

Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort

Ingegerd Johansson^{1,*}, Göran Hallmans¹, Åsa Wikman¹, Carine Biessy², Elio Riboli² and Rudolf Kaaks²

¹Department of Nutritional Research, Umeå University, Sweden: ²Unit of Analytical Epidemiology, Program of Nutrition and Cancer, International Agency for Research on Cancer, Lyon, France

Submitted 14 August 2001: Accepted 29 October 2001

Abstract

Objectives: To evaluate the reproducibility of, and to compare and calibrate, diet measures by the Northern Sweden 84-item food-frequency questionnaire (FFQ) with measures from 24-hour diet recalls (24-HDR).

Design: Randomly selected respondents ($n = 246$) from the EPIC (diet–cancer) and MONICA (diet–cardiovascular disease) study cohort in Northern Sweden were invited to answer the FFQ twice over a one-year interval (FFQ1 and FFQ2), and to complete ten 24-hour recalls (reference method) in the months between. Plasma β -carotene concentrations were determined from a subset of 47 participants.

Setting: Västerbotten and Norrbotten, Northern Sweden.

Participants: Ninety-six men and 99 women, who completed the study.

Results: The reproducibility of the FFQ was high in terms of both mean energy and nutrient intakes and relative ranking of participants by intake levels (median Pearson correlation of 0.68). Moderately higher food intake frequencies were recorded by FFQ1 compared with 24-hour recalls for dairy products, bread/cereals, vegetables, fruits and potato/rice/pasta, whereas meat, fish, sweet snacks and alcoholic beverage intakes were lower. The median Spearman coefficient of correlation between FFQ1 and the average of ten 24-HDR measurements was 0.50. Daily energy and nutrient intakes were similar for FFQ1 and 24-HDR measurements, except for fibre, vitamin C, β -carotene and retinol (FFQ1 > 24-HDR) and sucrose and cholesterol (FFQ1 < 24-HDR). Pearson coefficients of correlation between FFQ1 and 24-HDR corrected for attenuation due to residual day-to-day variation in the 24-HDR measurements ranged from 0.36 to 0.79 (median 0.54). Adjustment for energy had only very moderate effects on the correlation estimates. Calibration coefficients estimated by linear regression of the 24-HDR on the FFQ1 measurements varied between 0.30 and 0.59 for all nutrients except alcohol, which had calibration coefficients close to 1.0. These low calibration coefficients indicate that relative risk estimates corresponding to an absolute difference in dietary intake levels measured by the FFQ will generally be biased towards 1.0. Plasma β -carotene levels had a Pearson coefficient of correlation of 0.47 with the 24-HDR measurements, and of 0.23 with FFQ1 measurements.

Conclusions: The Northern Sweden FFQ measurements have good reproducibility and an estimated level of validity similar to that of FFQ measurements in other prospective cohort studies. The results from this study will form the basis for the correction of attenuation and regression dilution biases in relative risk estimates, in future studies relating FFQ measurements to disease outcomes.

Keywords

EPIC
MONICA
Sweden
Food-frequency questionnaire
24-Hour recalls
 β -Carotene
Validation
Calibration

Dietary intake estimates from self-administered food-frequency questionnaires (FFQs) are commonly used in epidemiological studies^{1–8}, but, like any other type of dietary intake measurement, suffer from systematic and random errors. As such errors generally cause bias in relative risk estimates, it is crucial to estimate the validity of the intake assessments, so as to enhance the interpretation

of estimated diet–disease associations and to improve the translation of such associations into dietary recommendations. Using the results from validation or calibration sub-studies it may also be possible to adjust relative risk estimates for (at least part of) biases induced by errors in the dietary exposure assessments. Thus, the validation and calibration of FFQ measurements of habitual dietary

*Corresponding author. Email ingeagerd.johansson@odont.umu.se

intakes are an essential part of large-scale epidemiological studies, and especially for prospective cohort studies in which the dietary intake assessments will eventually be related to a variety of chronic disease outcomes.

The population of the most arctic counties of Sweden (Västerbotten and Norrbotten) is part of the global cohort for the prospective European Prospective Investigation into Cancer and Nutrition (EPIC)⁹ as well as for the WHO MONICA cardiovascular study¹⁰. A self-administered FFQ was chosen for monitoring of dietary intakes in the Northern Sweden EPIC and MONICA centres (Northern Sweden Health–Diet Study, NSHDS). Data collection started in 1985 within the Västerbotten cardiovascular intervention project^{11,12}, in which all adults within the age span of 30–75 years in Västerbotten county are invited every 10 years for a general health examination. On each visit, dietary intake assessments are collected from all participants, and at present about 130 000 FFQ measurements of habitual dietary intake have been obtained from about 90 000 cohort participants.

The aim of the present study was to evaluate the reproducibility of the self-administered FFQ used in these northern Swedish cohorts, and to compare and calibrate FFQ measurements with 24-hour diet recalls (24-HDR) and plasma β -carotene levels. The study was conducted on a representative sub-sample of the EPIC and MONICA cohorts.

Study population and methods

Study population

A random sub-sample in individuals attending the ongoing Västerbotten county cardiovascular disease (CVD) study¹² in 1992 was requested by mail to participate in the present study. Out of a total of 246 individuals invited, 43 gave a negative response because they could not find the free time to participate in the study. Thus, 102 men and 101 women, equally distributed over the ages 30, 40, 50 and 60 years, consented to participate. The participants were recruited in equal proportions from the coastal (more urban) and the inland/mountain (more rural) areas of Västerbotten county.

The study was approved by the Ethics Committee for Human Experiments at Umeå University, Sweden.

Food-frequency questionnaire (test method)

The participants were requested to complete one set of a self-administered FFQ immediately before initiation of the 24-HDR period (in the first quarter of 1993, FFQ1) and a second identical set at the end (in the first quarter of 1994, FFQ2). The FFQ was designed to be semi-quantitative and optically readable for data input. Frequencies of consumption of 84 food items were reported on an increasing, nine-level scale, including never, maximum once a month, 1–3 times per month, once a week, 2–3 times a week, 4–6 times a week, once a day, 2–3 times a day, and 4 or more

times a day. The questionnaire included eight questions on the frequency of consumption of various types of fats used for spreading on bread or cooking, nine on milk and other dairy products, eight on bread and cereals, 10 on fruit, greens and root vegetables, and nine on soft drinks and sugar-containing snacks. Five questions on spirits, wine and beer consumption were included in a list of beverages. Twenty-nine of the remaining 35 questions recorded intake of potato, rice, pasta, meat and fish, and six were on varied items, such as salty snacks, coffee, tea and juice. The respondents indicated their average portion of (1) potato/pasta/rice, (2) vegetables and (3) meat/ground meat/sausages by comparison with four colour photos illustrating four plates with increasing portion sizes of potato, vegetables and meat. For the other food items, we assumed a standard portion size value (see below). The questionnaires were returned by mail, and the interviewer reviewed it and resolved any ambiguities by telephone.

The reported frequencies of consumption were converted to number of intakes per day, and energy and nutrient intakes were calculated by multiplying these frequencies by a portion size value and by the energy or nutrient content from a food composition database from the National Food Administration¹³. Portion sizes used were those indicated on the photos, or were natural portion sizes such as an apple, or average portion sizes for gender and age according to a national survey^{14,15}. The energy and nutrient contents were calculated using the software MATs (Rudans Lättdata, Sweden).

