

# Calibration of the dietary questionnaire for the Canadian Study of Diet, Lifestyle and Health cohort

Meera G Jain<sup>1,\*</sup>, Thomas E Rohan<sup>2</sup>, Colin L. Soskolne<sup>3</sup> and Nancy Kreiger<sup>4</sup>

<sup>1</sup>Department of Public Health Sciences, 12 Queen's Park Crescent West, University of Toronto, Toronto, Ontario, M5S 1A8, Canada: <sup>2</sup>Department of Epidemiology and Social Medicine, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA: <sup>3</sup>Department of Public Health Sciences, 13–103 Clinical Sciences Building, University of Alberta, Edmonton, Alberta, T6G 2G3, Canada: <sup>4</sup>Research Unit, Division of Preventive Oncology, Cancer Care Ontario, 620 University Avenue, Toronto, Ontario, M5G 2L7, Canada

Submitted 1 August 2001: Accepted 6 June 2002

## Abstract

**Objective:** For proper interpretation of results from epidemiological studies that use food-frequency questionnaires (FFQs), it is necessary to know the relationship between reported intakes from the FFQ and true usual intake. In this paper, we report a calibration study conducted to investigate the performance of the FFQ used in a cohort study, the Canadian Study of Diet, Lifestyle and Health.

**Methods:** Over a 1-year period, 151 men and 159 women completed a full set of questionnaires including a self-administered baseline FFQ, three 24-hour diet recalls administered by telephone, and a second FFQ self-administered subsequently. The association between the nutrient estimates derived from the FFQs and the diet recalls was evaluated by calculating deattenuated Pearson's correlation coefficients.

**Results:** The FFQs estimated mean daily nutrient intakes higher than the diet recalls. When the log-transformed and energy-adjusted nutrient intakes from the average of three 24-hour recalls were compared against the baseline FFQ, the following deattenuated correlations were obtained in men and women, respectively: total energy 0.44 and 0.32, total fat 0.64 and 0.68, saturated fat 0.68 and 0.70, dietary fibre 0.65 and 0.44, vitamin E 0.32 and 0.37, vitamin C 0.40 and 0.37,  $\beta$ -carotene 0.34 and 0.29, alcohol 0.74 and 0.67, caffeine 0.81 and 0.76, with a median correlation of 0.49 and 0.53. Correlations between the second FFQ and diet recalls were similar. The correlations between the two FFQs as a test of reliability had a median value 0.64 for men and 0.63 for women for selected nutrients.

**Conclusions:** The study suggests that the FFQ method gives acceptable levels of nutrients or food component estimates, as assessed by this calibration study against diet recalls, when limited to energy-adjusted and deattenuated values.

**Keywords**  
Calibration  
Diet methods  
Food-frequency questionnaires  
Validation

Diet-associated disease risk estimates obtained from case-control or cohort studies can be attenuated substantially by the measurement error that arises when assessing individual dietary intakes<sup>1</sup>. The food-frequency questionnaire (FFQ) is the preferred dietary assessment instrument in large nutritional epidemiology studies because of its convenience, relatively low cost and feasibility, compared with dietary records or personal interview methods. However, although dietary intakes estimated by FFQ are correlated with true usual intake, such estimates are often beset by systematic errors: underreporting or overreporting at the level of the individual<sup>2,3</sup>. Therefore, for proper interpretation of

the results of epidemiological studies that use FFQs, it is necessary to know the relationship between reported intakes from the FFQ and true usual intake. Food records and diet recalls are generally considered to be superior to FFQs, making them the current practical gold standard for dietary assessment. Studies have shown that multiple 24-hour diet recalls, conducted over the telephone, can be an efficient way to measure validity<sup>4,5</sup>. In the present paper, we report the results of a calibration study conducted to characterise the performance of the food-frequency questionnaire used in a cohort study, the Canadian Study of Diet, Lifestyle and Health (CSDLH). The estimates of measurement

\*Corresponding author. Email meera.jain@utoronto.ca

error thus generated will ultimately allow calculation of deattenuated rate ratios. The study also tested the reliability of the FFQ.

## Materials and methods

The CSDLH is a cohort study of diet, lifestyle and cancer risk that involves about 75 000 Canadians from across the country. The cohort was recruited mostly between 1992 and 1998 and predominantly from the alumni of several universities (Toronto, British Columbia, Western Ontario, and Alberta), while the remainder (approximately 5000) was recruited with the assistance of the Canadian Cancer Society (CCS). The study involves periodic contact with the study participants, and information on dietary intake is updated every four years.

