Reproducibility and Validity of Major Dietary Patterns among Swedish Women Assessed with a Food-Frequency Questionnaire

Bahram Rashid Khani,[2] Weimin Ye,* Paul Terry,† and Alicja Wolk

*Institute of Environmental Medicine and *Department of Medical Epidemiology, Karolinska Institutet, Stockholm, Sweden, and †Department of Epidemiology and Population Health, Albert Einstein College of Medicine, New York, NY*

ABSTRACT Defining dietary patterns by factor analysis is an alternative approach to dietary assessment that has been used recently to examine diet-disease relations. However, only 1 study evaluated the reproducibility and validity of this method. Our aim was to assess both the validity and reproducibility of major dietary patterns based on data from a 60-item FFQ. We chose 2 independent random samples among over 60,000 women aged 40–74 y participating in the Swedish Mammography Cohort (SMC). In the validation study, the FFQ was compared with 4 7-d dietary records (DRs) among 129 women. For the reproducibility study, the FFQ was administered twice, 1 y apart in 212 women. By conducting factor analysis, 3 major dietary patterns were identified: healthy (high in vegetables, fruits, fish, poultry, tomato, cereal, and low-fat dairy products), Western (processed meat, meat, refined grains, sweets, and fried potatoes), and drinker (beer, wine and liquor, snacks) pattern. These 3 patterns explained 29–34% of the total variance in these 2 studies. The Spearman correlation coefficients between FFQ1 and FFQ2 (reproducibility) for healthy, Western, and drinker pattern were 0.63, 0.68, and 0.73, respectively (all $P < 0.0001$). Correlation coefficients between the FFQ and DRs (validity) for these patterns were 0.59, 0.50, and 0.85, respectively (all $P < 0.0001$). Our results indicate that identification of dietary patterns through factor analysis is a reproducible and valid method. The dietary patterns approach might be used in nutritional epidemiology as an alternative method of dietary assessment. J. Nutr. 134: 1541–1545, 2004.

KEY WORDS: • *dietary patterns* • *reproducibility* • *validity* • *food-frequency questionnaire* • *factor analysis*

Throughout the nutrition literature, diet in its relation to disease has been described most often in the terms of food groups, single foods, or nutrient intakes. The food group or single-food approach may be inadequate to examine the health effects of these foods. The reasons for this are manifold. First, because consumption of a single food is commonly associated with a certain individual behavioral eating pattern, single-food analysis may be potentially confounded by the effect of that eating pattern (1,2). Second, the single-food approach may be inadequate for taking into account the biologic interactions among nutrients (for example, enhanced iron absorption in the presence of vitamin C) (3,4). Third, numerous analyses based on several food groups or specific food items may produce statistically significant associations simply by chance (5).

One approach to overcome these limitations and take into account the cumulative effect of multiple foods is to use "dietary pattern analysis" as proposed by Jacobson and Stanton (6). Dietary patterns may be defined by factor analysis that models interrelated variables (foods) as manifestations of composite factors. These factors represent eating patterns in the study population and help to distinguish individuals according to the combination of foods they choose to eat. Thus, the analysis of dietary patterns can be used further toward explaining disease occurrence. Major dietary patterns are likely to vary among different populations (3); therefore, use of these composite dietary exposures in epidemiologic studies requires evaluation of the validity and reproducibility of FFQ assessing identification of dietary patterns in a specific study population.

The purpose of our methodological study was to evaluate the validity and reproducibility of our FFQ regarding identification of major dietary patterns in the population of middle-aged and elderly women in central Sweden.

SUBJECTS AND METHODS

Subjects. The Swedish Mammography Cohort (SMC) is a population-based cohort established during a mammography screening program that was introduced in 2 counties in central Sweden from 1987 to 1990. All 90,303 women in Västmanland county born between 1917 and 1948 and in Uppsala county born between 1914 and 1948, received a mailed invitation to be screened by mammography between March 1987 and December 1990 together with a 6-page questionnaire; 66,651(73.8%) returned a completed questionnaire. The questionnaire included items about age, weight, height, education, family history of breast cancer, parity, age at first birth, and diet. An invitation to participate in a methodological (validation) study was mailed to a random subgroup of 362 women in the cohort.