24-hour recalls (reference method)

Because many participants lived at a large distance from the research centre in Umeå, 24-hour diet recall interviews were conducted by telephone^{16,17}. The interviews were done by nine trained interviewers, i.e. all interviewers had academic education in nutrition, had participated in nutritional research previously, and were trained to structure the interviews in a standardised fashion. The interviews recorded intake of food, beverages and supplements during the preceding 24 hours, at 10 unannounced occasions per subject. The ten 24-HDR were equally dispersed over the year and covered all weekdays. The estimation of food portions was facilitated by the use of a booklet with full-size portion size pictures (National Food Administration, Uppsala, Sweden^{18,19}), mailed in advance to the participants. These pictures included five options for portions of food on plates, five options for spread on a knife, and a set of schematic drawings ($n = 35$) to indicate thickness and sizes of various types of food items. Standard household measures were used for the estimation of portion sizes of food items not included in the booklet.

The daily energy and nutrient intake and intake frequencies of food items in 10 food groups were calculated using the software MATs (Rudans Lättdata,

Sweden) employing the database from the National Food Administration¹⁵.

Exclusions and final participation rate

Six participants completed less than three of the ten 24-HDR and two more did not return the baseline frequency questionnaire; these participants were excluded from further analyses. Thus, in total, 96 men (94%) and 99 women (99%) of the included participants completed the 24-HDR as well as the baseline food-frequency questionnaire. These 195 participants represented 79% of those initially invited to participate in the validation study.

Plasma β -carotene measurements

Forty-seven participants living within commuting distance of the city of Umeå volunteered to donate a blood sample within three weeks after returning FFQ1. None of these took β -carotene supplementation. Samples, drawn after overnight fasting and two days after ceasing any vitamin supplementation, were protected from light and air exposure and stored in liquid nitrogen. Plasma concentration of β -carotene was determined by reversed-phase high-performance liquid chromatography (HPLC) separation²⁰. Briefly, after hexane extraction (0.15 g l⁻¹ of butylated hydroxytoluene (BHT) added), evaporation under nitrogen and re-dissolving in ethanol, lipid soluble vitamins were separated on a Spherisorb ODS-2 column (C-18, 5 μ m, HPLC Teknik AB, Umeå, Sweden) using a mobile phase of 95% methanol and 5% tetrahydrofuran. Tocopherylacetate (Sigma Chemicals, St Louis, USA) was used as an internal standard and β -carotene (Sigma Chemicals) as an external standard.

Estimation of underreporting relative to daily energy requirements

In order to identify participants who might have been systematically underestimating their total food consumption, we estimated the participants' basal metabolic rate (BMR) from their age and body weight, according to the Food and Agriculture Organization (FAO)/World Health Organization (WHO) equations for adult men and women²¹, and calculated the ratio of total energy intake to BMR. For participants in approximate energy balance, this ratio corresponds to the physical activity level (PAL), defined as the ratio of total energy expenditure relative to basal metabolic rate²². A PAL value below 1.27²¹, corresponding to a value slightly higher than the survival limit when bed-bound, was applied as cut-off for evident underreporting.

Handling of missing values

None of the respondents had left more than 10 missing or illegible frequency questions. Missing answers were mainly present for food groups with several alternative items where only one of these items was eaten frequently,

such as in the list of milk products or types of fats for bread or cooking. Missing values were therefore interpreted as being never eaten.

Statistical analyses

The Statistical Analyses System (SAS, version 6.12, SAS Institute, Carry, MO, USA) was used for statistical evaluations. Evaluations were done separately for men and women. Nutrient and food intake data were log-transformed to normalise population frequency distributions. In addition, nutrient densities were calculated as amount (g, mg) per 1000 kcal.

Geometric mean intakes and 95% confidence intervals were calculated for food frequencies, absolute nutrient intakes and nutrient densities, for both the FFQs and the individuals' average 24-HDR. The geometric means were calculated by taking the exponential of the arithmetic mean of log-transformed values. Ratios of geometric means were calculated to estimate changes in mean intake levels between FFQ1 and FFQ2, or to estimate mean intakes of FFQ1 compared with the 24-HDR measurements. These ratios were calculated as the exponential of the mean difference between log-transformed FFQ1 and 24-HDR measurements.

For the food-frequency data, which formed somewhat irregular frequency distributions, with peaks of non-consumers (zero values) and values clustered around the eight additional frequency categories, the accuracy of relative ranking by intake level was evaluated by Spearman coefficients of correlation between the measurements by FFQ1 and FFQ2, as well as between FFQ1 and the individuals' average 24-HDR measurements. For the (log-transformed) nutrient intakes and nutrient densities, which followed smooth and approximately normal distributions, we calculated Pearson's product moment correlations (r_{QR}) between measurements by FFQ1 and by 24-HDR, and corrected these for attenuation due to residual day-to-day variations in the individuals' average 24-hour diet recalls. The de-attenuated Pearson correlation coefficients (ρ_{QT}) were calculated by multiplying the crude correlation coefficient r_{QR} with the factor $[1 + (\sigma_w^2/\sigma_b^2)/k]^{0.5}$, where k represents the number of repeated recall days, σ_w^2 the intra-individual and σ_b^2 the inter-individual variance²³. The intra- and inter-subject variance components were determined by a random effects model with the recall intake as dependent and subject identification number as the independent variable. The classification of participants by relative intake levels was also compared between FFQ1 and 24-HDR by examining the cross-classification of participants over quartiles of intake.

Calibration analyses were performed by linear regression models with FFQ1 measurements as independent (predictor) variable and 24-HDR intakes as the dependent variable, estimating a linear 'calibration factor' (λ) as the slope of this regression line^{24,25}. These analyses

were performed on both the log-transformed absolute intake levels and the nutrient densities. Calibrated FFQ measurements (FFQ') were calculated as:

$$\text{FFQ}' = 24\text{-HDR}_{\text{mean}} + \lambda(\text{FFQ} - \text{FFQ}_{\text{mean}}),$$

where $24\text{-HDR}_{\text{mean}}$ and FFQ_{mean} are the population mean intake levels as measured by the 24-HDR and FFQ, respectively. As a less parametric form of 'calibration' approach, we also compared the geometric mean 24-HDR frequencies of food consumption and the mean 24-HDR intakes of nutrients for quartiles of the corresponding FFQ1 measurement.

Results

Representativeness of the study population

Prevalence of tobacco use and averages of selected medical CVD risk factors were similar among participants in the sub-sample and the global study cohort^{12,26}. Thus, among participants in the study sub-sample, 20 and 27% of the men and women, respectively, smoked daily compared with 19 and 27% in the global cohort; average body mass index (BMI) was 25.3 and 24.9 kg m⁻² compared with 26.0 and 25.3 kg m⁻²; total cholesterol 5.6 and 5.5 mmol l⁻¹ compared with 6.0 and 5.9 mmol l⁻¹; systolic blood pressure 129 and 126 mmHg compared with 129 and 124 mmHg, and diastolic blood pressure 80 and 78 mmHg compared with 82 and 76 mmHg.