### *The food-frequency questionnaire*

The CSDLH uses an expanded version of a validated 131-item FFQ developed at the former Epidemiology Unit of the National Cancer Institute of Canada (NCIC)<sup>6-8</sup>. For the CSDLH, the FFQ was expanded to include additional contributors to energy intake (e.g. various sweets), some newer cheese products (e.g. low-fat cheese), and some important contributors to micronutrient intake (additional fruits and vegetables). The self-administered questionnaire contains questions on the frequency of consumption and usual portion size of 166 food items in the 1-year period prior to completion of the questionnaire. Frequency is reported as the number of times the food is consumed per day, per week or per month or none at all. The participants also report on how many months of the year the food is consumed to account for seasonal variation in intake. It has an open-ended section to report other food items normally eaten by the participant. Three photographs (rice, meat and chicken), each with a small, medium and large portion size, are included in the questionnaire to assist participants with quantification of intake. These photographs represent the shapes and portion sizes of a substantial number of foods consumed.

### *Calibration study sample*

An attempt was made to recruit cohort members into the calibration study in proportion to their representation by source of initial recruitment into the overall cohort. Therefore, 900 men and women were randomly selected from the cohort: 480 from the University of Toronto/CCS participants who completed the first follow-up FFQ between July 1999 and September 1999, 240 from University of Alberta participants who completed the first follow-up FFQ between June 2000 and August 2000, and 180 from the University of Western Ontario participants who were recruited into the study between September 1998 and November 1998. The participants' contact details and study identification number were extracted into a database for use in the calibration study,

but only 491 of the 900 were approached for the study as we were successful in recruiting sufficient numbers at this point. For each randomly selected participant, three randomly selected days of the week were generated for each of the three proposed 24-hour recalls, so that if an interviewer was not successful in interviewing a participant on the first randomly selected day, an attempt was made to conduct the interview on the next randomly selected day.

### *Procedures*

The calibration study involved completion of an initial food-frequency questionnaire, namely the questionnaire completed at the first 4-year follow-up (FFQ1) (except for University of Western Ontario), three 24-hour recalls administered at intervals of about one month on randomly selected days of the week, and a final food-frequency questionnaire completed about one month after the final 24-hour recall (FFQ2). For University of Western Ontario participants this was not possible given that the first 4-year follow-up was scheduled to take place about two years later. Therefore, for these persons, the FFQ completed upon recruitment into the study was used. Participants were recruited by mailing them a letter (together with an instruction page on food measurements to assist with completion of the 24-hour recall), and then calling them a few days later to invite them to participate and to administer the first 24-hour recall over the telephone. The procedure was the same for subsequent 24-hour recalls (i.e. a letter plus instructions followed by a phone call and interview). The participants were not informed in advance when the 24-hour dietary recall would occur. The second FFQ (together with a reply-paid envelope) was mailed to the participants about one month after completion of the third 24-hour recall (the final FFQ was identical to the initial one). Reminder phone calls were used to prompt participants to complete their FFQs and return them to the University of Toronto. Mailing and interviewing took place between May 2000 and April 2001.

The interviewers were trained through written instructions and practice interviews on how to conduct 24-hour recalls and record information in a standard manner. The interviewers had previous experience in conducting face-to-face two-hour diet history interviews using food models and manuals describing probes specific to each category of food. For the present study, various standardised probes were used to help improve the quality of information collected by recalls. For example, referral to instruction pages (mailed to the participants beforehand) for portion sizes and amounts, additions to foods such as butter, dressings and 'anything else' after each meal, and listing the ingredients and amounts for recipes. At each phone call the interviewer introduced herself and indicated that she was associated with the Canadian Study of Diet, Lifestyle and Health, which was being run from the University of Toronto. On telephone contact,

potential participants were given an opportunity to ask any questions they had about the study, and the letters mailed prior to the calls contained telephone numbers that the participants could call to obtain information about any aspect of the study. The calibration study was approved by the Human Subjects Review Committee of the University of Toronto.

### Coding, data management, and food composition tables

The food-frequency questionnaires and 24-hour recalls were coded by trained coders, using coding manuals prepared especially for the study. The data were then entered into computer files and re-entered for verification by trained data-entry staff.