[1] Supported by World Cancer Research Fund International and The Swedish Cancer Foundation.

[2] To whom correspondence should be addressed.
E-mail: Bahram.Rashidkhani@imm.ki.se.

Manuscript received 24 October 2003. Initial review completed 17 December 2003. Revision accepted 26 February 2004.

They were asked to fill in the FFQ again and to complete four 7-d weighted dietary records (DRs) over a 1-y period. Among them, 124 women completed 4, 1 completed 5, 1 completed 2, and 3 completed 1 7-d weighted DRs. A second random sample of 265 women from the cohort (independent from the validity subgroup) was chosen for the reproducibility study. They were asked to complete the same FFQ twice (FFQ1 and FFQ2), 1 y apart. Among them, 212 women (participation rate = 80%) completed FFQ2.

Dietary assessment. The self-administered FFQ included questions on 60 commonly eaten foods covering the whole diet. The questionnaire inquired only about the frequency of consumption without specification of portion size. Participants were asked how often, on average, they had consumed these foods over the past 6 mo. Eight predefined frequency categories ranging from "never/seldom" to "4 or more times per d" were used. There were also open questions about daily consumption of 4 different types of bread and glasses of milk. All self-reported frequencies were transformed to monthly consumption, considering 1 mo equal to 4 wk (i.e., if the subject reported consumption of 4 servings/wk, the monthly consumption was estimated as 16 servings). For the precoded frequency categories, the midpoint of each category was assumed as the most likely consumption (i.e., when reporting "1–3 portions/mo," the subject's monthly consumption was calculated as "2").

For dietary items that are not commonly consumed [butter on sandwiches, high-fat milk, French fries, liver and kidney, chips and pop corns, sweet soups, lemonade, soda, sugar, beer (2.8%), beer (4.5%), and hard liquor], missing frequency responses were considered as "never/ seldom" answers (7) and arbitrarily treated as very low consumption (0.5 times/mo). After that, we excluded women who had >19 missing food items on the FFQ. For energy calculation, we used age-specific portion sizes (40–52, 53–65, and 66–74 y) based on mean values from 5922 d of weighed food records among 213 women randomly selected from the study population. Total energy intake was calculated by summing up energy intakes from all foods. Furthermore we excluded outliers regarding energy intake (below or above mean ± 3 SD). Finally, we included 111 women in the validity and 197 women in the reproducibility analyses.

Dietary records. The women completed 4 7-d open-ended weighted DRs ~3 mo apart to cover variability in food consumption during different seasons. They were provided with an electronic scale, a set of household measures of volume, and a food diary.

Following instructions given by a research dietician, they measured and recorded their diet, and provided a recipe for unusual dishes. To obtain daily food intake measurements based on the DR comparable with those based on the FFQs, we matched the 1181 unique food/dishes codes recorded in dietary records to specific food items on the questionnaire. A total of 543 diet-record food codes were matched with the food items on the questionnaire and finally collapsed into the same 60 food items as on the questionnaire.

We summarized all occasions when specific foods were consumed to obtain an average frequency of consumption of a specific food item per month. The remaining 638 diet-record foods that did not match any of the questionnaire items were not used in our analysis because they were not asked for in the FFQ.

Food groupings. To reduce the complexity of the data, food items were grouped together (**Table 1**). The food grouping was based on similarity of nutrient profiles or culinary usage of the foods and was somewhat similar to that used in previous studies (4,8–10). Some individual food items were kept separately, either because it was inappropriate to incorporate them into a certain food group (e.g., eggs, tea, coffee, tomato, and pea soup) or because they were assumed to represent distinct dietary patterns (e.g., wine, liquor, beer, and soda). Finally 26 separate food groups were used in analyses to describe eating patterns.

