minerals were observed in people not expected to have changed their diet over time (23). As expected, the reproducibility of FFQs is higher than for recalls (r = 0.63 vs. 0.43) because of the expected variability in intake measured on a day-to-day basis (recalls) versus a summary of intake for an extended time (FFQs). This is true despite analysis of energy correlations among the four recalls collected within a 3-week period that showed no decrease in energy intake among the four recalls (data not shown).

Although not presented here, correlation coefficients for dietary variables measured by the AFFQ versus recalls among intervention group subjects at year 1 were found to be slightly lower but similar to correlations reported in the literature for healthy adults (17, 19, 20, 24–28); however, they were not as highly correlated as the National Cancer Institute Questionnaire (29). Correlations between the two dietary instruments for nutrients associated with the dietary intervention were each significant and were stronger than reported previously for total fat (17, 20, 24–26) and comparable for fiber (18, 24, 25), vitamin A, and vitamin C (20, 24, 25). The facts that the AFFQ was not specifically developed using food records from a population of breast cancer survivors and that the overall number of foods listed may not have been extensive enough to describe habitual diet may have reduced our ability to adequately describe intake. Cognitive interviewing as used when collecting diet recall data and adaptation and expansion of the FFQ to include specific food items commonly consumed based on recall reports would likely improve agreement in future studies (28, 30, 31). It has been suggested that agreement between recalls and FFQs is somewhat limited, given that FFQs and recalls/records measure different aspects of a person's diet (current vs. usual intake) (32, 33); underreporting is common, particularly with recalls collected from overweight subjects, including breast cancer survivors (34); and the methodology for instrument administration is not the same (35). Despite these limitations, both instruments detected a change in intake over time, particularly with regard to intervention-associated nutrient consumption among intervention group subjects.

There is no "gold standard" for measuring dietary change, but Guyatt et al. (15) and Kristal et al. (16) have suggested using responsiveness as a measure of an instrument's sensitivity to dietary change. Using this approach (table 5), these data suggest that in our study the two methods produced similar values for changes in dietary behavior made by the intervention group. Kristal et al. also evaluated responsiveness in terms of the percentage of energy from fat and showed similar responsiveness for the FFQ and 4-day food records (responsiveness = 1.2 and 1.4, respectively). Buzzard et al. (36) showed that, overall, telephone recalls yielded lower energy and fat intake than did diet records among those participating in a low fat intervention trial. Measurement of responsiveness in our study indicates that the recalls were slightly more responsive than were the AFFQs; however, both instruments demonstrated responsiveness. The slightly lower responsiveness of the AFFQ may be related to the difficulty in sufficiently describing fat intake, particularly in a sample population consuming large amounts of low- or modified-fat food items. This same loss of specificity could also account for the lower responsiveness for α-carotene and β-carotene given the wide variety of fruits and vegetables promoted in the intervention versus a somewhat selective list recorded on the AFFQ, although a database difference might also contribute to this finding. Both instruments appeared to show responsiveness to the reported dietary change among intervention group participants.

This study indicates that both the AFFQ and 24-hour dietary recalls were able to capture significant differences in dietary intervention target nutrients and carotenoids over time among women previously treated for breast cancer and participating in a dietary intervention trial. However, there is wide variability in the dietary data collected. The reproducibility of the AFFQ was greater than recalls as is expected given the open-ended inquiry and broad range of food items potentially reported on the recalls versus the AFFQ. The application of these approaches in other dietary intervention trials will be dependent on the similarities and differences between the sample populations under study. Both instruments were able to demonstrate responsiveness to dietary intervention for the key nutrients and carotenoids associated with the intervention efforts. Given the unique contributions of both FFQs and recall data in describing dietary intake and given the wide variability in self-reported dietary data, a multimethod approach to measuring dietary change that includes more than a single measure of self-reported intake may be useful in future dietary intervention trials. When several independent measures support change in the same direction, we can feel confident that the desired change in eating pattern has occurred, thus ensuring a more precise interpretation of the final research outcomes.