Reproducibility of FFQ measurements

Table 1 shows the reproducibility of the FFQ measurements (FFQ1 vs. FFQ2) in terms of mean intake level and Pearson correlation coefficients. Mean energy and nutrient intake levels were similar for FFQ1 and FFQ2 (geometric mean ratios between 0.93 and 1.05) except for alcohol intake in women, where FFQ1 showed lower values than FFQ2. Pearson coefficients of correlation between FFQ1 and FFQ2 measurements ranged from 0.53 to 0.99, and were mostly above or equal to 0.65.

Comparisons of levels of consumption and underreporting between FFQ1 and 24-HDR

In both genders FFQ1 measurements of dairy products, bread/cereals, vegetables, fruits and potato/rice/pasta were higher than intakes measured by 24-HDR (geometric mean ratio > 1.10) (Table 2), whereas meat, fish, sweet snacks and alcoholic beverage intake frequencies were lower (geometric mean ratio < 0.90). Fewer respondents (three men and nine women) reported zero intake of alcohol by FFQ1 than by the ten 24-HDR (22 men and 25 women). This can be explained by the fact that the 24-HDR covered only one weekend, during which alcohol consumption is usually concentrated. Further evaluations on alcohol consumption were restricted to participants with non-zero consumption levels, i.e. comparisons of

Table 1 Comparison of energy and nutrient intake measured by the identical, but distributed one year apart, 'Northern Sweden' FFQ1 and FFQ2

Nutrient*	Gender†	Ratio FFQ1/FFQ2	Pearson correlation§
		Mean (95% CI)‡	<i>r</i> _{QR} (95% CI)
Energy	Men	1.01 (0.96–1.00)	0.67 (0.55–0.77)
	Women	0.99 (0.95–1.02)	0.76 (0.66–0.83)
Carbohydrates	Men	1.03 (0.97–1.08)	0.67 (0.54–0.77)
	Women	1.01 (0.96–1.05)	0.76 (0.66–0.83)
Sucrose	Men	1.00 (0.93–1.07)	0.81 (0.73–0.87)
	Women	0.94 (0.88–1.01)	0.75 (0.65–0.82)
Fat _{total}	Men	1.01 (0.95–1.07)	0.62 (0.48–0.73)
	Women	0.98 (0.93–1.02)	0.71 (0.60–0.80)
Fat _{saturated}	Men	1.00 (0.93–1.07)	0.59 (0.44–0.71)
	Women	0.96 (0.92–1.01)	0.70 (0.58–0.79)
Cholesterol	Men	1.00 (0.93–1.07)	0.53 (0.37–0.66)
	Women	0.98 (0.93–1.02)	0.68 (0.55–0.77)
Protein	Men	0.99 (0.92–1.05)	0.62 (0.48–0.73)
	Women	0.96 (0.94–1.01)	0.69 (0.58–0.79)
Alcohol	Men	0.99 (0.92–1.06)	0.99 (0.98–0.99)
	Women	0.79 (0.59–1.05)	0.88 (0.82–0.92)
Fibre	Men	1.03 (0.97–1.10)	0.70 (0.58–0.79)
	Women	1.04 (0.99–1.09)	0.74 (0.64–0.82)
Vitamin C	Men	1.03 (0.94–1.13)	0.65 (0.52–0.75)
	Women	1.02 (0.94–1.11)	0.68 (0.56–0.78)
Vitamin E	Men	1.04 (0.98–1.10)	0.65 (0.52–0.75)
	Women	1.02 (0.98–1.07)	0.69 (0.57–0.78)
β-Carotene	Men	1.02 (0.91–1.13)	0.71 (0.60–0.80)
	Women	1.05 (0.96–1.15)	0.84 (0.77–0.89)
Retinol	Men	1.04 (0.95–1.13)	0.58 (0.43–0.70)
	Women	1.02 (0.95–1.09)	0.73 (0.62–0.81)
Calcium	Men	0.98 (0.90–1.05)	0.62 (0.48–0.73)
	Women	0.93 (0.86–1.00)	0.68 (0.55–0.77)
Iron	Men	1.01 (0.95–1.06)	0.63 (0.49–0.74)
	Women	1.01 (0.96–1.05)	0.68 (0.56–0.77)

* Mean intake values based on FFQ1 are presented in Tables 3 and 4.

† Ninety-six men and 99 women were included in all calculations. The proportions reporting zero intake of alcohol were the same by FFQ1 and FFQ2 for both genders (cf. footnote in Table 3).

‡ Geometric mean (95% CI) of individual ratios between intakes as recorded by FFQ1 and FFQ2.

§ Pearson correlation coefficients (log-transformed data).

FFQ1 with 24-HDR. The proportions classified as under-reporters were high by FFQ1 (42.7% and 49.5%, men and women, respectively) as well as 24-HDR (55.7% and 51.5%, men and women, respectively).

Spearman correlation coefficients between food intake frequencies assessed by FFQ1 and by 24-HDR were between 0.42 and 0.69 for most food groups except potato/pasta, meat and fish, which had values below 0.40 (Table 2). Cross-classification into 24-HDR quartiles of participants in the highest and lowest FFQ1 quartiles, respectively, placed more than 68% of the participants into either the two lowest or the two highest 24-HDR quartiles, and for none of the food groups were there more than four participants grossly misclassified, e.g. into the most opposite quartile (data not shown). Consumption frequencies of coffee and tea were similarly measured by FFQ1 and 24-HDR regardless of preparation method (two variants of coffee preparation are dominant in the study area – extraction by filtering or by boiling), and correlations between FFQ1 and 24-HDR were between

Table 2 Comparison of food intake frequencies measured by the 'Northern Sweden' FFQ1 and 24-hour dietary recalls (24-HDR)