Statistical methods. To identify behavioral food patterns in our study population, we used factor analysis (3,11) of 26 food groups (expressed as frequency of consumption per month). We conducted the analysis using the FACTOR procedure in SAS software (release 8; SAS Institute). Although focusing on factors with eigenvalues > 1.0 (showing that the factor describes more of the variability in the data than the average variable for any individual item within the factor) is a common practice, we focused on eigenvalues > 1.8. We chose this way to limit the factors and at the same time better identify meaningful factors (12). The factors were rotated by an orthogonal transformation (Varimax rotation function in SAS) to achieve simpler structure with greater interpretability. Factor loadings represent correlation coefficients between individual food groups and dietary patterns. Food groups with positive loadings contribute to a dietary pattern; food groups with negative loadings are inversely associated with a dietary pattern. The proportion of variance explained by each factor was calculated by dividing the sum of the squares of the respective factor loadings by the number of variables (i.e., food groups). The factor score for each pattern and for each individual was determined by summing the intakes from each food group weighted by the factor loadings (13). For assessing a comparability between dietary patterns derived from different dietary data (the baseline FFQ in the whole cohort, the FFQ and DR in the validity subgroup as well as FFQ1 and FFQ2 in the reproducibility subgroup), Spearman correlation coefficients were used (because the distributions of factor scores were usually skewed). We corrected the validity coefficients for random within-person error in the FFQ-based estimates (14).

TABLE 1

Food groupings used in dietary pattern analysis

Food group	Food items
Vegetables	Roots vegetables (carrots or beats), white cabbage, salad (lettuce or cucumbers), spinach
Tomato	Tomato
Fruit	Apples or pears, citrus fruit (oranges or grapefruits), banana
Whole grains	Whole grain soft bread, crisp bread, oatmeal or other whole grain, hot cereals
Refined grains	White bread, rice, spaghetti, waffles or pancakes
Cereal	Assorted breakfast cereals, muesli
Low-fat dairy	Low-fat milk, reduced-fat (medium) milk, low-fat yogurt
High-fat dairy	Butter, cheese, whole milk, whole yogurt, ice cream
Fish	Salmon, mackerel, sardines, tuna, herring, other fish
Poultry	Chicken
Meat	Beef, chopped meat, minced meat, liver, liver pate
Processed meat	Bacon, sausage, black pudding
Egg	Eggs
Margarine	Margarine, butter
Pea soup	Pea soup or bean soup
Cooked potatoes	Boiled potatoes
Fried potatoes	Fried potatoes, French fries
Snack	Potato chips or other snack chips, popcorn, fried and salted nuts
Sweets	Assorted candy, caramels, chocolate, cookies, sweets soups, marmalade or jams, sugar (sugar cubes)
Fruit juice	Juice
Soda	Carbonated sweetened drinks, uncarbonated sweetened drinks
Tea	Tea
Coffee	Coffee
Beer	Beer (3 different alcohol proofs)
Wine	Wine
Liquor	Liquor

RESULTS

For evaluation of representativity of the patterns identified in the reproducibility and validity subgroups of the SMC, we show major patterns identified in the whole cohort of nearly 60,000 woman (**Table 2**).

TABLE 2

Factor-loading matrix for 3 major dietary patterns identified from FFQs among 57,881 Swedish women[1]

Food group	Pattern 1 (Healthy)	Pattern 2 (Western)	Pattern 3 (Drinker)
Vegetables	0.72	—	—
Tomatoes	0.61	—	—
Fish	0.53	—	0.17
Fruit	0.52	—	−0.17
Poultry	0.37	—	0.29
Whole grains	0.36	0.25	−0.46
Cereal	0.32	—	—
Egg	0.31	0.20	0.16
Low-fat dairy	0.29	—	−0.19
Fruit juice	0.27	—	—
Tea	0.19	—	—
Sweets	−0.16	0.56	—
Processed meat	—	0.55	—
Refined grains	—	0.54	0.16
Margarine	—	0.51	−0.25
High-fat dairy	—	0.49	−0.16
Fried potatoes	—	0.41	0.24
Soda	—	0.40	—
Meat	0.32	0.40	0.24
Cooked potato	—	0.33	−0.27
Pea soup	—	0.27	—
Coffee	—	0.17	—
Wine	—	0.15	0.61
Liquor	—	—	0.55
Snacks	—	0.16	0.44
Beer	—	—	0.42
Proportion of variability	9%	8%	7%

[1] Values are factor loadings; absolute values < 0.15 are not displayed.

