British Journal of Nutrition (2007), **98**, 194–200
doi: 10.1017/S0007114507691636

Long-term reproducibility of a food-frequency questionnaire and dietary changes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Heidelberg cohort

Gabriele Nagel[1,2]*, Dorothee Zoller[1], Tilla Ruf[1], Sabine Rohrmann[1] and Jakob Linseisen[1]

[1]*Division of Clinical Epidemiology, German Cancer Research Centre, DKFZ, Heidelberg, Germany*
[2]*Department of Epidemiology, University of Ulm, Helmholtzstr. 22, 89081 Ulm, Germany*

(Received 25 July 2006 – Revised 6 December 2006 – Accepted 5 January 2007)

Within a prospective cohort study, we explored the long-term reproducibility of the food-frequency questionnaire (FFQ) and dietary changes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Heidelberg cohort. After a mean follow-up time of 68·8 (SD 4·1) months the dietary assessment by means of a validated FFQ was repeated in 21 462 participants in the EPIC-Heidelberg cohort. The correlation and test–retest reproducibility of both dietary intake measurements was explored. The long-term correlation coefficients ranged from 0·41 (vegetables in men) to 0·77 (alcoholic beverage consumption in women). The median intake of potatoes, added fat, sugar/confectionary, cakes and alcoholic beverages was lower in the second than in the first FFQ, whereas the median intake of fruits, vegetables, cereals/cereal products and non-alcoholic beverages were higher. Consistently for food groups, 60–70 % of the participants in both genders were re-classified to the same or adjacent quintile of intake. The results of fairly high correlation coefficients indicate good agreement between both measurements. It is acknowledged that this result reflects to a substantial extent the measurement error of the FFQ and conclusions on real changes in the diet should be drawn very carefully. For some nutrients the dietary changes were consistent with results from independent national surveys. The performance of the FFQ gives confidence in the dietary data to be used as long-term exposure variables.

Long-term reproducibility: Food-frequency questionnaire: Adults: Diet: EPIC-Heidelberg

In most epidemiological studies on diet and disease risk, food-frequency questionnaire (FFQs) are used to assess dietary habits (Hu & Willett, 2002). There is an on-going debate about the accuracy of dietary measurements in large-scale cohort studies, since the observations on the association between diet and cancer were sometimes inconsistent (Kipnis *et al.* 2002; Bingham *et al.* 2003; Schatzkin & Kipnis, 2004). An adequate balance between accuracy and effort has to be achieved.

Due to the long latency period for the development of most chronic diseases, the validity and reproducibility of dietary intake data are important issues in nutritional epidemiology. The validation against standard reference methods is used to appraise the ability to discriminate between subjects with true exposure differences (Willett, 1998). The reproducibility describes the consistency of FFQ measurements at different times for the same individual. The European Prospective Investigation into Cancer and Nutrition (EPIC) is a prospective multicentre study to further investigate the association between diet, lifestyle and chronic diseases with emphasis on cancer (Riboli & Kaaks, 1997). At the time of the baseline assessment the problem of measurement error was addressed in the Heidelberg component of EPIC by performing validity and reproducibility studies (Boeing *et al.* 1997; Bohlscheid-Thomas *et al.* 1997*a*,*b*).

Diet–disease associations can be explored for food groups and nutrients often by applying a ranking system (e.g. quantiles). The ability to rank individuals along the distribution of intake and the stability of ranking over time is a prerequisite of FFQs. Measurement error in dietary exposure assessment and instability in individual dietary intake may hide true diet–disease associations.

The aims of the present analyses were (1) to assess the long-term reproducibility of the FFQ used in EPIC-Heidelberg and (2) to appraise the dietary changes in the EPIC-Heidelberg cohort.

Material and methods

Study population

The EPIC-Heidelberg cohort comprises 25 540 subjects, of whom women were aged 35–65 years and men 40–65 years at recruitment. Participants were recruited between 1994 and 1998 from the general population in the Heidelberg region (Boeing *et al.* 1999*a*,*b*).

Abbreviations: EPIC, European Prospective Investigation into Cancer and Nutrition.

