American Journal of Epidemiology
Vol. 148, No. 5
Printed in U.S.A.

Adaptation of a Food Frequency Questionnaire to Assess Diets of Puerto Rican and Non-Hispanic Adults

Katherine L. Tucker, Lisa A. Bianchi, Janice Maras, and Odilia I. Bermudez

To study issues of diet and health among Hispanic adults living in the northeastern United States, the authors adapted a version of the National Cancer Institute (NCI)/Block food frequency questionnaire. Foods that contributed to nutrient intake of Puerto Rican adults in the Hispanic Health and Nutrition Examination Survey (HHANES) were ranked to identify items to be added to the food list. Portion sizes were compared across HHANES and the Second National Health and Nutrition Examination Survey (NHANES II) to assess the adequacy of the assumed values. Within line items, frequencies of consumption of individual foods were ranked and these data were used to adjust the weighting factors within the database. To test the revised form, 24-hour recalls were collected from 90 elderly Hispanics and 35 elderly non-Hispanic whites. These data were coded into the original and revised food frequency forms and nutrient intake results were compared with recall results by paired *t*-test, and by Pearson and intraclass correlations. Added foods include plantains, avocado, mango, cassava, empanadas, and custard. Portion sizes differed significantly between HHANES and NHANES II, and were left open-ended. Estimated mean nutrient intakes and correlations with recall data were lower with the original versus the revised form. The authors conclude that the use in minority populations of food frequency questionnaires developed for the general population is likely to result in biased estimates of intake unless modifications are made in the questionnaires. *Am J Epidemiol* 1998;148:507–18.

diet; epidemiologic methods; ethnic groups; questionnaires

With increased interest in the study of diet and disease in minority populations, the question of the applicability of existing food frequency questionnaires is of concern (1–3). Instruments in major use include the Harvard/Willett food frequency questionnaire and the National Cancer Institute (NCI)/Block health habits and history questionnaire. The Willett instrument was developed and validated with data from nurses (4) and subsequently expanded and validated in a population of male health professionals (5). The Block instrument was developed with data from the Second National Health and Nutrition Examination Survey (NHANES II) (6–9).

Because the food frequency method depends on a predetermined list of food items, it may be appropriately used only with the population for which it was developed or subsequently validated. Often, a comparison of dietary risk across sub-populations is of central interest. There is, therefore, a need for instruments which will capture the diets of differing groups without differential bias in ability to describe and rank individuals with regard to food and nutrient intake. In this paper, we describe the process we followed in order to develop an instrument for the Massachusetts Hispanic Elders Study, where diets of Hispanics, primarily of Puerto Rican origin, are assessed along with those of neighborhood matched non-Hispanic white subjects.

MATERIALS AND METHODS

Instrument development

The development of a food frequency questionnaire that can simultaneously provide valid measures of diet for Puerto Rican and non-Hispanic white subjects poses several problems. First, the Puerto Rican diet differs significantly from what may be considered dominant dietary patterns in the United States and includes foods not on any of the major existing instruments. Also of concern are the assumptions used for food item weights within the nutrient database. A single line item in the questionnaire often includes

Received for publication January 15, 1997 and accepted for publication February 19, 1998.

Abbreviations: HCRC, Hutchinson Cancer Research Center; HHANES, Hispanic Health and Nutrition Examination Survey; NCI, National Cancer Institute; NDS, Nutrition Data System (Minnesota); NHANES II, Second National Health and Nutrition Examination Survey; USDA, US Department of Agriculture.

From the Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA.

Reprint requests to Dr. Katherine L. Tucker, Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, 711 Washington Street, Boston, MA 02111.

multiple foods, which are generally weighted by frequency of use in the population (e.g., apples are weighted more than applesauce or pears in that line item of the Block questionnaire) in order to obtain nutrient estimates. These relative weights may not translate appropriately to other population groups. In addition, typical portion sizes differ for major food contributors to intake.

We used the Hutchinson Cancer Research Center (HCRC) adaptation of the NCI/Block food frequency form as a base for further adaptation. The HCRC version of the Block is written in SAS (10), is linked with the Minnesota Nutrient Data System (NDS) (11), and allows complete flexibility in adaptation (12, 13). Following the methodology originally used by Block (1986) with NHANES II data, we analyzed 24-hour recall data from the Puerto Rican subset (aged 19–74 years) of the Hispanic Health and Nutrition Examination Survey (HHANES) (14) to define foods that were major contributors to intake for the following nutrients: energy, fat, saturated fat, carbohydrate, protein, cholesterol, vitamins (A, C, E, thiamin, riboflavin, niacin, B6, B12, and folate), and minerals (calcium, magnesium, iron, and zinc). Because there are several thousand individual food items in the HHANES database, with many variations of the same food type, foods were coded into food categories based on the existing food frequency food list. Foods that could not be accommodated into this list were assigned to additional categories. Sample weights, provided within the HHANES data set, were used to adjust for nonresponse (14). The contribution of each food category to total population intake of each nutrient was ranked, using the Rank procedure in SAS (10). Foods that contribute 0.5 percent or more to the intake of energy or any single nutrient were added to the food list, if not already included. No foods already on the HCRC/Block questionnaire were removed.

Revision of the database

A second potential source of error may result from assumptions on the weighting of food items in the nutrient database. Foods are grouped on the questionnaire, based on similarity of food category and nutrient content. Each line item in the food frequency form contains either a single food, or multiple similar foods. The nutrient data for each line item reflects a weighted average of this combination of foods. In the HCRC questionnaire, these sub-foods are weighted according to their relative use in the general population. It is likely that relative use may differ in minority populations. We therefore examined the relative frequencies of intake of foods in the Puerto Rican adult sample of HHANES within each food frequency line and compared these with existing analysis weights. When the frequencies of intake differed, weighting assumptions in the nutrient database were revised to an intermediate position by averaging the two percentage contributions to total frequency. Foods that contributed less than 5 percent to the frequency of consumption within line item for either group were not included in the nutrient database. Because differences in food preparation may also affect results, foods prepared with very different recipes were considered and included in these weights separately. If fat or other nutrient composition differed and the food was commonly consumed by the Puerto Rican population, it was removed to a separate line item or an adjustment question about preparation was included.