In the whole cohort and in the validation and reproducibility studies of subgroups within the SMC, we identified 3 major dietary patterns that we termed the "Healthy," "Western," and "Drinker" patterns. In the whole cohort, the healthy dietary pattern, which reflected the correlated intakes of foods commonly considered to be healthy, was loaded heavily by vegetables, fruits, fish, poultry, tomato, whole grains, cereal and low-fat dairy products. The Western pattern reflected mainly consumption of processed meat, meat, refined grains, sweets, margarine, high-fat dairy, potatoes, and soda. Alcoholic beverages such as wine, liquor, and beer and snacks contributed heavily to the drinker pattern. These factors accounted for 24% of total variance in the whole cohort of >60,000 women, 30% in the weighted food records, and 34% in the FFQ in the validation subgroup. In the reproducibility study, they explained 29% of total variance in the FFQ1 and 30% in the FFQ2. Other remaining factors explained <7% of the variance each and are not presented in the paper. Dietary patterns were loaded almost similarly by food groups in the validity (between FFQ and DR) and reproducibility (between FFQ1 and FFQ2) study (**Table 3**).

Mean monthly frequency of consumption for 26 food groups in the validity and the reproducibility study is shown in **Table 4**. Food overestimation > 20% by the FFQ compared with the DR included sweets, refined grains, wine, liquor, and snacks; in contrast, pea soup, cooked and fried potato, vegetables, processed meat, poultry, tea, fruit, and whole grain were underestimated. The Spearman correlation coefficient for the comparison of monthly consumption of food groups derived from the 2 FFQ and DRs was lowest for refined grains and highest for wine (Table 4). The Spearman correlation coefficients for the comparison of monthly consumption of food groups derived from the 2 FFQs (reproducibility) was lowest for egg and highest for wine.

The Spearman correlations between FFQ1 and FFQ2 (filled in 1 y apart) ranged from 0.63 for the "Healthy" to 0.73 for the "Drinker" pattern (**Table 5**). The FFQ and the DR (made during the year after filling in the FFQ) were also reasonably correlated ($r = 0.47$ to 0.73) for the Healthy, Western and Drinker patterns. The correlation coefficients were strengthened after adjusting for unreproducibility of the FFQ.

DISCUSSION

In the study population of middle-aged and elderly Swedish women, we identified 3 major dietary patterns that we named "Healthy," "Western," and "Drinker." The correlation coefficients between these major patterns estimated from 2 FFQs filled in 1 y apart were relatively high, indicating a good reproducibility. The correlation coefficients between the FFQ and the DRs used as a gold standard were also relatively high for these 3 major patterns, suggesting a reasonable validity of dietary patterns identified by factor analysis of the FFQ. The study subsamples in the reproducibility and validity study were representative of our population-based cohort because we found similar patterns in both.

For the 3 patterns, there were some differences in the factor loadings for the food items between the FFQ and DRs, probably because of methodological differences between the dietary assessment methods (14) and random statistical variations. However, the major patterns generated from the FFQ and DRs were similar, and the correlations of the dietary patterns between the FFQ and the DRs ranged from 0.47 to 0.73, suggesting the usefulness of the FFQ in assessing dietary patterns.

Because of changes in the seasonal food availability of different fruits and vegetables, and also differences in seasonal food preferences, the eating habits of subjects could change over time. However, 4 7-d DRs ~3 mo apart should cover variability in food consumption during different seasons.