Food group	Gender*	Mean intake†		Ratio FFQ1/24-HDR	Spearman correlation§	Means by FFQ1 quartiles¶				Calibration coefficient
		Mean _{FFQ}	Mean _{24-HDR}	Mean (95% CI)‡	<i>r</i> _{QR}	1st	2nd	3rd	4th	λ -value (95% CI)
Fats on bread**	Men	9.5	10.8	0.88 (0.73–1.06)	0.56	7.4	11.3	16.4	16.0	0.63 (0.53–0.73)
	Women	12.3	11.8	1.05 (0.92–1.20)	0.50	7.1	12.4	12.2	16.8	0.47 (0.38–0.57)
Dairy products	Men	20.8	18.5	1.12 (1.01–1.24)	0.45	14.2	18.6	18.4	25.4	0.47 (0.31–0.62)
	Women	20.3	18.0	1.12 (1.01–1.25)	0.42	14.1	17.2	18.8	23.9	0.38 (0.23–0.53)
Bread/cereals	Men	25.1	22.7	1.11 (1.01–1.22)	0.53	17.4	18.8	24.2	30.5	0.40 (0.27–0.54)
	Women	25.5	21.3	1.19 (1.11–1.29)	0.56	16.2	21.0	22.5	27.0	0.39 (0.26–0.52)
Vegetables	Men	6.1	3.8	1.61 (1.27–2.05)	0.47	2.8	3.3	4.2	7.8	0.80 (0.45–1.15)
	Women	10.4	6.5	1.61 (1.40–1.84)	0.51	4.3	5.4	8.8	9.2	0.57 (0.36–0.77)
Fruits	Men	6.0	2.4	2.56 (1.91–3.43)	0.62	0.8	2.4	4.8	6.3	1.35 (1.02–1.69)
	Women	9.2	5.4	1.71 (1.38–2.12)	0.68	2.0	6.1	7.9	11.2	0.99 (0.71–1.28)
Potato/rice/pasta	Men	8.5	6.4	1.32 (1.19–1.46)	0.26	5.9	6.0	6.9	7.0	0.17 (0.02–0.33)
	Women	8.4	5.5	1.52 (1.39–1.65)	0.36	4.3	5.5	6.2	6.4	0.38 (0.19–0.58)
Meat	Men	4.8	6.5	0.74 (0.66–0.82)	0.27	6.7	5.3	6.2	7.8	0.32 (0.08–0.55)
	Women	3.9	4.5	0.87 (0.76–0.99)	0.44	3.8	4.8	5.7	7.1	1.01 (0.87–1.15)
Fish	Men	1.3	1.8	0.76 (0.63–0.93)	0.15	1.9	1.2	1.4	2.5	0.11 (–0.05–0.28)
	Women	1.4	1.5	0.99 (0.83–1.17)	0.21	1.3	1.3	1.6	1.6	0.19 (0.03–0.34)
Sweets	Men	12.7	20.5	0.62 (0.55–0.69)	0.69	11.3	19.3	26.4	33.2	0.57 (0.46–0.69)
	Women	8.3	18.4	0.45 (0.40–0.50)	0.62	14.0	20.7	22.8	34.7	0.48 (0.37–0.59)
Alcoholic beverages	Men	1.6	3.2	0.69 (0.56–0.84)	0.45	1.7	2.5	2.6	5.5	0.68 (0.35–1.01)
	Women	1.1	1.9	0.76 (0.64–0.91)	0.56	1.1	2.0	1.3	3.7	0.70 (0.45–0.95)

*Ninety-six men and 99 women are included in calculations of nutrients. For alcohol, zero reporters (three men and nine women for FFQ1; 22 men and 25 women for 24-hour recalls) were excluded.

† Geometric mean number of intakes per week as obtained by FFQ1 (mean_{FFQ}) or 24-HDR (mean_{24-HDR}).

‡ Geometric mean (95% CI) of individual ratios between number of intakes per day as obtained by FFQ1 and 24-HDR.

§ Spearman correlation coefficients between intake frequencies obtained by FFQ1 and by 24-HDR

¶ Geometric mean of 24-hour recall measured intakes in FFQ1 ranked quartiles.

|| Calibration coefficient equal to the slope (λ , with 95% CI) from the linear regression of the 24-HDR measured intake on the corresponding intake measured by FFQ1.

** All types of fats spread on bread.

0.72 and 0.84 (data not shown). Frequency of food consumption according to the 24-HDR measurements increased in a monotonic manner over quartiles of the corresponding FFQ1 measurements (Table 2).

Comparisons of energy and nutrient intakes by FFQ1 and 24-HDR

Gender-specific intakes of energy and nutrients (mean amount per day) as measured by FFQ1 and 24-HDR, as well as the mean ratio (95% confidence interval (CI)) between these measures, are shown in Table 3 for energy and macronutrients and in Table 4 for micronutrients; Tables 5 and 6 show the corresponding results for nutrient densities. Mean intake measurements by FFQ1 and by 24-HDR were similar (geometric mean ratios FFQ1/24-HDR = 0.90–1.10) for energy and most nutrients (Tables 3 and 4), with the exception of sucrose and cholesterol (FFQ1/24-HDR < 0.90) and fibre, vitamin C, β -carotene and retinol (FFQ1/24-HDR > 1.10). The pattern was very similar for nutrient densities (Tables 5 and 6).

Pearson correlation coefficients between FFQ1 and individuals' average 24-HDR measurements (r_{QR}) ranged from 0.61 (total fat in men, fibre in women) to 0.28 (retinol in women) (Tables 3 and 4). Half of the correlation coefficients were equal to or greater than 0.50, 12 values were between 0.40 and 0.49, and four values (for retinol in both genders, sucrose and cholesterol in women) were

below 0.35. The low correlation for retinol (vitamin A) can be explained by the fact that, to a large extent, average intake levels are determined by the irregular consumption of foods (e.g. liver) extremely rich in this nutrient, which may not have been captured accurately by only ten 24-HDR per subject²⁷. Cross-classification of the participants in the lowest or highest FFQ1 quartiles placed more than 70% into the two lowest or two highest 24-HDR quartiles, respectively. Extreme misclassification between the top and bottom quartiles was seen in no more than one to three participants for each nutrient. The 24-HDR measurements increased in a monotonic manner over quartiles of the corresponding FFQ1 measurements for most nutrients (Tables 3 and 4).

The ratio of within- to between-subject variation in the ten 24-hour diet recalls ranged from 1.05 to 1.21. This elevated within-subject variation in the 24-HDR measurements compared with the between-subject variance indicated that residual random day-to-day variations in the 24-HDR attenuated the associations between FFQ1 and the mean 24-HDR measures. Correction for this attenuation effect notably improved the Pearson correlations (ρ_{QT}) between FFQ1 and 24-HDR for alcohol by a factor of 1.26 and 1.36, respectively, for men and women, to 0.60 and 0.79. For the other nutrients, the attenuation correction led to much smaller improvements (Tables 3 and 4). For nutrient densities, the correlations between FFQ1 and 24-HDR measurements were similar to those for

the absolute intake estimates, and correction for attenuation due to variations over time in the 24-HDR also led to a similar degree of improvement in these correlation coefficients.

Among those participants who had provided a blood sample, but after exclusion of one vegetarian subject with extremely elevated intakes as well as serum values of β -carotene, plasma β -carotene levels had a Pearson coefficient of correlation of 0.47 with the 24-HDR measurements, and of 0.23 with FFQ1 measurements.

Calibration of nutrient intake levels

For all nutrients except alcohol, the linear regression of 24-HDR on FFQ1 resulted in slope estimates (linear calibration coefficients) between 0.30 for protein (men and women) and 0.59 for vitamin C (men) (Tables 3 and 4). For most of the nutrient densities, the slopes were of similar average magnitude, except for cholesterol (men), retinol (men) and iron (women), for which the calibration coefficients were lower (Tables 5 and 6). Only for alcohol intake were the regression slopes very close to 1.0.

Discussion

In the present study, we examined the validity of dietary questionnaire measurements in a large prospective cohort study of nutrition and chronic disease risk in the north of Sweden, and used 24-hour diet recalls to calibrate the FFQ measurements.

The validity of food-frequency questionnaire measurements generally cannot be assumed to be independent of the population or the specific context in which the measurements were collected. For example, the correlation between FFQ and true habitual intake levels (as well as reference measurements, e.g. average 24-HDR) would generally tend to improve if the measurements are collected in a population with greater between-subject heterogeneity in true habitual intake levels and, inversely, would be lower in populations with more homogeneous dietary intake patterns. Furthermore, the magnitude of systematic and random errors may depend on the quality of response to the questionnaire, and hence on the interest and backgrounds of participants studied. Estimates of validity therefore should not be extrapolated beyond the actual populations in which measurements were collected and validated; in the present case, the Northern Sweden EPIC/MONICA cohort. A corollary of this is that validation/calibration sub-studies should be conducted on a subset of participants that is representative for the global study cohort. In the present study, such representativeness was high, as the validation study was based on a random selection of cohort participants, with a high participation rate. Representativeness was further confirmed by the great similarity between values for tobacco use, BMI, blood pressures and blood cholesterol in the validation study subgroup and the rest of the cohort.