* **Corresponding author:** Dr Gabriele Nagel, Department of Epidemiology, University of Ulm, Helmholtzstr. 22, 89081 Ulm, Germany, fax +49 731 50 31069, email gabriele.nagel@uni-ulm.de

A self-administered FFQ was used to attain data on the habitual diet at baseline. The FFQ comprised questions on the frequency and portion size of 148 food items eaten during the year preceding enrolment (FFQ-1; Bohlscheid-Thomas *et al.* 1997*a*). The second dietary assessment was performed consecutively along with the date of study entry. After a mean follow-up time of 68·9 (SD 4·1) months in men and 68·7 (SD 4·0) months in women the second follow-up (2001–3) was performed. The repeated dietary measurement was performed with the baseline instrument (FFQ-2), with a supplement of forty-five questions on food preparation methods (Rohrmann & Becker, 2002).

From 25 540 participants with complete FFQ-1, 2505 (10 %) subjects omitted FFQ-2, 1003 (4 %) were lost to follow-up or withdrew, and 570 (2 %) had died. In total 21 462 subjects completed FFQ-2, of whom 729 subjects were excluded based on extreme energy intake (<1st or >99th percentile of total energy intake), leaving 20 733 subjects (9530 men and 11 203 women) for the analysis.

At recruitment and during the follow-up, data on lifestyle factors such as smoking habits, physical activity and anthropometric measures were collected (Klipstein-Grobusch *et al.* 1997; Bergmann *et al.* 1999).

Variable definition

Dietary intake estimates for food groups were based on the results of the self-administered, quantitative FFQ. Nutrient intake data were calculated based on the food intake data of the Bundesanstalt für gesundheitlichen Verbraucherschutz und Veterinärmedizin (1999). The validity of the questionnaire against 24 h dietary recalls as standard method was determined (Kroke *et al.* 1999). Foods were summarized into seventeen food groups based on the EPIC-SOFT classification (Slimani *et al.* 1999). Energy adjustment of food and nutrient intake was performed by means of the residual method regression analysis for both dietary assessments (Willett & Stampfer, 1986). The quintile cut-points were based on the individual consumption levels of the corresponding percentiles.

Educational level was categorized as low, middle and high. BMI was calculated as weight (kg)/height^2 (m^2) and classified in <18·5, 18·5–24·9, 25–29·9, 30–34·9 and ≥35 kg/m^2. Smoking status was summarized as non-smoker and current smoker at the time of the assessment.

Despite the visual (FFQ-1) and programmed (FFQ-2) data check at data-entry, missing values in the FFQ occurred (Bohlscheid-Thomas *et al.* 1997*a*). Missing values were replaced according to the following manner in both assessments: (1) food items with no marks were regarded as not being consumed; (2) if a food item was indicated as consumed, but the information on portion size or food frequency was missing, then the most frequently indicated portion size or food frequency given in the entire cohort was used; (3) if a food item was marked as not consumed, but either portion size or food frequency was indicated then this item was set as not consumed (Bohlscheid-Thomas *et al.* 1997*a*).

Statistical method

Comparative evaluation of absolute and energy-adjusted values of food and nutrient intake by gender was performed. Median, range and percentiles were calculated using the SAS procedure UNIVARIATE. The relative changes between both assessments were calculated and shown as percentage of difference to total intake. The distribution of the food group intake was not normally distributed, thus non-parametric tests were carried out. Spearman correlation coefficients between FFQ-1 and FFQ-2 for food groups and selected nutrients were calculated using the SAS procedure CORR (Masson *et al.* 2003).

Quintiles were calculated based on gender-specific percentile cut-points. The agreement of individual quintile classification by food groups between baseline and follow-up was evaluated and weighted κ values were calculated by means of the SAS procedure PROC FREQ (Fleiss & Cohen, 1973). The results of the κ statistic indicate excellent agreement for values over 0·80, good agreement for values between 0·61 and 0·80, moderate agreement for values between 0·41 and 0·60, fair agreement for values between 0·21 and 0·40, and poor agreement for values equal or less than 0·20 (Altman, 1991). The statistical software package SAS release 9·1 (SAS Institute, Cary, NC, USA) was used for all analyses. All statistical tests and corresponding P values were two-sided, and $P<0·05$ was considered statistically significant.