The same food composition data, obtained from The Canadian Nutrient File (CNF – 1997 version)<sup>9</sup>, were used to compute nutrients intakes from the 24-hour recalls and the FFQ. The CNF was supplemented for values for carotenoids from the US Department of Agriculture<sup>10</sup>. The food composition data were modified for grouped foods in the FFQ. To develop this database for the FFQ, the three 24-hour recalls from the 343 participants who completed all three diet recalls were utilised for grouped foods. Individual foods from the recalls and their frequencies were classified according to the listing of the food items in the FFQ. Final nutrient values for each composite item were weighted averages of the nutrients for the

component foods, where weights were determined by the frequencies and amounts of consumption reported in the recalls. This method has been used by other investigators to derive customised nutrient databases<sup>1</sup>; however, the representativeness of the subjects in the calibration sub-study can affect its application to the full cohort. The grouping scheme was based on the similarity of nutrient profile or culinary usage among the foods and was somewhat similar to that used in other studies<sup>11</sup>. Only nutrients based on foods, not vitamin supplements, were used for the calibration analysis.

### Statistical analysis

Means and standard deviations (SDs) were calculated for total nutrient intakes using the baseline FFQ (FFQ1), the three 24-hour diet recalls, and the second FFQ administered following the diet recalls (FFQ2) and completed approximately one year after the baseline FFQ. All nutrient estimates were log-transformed to reduce skewness and approximate the normal distribution better, and they were energy-adjusted by the residual method<sup>12</sup> to estimate the portion of nutrient intake that is uncorrelated with total energy intake. However, the means and SDs shown in Table 1 were calculated for untransformed, unadjusted nutrients, while the correlations in Table 2 were for log-transformed and energy-adjusted nutrients. A minimal value was assigned to nutrients with zero intake in a few participants, to facilitate log conversions. Pearson's

**Table 1** Daily nutrient/dietary component intakes of 310 participants, 151 men and 159 women. The CSDLH sub-study, 2000–2001. Values are given as mean and standard deviation (SD)

	Men						Women					
	FFQ1		FFQ2		Average of three 24-hour recalls		FFQ1		FFQ2		Average of three 24-hour recalls	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Energy (kcal)	2341	697	2243	667	2169	746	2091	771	2017	612	1805	653
Total fat (g)	79.9	34.9	77.8	29.4	77.9	38.1	70.6	38.3	68.4	27.2	64.0	32.9
Saturated fat (g)	26.4	15.0	25.8	11.7	26.3	14.8	23.7	14.9	23.2	11.0	22.0	12.7
Monounsaturated fat (g)	31.9	14.0	31.2	12.1	31.7	17.0	27.8	15.8	26.9	11.2	25.2	13.8
Polyunsaturated fat (g)	14.4	6.1	13.8	5.4	13.0	7.6	12.8	6.7	12.2	4.8	11.0	8.4
Protein (g)	94.4	27.7	92.7	30.7	84.8	38.6	88.4	31.3	87.7	29.6	73.3	35.8
Carbohydrate (g)	303.6	96.3	284.6	91.5	275.1	104.0	276.4	97.2	264.4	84.9	232.5	89.8
Cholesterol (mg)	271.0	128.2	264.1	116.5	260.5	208.7	228.3	111.8	235.0	109.4	221.9	187.2
Dietary fibre (mg)	24.6	9.5	23.4	10.0	18.8	10.4	25.0	10.2	23.8	8.4	17.1	9.3
Insoluble fibre (mg)	8.4	4.4	8.0	4.6	5.7	6.1	9.3	4.6	8.8	3.8	5.5	5.1
Calcium (mg)	960.8	424.0	941.2	419.9	807.8	407.2	1051.1	479.9	1046.1	476.5	809.6	416.4
Iron (mg)	18.2	6.9	17.1	6.8	15.9	7.2	16.9	6.9	15.8	5.4	14.2	7.2
Thiamin (mg)	2.2	0.9	2.1	0.8	2.0	1.0	2.1	0.8	1.9	0.7	1.7	1.0
Methionine (g)	2.0	0.6	2.0	0.7	1.8	1.0	1.8	0.7	1.9	0.7	1.5	0.9
Vitamin C (mg)	203.9	87.7	198.2	135.4	151.2	139.1	233.2	128.8	222.1	101.9	138.7	112.6
Vitamin E (mg)	8.1	6.9	7.2	4.8	5.0	7.2	7.5	6.8	6.9	4.6	3.6	3.9
Folate (µg)	354.9	111.2	340.9	131.4	289.2	159.1	367.4	131.1	356.3	115.0	277.0	182.4
Vitamin A (RE)	1712	907	1685	1250	1194.1	2224	2065	1242	1975	923	1291.3	1669
β-Carotene (µg)	5300	3305	4975	2969	3261.0	4814	6482	4256	6190	3270	3744.2	4650
Total carotenoids (µg)	16 163	8468	15 231	9367	10 024	10 708	19 056	11 006	18 117	9175	11 378.5	11 231
Alcohol (g)	13.0	16.1	13.1	14.9	10.7	16.9	8.2	11.3	7.6	9.4	6.8	11.8
Caffeine (mg)	247.0	196.1	238.7	189.2	256.0	658.4	230.5	194.0	215.6	173.8	218.3	280.2