A high level of reproducibility and validity of the pattern "Drinker" was accompanied by high correlations observed for specific alcoholic beverages. The validity and reproducibility was somewhat higher for wine ($r = 0.82$) than for beer or liquor. This is probably accounted for by the more regular pattern of wine consumption during the year compared with other alcoholic beverages, which are more strongly influenced by seasonal variations (15). These correlations are likely somewhat underestimating the true reproducibility. This is because over an interval of 1 y, some real changes in dietary intake may have occurred.

The dietary patterns derived from our data are similar to patterns identified in other studies using the same method (factor analysis) and performed in other populations. Hu et al.

TABLE 3

Differences in factor loadings for three major dietary patterns identified from the FFQ and dietary records (DRs) in the validity study and from FFQ1 and FFQ2 in the reproducibility study[1]

	Pattern 1	Pattern 2	Pattern 3
FFQ − DRs (validity)	−0.017 (0.203)	0.046 (0.249)	0.004 (0.144)
FFQ1 − FFQ2 (reproducibility)	0.019 (0.146)	0.010 (0.106)	0.034 (0.149)

[1] Values are means ± SD, $n = 26$.

TABLE 4

Monthly consumption of 26 food groups in the FFQ and the dietary records (DRs) in the validity study among 111 women, and in FFQ1 and FFQ2 in the reproducibility study among 197 women[1]

	Servings, *n/mo*				Spearman correlation coefficient	
	FFQ	DR	FFQ1	FFQ2	FFQ vs. DRF	FQ1 vs. FFQ2
Sweets	93 ± 52	43 ± 35	51 ± 44	55 ± 51	0.52	0.69
Refined grains	5 ± 15	21 ± 14	29 ± 22	31 ± 20	0.16	0.49
Low-fat dairy	34 ± 32	30 ± 29	31 ± 29	36 ± 30	0.45	0.62
High-fat dairy	60 ± 31	52 ± 31	50 ± 30	52 ± 31	0.43	0.54
Soda	5 ± 7	5 ± 9	8 ± 14	7 ± 13	0.35	0.67
Pea soup	1 ± 1	2 ± 1	2 ± 3	2 ± 3	0.33	0.54
Cooked potatoes	17 ± 7	21 ± 10	22 ± 14	20 ± 11	0.32	0.49
Fried potatoes	2 ± 2	3 ± 3	4 ± 5	4 ± 4	0.35	0.59
Coffee	75 ± 36	71 ± 27	75 ± 29	74 ± 31	0.61	0.69
Vegetables	33 ± 19	40 ± 25	40 ± 26	43 ± 33	0.37	0.58
Processed meat	9 ± 7	12 ± 12	16 ± 12	18 ± 13	0.47	0.65
Meat	21 ± 12	17 ± 10	25 ± 12	26 ± 14	0.60	0.56
Poultry	1 ± 1	2 ± 1	2 ± 3	3 ± 3	0.37	0.70
Wine	4 ± 5	2 ± 3	3 ± 4	3 ± 4	0.82	0.82
Fish	10 ± 6	9 ± 5	14 ± 8	14 ± 7	0.44	0.66
Tomatoes	15 ± 8	13 ± 14	16 ± 13	15 ± 12	0.30	0.59
Liquor	2 ± 4	1 ± 1	1 ± 2	1 ± 2	0.56	0.63
Egg	6 ± 5	5 ± 5	7 ± 5	6 ± 5	0.19	0.44
Beer	8 ± 9	8 ± 11	8 ± 12	8 ± 10	0.70	0.74
Fruit juice	7 ± 10	8 ± 11	8 ± 12	7 ± 10	0.38	0.52
Tea	12 ± 16	16 ± 22	15 ± 20	13 ± 18	0.71	0.77
Fruit	34 ± 20	48 ± 27	53 ± 35	54 ± 34	0.49	0.59
Whole grain	46 ± 26	79 ± 45	74 ± 39	78 ± 39	0.35	0.61
Snack	2 ± 2	1 ± 1	1 ± 1	1 ± 1	0.60	0.56
Margarine	57 ± 28	53 ± 37	55 ± 39	53 ± 37	0.48	0.59
Cereal	6 ± 8	9 ± 11	10 ± 12	9 ± 11	0.61	0.64
Mean					0.46	0.61

[1] Values are means ± SD or Spearman correlation coefficients, all $P < 0.0001$.