Nutrient data for the final listing of food and subfood items were updated from the HCRC database, using the Minnesota Nutrient Data System (NDS) (11). Due to its affiliation with the HHANES dietary data collection and processing, the NDS contains many Hispanic foods, including those used by Puerto Ricans. All of the most common foods required by our final food list were available directly, or through the ingredients of a recipe. If not already in the NDS, recipes for Hispanic foods were obtained from Puerto Rican cookbooks and verified and/or adapted with data from 24-hour recalls administered locally to Hispanic elders.

Determination of methodology for portion size assessment

A third potential source of error in the estimation of intake for a differing group, and particularly in comparing intake across groups, is the usual portion size reported. In the HCRC questionnaire, portion size options are given as small, medium, and large—with medium set as the median gram weight of portion sizes in NHANES II, with 50 percent of the medium defined as small and 150 percent as large (15). We compared portion sizes for foods recorded by 24-hour recall by Puerto Rican adults in HHANES with those in NHANES II using *t*-tests and frequency plots in SAS. The distributions of portion size intake were usually skewed, and log-transformed variables were used for two sample *t*-tests.

Because differences did exist, we decided to leave the portion sizes open-ended. With the aid of three-dimensional food models (NASCO, Fort Atkinson, Wisconsin) and common household utensils, we asked subjects to describe their usual portion size. This required the development of a computer program for portion size entry which allowed the use of multiples of differing measures, such as cups, tablespoons, slices, units, and so on. We adapted the HCRC form and program for optical scanning of intake frequency and adjustment questions. This information was then linked to the food portion entry program and to our revised nutrient database, by building on and adapting from the SAS program originally written at the HCRC.

Testing and calibration of the revised instrument

To confirm the completeness of the food list for Hispanic subjects in Massachusetts, and to perform an initial calibration of the questionnaire, we collected 24-hour dietary recalls from 90 Hispanic men and women (50 percent Puerto Rican, 39 percent Dominican, and 11 percent other Hispanics (Cuban, Central, and South American)) aged 52–91 years, who were participants in senior day care programs in Boston and Lawrence, Massachusetts. Additional recalls were collected from 35 non-Hispanic white elders aged 60–88 years, who lived in the same communities as the Hispanic elders. We then coded the data from these recalls into both the original HCRC/Block food frequency form and the newly developed form. In addition, we analyzed the data coded into the new form with both open-ended portion sizes and with standard portion sizes. The latter were set as small, medium, or large, using the original portions in the HCRC program. For foods not previously on the questionnaire, we based portion sizes on data from HHANES. Where the median portion size from HHANES differed greatly from that in NHANES II, a new portion was assigned as the average of the HCRC medium portion and the median portion in the HHANES data. Nutrient data for analysis of all versions of the food frequency form, as well as for the 24-hour recall analysis were from the same NDS nutrient database, version 25, 1995 (11).

Intakes of energy and key nutrients, estimated by 24-hour recall, were compared with those obtained from each of the three food frequency methods—the original HCRC questionnaire, the revised questionnaire with fixed portion sizes, and the revised questionnaire with open-ended portion sizes—using paired *t*-tests in SAS (10). Where the distribution of the differences across methods differed significantly from normality, significance levels were assessed using the Sign test in SAS (10). In addition to comparing the mean intakes obtained with differing versions of the questionnaire with the 24-hour recall, we compared the means across ethnic groups (Hispanic vs. non-Hispanic white) within questionnaire type for each of the nutrients (log-transformed due to skewed distributions), using the General Linear Models (GLM) procedure in SAS (10) with adjustment for age and sex. These analyses were also repeated with additional adjustment for total energy intake. Finally, estimates of energy intake from each of the three questionnaires were regressed on the 24-hour recall estimates to examine correspondence across the distribution.

To test validity of rankings, both Pearson and intraclass correlations were examined across distributions of log-transformed nutrient variables obtained from the 24-hour recall, and from each of the food frequency forms, using SAS (10). Initial results from these analyses were examined for content validity, and led to review of assumptions and programs for correction of errors in the database and programs. This process did not lead to changes in the food list, but did affect some of the recipes and weights for foods within line items of the questionnaire. For example, we originally used cookbook recipes for dishes such as chicken with rice, but found that the actual recipes that were used were much simpler. In addition, we found that the weight of a slice of fried plantain in actual use (average 30 g) differed from that listed in the Minnesota database (74.5 g), and adjusted that accordingly. The changes that were made are not likely to be specific to the elder population, and we believe that the questionnaire should be valid for younger Puerto Rican adults as well. However, this should be tested with a younger population.

RESULTS

Questionnaire development

Examination of the HHANES Puerto Rican adult data (with a sample of approximately 1,300 subjects) revealed several differences in dietary pattern when compared with the overall US adult diet pattern. Table 1 presents the results of the ranking of foods for the Puerto Rican population after grouping individual foods into food groups similar to those constructed by Block et al. (16) for NHANES II. The number one ranking food contributor for Puerto Rican adults living in the New York-New Jersey area was rice, which does not appear in the top 30 food contributors for the average US population. Other items which were more important to the Puerto Rican diet, but not used by the general population, include rice dishes, meat turnovers ("empanadas" and other "frituras"), and plantains. Compared with the general population, Puerto Ricans also appeared to consume considerably more fried chicken, dried beans, and pizza, but less fried potatoes, chips and snacks, pies, and corn-based products.

The process completed in table 1 was repeated for macronutrient components, vitamins, and minerals. Foods which contributed at least 0.5 percent to intake of any nutrient among Puerto Rican adults in the HHANES, and which could not be clearly coded into

TABLE 1. Major sources of energy in the Puerto Rican diet compared with the general US diet: Puerto Rican diet by the Hispanic Health and Nutrition Examination Survey (HHANES) vs. US diet by the Second National Health and Nutrition Examination Survey (NHANES II)