Results

Table 1 shows the cohort characteristics by gender of the 20 733 subjects included in the analyses. During the follow-up of the cohort, the age distribution of the cohort shifted to older age groups and the rate of retired persons increased from 9·8 to 20·3 % in women and from 16·5 to 36·5 % in men.

The absolute and the energy-adjusted distribution of daily intake for food groups and selected nutrients for men are given in Table 2 and for women in Table 3. In both genders, the distribution of the BMI did not change substantially. The increase in median daily consumption of non-alcoholic beverages in men (+11 % relative increase) and in women (+14 % relative increase) was most striking. The median consumption of alcoholic beverages declined in men by 1 % and in women by 2 %. Among foods, the median daily intake of fruits (+6 % in men, +6 % in women), vegetables (+3 % in men, +4 % in women) and cereals/cereal products (+5 % in men, +5 % in women) was higher in the FFQ-2 than in the baseline data for both genders. In contrast, the daily median consumption of potatoes (−5 % in men, −3 % in women), cakes (−9 % in men, −14 % in women), added fats and oils (−6 % in men, −11 % in women) and sugar/confectionary (−7 % in men, −8 % in women) was lower in FFQ-2 than in the FFQ-1. The median intake of dairy products was lower in FFQ-2 than FFQ-1 (−2 %) in women only. The observed shifts in consumption were more prominent in men than in women for intake of potatoes, cereals/cereal products and alcoholic beverages. A marginal reduction in the median meat/meat products consumption (−2 %) was observed only in men, whereas the shift for vegetable consumption was more pronounced in women. The median daily energy intake declined by 1 % in both men and women. Among the macronutrients the reduction in the daily fat intake was most prominent (−3 % in men, −5 % in women) whereas among the micronutrients, the average intakes of β-carotene (+2 % in men, +4 % in women), vitamin C (+6 % in men, +7 %

Table 1. Characteristics of the study cohort (*n* 20 733) at the baseline (1994–8) and follow-up assessment (2001–3)

	Men				Women			
	Baseline		Follow-up		Baseline		Follow-up	
	n	%	*n*	%	*n*	%	*n*	%
n	9530	46·0			11 203	54·0		
Age (years)								
30–39	4	0·0			2025	18·1		
40–49	3714	39·0	1684	17·7	3857	34·4	3879	34·6
50–59	4043	42·4	3695	38·8	3633	32·4	3616	32·3
60–69	1769	18·6	3966	41·6	1688	15·1	3524	31·5
70–79			185	1·9			184	1·6
BMI								
< 18·5	15	0·2	18	0·2	155	1·4	158	1·4
18·5–25	2933	30·8	2887	30·3	6102	54·5	5999	53·5
25–30	4946	51·9	4997	52·4	3360	30·0	3477	31·0
30–35	1390	14·6	1370	14·4	1143	10·2	1135	10·1
≥ 35	237	2·5	241	2·5	440	3·9	419	3·7
Missing	9	0·1	17	0·2	3	0·0	15	0·1
Current smoker								
No	7327	76·9	7585	79·6	8978	80·1	9109	81·3
Yes	2203	23·1	1864	19·6	2225	19·9	1981	17·7
Missing			81	0·8			113	1·0
Current employment								
Full-time	7340	77·0	5243	55·0	3116	27·8	2727	24·3
Part-time	235	2·5	366	3·8	4266	38·1	3625	32·4
Retired	1576	16·5	3478	36·5	1100	9·8	2274	20·3
Jobless	325	3·4	245	2·6	361	3·2	230	2·1
Not employed (household, parental leave, education)	54	0·6	89	0·9	2357	21·0	2149	19·2
Missing			109	1·1	3	0·0	198	1·8
Education level								
Low (none/primary school)	2830	29·7			2847	25·4		
Medium (professional/secondary school)	3057	32·1			5510	49·2		
High (university degree)	3642	38·2			2843	25·4		
Missing	1	0·0			3	0·0		

in women) and calcium (+3 % in men, +3 % in women) were slightly higher in the second assessment.

Analyses by 10-year age-strata revealed no clear differences for most dietary factors, except for the meat intake among men, which increased in younger age groups by 1 %, but decreased by 4 % among men in the age of 50 years or more, and the consumption of sugar and confectionery, for which a stronger decrease (−11 %) in men aged 40–49 years than in older men (−5 % and −1 %) was observed (data not shown). Concerning the consumption of sugar and confectionery also stronger reduction was observed among younger (age 35–39 years) women (−13 %) than in women of older age (−9 %, −3 % and −9 %).