RE – retinol equivalents.

**Table 2** Pearson's correlation coefficients between nutrient/dietary component intakes estimated by food-frequency questionnaires (FFQs) and average of three 24-hour diet recalls. The CSDLH sub-study, 2000–2001\*

	Non-energy adjusted, non-deattenuated				Energy-adjusted, deattenuated				Energy-adjusted, non-deattenuated	
	Men		Women		Men		Women		Men	Women
	FFQ1 and 24-hour recalls	FFQ2 and 24-hour recalls	FFQ1 and 24-hour recalls	FFQ2 and 24-hour recalls	FFQ1 and 24-hour recalls	FFQ2 and 24-hour recalls	FFQ1 and 24-hour recalls	FFQ2 and 24-hour recalls	FFQ1 and FFQ2	FFQ1 and FFQ2
Energy	0.41	0.44	0.27	0.26	0.44	0.46	0.32	0.33	0.61	0.62
Total fat	0.49	0.43	0.38	0.41	0.64	0.56	0.68	0.80	0.62	0.71
Saturated fat	0.56	0.54	0.44	0.50	0.68	0.65	0.70	0.75	0.74	0.76
Monounsaturated fat	0.44	0.36	0.39	0.42	0.52	0.49	0.52	0.62	0.57	0.68
Polyunsaturated fat	0.41	0.38	0.29	0.25	0.40	0.37	0.35	0.38	0.52	0.58
Protein	0.30	0.38	0.33	0.30	0.44	0.50	0.66	0.63	0.57	0.64
Carbohydrate	0.48	0.52	0.29	0.30	0.69	0.71	0.73	0.79	0.66	0.64
Cholesterol	0.39	0.36	0.44	0.42	0.54	0.64	0.57	0.66	0.68	0.55
Dietary fibre	0.53	0.52	0.30	0.35	0.65	0.66	0.44	0.50	0.73	0.67
Insoluble fibre	0.51	0.48	0.29	0.34	0.49	0.49	0.36	0.44	0.71	0.66
Calcium	0.49	0.62	0.53	0.53	0.46	0.61	0.76	0.71	0.55	0.59
Iron	0.49	0.49	0.30	0.22	0.48	0.53	0.53	0.45	0.65	0.65
Thiamin	0.52	0.41	0.25	0.22	0.50	0.45	0.41	0.46	0.68	0.60
Methionine	0.27	0.36	0.41	0.35	0.40	0.49	0.64	0.59	0.55	0.61
Vitamin C	0.33	0.38	0.37	0.43	0.40	0.44	0.37	0.45	0.68	0.60
Vitamin E	0.35	0.36	0.30	0.31	0.32	0.42	0.37	0.31	0.56	0.61
Folate	0.30	0.30	0.27	0.32	0.38	0.42	0.48	0.56	0.67	0.65
Vitamin A	0.21	0.18	0.29	0.19	0.23	0.23	0.30	0.21	0.58	0.49
$\beta$ -Carotene	0.27	0.24	0.28	0.17	0.34	0.32	0.29	0.20	0.62	0.53
Total carotenes	0.22	0.25	0.30	0.20	0.28	0.31	0.33	0.24	0.59	0.58
Alcohol	0.69	0.74	0.63	0.66	0.74	0.76	0.67	0.71	0.86	0.86
Caffeine	0.80	0.85	0.76	0.77	0.81	0.86	0.76	0.76	0.86	0.85

\* All data are log-transformed. Correlations for energy are not energy-adjusted. Correlations of 0.20 and higher have  $P < 0.01$ .