Rank	Puerto Rican diet by HHANES, 1982–1984		US diet by NHANES II, 1976–1980	
	Food	% of energy	Food	% of energy
1	Rice	9.53	White bread, rolls, crackers	9.59
2	Whole milk	7.69	Doughnuts, cookies, cake	5.70
3	White bread, rolls, crackers	6.63	Alcoholic beverages	5.60
4	Hamburgers, meat loaf	4.59	Whole milk	4.72
5	Beef steaks, roasts	4.47	Hamburgers, meat loaf	4.39
6	Doughnuts, cookies, cakes	4.28	Beef steaks, roasts	4.14
7	Soft drinks	4.19	Soft drinks	3.63
8	Pork	3.64	Hot dogs, ham, lunch meats	3.19
9	Alcoholic beverages	3.39	Eggs	2.53
10	Fried chicken	2.87	Fried potatoes	2.53
11	Sugar	2.86	Cheese	2.45
12	Hot dogs, ham, lunch meats	2.58	Pork	2.28
13	Dried beans	2.44	Ice cream, frozen desserts	1.71
14	Cheese	2.20	Dark breads	1.70
15	Eggs	2.16	Mayonnaise, salad dressing	1.67
16	Mayonnaise, salad dressing	2.11	2% milk	1.67
17	Orange juice	2.06	Margarine	1.64
18	Rice with meat	1.96	Spaghetti	1.64
19	Pizza	1.82	Sugar	1.48
20	Poultry, not fried	1.74	Potatoes, not fried	1.47
21	Spaghetti	1.41	Salty snacks	1.41
22	Fried potatoes	1.24	Orange juice	1.38
23	Plantains	1.20	Coffee, tea	1.35
24	Potatoes, not fried	0.93	Pies, excluding pumpkin	1.31
25	Fried fish	0.92	Dried beans	1.17
26	Soups	0.88	Poultry, not fried	1.12
27	Meat turnovers	0.85	Cornbread, grits, tortillas	0.99
28	Butter	0.83	Salad and cooking oil	0.94
29	Rice and beans	0.81	Fried fish	0.91
30	Peanuts, peanut butter	0.77	Peanuts, peanut butter	0.89

an existing line item on the HCRC/Block food frequency questionnaire are listed in table 2. Plantains, rice dishes, and frituras are commonly used by Hispanics, but not by the general US population. Others, such as low-fat milk added to coffee and non-fortified fruit drinks, were also identified and added as separate line items because of their contribution to intake among the Puerto Rican HHANES sample.

Examination of the frequency of consumption of specific foods within food frequency line item also revealed differences between HHANES Puerto Rican food patterns and weights in the original questionnaire. For example, the original database included only grape juice, apple juice, and mixed fruit juice for "other fruit juices." HHANES data revealed lower proportional use of grape juice, but significant use of pineapple juice and peach and guava nectars. The weights were, therefore, changed from the original (grape juice, 33 percent; apple juice, 34 percent; and fruit juice, 33 percent) to the average of the two sets (grape juice, 27 percent; apple juice, 42 percent; fruit juice, 16 percent; and pineapple juice, peach nectar, and guava nectar, 5 percent each). Similarly, the types of dried beans consumed differed, resulting in the addition of pink beans and reduction in percent contribution from "pork and beans" and lima beans. Other line items in which changes were made include apples, pears and applesauce, breakfast cereals, eggs, tomatoes, coleslaw and cabbage, salad dressing, "other vegetables," hamburgers and ground beef, pork, fried fish, shellfish, non-fried fish, biscuits and muffins, and breads.

Examination of the HHANES data also suggested a need for revision of adjustment questions on the food frequency form. The HCRC/Block form included questions about the leanness of meat and poultry, type of canned tuna used, and the type of fat usually added to vegetables, and in cooking. Based on the answers to these questions, the nutrient intake of individuals is adjusted to more accurately reflect usual intake of total fat and type of fat consumed. Based on evidence of use by Puerto Ricans, along with current interest in differ-

TABLE 2. Foods added to the Hutchinson Cancer Research Center/Block food frequency questionnaire for assessment of Puerto Rican diets

Food category and food name
Fruits and vegetables
Mangoes
Avocado, raw
Winter squash (including butternut, hubbard)
Other starchy roots/fruits (tannier, cassava, breadfruit)
Green plantains, boiled or baked
Green plantains, fried
Ripe plantains, boiled or baked
Ripe plantains, fried
Mixed dishes
Rice with meat
Rice with chicken
Rice with pigeon peas
Rice with beans
Homemade soups with meat/chicken
Meat turnovers, fritters, egg rolls
Beverages
Fortified fruit drinks (as separate line item)
Non-fortified fruit drinks
Whole milk in coffee or tea
2% milk in coffee or tea
Skim or 1% milk in coffee or tea
Desserts
Custard, pudding, cheesecake

entiating mono- and polyunsaturated oils, we separated vegetable oil into olive oil and "other vegetable oil." We also noted that Puerto Rican recipes usually include the addition of corn oil to rice and to beans, which, given the large amount of these items consumed, could affect fat intake estimates. We therefore added questions about whether fat or oil was used when cooking these items, and the type of fat or oil.

Portion size comparisons

A *t*-test comparison of log-transformed portion sizes for 93 line items in the food frequency questionnaire for which there were data in both HHANES and NHANES II showed significant differences ($p < 0.05$) across the two surveys for 55 items. Food items/groups for which the significance level was <0.01 and with a frequency of use sample size of at least 100 per group are presented in table 3. These data suggest that Puerto Ricans tend to consume larger portions of some fruits, including banana and fruit juices, and smaller portions of some vegetables, such as lettuce and tomato than the general population. They consume significantly larger portions of rice, beans, beef, pork, chicken, bread, and sweet baked goods, but smaller cups of coffee and tea and glasses of milk. In some cases, the

TABLE 3. Portion sizes: Hispanic Health and Nutrition Examination Survey (HHANES) Puerto Ricans vs. Second National Health and Nutrition Examination Survey (NHANES II) US adults*

Food name	HHANES Puerto Rican adults, 1982–1984		NHANES II US adults, 1976–1980	
	Mean (g) ± SE†	Median	Mean (g) ± SE	Median
Banana	121 ± 4.7	114	94 ± 1.4	102
Citrus juices	254 ± 9.2	256	212 ± 2.3	155
Other fruit juices	222 ± 10.1	209	169 ± 5.3	155
Tomatoes	66 ± 2.7	60	94 ± 1.4	75
Green salad	44 ± 2.0	37	55 ± 0.8	48
Salad dressing	16 ± 0.8	10	13 ± 0.2	9
Rice	231 ± 5.7	206	128 ± 2.8	126
Beans	129 ± 4.0	117	102 ± 2.2	80
Eggs	80.5 ± 1.8	92	78.2 ± 0.7	64
Hamburger	251 ± 17.4	206	160 ± 3.7	157
Beef	145 ± 7.3	113	120 ± 1.7	84
Pork	131 ± 7.4	98	92 ± 2.0	67
Fried chicken	101 ± 4.0	88	63 ± 1.8	51
Other chicken	98 ± 5.0	77	82 ± 1.4	66
Soups	406 ± 20.6	360	305 ± 4.9	298
Dark bread	42 ± 1.5	52	38 ± 0.4	38
Cake, cookies	68 ± 3.7	51	52 ± 0.6	42
Whole milk	160 ± 3.3	93	185 ± 2.2	150
Low-fat milk	108 ± 10.2	47	193 ± 3.5	165
Coffee	170 ± 4.2	140	363 ± 3.7	280
Tea	252 ± 12.1	186	312 ± 4.2	280
Sugar	11.9 ± 0.3	9.0	9.7 ± 0.2	7.5