ACKNOWLEDGMENTS

This study was started with the support of the Walton Family Foundation, and it continued with funding from National Cancer Institute grant CA69375 and by National Institutes of Health grants M01-R0079, M01-RR00827, and M01-R00070. This publication was also made possible in part by funds received from the Cancer Research Fund (grant 99-00548V-10147), under Interagency Agreement 97-12013 (University of California, Davis, contract 98-00924V) with the Department of Health Services Cancer Research Section.

The WHEL Study Group comprises the following: *WHEL Study coordinating center*: University of California, San Diego, Cancer Prevention and Control Program, San Diego, California: Dr. John P. Pierce (Principal Investigator), Dr. Cheryl L. Rock, Susan Faerber, Vicky Newman, Shirley W. Flatt, Sheila Kealey, Elizabeth Gilpin, Dr. Linda

Wasserman; *WHEL Study clinical sites*: Center for Health Research, Portland, Oregon: Dr. Cheryl Ritenbaugh, Dr. Mark Rarick; Kaiser Permanente Northern California, Oakland, California: Dr. Bette J. Caan, Dr. Lou Ferenbacher; Northern California Cancer Center, Union City, California: Dr. Marcia L. Stefanick, Dr. Robert Carlson; University of Arizona, Tucson and Phoenix, Arizona: Dr. James R. Marshall, Dr. Cynthia Thomson, Dr. James Warnecke; University of California, Davis, Davis, California: Dr. Ellen B. Gold, Dr. Mary N. Haan, Dr. Sidney Scudder; University of California, San Diego Cancer Center, San Diego, California: Dr. Vicky E. Jones, Dr. Kathryn A. Hollenbach; University of Texas M. D. Anderson Cancer Center, Houston, Texas: Dr. Lovell A. Jones, Dr. Richard Theriault.