correlation coefficients were used to estimate correlations between intakes as measured by the FFQ and intakes as measured by the mean of the 24-hour recalls, for both the unadjusted and the energy-adjusted nutrient intakes. Correlations obtained with the unadjusted nutrients were generally lower than those obtained using the energy-adjusted nutrients but are not presented in this paper since they will have little application in the main cohort analysis. To reduce within-person variation in food intake derived from the diet recalls, we used the average nutrient intake across three 24-hour recalls. We also calculated deattenuated correlation coefficients when evaluating agreement between nutrients estimated from the FFQ and 24-hour recalls, corrected for month-to-month variation in diet recalls by using the following formula<sup>13,14</sup>:

$$r_t = r_0(1 + r/k),$$

where  $r_t$  is the corrected correlation between the nutrient derived from the FFQ and diet recalls,  $r_0$  is the observed correlation,  $r$  is the ratio of estimated within-person and between-person variation in nutrient intake derived from the three 24-hour recalls, and  $k$  is the number of repeated observations of diet recall (in this study  $k = 3$ ).

## Results

Of the 491 men and women who had completed the first FFQ and were approached in the calibration study, 362 (88% of the 410 actually contacted) completed the first diet recall, 348 (85%) completed the second diet recall and 343 (84%) completed the third diet recall. The second FFQ was completed by 310 participants (90% of the 343 contacted). Of those contacted, 48 (12%) refused to complete the first diet recall, three refused to complete the second and two refused to complete the third diet recall. The remaining non-response was mainly because of an inability to contact the person. The background characteristics of the 151 men and 159 women completing the second FFQ were as follows: respective mean age (SD) 54.7 (13.7) and 54.2 (14.1) years, weight 82.4 (14.1) and 64.5 (11.1) kg, body mass index 27.2 (4.3) and 25.6 (7.4) kg m<sup>-2</sup>, current smokers 3.3 and 7.6%, and never smoked 57.6 and 62.7%. The proportion of current smokers was much lower than the average prevalence of 26% in the general Canadian population aged 45–64 years<sup>15</sup>. The mean age (SD) of the men and women participants at recruitment in the main study cohort was 52.4 (15.6) and 46.6 (15.2) years, respectively, showing a higher percentage of older women in the calibration study. Other demographic data are not available on the full cohort, except that data from a pilot study on this cohort ( $n = 195$ ) had shown a similar proportion of ever/never smokers. There were no marked differences by demographic characteristics between the 343 men and women who completed the baseline FFQ and the 310 participants who completed the second FFQ.

Mean daily intakes of nutrients from the two FFQs and

from the average of three diet recalls for the 310 participants are shown in Table 1. Values for macronutrients estimated by the second FFQ were only slightly higher than those estimated from the diet recalls. However, estimates for vitamins were substantially higher by both FFQs compared with diet recalls, by a ratio of 1.2 to 1.7. All nutrient intakes were within generally acceptable ranges<sup>16</sup>. The average daily total energy intakes from FFQ1, FFQ2 and the average of 24-hour recalls among men were 2341, 2243 and 2169 kcal, respectively, and among women they were 2091, 2017 and 1805 kcal, respectively. There were no substantial differences between energy and nutrient estimates from FFQ1 and FFQ2.

Correlations between the FFQs and the mean 24-hour diet recall for selected dietary components are given in Table 2 for men and women. The correlations are given as non-energy adjusted and non-deattenuated, as well as energy-adjusted and deattenuated, the latter being generally higher. Since nutrient densities expressed as percentage of calories gave results very similar to those obtained using the energy-adjusted nutrient values, only the latter are presented<sup>1</sup>. The following results were obtained for the energy-adjusted nutrients and deattenuated correlations. As is generally expected, caffeine and alcohol gave relatively high correlations in men and women (respective correlations between FFQ2 and diet recalls for alcohol: 0.71 and 0.76, for caffeine: 0.85 and 0.86). The correlation coefficients between FFQ1 and the mean of 24-hour diet recalls for selected dietary components ranged from 0.23 (vitamin A) to 0.81 (caffeine) in men and between 0.29 ( $\beta$ -carotene) to 0.76 (caffeine) in women, with a median correlation value of 0.47 in men and 0.50 in women. The correlation coefficients between FFQ2 and the mean of 24-hour diet recalls for selected dietary components ranged from 0.23 (vitamin A) to 0.86 (caffeine) in men and between 0.20 (vitamin A) to 0.80 (total fat) in women, with a median correlation value of 0.49 in men and 0.53 in women. Except for vitamin A in men ( $P = 0.006$ ) and total carotenes in women ( $P = 0.005$ ), all correlations were statistically significant with  $P < 0.0005$ . For some nutrients where the ratio of intra- to inter-individual variability was low, little improvement in deattenuated correlation coefficients was observed.