* Items selected for inclusion based on sample sizes >100 in each group and significant *t*-tests, using log-transformed variables: $p < 0.01$. Means and medians are presented in their original scale.
† SE, standard error.

apparent difference in portion size may be due to differing coding assumptions used in NHANES II versus HHANES. For example, the median and mode portion size for banana in NHANES II is 102 g, while in HHANES it is 114 g, which is also the amount in the *USDA Composition of Foods Handbook No. 8* (17). Because this is the assumed weight for one banana, the most common portion size, the *t*-test is significantly different due to this assumption, not to differing actual practice. Differences in most items, however, are probably due to differences in food patterns across these groups.

We plotted the distribution of portion sizes for the HHANES and NHANES II data (not shown). These confirmed differing patterns of portion size in several foods. Examples of two key foods were rice and chicken soup. For rice, the distribution of intake was clearly shifted upward for Puerto Ricans in HHANES compared with the general population in NHANES II, with modal points (in order of descending height) at approximately 500, 300, and 700 g (Puerto Ricans) compared with 200 and 400 g (non-Hispanic whites). Similarly, portion sizes for chicken soup were clearly greater and more varied among the Puerto Ricans (modal points at 120, 250, 190, 300, and 410 g), reflecting their use of this item as more of a main meal than the more typical side cup of soup in the general US meal pattern (150 and 80 g). Use of standard portion sizes would, therefore, systematically underestimate both intake for the Hispanic group and true differences in intake across groups.

Questionnaire comparisons of mean estimation

Mean energy and nutrient intakes among Hispanics, as measured with the 24-hour recall and the three versions of the food frequency questionnaire, are presented in table 4. Because the data from the food frequency questionnaires were generated from the 24-hour recall, the recall data serve as the gold standard from which to compare the differing food frequency tools. As hypothesized, the energy and nutrient intakes of these Hispanic subjects were significantly underestimated by the original HCRC questionnaire for every nutrient except fat, and vitamins B12 and D. Energy was underestimated by close to 300 kcal (1.26 MJ), and nutrient underestimates were particularly great for protein, carbohydrate, vitamin A, folate, vitamin C, calcium, iron, and magnesium. The revised questionnaire with fixed portion sizes, which has a more complete food list, resulted in mean values which are closer to the recall data (exceptions were fat, vitamins D and E, and calcium). With the exception of vitamin B12, where the closest estimate was from the revised questionnaire with fixed portions, the revised questionnaire with open-ended portions yielded estimates closest to the 24-hour recall means for all nutrients.

TABLE 4. Comparison of mean intakes by differing methods in Hispanics aged 52–91 years (*n* = 90), northeast United States, August 1992 to May 1994

	24-hour recall	Original FFQ†	Revised FFQ, fixed portions	Revised FFQ, open portions
	Mean ± SE†	Mean ± SE	Mean ± SE	Mean ± SE
Energy (kcal)‡	1,464 ± 64	1,163 ± 45****	1,202 ± 44****	1,409 ± 64
Protein (g)	60.0 ± 3.0	47.6 ± 2.4****	53.5 ± 2.5***	62.0 ± 3.4*
Fat (g)	52.5 ± 2.9	48.7 ± 2.4	43.3 ± 2.1****	51.4 ± 2.9
Carbohydrate (g)	191 ± 8.7	134 ± 5.6****	151 ± 5.6****	176 ± 8.2**
Vitamin A (μg RE†)	732 ± 120	330 ± 44****	454 ± 75****	487 ± 76*
Thiamin (mg)	1.2 ± 0.06	0.9 ± 0.04****	1.0 ± 0.04****	1.2 ± 0.06**
Riboflavin (mg)	1.2 ± 0.05	1.0 ± 0.06****	1.0 ± 0.05****	1.2 ± 0.06
Niacin (mg)	14.3 ± 0.8	11.2 ± 0.6****	12.4 ± 0.6	14.4 ± 0.9*
Vitamin B6 (mg)	1.4 ± 0.08	1.0 ± 0.06****	1.2 ± 0.06****	1.4 ± 0.07
Vitamin B12 (μg)	2.4 ± 0.2	2.7 ± 0.2	2.3 ± 0.2	2.7 ± 0.2***
Folate (μg)	256 ± 15	125 ± 9****	206 ± 12****	214 ± 13****
Vitamin C (mg)	93 ± 8.3	65 ± 6.6****	75 ± 7.1***	81 ± 7.6*
Vitamin D (μg)	3.8 ± 0.3	3.7 ± 0.3	2.9 ± 0.2****	3.8 ± 0.3
Vitamin E (mg α-TE†)	5.1 ± 0.3	4.0 ± 0.2**	3.8 ± 0.2**	4.4 ± 0.3*
Calcium (mg)	573 ± 31	454 ± 32****	431 ± 26****	549 ± 31**
Iron (mg)	10.2 ± 0.6	7.3 ± 0.3****	8.6 ± 0.4*	9.8 ± 0.5
Magnesium (mg)	221 ± 9.3	141 ± 6.5****	182 ± 7.1****	209 ± 8.9
Zinc (mg)	7.6 ± 0.4	6.5 ± 0.5****	7.2 ± 0.4	8.2 ± 0.5****

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$, by paired *t*-test statistics for H_0: mean from food frequency = mean from 24-hour recall. For differences with non-normal distributions, significance levels were determined with the non-parametric sign test.

† SE, standard error; FFQ, food frequency questionnaire; RE, retinol equivalents; α-TE, alpha-tocopherol equivalents.

‡ To convert kcal to MJ, multiply by 0.004184.