Spearman's correlation coefficients between FFQ-1 and FFQ-2 ranged for food groups and selected nutrients from 0·41 (vegetables) to 0·73 (alcoholic beverages consumption) in men (Table 2). The values in women ranged from 0·44 (cereals/cereal products and fat) to 0·77 (alcohol beverages consumption) (Table 3). In both genders, the correlations reached statistical significance. After adjustment for energy intake, most of the correlation coefficients remained similar or were slightly attenuated. An exception was total fat intake, with an attenuation of the correlation coefficient by energy adjustment from 0·53 to 0·43 in men and from 0·53 to 0·41 in women. Overall, the long-term reproducibility of the FFQ appears best for alcoholic beverages and moderate to good for food groups and selected nutrients. Notably, after energy adjustment the correlation coefficients for the macronutrients protein (in both genders) and carbohydrates (in women) did not improve.

In order to appraise the test–retest reproducibility, the ranking of individuals according to quintile cut-points in both assessment periods was compared (Table 4). The proportions of subjects classified into the comparable quintile ranged from 31·2 % for vegetables to 50·3 % for the consumption of alcoholic beverages in men and from 32·1 % for added fats and oils to 50·3 % for alcoholic beverages in women. Concerning nutrients, reclassification to the same quintile was lowest for vitamin C (31·7 % in men, 32·4 % in women) and highest for ethanol intake (50·1 % in men, 50·5 % in women). The percentages for the re-classification into an opposite quintile ranged from 1·0 % (alcoholic beverages) to 3·2 % (miscellaneous) in men and from 0·7 % (alcoholic beverages) to 2·6 % (vegetables, added fats and oils) in women. Overall, 60–70 % of the participants in both genders were re-classified to the same or adjacent quintile, indicating that the agreement for all food groups and selected nutrients is high.

In order to evaluate the test–retest agreement, the weighted κ statistic was calculated. The lowest weighted κ value with 0·28 was observed for vegetable intake in men and 0·29 for cereals/cereal products intake in women. The highest weighted κ values in both genders (0·57 in men and 0·60 in women) were calculated for alcoholic beverage or ethanol intake, respectively. For most food groups or nutrients, the weighted κ ranged around 0·35 in both genders, indicating

Table 2. Spearman rank correlations between repeated measurements (baseline (FFQ-1) and after 5·8 years mean follow-up (FFQ-2)) with the same FFQ for selected food groups and nutrients among 9530 men in the European Prospective Investigation into Cancer and Nutrition-Heidelberg cohort

	Baseline Median	Follow-up Median	Absolute difference Median	Relative change (%) Median	Spearman correlation coefficient (r_s)*	
					Crude	Energy adjusted
Food groups						
Potatoes (g/d)	87·8	82·9	−4·3	−5·2	0·51	0·50
Vegetables (g/d)	103·5	106·8	3·3	3·3	0·41	0·41
Legumes (g/d)	4·0	4·3	0·0	1·5	0·45	0·45
Fruits (g/d)	95·9	99·8	5·2	6·0	0·50	0·50
Dairy products (g/d)	177·6	179·1	0·5	0·4	0·61	0·62
Cereals/cereal products (g/d)	202·6	217·4	10·4	5·1	0·43	0·45
Meat/meat products (g/d)	107·5	104·4	−1·7	−2·0	0·60	0·57
Fish/shellfish (g/d)	16·4	19·7	0·0	0·1	0·52	0·50
Eggs/egg products (g/d)	10·0	10·3	0·1	1·6	0·53	0·51
Added fat (g/d)	21·6	20·0	−1·2	−6·3	0·47	0·44
Sugar/confectionery (g/d)	32·5	30·5	−1·5	−6·7	0·60	0·56
Cakes (g/d)	50·4	44·4	−3·1	−9·2	0·59	0·57
Non-alcoholic beverages (g/d)	1324·2	1489·7	137·2	10·5	0·48	0·47
Alcoholic beverages (g/d)	319·2	275·2	−3·0	−1·5	0·73	0·73
Condiments/sauces (g/d)	12·5	11·7	−0·7	−7·1	0·50	0·48
Soups/bouillon (g/d)	35·9	34·3	−1·3	−4·2	0·52	0·51
Miscellaneous (g/d)	0·1	0·2	0·0	3·3	0·48	0·48
Nutrients						
Energy (kJ/d)	8748·3	8672·6	−20·4	−1·0	0·53	
Protein (g/d)	73·8	73·8	−0·2	−0·2	0·52	0·50
Fat (g/d)	77·4	74·8	−2·3	−3·2	0·53	0·43
Carbohydrates (g/d)	221·8	222·9	1·7	0·8	0·51	0·52
Ethanol (g/d)	18·9	17·8	−0·2	−2·4	0·74	0·74
β-Carotene (μg/d)	2103·4	2145·5	50·7	2·5	0·46	0·45
Vitamin C (mg/d)	87·5	92·5	4·4	5·5	0·43	0·43
Calcium (mg/d)	723·0	749·2	21·3	3·2	0·54	0·57