One of the characteristics of the Northern Sweden EPIC/MONICA cohort is the repeated measurement of dietary intakes and other lifestyle variables, over 10-year intervals¹². Thus, not only will it be possible to relate the

Table 3 Comparison of daily energy and macronutrient intake as measured by the 'Northern Sweden' FFQ and 24-hour dietary recalls (24-HDR)

Nutrient	Gender*	Mean intake†		Ratio	Pearson correlation§		Means by FFQ1 quartiles¶				Calibration
		Mean _{FFQ}	Mean _{24-HDR}	FFQ1/24-HDR‡	<i>r</i> _{QR}	<i>p</i> _{QT}	1st	2nd	3rd	4th	coefficient
Energy (kcal)	Men	2018	2108	0.96 (0.90–1.02)	0.50 (0.34–0.64)	0.53	1753	2024	2138	2603	0.41 (0.27–0.55)
	Women	1609	1588	1.01 (0.96–1.07)	0.45 (0.27–0.59)	0.48	1346	1571	1552	1935	0.45 (0.27–0.63)
Carbohydrate (g)	Men	244.2	240.8	1.01 (0.95–1.09)	0.44 (0.26–0.59)	0.46	202.6	232.8	253.3	281.5	0.35 (0.21–0.49)
	Women	192.6	188.0	1.02 (0.97–1.08)	0.51 (0.35–0.64)	0.54	156.6	175.4	194.0	234.7	0.46 (0.30–0.62)
Sucrose (g)	Men	30.9	40.1	0.77 (0.69–0.86)	0.60 (0.45–0.72)	0.65	24.2	36.2	52.9	56.7	0.55 (0.41–0.69)
	Women	23.3	34.0	0.69 (0.62–0.76)	0.34 (0.15–0.50)	0.37	27.3	30.2	41.2	39.5	0.35 (0.15–0.55)
Fat _{total} (g)	Men	75.5	81.1	0.93 (0.87–0.99)	0.61 (0.47–0.72)	0.66	64.3	74.2	85.9	105.4	0.54 (0.40–0.68)
	Women	56.0	61.8	0.91 (0.85–0.96)	0.54 (0.38–0.66)	0.59	50.6	57.7	65.9	75.7	0.55 (0.37–0.73)
Fat _{saturated} (g)	Men	31.6	34.8	0.91 (0.84–0.97)	0.58 (0.43–0.70)	0.62	26.7	32.4	37.3	45.6	0.52 (0.36–0.68)
	Women	23.4	26.6	0.88 (0.83–0.94)	0.54 (0.38–0.66)	0.59	23.1	22.2	28.9	33.3	0.53 (0.37–0.69)
Cholesterol (mg)	Men	0.25	0.32	0.77 (0.71–0.83)	0.45 (0.27–0.60)	0.50	0.25	0.31	0.33	0.41	0.43 (0.25–0.61)
	Women	0.19	0.24	0.77 (0.72–0.82)	0.31 (0.12–0.48)	0.37	0.22	0.22	0.24	0.29	0.33 (0.13–0.53)
Protein (g)	Men	74.7	82.2	0.91 (0.85–0.97)	0.41 (0.23–0.56)	0.44	73.8	79.0	81.1	96.6	0.30 (0.16–0.44)
	Women	59.4	59.9	0.99 (0.94–1.05)	0.40 (0.23–0.56)	0.44	56.0	55.5	59.6	69.6	0.30 (0.16–0.44)
Alcohol (g)	Men	4.53	6.36	0.94 (0.70–1.26)	0.49 (0.29–0.65)	0.60	1.77	4.25	9.58	11.45	1.13 (0.66–1.60)
	Women	2.51	3.13	0.92 (0.73–1.14)	0.58 (0.41–0.71)	0.79	0.89	2.87	2.31	8.28	0.96 (0.65–1.27)

* Ninety-six men and 99 women are included in calculations of nutrients. For alcohol, zero reporters (three men and nine women for FFQ1; 22 men and 25 women for 24-HDR) were excluded.

† Geometric mean intakes (amount per day) obtained by FFQ1 (mean_{FFQ}) or 24-HDR (mean_{24-HDR}).

‡ Geometric mean (95% CI) of individual ratios between intake (amount per day) as obtained by FFQ1 and 24-HDR.

§ Pearson correlation coefficients – crude coefficients (*r*_{QR}) and after de-attenuation (*p*_{QT}) – between intakes measured by FFQ1 and 24-HDR.

¶ Geometric mean of 24-HDR measured intake (amount per day) in FFQ1 ranked quartiles.

|| Calibration coefficient equalling the slope (λ , with 95% CI) from the linear regression of the 24-hour recall measured intake on the corresponding intake measured by FFQ1.

Table 4 Comparison of daily micronutrient intake as measured by the 'Northern Sweden' FFQ and 24-hour dietary recalls (24-HDR)

Nutrient	Gender*	Mean intake†		Ratio	Pearson correlation§		Means by FFQ1 quartiles¶				Calibration coefficient
		Mean _{FFQ}	Mean _{24-HDR}	FFQ1/24-HDR‡	<i>r</i> _{QR}	<i>p</i> _{QT}	1st	2nd	3rd	4th	λ -value (95% CI)
Fibre (g)	Men	19.5	16.8	1.19 (1.10–1.29)	0.48 (0.31–0.62)	0.54	13.4	15.2	17.4	20.3	0.39 (0.25–0.53)
	Women	18.0	14.5	1.24 (1.18–1.32)	0.61 (0.47–0.72)	0.67	11.9	13.1	14.9	19.0	0.56 (0.42–0.70)
Vitamin C (mg)	Men	69.3	50.6	1.37 (1.22–1.53)	0.52 (0.36–0.66)	0.62	32.7	46.7	63.0	68.4	0.59 (0.39–0.79)
	Women	85.5	60.0	1.42 (1.28–1.58)	0.51 (0.35–0.64)	0.59	39.3	59.0	65.9	86.2	0.52 (0.34–0.70)
Vitamin E (mg)	Men	6.76	6.80	1.00 (0.93–1.06)	0.59 (0.45–0.71)	0.64	5.34	6.05	7.56	8.75	0.54 (0.38–0.70)
	Women	5.90	5.98	0.99 (0.92–1.05)	0.45 (0.28–0.60)	0.50	4.80	5.93	6.53	6.88	0.45 (0.27–0.63)
β -Carotene (mg)	Men	1.86	1.51	1.23 (1.05–1.45)	0.50 (0.34–0.64)	0.56	1.13	1.33	1.49	2.64	0.49 (0.31–0.67)
	Women	3.49	2.18	1.60 (1.38–1.85)	0.57 (0.42–0.69)	0.67	1.31	2.22	2.48	3.36	0.46 (0.32–0.60)
Retinol (mg)	Men	1.19	0.82	1.46 (1.30–1.64)	0.32 (0.13–0.49)	0.36	0.65	0.84	0.71	1.13	0.39 (0.15–0.63)
	Women	0.84	0.66	1.28 (1.14–1.43)	0.28 (0.09–0.45)	0.33	0.60	0.62	0.58	0.87	0.36 (0.11–0.61)
Calcium (mg)	Men	942	1027	0.92 (0.84–1.00)	0.41 (0.23–0.56)	0.44	825	1006	1101	1216	0.32 (0.18–0.46)
	Women	761	797	0.96 (0.89–1.03)	0.47 (0.30–0.61)	0.51	690	736	828	956	0.34 (0.22–0.46)
Iron (mg)	Men	15.7	15.8	0.99 (0.93–1.06)	0.41 (0.23–0.56)	0.46	13.8	15.3	15.8	18.8	0.33 (0.18–0.48)
	Women	12.5	11.8	1.06 (1.00–1.12)	0.43 (0.24–0.57)	0.48	9.9	11.9	12.0	13.8	0.48 (0.23–0.56)