Because modifications in the questionnaire were based mostly on Puerto Rican data, there was concern that some validity may have been lost for use with non-Hispanic whites. Table 5 presents the comparison of the 24-hour recall data for the non-Hispanic white comparison group with the three versions of the food frequency. Although the original HCRC questionnaire overestimated intake of some nutrients and underestimated intake of others for this group as well, the differences are smaller and less significant than those seen with the Hispanics. The difference in estimated average energy intake was 32 kcal (0.13 MJ) for non-Hispanic whites, compared with approximately 300 kcal (1.26 MJ) for Hispanics. In relation to the 24-hour recall estimates, the revised questionnaire with fixed portion sizes underestimated energy intake, carbohydrate, and vitamin A, more than the original questionnaire for non-Hispanic whites, but yielded closer estimates for fat and vitamin D. The revised questionnaire with open portions measured most nutrients more accurately for non-Hispanic whites than did the original or revised with fixed portions versions of the questionnaire. The original questionnaire remained most accurate for protein and zinc. In contrast to concerns that the revisions may lessen accuracy for the non-Hispanic whites, the new instrument appears to perform at least as well and, for some nutrients, better than the original when open-ended portion sizes are used. In comparison with the 24-hour recalls, the mean energy intake from the revised questionnaire with open-ended portions differs by only 11 kcal (0.05 MJ).

Figure 1 illustrates, as examples, the differences in mean estimation for energy, folate, and iron, both across methods and across the two ethnic groups. As hypothesized, the actual small difference in energy intake between the Hispanic and non-Hispanic white samples was greatly exaggerated. A comparison of means across tables shows that an average difference of 200 kcal (0.84 MJ) across groups as measured by the 24-hour recall, would be exaggerated with the original questionnaire to a 470 kcal (1.97 MJ) difference if the original food frequency form was used. With the revised questionnaire with open-ended portions, the difference is 256 kcal (1.07 MJ). Comparison of nutrient intakes across ethnic groups with the original questionnaire revealed significant differences in all of the 18 nutrients measured, although only eight of the nutrients (vitamin A, thiamin, riboflavin, niacin, vitamin B12, iron, magnesium, and zinc) were significantly different with the 24-hour recall. Use of the revised, open-ended questionnaire identified differences in these eight nutrients, but also showed differences in energy, protein, and calcium not seen with the recalls. When analyses were adjusted for energy intake, vitamin A, riboflavin, niacin, vitamin B12, iron, and zinc intakes differed significantly across ethnic groups by 24-hour recall. With the exception of mag-

TABLE 5. Comparison of mean intakes by differing methods in non-Hispanic whites aged 60–88 years (*n* = 35), northeast United States, August 1992 to May 1994

	24-hour recall	Original FFQ†	Revised FFQ, fixed portions	Revised FFQ, open portions
	Mean ± SE†	Mean ± SE	Mean ± SE	Mean ± SE
Energy (kcal)‡	1,527 ± 95	1,559 ± 95	1,364 ± 83*	1,538 ± 100
Protein (g)	61.4 ± 4.4	60.3 ± 3.9	57.2 ± 3.5	65.5 ± 4.5
Fat (g)	54.3 ± 5.7	67.2 ± 5.6**	51.0 ± 4.7	57.1 ± 5.3
Carbohydrate (g)	201 ± 11.3	180 ± 10.8*	171 ± 11.0***	191 ± 12.6
Vitamin A (μg RE†)	1,188 ± 231	959 ± 250	919 ± 270**	1,052 ± 287
Thiamin (mg)	1.4 ± 0.08	1.2 ± 0.08*	1.2 ± 0.07*	1.4 ± 0.09
Riboflavin (mg)	1.7 ± 0.12	1.7 ± 0.16	1.5 ± 0.14	1.8 ± 0.16
Niacin (mg)	17.0 ± 1.3	16.1 ± 1.1	15.1 ± 0.9	17.3 ± 1.2
Vitamin B6 (mg)	1.5 ± 0.1	1.4 ± 0.1	1.3 ± 0.1	1.6 ± 0.1
Vitamin B12 (μg)	3.9 ± 0.6	6.3 ± 1.5**	5.0 ± 1.5	5.8 ± 1.5*
Folate (μg)	272 ± 29	221 ± 23*	232 ± 26*	247 ± 24
Vitamin C (mg)	112 ± 18	83 ± 11	98 ± 18	89 ± 11
Vitamin D (μg)	4.6 ± 0.5	5.6 ± 0.7*	4.1 ± 0.7	5.2 ± 0.7
Vitamin E (mg α-TE†)	4.6 ± 0.5	4.7 ± 0.4	4.2 ± 0.4	4.7 ± 0.4
Calcium (mg)	684 ± 59	647 ± 60	590 ± 64*	705 ± 69
Iron (mg)	13.5 ± 1.6	12.2 ± 1.1	10.7 ± 0.8	12.7 ± 1.2
Magnesium (mg)	261 ± 22	235 ± 18	238 ± 14	278 ± 20
Zinc (mg)	8.5 ± 0.7	8.5 ± 0.7	8.2 ± 0.6	9.4 ± 0.8*

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$, by paired *t*-test statistics for H_0: mean from food frequency = mean from 24-hour recall. For differences with non-normal distributions, significance levels were determined with the non-parametric sign test.

† SE, standard error; FFQ, food frequency questionnaire; RE, retinol equivalents; α-TE, alpha-tocopherol equivalents.

‡ To convert kcal to MJ, multiply by 0.004184.