Table 2 also shows correlations between energy-adjusted nutrients estimates from the two FFQs. The two FFQs showed significant correlations with  $P < 0.0001$  for all nutrients and a median correlation value of 0.64 for men and 0.63 for women.

## Discussion

The correlations observed in our study are comparable to those reported by others and to our previous work<sup>8</sup>. Values for energy-adjusted protein, fat, saturated fat and

carbohydrate ranged from 0.47 to 0.59 in the Nurses' Health Study<sup>17</sup>, compared with our range of 0.63 to 0.80 in women. In the Hawaii–Los Angeles multiethnic cohort study, deattenuated correlations in the range of 0.67 to 0.80 were reported for these nutrients for white women<sup>1</sup>. Deattenuated correlations of 0.54 to 0.84 for macronutrients (protein, carbohydrates, total fat) were reported for 134 men and women in the EPIC–Potsdam Study<sup>18</sup>. Our correlations (based on FFQ2) for men ranged from 0.50 to 0.71 for the same nutrients, which compares favourably with those reported in the Hawaii study for white men (correlations 0.44 to 0.71). The correlations for energy-adjusted fat ranged from 0.09 to 0.87 for men and women in the EPIC study<sup>19</sup>, compared with 0.56 and 0.80 found in our study. Our correlations for energy were similar to those in other reports<sup>1,20,21</sup>. With a standard measurement error model, correlations for energy between four telephone-administered 24-hour recalls and three different FFQs (US National Cancer Institute's Diet History, the Block FFQ and the Willett FFQ) were 0.48, 0.45 and 0.18, respectively, for women and 0.49, 0.45 and 0.21, respectively, for men<sup>22</sup>. The deattenuated correlations for 139 participants in a chemoprevention trial in Arizona ranged from 0.25 to 0.67 for various nutrients<sup>21</sup>. Although our correlations for vitamin A of 0.21 and 0.23 for men and women, respectively, are low, they are significant with  $P < 0.01$ , suggesting reasonably good agreement between FFQ and 24-hour recalls. Other studies have also reported somewhat low correlations for this vitamin (correlations of 0.25, 0.36, 0.48 and 0.61)<sup>17,21,23,24</sup>.

Several limitations underlying the use of a calibration study for assessing the performance of a questionnaire have been recognised, the most important being the assumption that the 24-hour recalls provide unbiased estimates of true diet and that errors in 24-hour recalls are uncorrelated with errors in the FFQs<sup>1</sup>. The variability between multiple 24-hour recalls is expected to capture the long-term intake of a person. However, repeat interviews are subject to respondents' fatigue and increasing lack of interest, which might affect their reporting of intake. The linear relationship between 24-hour recalls and FFQ was examined by plotting nutrients in absolute and in quadratic terms in our study<sup>25,26</sup>. None of the nutrients appeared to violate the linearity assumption.

As with other reports<sup>1</sup>, our 24-hour recalls seem to underreport energy and macronutrient intakes compared with the FFQs. Energy expenditure measurements using doubly labelled water methods suggest that self-assessment of diet by weighed food records underestimates energy intakes<sup>20,27</sup>. This may imply that nutrient estimates from 24-hour recalls may underestimate true intakes while intakes estimated using FFQs may be less of an overestimate of true intake when compared with diet recalls. A study in 19 elderly people consuming diets of known composition showed that their FFQ

underestimated fat, saturated fat, monounsaturated fat and protein intakes, and overestimated carbohydrate intake<sup>28</sup>, suggesting that the assessment varies by the type of energy source. Estimates from diet records better agreed with the actual intakes than did estimates from the FFQ in that study. Although the values from recalls were only slightly lower than those from FFQs for the major macronutrients in our study, FFQs measured micronutrient intakes 1.2 to 1.7 times higher than recalls, suggesting that fruit and vegetable intakes might be overreported by FFQs. The high correlation between the two FFQs for vitamin A and other vitamins but the low correlation with diet recalls suggests that there might be a systematic overestimate of vitamins by FFQs. If so, methods to correct this should be explored such as adjusting the nutrient values or amounts per serving of food item on the FFQ itself. This overestimation, however, should have a minimal impact when the entire study population is subjected to the same method.