*Ninety-six men and 99 women are included in calculations of nutrients. For alcohol, zero reporters (three men and nine women for FFQ1; 22 men and 25 women for 24-HDR) were excluded.

† Geometric mean intakes (amount per day) obtained by FFQ1 (mean_{FFQ}) or 24-HDR (mean_{24-HDR}).

‡ Geometric mean (95% CI) of individual ratios between intake (amount per day) as obtained by FFQ1 and 24-HDR.

§ Pearson correlation coefficients – crude coefficients (*r*_{QR}) and after de-attenuation (*p*_{QT}) – between intakes measured by FFQ1 and 24-HDR.

¶ Geometric mean of 24-HDR measured intake (amount per day) in FFQ1 ranked quartiles.

|| Calibration coefficient equalling the slope (λ , with 95% CI) from the linear regression of the 24-hour recall measured intake on the corresponding intake measured by FFQ1.

dietary intake assessments to subsequent disease risk, but the replicate data can also be used to monitor changes of diet over more prolonged time periods. The reproducibility of the FFQ data over a one-year period was high, in terms of both estimated mean intake levels and relative classification of participants from low to high levels. Thus, although mean intake estimates from the FFQ may not be absolutely valid (in fact, there is clear underreporting for at least part of the participants), the differences in mean FFQ intake levels over time may be indicative of true changes

in food consumption behaviour. This justifies the use of the FFQ measurements for monitoring of changes that may be induced by the local Västerbotten CVD community intervention programme¹². Mean intake levels are also in reasonable accordance with those of other published studies, including the European EPIC study, of which the NSHD cohort is a part^{1–8}.

Since perfectly error-free reference measurements for diet do not exist in practice, the validation and calibration of FFQ measurements depend crucially on statistical

Table 5 Comparison of macronutrient intake (density as amount per 1000 kcal) as measured by the 'Northern Sweden' FFQ and 24-hour dietary recalls (24-HDR)

Nutrient	Gender*	Mean intake†		Ratio	Pearson correlation§		Means by FFQ1 quartiles¶				Calibration coefficient
		Mean _{FFQ}	Mean _{24-HDR}	FFQ1/24-HDR‡	<i>r</i> _{QR}	<i>p</i> _{QT}	1st	2nd	3rd	4th	λ -value (95% CI)
Carbohydrate (g)	Men	121	114	1.06 (1.04–1.08)	0.57 (0.41–0.69)	0.64	104	113	117	124	0.60 (0.42–0.78)
	Women	120	118	1.01 (0.99–1.04)	0.50 (0.34–0.64)	0.61	112	118	117	113	0.33 (0.22–0.44)
Sucrose (g)	Men	15.3	19.0	0.80 (0.75–0.87)	0.65 (0.51–0.75)	0.71	13.9	18.0	21.7	26.0	0.58 (0.44–0.72)
	Women	14.5	21.4	0.68 (0.62–0.74)	0.31 (0.12–0.48)	0.35	19.5	20.4	20.3	26.3	0.28 (0.11–0.45)
Fat _{total} (g)	Men	37.4	38.5	0.97 (0.94–1.00)	0.42 (0.24–0.57)	0.51	35.1	39.0	39.1	40.8	0.38 (0.21–0.55)
	Women	34.8	38.9	0.89 (0.86–0.92)	0.57 (0.41–0.69)	0.70	35.8	38.2	40.5	41.8	0.34 (0.24–0.44)
Fat _{saturated} (g)	Men	15.6	16.5	0.95 (0.91–0.98)	0.52 (0.36–0.66)	0.58	14.5	16.5	17.2	18.3	0.47 (0.32–0.62)
	Women	14.6	16.7	0.87 (0.84–0.91)	0.54 (0.39–0.67)	0.62	15.2	15.9	17.5	18.5	0.36 (0.25–0.47)
Cholesterol (mg)	Men	0.1	0.2	0.80 (0.76–0.85)	0.21 (0.01–0.40)	0.26	0.15	0.14	0.16	0.16	0.21 (0.02–0.40)
	Women	0.1	0.2	0.76 (0.72–0.80)	0.35 (0.17–0.52)	0.42	0.13	0.14	0.16	0.17	0.48 (0.23–0.73)
Protein (g)	Men	37.0	39.0	0.95 (0.92–0.98)	0.35 (0.16–0.52)	0.40	37.6	37.9	37.4	43.7	0.41 (0.19–0.63)
	Women	36.9	37.7	0.98 (0.95–1.01)	0.39 (0.22–0.55)	0.45	34.8	37.4	37.9	41.3	0.37 (0.20–0.54)
Alcohol (g)	Men	2.2	3.0	0.98 (0.74–1.29)	0.48 (0.28–0.64)	0.57	1.0	2.2	3.2	7.1	1.03 (0.59–1.47)
	Women	1.6	1.9	0.94 (0.76–1.17)	0.60 (0.43–0.73)	0.83	0.74	1.3	1.9	4.8	0.99 (0.68–1.30)

*Ninety-six men and 99 women are included in calculations of nutrients. For alcohol, zero reporters (three men and nine women for FFQ1; 22 men and 25 women for 24-HDR) were excluded.

† Geometric mean intakes (amount per 1000 kcal) obtained by FFQ1 (mean_{FFQ}) or 24-HDR (mean_{24-HDR}).

‡ Geometric mean (95% CI) of individual ratios between intake (amount per 1000 kcal) as obtained by FFQ1 and 24-HDR.

§ Pearson correlation coefficients – crude coefficients (*r*_{QR}) and after de-attenuation (*p*_{QT}) – between intakes adjusted for energy (residuals) measured by FFQ1 and 24-HDR.

¶ Geometric mean of 24-hour recall measured intake (amount per 1000 kcal) in FFQ1 ranked quartiles.

|| Calibration coefficient equalling the slope (λ with 95% CI) from the linear regression of the 24-HDR measured intake on the corresponding intake measured by FFQ1.