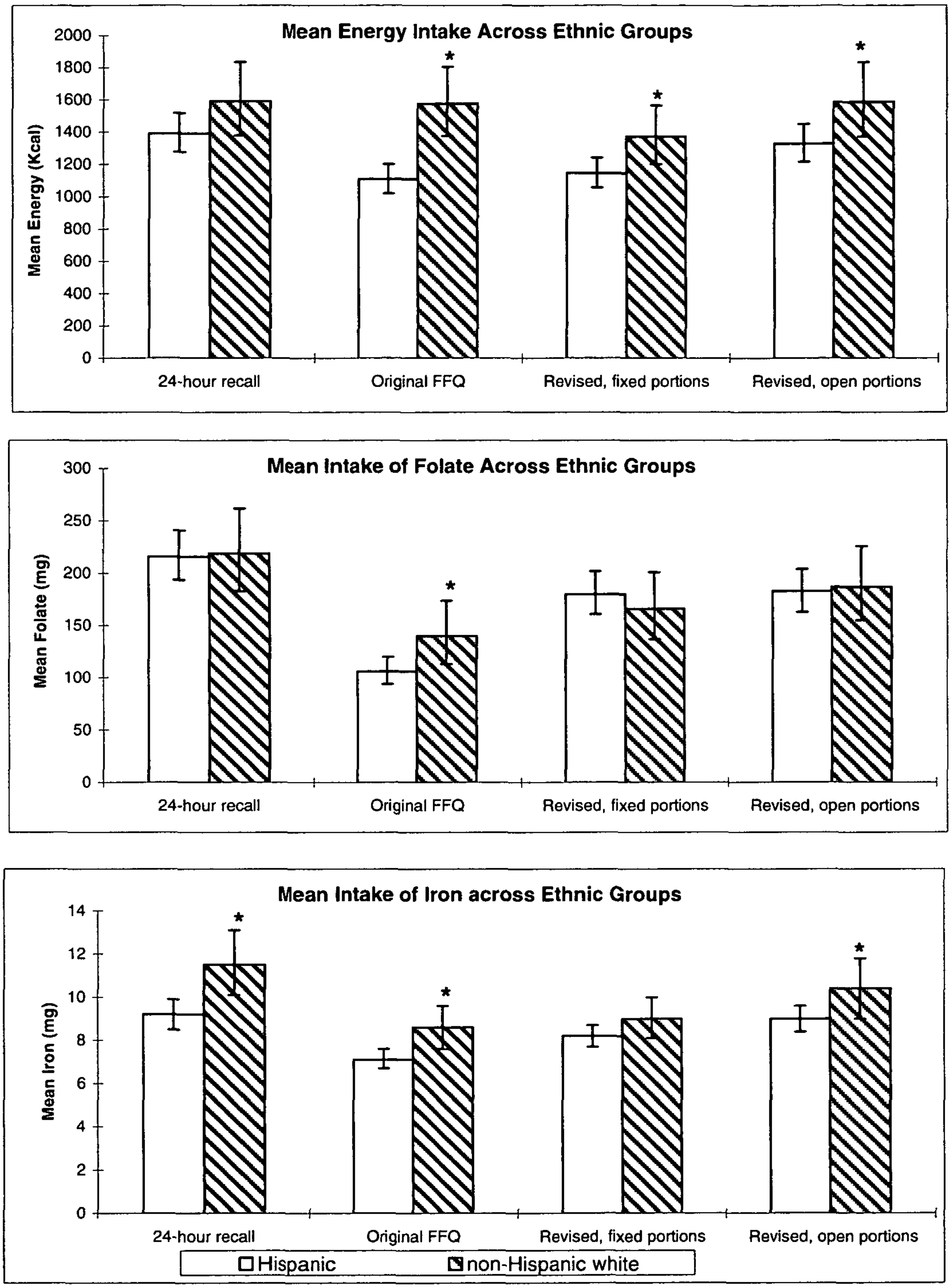

FIGURE 1. Mean estimates (and standard errors) of intake across estimates of intake for Hispanics aged 52–91 years (n = 90) and non-Hispanic whites aged 60–88 (n = 35) using various instruments, northeast United States, August 1992 to May 1994. Data are presented in their original scale, but significance testing was done with log-transformed variables and with adjustment for age, sex, and (for folate and iron only) total energy intake. *Significantly different ($p < 0.05$) across the two groups using the General Linear Models (GLM) procedure in SAS (10).

nesium rather than zinc, this same set of nutrients differed significantly using the revised, open-ended questionnaire.

Questionnaire comparisons of ranking

Tables 6 and 7 present the results of correlations between log-transformed nutrient intake data obtained from each of the three food frequency methods and the 24-hour recall. Because these compare original recall data coded directly into the food frequency questionnaires, they reflect the relative performance of the questionnaire and do not contain the error which would be associated with subject completion of the questionnaires. Among Hispanics, the correlations with the revised questionnaire, and particularly with the open-ended revised questionnaire show considerable improvement in comparison with the original questionnaire. The Pearson correlation for energy intake increased from 0.83 using the original to 0.95 using the revised open-ended questionnaire. Differences were even more dramatic for several micronutrients, including vitamin A (Pearson correlations increased from 0.62 to 0.85), vitamin B6 (0.69 to 0.95), folate (0.57 to 0.88), and iron (0.68 to 0.91). The only nutrients which did not show improvement with the revised questionnaire were vitamins D and E. The intraclass correlations tended to be higher, reflecting the greater inter- versus intra-person consistency in these paired data; they ranged from 0.84 for vitamin E to 0.98 for protein and riboflavin.

Among the Hispanic sample, results for the regression of energy intake from each of the questionnaire versions on the energy intake from the 24-hour recall were as follows: original questionnaire—intercept 380 kcal and slope 0.53; revised questionnaire with fixed portions—intercept 350, slope 0.58; and with open-ended portions—intercept 42, slope 0.93. Plots (not shown) reveal that estimates are quite good at the low intake end, but tend to be underestimated at higher intake levels, with closest overall agreement for the revised form with open-ended portions.

Correlations for the non-Hispanic whites are presented in table 7. Again, in contrast to concerns that the improvements for Hispanics may lead to decreased validity for non-Hispanic whites, we saw improvements in correlations with the revised, open-ended, food frequency questionnaire for all nutrients except vitamins C and E. Correlations with the revised questionnaire with fixed portion sizes were frequently lower than the original questionnaire (while they tended to be higher among the Hispanics), suggesting that the use of open-ended portions is important when both ethnic groups are being measured. In general, the correlations were slightly lower for the non-Hispanic whites than for the Hispanics. For the questionnaire with open-ended portions, intraclass correlations

TABLE 6. Correlations between food frequency questionnaires and the 24-hour recall in Hispanics aged 52–91 years (*n* = 90), northeast United States, August 1992 to May 1994*