One would have expected better correlations between FFQ2 and diet recalls, compared with FFQ1 and recalls, because of the possible training effect experienced by participants as a result of completing diet recalls in between the two FFQs. However, this was not seen consistently across nutrients or sex of the participants. A number of other studies showed that a second FFQ nearly always produces lower nutrient estimates and the reason for this is not clear<sup>21,23</sup>. In a methodological review, Kaaks *et al.*<sup>29</sup> concluded that when relative risks are estimated for scaled, absolute intake differences rather than for quantile categories, a 'calibration' study based on only a single day's food intake record (but generally on a larger number of subjects) can provide sufficient reference information to (1) correct relative risk estimates for biases due to measurement error and (2) account for statistical power losses when estimating the sample size requirements of the cohort. A major advantage of calibration studies based on this single-day-per-subject design is that they can be conducted on a representative sample of cohort participants more easily than validity studies in which reference measurements are repeated.

The implications of these findings for epidemiological studies of diet and cancer have been considered by calculating hypothetical 'corrected' odds ratios (ORs), which demonstrate the potential impact of such measurement errors on relative risk estimates (using the method of Rosner *et al.*<sup>30</sup> but not shown here)<sup>31</sup>. The correction process demonstrated that, after accounting for error in exposure measurement, an OR observed in a particular study would depart from unity than an uncorrected estimate but would also be more unstable as indicated by the widening of the associated confidence intervals (CIs). For example, if a study used the FFQ and found an OR of 1.5 (95% CI: 0.9–2.5) for a high intake of total fat (for which the diet recall–FFQ correlation coefficient was 0.61), then the true association may have been much

stronger as indicated by a 'corrected' OR of 1.8 (as would have been detected if the diet recalls were employed) but would have been less precise as indicated by wider confidence intervals (95% CI: 0.8–3.9). While the impact of exposure misclassification is generally well understood, this correction process demonstrates that substantial attenuating effects can arise in the presence of a large amount of measurement error, particularly when relatively strong associations are detected (i.e. ORs different from 1.0). The magnitude of the bias was directly related to the extent of misclassification: the higher the diet recall–FFQ correlations, the lower the misclassification errors. The differences between these hypothetical observed and corrected ORs provide an indication of the magnitude of bias and loss of precision that may result from the measurement error that is known to arise in dietary studies. Kipnis *et al.* proposed a new measurement error model to accommodate person-specific bias in the reference measure and its correlation with systematic error in the FFQ<sup>32,33</sup>. Using a urinary nitrogen biomarker measurements and dietary report measurements from a study in the United Kingdom, Kipnis *et al.* suggested that, for the data studied by them, measurement error in the FFQ could lead to a 51% greater attenuation of the true nutrient effect and the need for a 2–3 times larger study than would be estimated by the standard approach.

In conclusion, the FFQ method gives acceptable levels of nutrient/food component estimates, as assessed in this calibration study against the mean of three 24-hour diet recalls, when limited to energy-adjusted and deattenuated values. The correlations are somewhat low prior to deattenuation. The implications of measurement errors in interpreting the results of FFQ-based epidemiological studies to detect important diet–disease associations are many and should be considered when reporting results. The calibration and correlation characteristics of FFQs collected for this cohort were similar to those in other, similar studies.

### Acknowledgements

The study was supported by funds from the Canadian Institute of Health Research, formerly the Medical Research Council of Canada. Fieldwork co-ordination support was provided by Joan Antal.