(4) assessed the reproducibility and validity of dietary patterns among men, using dietary data collected by an FFQ and DRs among participants in the Health Professionals' Follow-up Study. These investigators identified 2 major eating patterns that were named "Prudent" (vegetables, fruits, legumes, whole grains, and fish) and "Western" (processed meat, red meat, butter, high-fat dairy products, eggs, and refined grains). The correlation coefficients between each of the patterns based on the FFQs and on the DRs were 0.45–0.74 for the 2 patterns, suggesting reasonable comparability between the FFQs and the DRs in characterizing dietary patterns. Slattery et al. (12) found similar major dietary patterns in American women aged 30–79 y.

Our study results are generally comparable to those reported by Harvard researchers (4,16) regarding the type of major patterns identified ("Healthy" and "Western") as well as reproducibility and validity of the FFQ to identify and replicate those patterns. Terry et al. (8,9), using the same data from the SMC but with slightly different food groups (24 food groups instead of 26 in the present study), identified 3 similar major dietary patterns: "Healthy" (fruit and vegetables, fish and poultry, low-fat dairy, and whole grains), "Western" (characterized by such foods as red and processed meats, refined grains, fat, and sweets) and "drinker" (wine, liquor, and beer) pattern. Slattery et al. (12) identified 5 major eating patterns in American men and women (eigenvalues > 1.25) that were labeled Western, prudent, high fat/sugar dairy, substituters, and drinker. Although dietary pattern analyses should be interpreted with caution because they depend on geographical, cultural, and methodological variations [sampling, food grouping, number of variables used in factor analysis (4), deciding on the number of factors, the rotations employed], 2 major patterns (Healthy/prudent and Western) were common in the American and Swedish populations. In other words, some foods commonly thought to be healthy are correlated with each other, and less healthy foods (Western diet) are also correlated with each other in general eating patterns.

TABLE 5

Spearman correlation coefficients for healthy, Western, and drinker dietary pattern scores between the FFQ1 and FFQ2 (reproducibility) and between the baseline FFQ and DRs (validity)[1]

	Healthy	Western	Drinker
FFQ1 vs. FFQ2	0.63	0.68	0.73
FFQ vs. DRs	0.47	0.41	0.73
FFQ vs. DRs[2]	0.59	0.50	0.85

[1] Values are Spearman correlation coefficients, all $P < 0.0001$.
[2] Adjusted for reproducibility of the FFQ.

There are some limitations in our data. First, only ~36% of the randomly selected subjects completed DRs and were included in the validation study. Those who did not participate in the study may differ in some way from those who did. The participants may be more health and diet conscious and more attentive when filling in the FFQ. This may lead to a slight overestimation of the observed validity. Second, we assumed

that the patterns generated from the food records were the "gold standard." However, diet records are also susceptible to measurement error due to erroneous recording and potential changes in eating behaviors (14). Third, although the FFQ used in this study contained 60 commonly eaten foods, it was shorter than other FFQs that were used to derive dietary patterns (4). Fourth, 3 patterns were not representative of all of our available patterns, as was indicated by the proportion of variability (30 and 34% of total variance in DR and FFQ, respectively and 29 and 30% of total variance in the FFQ1 and FFQ2). Other minor dietary patterns were less interpretable and were highly variable in the 4 sources of data. Furthermore, our study included only women. Even in the same population, eating patterns may be different in men. In a study of dietary patterns in 939 Swiss adults (17), the major difference between men and women related to the satiating capacity (heavy and basic foods such as potatoes, fatty pork, and sausages) of their diets. Women ingested smaller amounts of rich and heavy foods and their daily energy intake was lower than that of men. In a study from the United States (12), the patterns identified for men and women were similar, although the order of their importance varied. For both men and women, the first 3 patterns were similar but the "Drinker" pattern, in which alcoholic beverages loaded highest, was the 4th dietary pattern in men and the 6th pattern in women.