	Pearson correlations			Intraclass coefficients		
	Original FFQ†	Revised FFQ, fixed portions	Revised FFQ, open portions	Original FFQ	Revised FFQ, fixed portions	Revised FFQ, open portions
Energy (kcal)‡	0.83	0.89	0.95	0.90	0.94	0.97
Protein (g)	0.80	0.91	0.97	0.89	0.95	0.98
Fat (g)	0.79	0.85	0.87	0.87	0.92	0.93
Carbohydrate (g)	0.76	0.84	0.92	0.86	0.91	0.96
Vitamin A (μg RE†)	0.62	0.83	0.85	0.75	0.90	0.91
Thiamin (mg)	0.74	0.85	0.93	0.85	0.91	0.96
Riboflavin (mg)	0.83	0.88	0.95	0.90	0.94	0.98
Niacin (mg)	0.81	0.88	0.95	0.89	0.93	0.97
Vitamin B6 (mg)	0.69	0.89	0.95	0.82	0.93	0.97
Vitamin B12 (μg)	0.77	0.82	0.85	0.87	0.90	0.92
Folate (μg)	0.57	0.85	0.88	0.72	0.92	0.93
Vitamin C (mg)	0.66	0.88	0.91	0.75	0.93	0.95
Vitamin D (μg)	0.92	0.82	0.84	0.96	0.88	0.89
Vitamin E (mg α-TE†)	0.74	0.67	0.72	0.83	0.79	0.84
Calcium (mg)	0.83	0.86	0.93	0.89	0.92	0.96
Iron (mg)	0.68	0.83	0.91	0.81	0.90	0.95
Magnesium (mg)	0.70	0.86	0.92	0.82	0.92	0.96
Zinc (mg)	0.82	0.91	0.94	0.90	0.95	0.97

* All correlations were done with log-transformed variables.
† FFQ, food frequency questionnaire; RE, retinol equivalents; α-TE, alpha-tocopherol equivalents.
‡ To convert kcal to MJ, multiply by 0.004184.

TABLE 7. Correlations between food frequency questionnaires and the 24-hour recall in non-Hispanic whites aged 60–88 years (*n* = 35), northeast United States, August 1992 to May 1994

	Pearson correlations			Intraclass coefficients		
	Original FFQ†	Revised FFQ, fixed portions	Revised FFQ, open portions	Original FFQ	Revised FFQ, fixed portions	Revised FFQ, open portions
Energy (kcal)‡	0.80	0.79	0.92	0.88	0.88	0.96
Protein (g)	0.77	0.72	0.91	0.87	0.84	0.95
Fat (g)	0.81	0.75	0.84	0.86	0.84	0.90
Carbohydrate (g)	0.84	0.85	0.92	0.91	0.91	0.96
Vitamin A (μg RE†)	0.67	0.75	0.77	0.81	0.85	0.86
Thiamin (mg)	0.70	0.56	0.70	0.82	0.72	0.82
Riboflavin (mg)	0.70	0.72	0.86	0.82	0.83	0.92
Niacin (mg)	0.59	0.49	0.71	0.74	0.65	0.83
Vitamin B6 (mg)	0.70	0.66	0.78	0.83	0.79	0.87
Vitamin B12 (μg)	0.64	0.82	0.91	0.78	0.89	0.95
Folate (μg)	0.77	0.73	0.84	0.87	0.83	0.91
Vitamin C (mg)	0.76	0.63	0.72	0.86	0.77	0.84
Vitamin D (μg)	0.65	0.70	0.76	0.78	0.80	0.84
Vitamin E (mg α-E†)	0.85	0.65	0.70	0.90	0.77	0.82
Calcium (mg)	0.75	0.81	0.83	0.86	0.89	0.90
Iron (mg)	0.65	0.56	0.66	0.79	0.71	0.80
Magnesium (mg)	0.82	0.77	0.88	0.90	0.87	0.94
Zinc (mg)	0.75	0.74	0.88	0.85	0.85	0.93

* All correlations were done with log-transformed variables.
† FFQ, food frequency questionnaire; RE, retinol equivalents; α-TE, alpha-tocopherol equivalents.
‡ To convert kcal to MJ, multiply by 0.004184.

ranged from 0.80 for iron to 0.96 for energy and carbohydrate among non-Hispanic whites.

Regressions of the food frequency estimates on the recall estimates of energy intake for the non-Hispanic whites resulted in the following: original questionnaire—intercept 430 kcal and slope 0.74; revised questionnaire with fixed portions—intercept 419, slope 0.62; and revised open-ended questionnaire—intercept 108, slope 0.94. As with the Hispanics, plots (not shown) reveal underestimation of intake at the high intake range with closest agreement for the revised form with open-ended portions.

DISCUSSION

Examination of the dietary patterns of Puerto Rican adults reveals several differences from the general US population. They consume some foods which are not widely consumed by others, and consume many foods in differing patterns of frequency and portion size. These differences in dietary pattern invalidate the use of existing food frequency instruments with this group. In order to obtain valid data on nutrient intake from a food frequency instrument, substantial modification was necessary, including the addition of foods to the food list, adaptation of the nutrient database, weighting of individual contributors to line items, and a reassessment of portion sizes.

Because of significant differences in portion size for many food items, and because of irregular and wide distributions of portion size, the instrument was tested using both open-ended as well as fixed portion sizes. Results demonstrate significant gains in both mean estimation and correlational ranking with open-ended portions, compared with the fixed portion method. This finding differs from other reports of fixed versus open portion sizes. Willett (18) has argued that the added benefit of portion size on ranking of subjects and correlation coefficients or regression coefficients does not justify the expense of collecting the data. He cites several studies (8, 19–21), which saw increments in correlations with diet records of only 0.05 or less. Samet et al. (22) found strong correlations between instruments using fixed vs. open portions for vitamin A assessment and also concluded that standard portion sizes were appropriate for their population. However, as Willett (18) notes, this may not be true when differing portion sizes are culturally based. In the usual Puerto Rican diet, for example, rice—the major source of energy—is often consumed in very large amounts at one meal, and may routinely exceed the "large" serving size, estimated in the HCRC questionnaire, as 1.125 cups. As rice may be consumed several times per day, any comparison of intakes across these two groups would appear to show that the Puerto Ricans were consuming significantly less energy and associated nutrients than they actually were.

Our results support the use of open-ended portion sizes with this population. Substantial improvements in mean estimates and in correlations with the 24-hour recall data source were seen for most nutrients. With further testing of this revised questionnaire, it may be possible to determine a wider range of standardized portions which would allow more automation, or to limit the open-ended portions to those key foods, such as rice, soups, juices, and milk, for which there are large variations which affect results.