Table 6 Comparison of micronutrient intake (density as amount per 1000 kcal) as measured by the 'Northern Sweden' FFQ and 24-hour dietary recalls (24-HDR)

Nutrient	Gender*	Mean intake†		Ratio FFQ1/24-HDR‡	Pearson correlation§		Means by FFQ1 quartiles¶				Calibration coefficient
		Mean _{FFQ}	Mean _{24-HDR}	Intake (95% CI)	<i>r</i> _{QR}	<i>p</i> _{QT}	1st	2nd	3rd	4th	λ -value (95% CI)
Fibre (g)	Men	9.7	7.8	1.24 (1.19–1.30)	0.58 (0.43–0.70)	0.63	6.5	7.3	8.1	9.5	0.56 (0.40–0.72)
	Women	11.2	9.1	1.23 (1.19–1.27)	0.70 (0.58–0.79)	0.79	7.8	8.7	9.6	10.7	0.65 (0.52–0.78)
Vitamin C (mg)	Men	34.3	24.0	1.43 (1.29–1.58)	0.47 (0.29–0.61)	0.57	18.0	20.2	26.1	35.5	0.64 (0.40–0.86)
	Women	53.1	37.8	1.41 (1.29–1.54)	0.54 (0.39–0.67)	0.62	25.6	36.5	44.9	48.2	0.58 (0.40–0.76)
Vitamin E (mg)	Men	3.4	3.2	1.04 (0.99–1.09)	0.26 (0.06–0.44)	0.30	3.0	3.3	3.1	3.5	0.30 (0.07–0.53)
	Women	3.7	3.8	0.97 (0.93–1.02)	0.35 (0.17–0.52)	0.42	3.4	3.7	4.0	4.1	0.33 (0.16–0.50)
β -Carotene (mg)	Men	0.9	9.7	1.29 (1.10–1.50)	0.43 (0.25–0.58)	0.50	0.53	0.60	0.77	1.09	0.46 (0.27–0.65)
	Women	2.2	1.4	1.58 (1.37–1.82)	0.58 (0.43–0.70)	0.72	0.84	1.32	1.50	2.18	0.47 (0.34–0.60)
Retinol (mg)	Men	0.6	0.4	1.53 (1.37–1.69)	0.22 (0.02–0.41)	0.27	0.36	0.36	0.44	0.41	0.22 (0.03–0.41)
	Women	0.5	0.4	1.26 (1.15–1.39)	0.29 (0.10–0.46)	0.42	0.40	0.38	0.35	0.56	0.39 (0.14–0.64)
Calcium (mg)	Men	467	487	0.96 (0.91–1.01)	0.50 (0.34–0.64)	0.57	410	492	531	529	0.42 (0.27–0.57)
	Women	473	502	0.94 (0.89–1.00)	0.59 (0.45–0.71)	0.66	437	471	511	608	0.41 (0.30–0.52)
Iron (mg)	Men	7.8	7.5	1.04 (1.00–1.08)	0.31 (0.12–0.49)	0.45	7.1	7.2	7.8	7.9	0.32 (0.12–0.51)
	Women	7.8	7.4	1.05 (1.00–1.09)	0.02 (–0.18–0.22)	0.03	7.6	7.1	7.4	7.8	0.02 (–0.18–0.22)

* Ninety-six men and 99 women are included in calculations of nutrients. For alcohol, zero reporters (three men and nine women for FFQ1; 22 men and 25 women for 24-HDR) were excluded.

† Geometric mean intakes (amount per 1000 kcal) obtained by FFQ1 (mean_{FFQ}) or 24-HDR (mean_{24-HDR}).

‡ Geometric mean (95% CI) of individual ratios between intake (amount per 1000 kcal) as obtained by FFQ1 and 24-HDR.

§ Pearson correlation coefficients – crude coefficients (*r*_{QR}) and after de-attenuation (*p*_{QT}) – between intakes adjusted for energy (residuals) measured by FFQ1 and 24-HDR.

¶ Geometric mean of 24-hour recall measured intake (amount per 1000 kcal) in FFQ1 ranked quartiles.

|| Calibration coefficient equalling the slope (λ with 95% CI) from the linear regression of the 24-HDR measured intake on the corresponding intake measured by FFQ1.

model assumptions^{15,28}. One assumption generally made is that at the population level there is a linear relationship between dietary intake assessments obtained by FFQ (Q) or by 24-HDR (R) with individuals' true habitual intake levels²⁹. Then, to estimate the correlation ρ_{QT} between individuals' FFQ measurements and their true intake levels, it can be shown that at least two additional measurements, X_1 and X_2 , are required, for which the random errors (ϵ_{X1} and ϵ_{X2}) can be assumed to be uncorrelated both between themselves and with random errors of the FFQ measurements²⁹. In practice, it is generally felt that the best possible attempt to meet these requirements is to compare with replicate recordings (i.e. $X_1, X_2, \dots = R_1, R_2, \dots$) of *actual* food consumption, obtained by food consumption diaries, weighed food consumption records or a 24-HDR. If indeed it can be assumed that the random errors $\epsilon_{R1}, \epsilon_{R2}, \dots, \epsilon_{R10}$ of the ten replicate 24-HDR measurements are uncorrelated between themselves and with the random errors ϵ_Q in the FFQ measurements, the de-attenuated correlation coefficients in Tables 3–6 (indicated by ' ρ_{QT} ') would provide an unbiased estimate of the correlation between the FFQ and individuals' true habitual intake estimates. We decided to use replicate 24-HDR measurements, because the method generally has high response rates (relevant for the representativeness of the validation sub-study) and because it could be administered by telephone for participants living far from the research centre.

Unfortunately, it is questionable whether the assumption of uncorrelatedness of random errors $\epsilon_{Ri}, \epsilon_{Rj}, \epsilon_Q$ holds in practice. Many studies have shown that 24-HDR data, but also weighed food records or FFQ measurements, may

be systematically underestimated, and that participants may vary systematically in their tendency to underestimate by each of these methods^{22,28}. As explained in greater detail elsewhere³⁰, the consequence of this is that the de-attenuated correlation coefficients of Tables 3–6 can be either an underestimation or an overestimation of the correlation of FFQ measurements with true intake levels, and it is unclear which of these two biases – upwards or downwards – would predominate in practice.

It is often assumed that correlations between random errors of FFQ and 24-HDR measurements would be mostly a consequence of individuals' systematic differences in tendency to globally underreport intake levels, irrespective of the type of assessment method used. In this context, a crucial question is whether such underreporting is generally of an approximately equal magnitude for any type of food consumed – that is, whether or not the magnitude of underreporting is food- or nutrient-specific. If the latter can be assumed to be true, then the correction for total energy intake level, e.g. by calculating nutrient densities, would strongly diminish the error correlations and increase the validity of validation and calibration study results. This type of assumption deserves further investigation.