In conclusion, our data indicate the reproducibility and validity of the major dietary patterns defined by factor analysis using data from the FFQ. Identification of dietary patterns through factor analysis might be used in epidemiology as an alternative dietary assessment method and suitable approach for studying the diet-disease association.

LITERATURE CITED

1. Kant, A. K., Schatzkin, A., Block, G., Ziegler, R. G. & Nestle, M. (1991) Food group intake patterns and associated nutrient profiles of the US population. J. Am. Diet. Assoc. 91: 1532–1537.

2. Randall, E., Marshall, J. R., Graham, S. & Brasure, J. (1990) Patterns in food use and their associations with nutrient intakes. Am. J. Clin. Nutr. 52: 739–745.

3. Hu, F. B. (2002) Dietary pattern analysis: a new direction in nutritional epidemiology. Curr. Opin. Lipidol. 13: 3–9.

4. Hu, F. B., Rimm, E., Smith-Warner, S. A., Feskanich, D., Stampfer, M. J., Ascherio, A., Sampson, L. & Willett, W. C. (1999) Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. Am. J. Clin. Nutr. 69: 243–249.

5. Farchi, G., Mariotti, S., Menotti, A., Seccareccia, F., Torsello, S. & Fidanza, F. (1989) Diet and 20-y mortality in two rural population groups of middle-aged men in Italy. Am. J. Clin. Nutr. 50: 1095–1103.

6. Jacobson, H. N. & Stanton, J. L. (1986) Pattern analysis in nutrition. Clin. Nutr. 5: 249–253.

7. Hansson, L. M. & Galanti, M. R. (2000) Diet-associated risks of disease and self-reported food consumption: how shall we treat partial nonresponse in a food frequency questionnaire? Nutr. Cancer 36: 1–6.

8. Terry, P., Suzuki, R., Hu, F. B. & Wolk, A. (2001) A prospective study of major dietary patterns and the risk of breast cancer. Cancer Epidemiol. Biomark. Prev. 10: 1281–1285.

9. Terry, P., Hu, F. B., Hansen, H. & Wolk, A. (2001) Prospective study of major dietary patterns and colorectal cancer risk in women. Am. J. Epidemiol. 154: 1143–1149.

10. Fung, T., Hu, F. B., Fuchs, C., Giovannucci, E., Hunter, D. J., Stampfer, M. J., Colditz, G. A. & Willett, W. C. (2003) Major dietary patterns and the risk of colorectal cancer in women. Arch. Intern. Med. 163: 309–314.

11. Kleinbaum, D. G., Kupper, L. L. & Muller, K. E. (1988) Variable reduction and factor analysis. In: Applied Regression Analysis and Other Multivariable Methods, pp. 595–640. PWS–Kent Publishing Company, Boston, MA.

12. Slattery, M. L., Boucher, K. M., Caan, B. J., Potter, J. D. & Ma, K. N. (1998) Eating patterns and risk of colon cancer. Am. J. Epidemiol. 148: 4–16.

13. Kim, J.-O. & Mueller, C. W. (1978) Factor Analysis: Statistical Methods and Practical Issues. Sage Publications, Thousand Oaks, CA.

14. Willett, W. C. (1998) Nutritional Epidemiology. Oxford University Press, New York, NY.

15. Ferraroni, M., Decarli, A., Franceschi, S., La Vecchia, C., Enard, L., Negri, E., Parpinel, M. & Salvini, S. (1996) Validity and reproducibility of alcohol consumption in Italy. Int. J. Epidemiol. 25: 775–782.

16. Fung, T. T., Willett, W. C., Stampfer, M. J., Manson, J. E. & Hu, F. B. (2001) Dietary patterns and the risk of coronary heart disease in women. Arch. Intern. Med. 161: 1857–1862.

17. Gex-Fabry, M., Raymond, L. & Jeanneret, O. (1988) Multivariate analysis of dietary patterns in 939 Swiss adults: sociodemographic parameters and alcohol consumption profiles. Int. J. Epidemiol. 17: 548–555.