Because the data coded into the questionnaires came directly from the 24-hour recall with which it is being compared, this study tests only the relative validity of the final questionnaire. It does not include differences which may be seen with actual implementation of the questionnaire or with comparison of the questionnaire with multiple diet records in this group of Hispanics. However, our results for both the Hispanic and non-Hispanic white groups suggest that it should, at minimum, perform significantly better than existing questionnaires. The completion of usual portion size may be a more complex task, and certainly demands more from respondents. Hunter et al. (23) observed high intra-individual variation in actual portion size and concluded that individuals may, therefore, have difficulty in specifying usual portion size. To the extent that this is true, the gains from collecting data on portion size will be diminished. Because of high respondent demands, it is likely that interviewer administration will be required in order to obtain valid and complete portion size data. The target population for which this questionnaire was developed, Hispanic elders, contain large proportions with low education levels, making interviewer administration of the questionnaire necessary.

Results from cross-group comparisons demonstrate that caution should be taken whenever food frequency data are used to determine group differences in nutrient intake. The 24-hour recall has been validated for estimating mean intakes and comparing across groups (24). We found that use of the original questionnaire would suggest false differences for most of the nutrients. With the use of the revised, open-ended questionnaire with energy adjustment, results of cross-group comparisons were the same as those seen with the 24-hour recall for all but a few nutrients.

In summary, a systematic approach to the adaptation of a widely used food frequency instrument has led to the development of a questionnaire with increased validity for the specific population it is targeting without loss of validity for the comparison population. This required changes in the food list, food item weights within the database, adjustments in or use of open-ended portion sizes and the addition of fat adjustment questions and options. The large relative gains in validity clearly show that the use of general instruments for minority populations with differing diets may lead not only to error, but quite possibly to biased results, reinforcing the need for new validity studies with use in each new population.

ACKNOWLEDGMENTS

This project has been funded by the National Institute on Aging under grant no. R01 AG10425 and, in part, by the US Department of Agriculture, Agricultural Research Service under contract no. 53-3K06-01.

The authors acknowledge the programming contributions of Alex Pedan for the analysis of HANES data; Thomas Schaetzel and Peter Zhu for the adaptation of the Hutchinson Cancer Research Center SAS food frequency program, previously adapted from the NCI/Block questionnaire by Al Williams with Alan Kristal at the Hutchinson Cancer Research Center; Peter Zhu for the creation of the SAS SQL portion size entry program; and Gerard Dallal for statistical advice.

REFERENCES

1. Sempos CT. Invited commentary: Some limitations of semiquantitative food frequency questionnaires. Am J Epidemiol 1992;135:1127–32.
2. Briefel RR, Flegal KM, Winn DM, et al. Assessing the nation's diet: limitations of the food frequency questionnaire. J Am Diet Assoc 1992;92:959–62.
3. Loria CM, McDowell MA, Johnson CL, et al. Commentary: nutrient data for Mexican-American foods: are current data adequate? J Am Diet Assoc 1991;91:919–22.
4. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semi-quantitative food frequency questionnaire. Am J Epidemiol 1985;122:51–65.
5. Rimm EB, Giovannucci EL, Meir JS, et al. Reproducibility and validity of an expanded self-administered semi-quantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992;135:1114–27.
6. National Center for Health Statistics. Dietary intake source data: United States, 1976–80. Series 11. no. 231. Washington, DC: US GPO, 1983.
7. Block G, Hartman AM, Dresser CM, et al. A data-based approach to diet questionnaire design and testing. Am J Epidemiol 1986;124:453–69.
8. Block G, Woods M, Potosky A, et al. Validation of a self-administered diet history questionnaire using multiple diet records. J Clin Epidemiol 1990;12:1327–35.
9. Block G, Thompson FE, Hartman AM, et al. Comparison of two dietary questionnaires validated against multiple dietary records collected during a 1-year period. J Am Diet Assoc 1992;92:686–93.
10. SAS Institute Inc. SAS user guide: basics. Version 6 ed. Cary, NC: SAS Institute Inc, 1990.
11. Minnesota Nutrition Data System (NDS) software, developed by the Nutrition Coordinating Center (NCC), University of Minnesota. Food database version 10A; nutrient database version 25. Minneapolis, MN: Nutrition Coordinating Center (NCC), University of Minnesota, 1995.
12. Kristal AR, Shattuck AL, Henry HJ, et al. Rapid assessment of dietary intake of fat, fiber, and saturated fat: validity of an instrument suitable for community intervention research and nutritional surveillance. Am J Health Promot 1990;4:288–95.
13. Thompson FE, Byers T, Kohlmeier L. Dietary assessment resource manual. J Nutrition 1994;124:2245S–2317S.

14. US Department of Health and Human Services, National Center for Health Statistics. Hispanic Health and Nutrition Examination Survey, 1982–1984. [Computer file]. 7th release. Washington, DC: US DHHS, NCHS, 1992.
15. US Department of Health and Human Services, National Center for Health Statistics. Health and Nutrition Examination Survey, 1976–1980. [Computer file no. 5704]. Washington, DC: US DHHS, NCHS.
16. Block G, Dresser CM, Hartman AM, et al. Nutrient sources in the American diet: quantitative data from the NHANES II survey. Am J Epidemiol 1985;122:27-39.
17. US Department of Agriculture, Human Nutrition Information Service. Handbook 8–9. Composition of foods: fruits and fruit juices. Washington, DC: US GPO, 1982.
18. Willett WC. Future directions in the development of food frequency questionnaires. Am J Clin Nutr 1994;59(suppl): 171S–174S.
19. Hankin JH, Rhoades GG, Glober GA. A dietary method for an epidemiologic study of gastrointestinal cancer. Am J Clin Nutr 1975;28:1055–61.
20. Hernandez-Avila M, Master C, Hunter DJ. Influence of additional portion size data on the validity of a semiquantitative food frequency questionnaire. (Abstract). Am J Epidemiol 1988;128:891.
21. Cohen NL, Laus MJ, Ferris AM. The contribution of portion data to estimating nutrient intake by food frequency questionnaire. Research bulletin no. 730. Amherst, MA: Massachusetts Agricultural Experiment Station, University of Massachusetts, 1990.
22. Samet JM, Humble CG, Skipper BE. Alternatives in the collection and analysis of food frequency interview data. Am J Epidemiol 1984;120:572–81.
23. Hunter DJ, Sampson L, Stampfer MJ, et al. Variability in portion sizes of commonly consumed foods among a population of women in the United States. Am J Epidemiol 1988; 127:1240–9.
24. Gersovitz M, Madden JP, Smickilas-Wright H. Validity of the 24-hr dietary recall and seven-day record for group comparisons. J Am Diet Assoc 1978;73:48–55.