An approach closely related to validation is the calibration of FFQ measurements. Calibration can be defined as a scale adjustment to the FFQ measurements, such that after the adjustments estimates of relative risk corresponding to a measured absolute intake level are approximately unbiased^{24,25}. We estimated the calibration factor λ by linear regression of 24-HDR on the FFQ measurements and, as shown in Tables 2–6, the estimated

calibration factors generally have values below 1.0, reflecting a large degree of attenuation bias due to random errors in the FFQ measurements²⁵. As discussed elsewhere^{24,25}, these estimated calibration factors may be used to correct relative risks estimated for quantitative differences in food or nutrient intake levels, as measured by the FFQ. However, the validity of this correction procedure again requires the assumption that random errors of the 24-HDR (ϵ_R) and random errors in the FFQ measurements (ϵ_Q) are uncorrelated. In addition, it must be assumed that the 24-HDR measurements have no systematic scaling biases; i.e. they must be assumed to provide accurate mean intake measurements for any randomly chosen subgroup of the total study population. In practice, as mentioned, it is likely that random errors ϵ_R and ϵ_Q are positively correlated, and this would lead to an overestimation of the calibration factor λ and hence to an underestimation of regression dilution bias³¹. As mentioned earlier, the problem of correlated random errors might be alleviated by using energy-adjusted intake measurements such as nutrient densities.

In conclusion, The Northern Sweden FFQ measurements have good reproducibility and an estimated level of validity similar to that of FFQ measurements in other prospective cohort studies. The results from this study will form the basis for the correction of attenuation and regression dilution biases in relative risk estimates, in future studies relating FFQ measurements to disease outcomes in the Northern Sweden cohort. However, in view of possible biases in the estimates of ρ_{QT} and calibration factors λ , due to correlatedness of errors in FFQ measurements and reference measurements, such corrections must be used cautiously, and may be more easily applicable for nutrient densities than for absolute nutrient intakes.

Acknowledgements

This study was supported by the Swedish Cancer Society and the Swedish Research Council.

References

- 1 Bingham SA, Gill C, Welch A, Cassidy A, Runswick SA, Oakes S, Lubin R, Thurnham DI, Key TJ, Roe L, Khaw KT, Day NE. Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. *Int. J. Epidemiol.* 1997; **26**: S137–51.
- 2 Ocke MC, Bueno de Mesquita HB, Goddijn HE, Jansen A, Pols MA, van Staveren WA, Kromhout D. The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int. J. Epidemiol.* 1997; **26**: S37–48.
- 3 Ocke MC, Bueno de Mesquita HB, Pols MA, Smit HA, van Staveren WA, Kromhout D. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int. J. Epidemiol.* 1997; **26**: S49–58.
- 4 Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J.

- Reproducibility and relative validity of energy and macronutrient intake of a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int. J. Epidemiol.* 1997; **26**: S71–81.
- 5 Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am. J. Epidemiol.* 1985; **122**: 51–65.
 - 6 Kroke A, Klipstein-Grobusch K, Voss S, Moseneder J, Thielecke F, Noack R, Boeing H. 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**: 439–47.
 - 7 Pietinen P, Hartman AM, Haapa E, Räsänen L, Haapakoski J, Palmgren J, Albanes D, Virtamo J, Huttunen JK. Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. *Am. J. Epidemiol.* 1988; **128**: 655–66.
 - 8 Klipstein-Grobusch K, den Breeijen JH, Goldbohm RA, Geleijnse JM, Hofman A, Grobbee DE, Witteman JC. Dietary assessment in the elderly: validation of a semiquantitative food frequency questionnaire. *Eur. J. Clin. Nutr.* 1998; **52**: 588–96.
 - 9 Riboli E, Kaaks R. The EPIC project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int. J. Epidemiol.* 1997; **26**: S6–14.
 - 10 MONICA Principal Investigators. WHO MONICA Project: geographic variation in mortality from cardiovascular diseases. Baseline data on selected population characteristics and cardiovascular mortality. *World Health Statist. Quart.* 1987; **40**: 171–84.
 - 11 Brännström I, Rosén M, Wall S, Weinehall L. Local health planning and intervention – the case of a Swedish municipality. *Scand. J. Prim. Health Care* 1988; (Suppl. 1) 57–64.
 - 12 Weinehall L. *Partnership for health. On the role of primary health care in a community intervention programme.* Thesis, Umeå University, Sweden, 1997.
 - 13 Bergström L, Kylberg E, Hagman U, Eriksson H, Bruce Å. The food composition data base system (KOST-systemet) – its use for nutrient values [in Swedish]. *Vår. Föda* 1991; **43**: 439–47.
 - 14 Livsmedelsverket. *Weight Tables* [in Swedish]. Uppsala: Livsmedelsverkets repro, 1999.
 - 15 Johansson G, Wikman Å, Åhrén A-M, Hallmans G, Johansson I. Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living. *Public Health Nutr.* 2001; **4**(4): 919–27.
 - 16 Krantzler NJ, Mullen BJ, Schutz HG, Grivetti LE, Holden CA, Meiselman HL. Validity of telephone diet recalls and records for assessment of individual food intake. *Am. J. Clin. Nutr.* 1982; **36**: 1234–42.
 - 17 Tran KM, Johnson RK, Soutlanakis RP, Matthews DE J. In-person vs telephone-administered multiple-pass 24-hour recalls in women: validation with doubly labeled water. *J. Am. Diet. Assoc.* 2000; **100**: 777–83.
 - 18 Bergström L. Illustrations of food items [in Swedish]. *Vår Föda* 1979; **31**(Suppl. 4): 401–3.
 - 19 Håglin L, Hagman U, Nilsson M. Evaluation of the meal model 'matmallen'. A means of estimating consumed amounts of food. *Scand. J. Nutr.* 1995; **39**: 79–83.
 - 20 Epler KS, Ziegler RG, Craft NE. Liquid-chromatographic method for the determination of carotenoids, retinoids, tocopherols in human serum and food. *J. Chromatogr.* 1993; **619**: 37–48.

- 21 FAO/WHO/UNU. *Energy and Protein Requirements*. Report of a Joint Expert Consultation. WHO Technical Report Series No. 724. Geneva: World Health Organization (WHO), 1985.
- 22 Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-reporting. *Eur. J. Clin. Nutr.* 1991; **45**: 569–81.
- 23 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.
- 24 Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat. Med.* 1989; **8**: 1051–69.
- 25 Kaaks R, Riboli E, van Staveren W. Calibration of dietary intake measurements in prospective cohort studies. *Am. J. Epidemiol.* 1995; **142**: 548–56.
- 26 Peltonen M, Huhtasaari F, Stegmayr B, Lundberg V, Asplund K. Secular trends in cardiovascular risk factor levels in Sweden. The Northern Sweden MONICA study. *J. Intern. Med.* 1998; **244**: 1–9.
- 27 Basiotis PP, Welsh SO, Cronin FJ, Kelsay JL, Mertz W. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. *J. Nutr.* 1987; **117**: 1638–41.
- 28 Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *Eur. J. Clin. Nutr.* 1991; **45**: 583–99.
- 29 Kaaks R, Riboli E, Esteve J, van Kappel AL, van Staveren WA. Estimating the accuracy of dietary questionnaire assessments: validation in terms of structural equation models. *Stat. Med.* 1994; **13**: 127–42.
- 30 Chow CK, Thacker RR, Changchit C, Bridges RB, Rehm SR, Humble J, Turbek J. Lower levels of vitamin C and carotenes in plasma of cigarette smokers. *J. Am. Coll. Nutr.* 1986; **5**: 305–12.
- 31 Kipnis V, Freedman LS, Brown CC, Hartman AM, Schatzkin A, Wacholder S. Effect of measurement error on energy-adjustment models in nutritional epidemiology. *Am. J. Epidemiol.* 1997; **146**: 842–55.