REFERENCES

1. Willett W. Nutritional epidemiology. 2nd ed. New York, NY: Oxford University Press, 1998.
2. Thompson FE, Subar AF. Dietary assessment methodology. In: Coulston AM, Rock CL, Monsen ER, eds. Nutrition in the prevention and treatment of disease. San Diego, CA: Academic Press, 2001:3–30.
3. Spiegelman D, Schneeweiss S, McDermott A. Measurement error correction for logistic regression models with an "alloyed gold standard." Am J Epidemiol 1997;145:184–96.
4. Beaton GH. Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. Am J Clin Nutr 1994;59(suppl):253S–61S.
5. Carroll RJ, Feedman LS, Kipnis V. Measurement error and dietary intake. Adv Exp Med Biol 1998;445:139–45.
6. Briefel RR. Assessment of the US diet in national nutrition surveys: national collaborative efforts and NHANES. Am J Clin Nutr 1994;59(suppl):164S–7S.
7. Pierce J, Faerber S, Wright FA, et al. Feasibility of a randomized trial of a high-vegetable diet to prevent breast cancer recurrence. Nutr Cancer 1997;28:282–8.
8. Pierce JP, Faerber S, Wright FA, et al. A randomized trial of the effect of a plant-based dietary pattern on additional breast cancer events and survival: the Women's Healthy Eating and Living (WHEL) Study. Control Clin Trials 2002;23:728–56.
9. Block G, Woods M, Potosky A, et al. Validation of a self-administered diet history questionnaire using multiple diet records. J Clin Epidemiol 1990;43:1327–35.
10. Martinez ME, Marshall JR, Graver E, et al. Reliability and validity of a self-administered food frequency questionnaire in a chemoprevention trial of adenoma recurrence. Cancer Epidemiol Biomarkers Prev 1999;8:941–6.
11. Agricultural Research Service, US Department of Agriculture. USDA nutrient database for standard reference, releases 11–13. Beltsville, MD: Nutrient Data Laboratory, Beltsville Human Nutrition Research Center, 2001. (http://www.nal.usda.gov/fnic/foodcomp).
12. Agricultural Research Service, US Department of Agriculture. Design and operation: the continuing survey of food intakes by individuals and the diet and health knowledge survey, 1994–96. Washington, DC: US Department of Agriculture, 1997.
13. Subar AF, Midthune D, Kulldorff M, et al. Evaluation of alternative approaches to assign nutrient values to food groups in food frequency questionnaires. Am J Epidemiol 2000;152:279–86.
14. Schakel SF, Sievert YA, Buzzard IM. Sources of data for developing and maintaining a nutrient database. J Am Diet Assoc 1988;88:1268–71.
15. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis 1987;40:171–8.
16. Kristal AR, Beresford SAA, Lazovich D. Assessing change in diet-intervention research. Am J Clin Nutr 1994;59(suppl): 185S–9S.
17. Mannisto S, Virtanen M, Mikkonen T, et al. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. J Clin Epidemiol 1996;49:401–9.
18. Stram DO, Hankin JH, Wilkens LR, et al. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. Am J Epidemiol 2000;151:358–70.
19. Rimm EB, Giovannucci EL, Stampfer MJ, et al. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992;135:1114–24.
20. Pisani P, Faggiano F, Krogh V, et al. Relative validity and reproducibility of a food frequency dietary questionnaire for use in the Italian EPIC centres. Int J Epidemiol 1997;26(suppl): S152–60.
21. Jain M, Howe GR, Rohan T. Dietary assessment in epidemiology: comparison on food frequency and a diet history questionnaire with a 7-day food record. Am J Epidemiol 1996;143:953–60.
22. Thomson CA, Flatt SW, Rock CL, et al. Increased fruit, vegetable and fiber intake and lower fat intake reported among women previously treated for invasive breast cancer. J Am Diet Assoc 2002;102:801–8.
23. Flagg EW, Coates RJ, Calle EE, et al. Validation of the American Cancer Society Cancer Prevention Study II Nutrition Survey Cohort Food Frequency Questionnaire. Epidemiology 2000;11:462–8.
24. Pietinen P, Hartman AM, Haapa E, et al. Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. Am J Epidemiol 1988;128:655–66.
25. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51–65.
26. Larkin FA, Metzner HL, Thompson FE, et al. Comparison of estimated nutrient intakes by food frequency and dietary records in adults. J Am Diet Assoc 1989;89:215–23.
27. Kristal AR, Feng Z, Coates RJ, et al. Associations of race/ethnicity, education, and dietary intervention with the validity and reliability of a food frequency questionnaire: the Women's Health Trial Feasibility Study in Minority Populations. Am J Epidemiol 1997;146:856–69.
28. Tjonneland A, Overad K, Haraldsdottir J, et al. Validation of a semiquantitative food frequency questionnaire developed in Denmark. Int J Epidemiol 1991;20:906–12.
29. Subar AF, Thompson FE, Kipnis V, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. Am J Epidemiol 2001;154:1089–99.
30. Subar AF, Ziegler RG, Thompson FE, et al. Is shorter always better? Relative importance of questionnaire length and cognitive ease on response rates and data quality for two dietary questionnaires. Am J Epidemiol 2001;153:404–9.
31. Byers T. Food frequency dietary assessment: how bad is good enough? Am J Epidemiol 2001;154:1087–8.
32. Willett W. Invited commentary: a further look at dietary questionnaire validation. Am J Epidemiol 2001;154:1100–2.
33. Block G. Invited commentary: another perspective on food frequency questionnaires. Am J Epidemiol 2001;154:1103–4.

34. Caan BJ, Flatt SW, Rock CL, et al. Low-energy reporting in women at risk for breast cancer recurrence. Cancer Epidemiol Biomarkers Prev 2000;9:1091–7.
35. Caan B, Hiatt RA, Owen AM. Mailed dietary surveys: response rates, error rates, and the effect of omitted food items on nutrient values. Epidemiology 1991;2:430–6.
36. Buzzard IM, Faucett CL, Jeffrey RW, et al. Monitoring dietary change in a low-fat diet intervention study: advantages of using 24-hour dietary recalls vs food records. J Am Diet Assoc 1996; 96:574–9.