### References

- 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.
- Carroll RJ, Pee D, Freedman LS, *et al.* Statistical design of calibration studies. *Am. J. Clin. Nutr.* 1997; **65**(Suppl.): 1187S–9S.
- Willett W. *Nutritional Epidemiology*. New York: Oxford University Press, 1998.
- Stram DO, Longnecker MP, Shames L, *et al.* Cost effective design of a diet validation study. *Am. J. Epidemiol.* 1995; **142**: 353–62.
- Kaaks R, Riboli E. Validation and calibration of dietary intake measurements in the EPIC project: methodological considerations. *Int. J. Epidemiol.* 1997; **26**(Suppl. 1): S15–25.
- Jain M, Harrison L, Howe GR, *et al.* Evaluation of a self-administered dietary questionnaire for use in a cohort study. *Am. J. Clin. Nutr.* 1982; **36**: 931–5.
- Miller AB, Howe GR, Wall C. The national study of breast cancer screening. *Clin. Invest. Med.* 1981; **4**: 227–58.
- Jain M, Howe GR, Rohan T. Dietary assessment in epidemiology: comparison of a food frequency and a diet history questionnaire with 7-day food records. *Am. J. Epidemiol.* 1996; **143**: 953–60.
- Health Canada. *Canadian Nutrient File*. Ottawa: Nutrition Research Division, Health Canada, 1997.
- Mangels AR, Holden JM, Beecher GR, *et al.* The carotenoid content of fruits and vegetables: an evaluation of analytic data. *J. Am. Diet. Assoc.* 1993; **93**: 284–96.
- National Cancer Institute. *NFCS Classification System. National Food Consumption Survey. US Department of Agriculture*. Bethesda, MD: National Cancer Institute, 1993.
- Willett WC, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *Am. J. Epidemiol.* 1986; **124**: 17–27.
- Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am. J. Epidemiol.* 1988; **127**: 377–86.
- Hu FB, Rimm E, Smith-Warner SA, *et al.* Reproducibility and validity of dietary patterns associated with a food-frequency questionnaire. *Am. J. Clin. Nutr.* 1999; **69**: 243–9.
- Statistics Canada. Personal health practices: smoking, drinking, physical activity and weight. *Health Reports* 1999; **11**: 83–90.
- Bright-See E, Catlin G, Godin G. Assessment of the relative validity of the Ontario Health Survey Food Frequency Questionnaire. *J. Can. Diet. Assoc.* 1994; **55**: 33–8.
- Willett WC, Sampson L, Stampfer MJ, *et al.* Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol.* 1985; **122**(1): 51–65.
- Kroke A, Klipstein-Grobusch K, Voss S, *et al.* Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am. J. Clin. Nutr.* 1999; **70**(4): 439–47.
- Kaaks R, Slamani N, Riboli E. Pilot phase studies on the accuracy of dietary intake measurements in the EPIC project: overall evaluation of results. *Int. J. Epidemiol.* 1997; **26**: S26–36.
- Black AE. Physical activity levels from a meta-analysis of doubly labeled water studies for validating energy intake as measured by dietary assessment. *Nutr. Rev.* 1996; **54**: 170–4.
- Martínez 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. Biomark. Prev.* 1999; **8**: 941–6.
- 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**(12): 1089–99.
- 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**: 1114–26.
- Hankin JH, Wilkens LR, Kolonel LN, *et al.* Validation of a

- quantitative diet history method in Hawaii. *Am. J. Epidemiol.* 1991; **133**: 616–28.
- 25 Lee J. Use of regression adjustment for ascertaining nutrient intakes. *J. Am. Diet. Assoc.* 1981; **78**: 245–9.
- 26 Thomas D, Stram D, Dwyer J. Exposure measurement error: influence on exposure–disease relationships and methods of correction. *Ann. Public Health* 1993; **14**: 69–93.
- 27 Black AE, Prentice AM, Goldberg G, *et al.* Measurements of total energy expenditure provide insights into the validity of dietary measurements of energy intake. *J. Am. Diet. Assoc.* 1993; **93**: 572–9.
- 28 Schaefer EJ, Augustin JL, Schaefer MM, *et al.* Lack of efficacy of a food-frequency questionnaire in assessing dietary macronutrient intakes in participants consuming diets of known composition. *Am. J. Clin. Nutr.* 2000; **71**: 746–51.
- 29 Kaaks R, Riboli E, van Staveren W. Calibration of dietary intake measurements in prospective cohort studies. *Am. J. Epidemiol.* 1995; **142**(5): 548–56.
- 30 Rosner B, Willett W, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat. Med.* 1989; **8**: 1051–69.
- 31 Jain M, McLaughlin J. Validity of nutrient estimates by food frequency questionnaires based either on exact frequencies or categories. *Ann. Epidemiol.* 2000; **10**: 354–60.
- 32 Kipnis V, Carroll RJ, Freedman LS, *et al.* Implications of a new dietary measurement error model for estimation of relative risk: application to four calibration studies. *Am. J. Epidemiol.* 2000; **152**(5): 494–6.
- 33 Kipnis V, Midthune D, Freedman LS, *et al.* Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am. J. Epidemiol.* 2001; **153**(4): 394–403.