* All correlations reached significance level $P<0·0001$.

fair concordance. Moderate agreement was observed for the intake of meat/meat products (0·44), sugar/confectionery (0·42) and cakes (0·41) in women and for the intake of dairy products (0·44), sugar/confectionery (0·43) and cakes (0·42) in men. Overall, the present results indicate fair test–retest agreement for most food groups and the selected nutrients.

Discussion

In both genders, the median food and nutrient intake levels remained relatively stable over time. An increase in consumption was found for non-alcoholic beverages, cereals/cereal products, fruits and vegetables, which was more pronounced in women than in men. The median daily consumption of potatoes, dairy products (in women), alcohol and meat (in men) was lower according to the FFQ-2 data. To some extent the observed dietary changes may be attributed to real changes of dietary habits. On the ecological level, the observed trends are consistent with other reports on dietary trends based on food consumption data (Mensink, 2002; Deutsche Gesellschaft für Ernährung, 2004; Jungjohann *et al.* 2005), indicating a slight increase in fruit, vegetable, fish and cereals/cereal products consumption, and a decrease in the consumption of potatoes, meat/meat products and alcoholic beverages between 1995 and 2002 in western Germany (Deutsche Gesellschaft für Ernährung, 2004). The trend for increasing dairy product intake in Germany was in contrast to the present observations in women. The present result may be influenced by different selection patterns of dairy foods such as cheese or yoghurt. Data on dietary consumption of middle-aged men in 1984–5 and 1994–5 from the MONICA-Project in Augsburg are available (Döring *et al.* 1998). In the MONICA data, the mean intake of fat and alcohol were lower in the second survey and the mean carbohydrate intake was higher than in the first survey (Döring *et al.* 1998). The present observations on the intake of carbohydrates, protein, calcium, β-carotene and vitamin C are in line with the national trends for these nutrients (Deutsche Gesellschaft für Ernährung, 2004). Regarding total fat intake, the present data suggest a reduction of the median daily intake, which is also observed in some other reports (Döring *et al.* 1998; Mensink, 2002; Deutsche Gesellschaft für Ernährung, 2004; Jungjohann *et al.* 2005). The increased consumption of non-alcoholic beverages may reflect a trend to more health-conscious choices (Jungjohann *et al.* 2005; Roddam *et al.* 2005) but may also be affected by the past hot summers. Changes in dietary habits depend on the age group, sociodemographic factors and the cultural setting (Colditz *et al.* 1987). This limits a direct comparison with the results from other surveys. Looking at the changes in median intakes it has to be considered that if the measurement error in both measurements was the same, then the effect of the flattened slope would underestimate the true changes in median intakes (Kipnis *et al.* 2003).

In the present data, correlation coefficients for food and nutrient intakes ranged from 0·5 to 0·7, indicating a moderate to high correlation for the food groups and nutrients. This is in line with other studies on long-term reproducibility measured

Table 3. Spearman rank correlations between repeated measurements (baseline (FFQ-1) and after 5·8 years mean follow-up (FFQ-2)) with the same FFQ for selected food groups and nutrients among 11 203 women in the European Prospective Investigation into Cancer and Nutrition-Heidelberg cohort

	Baseline Median	Follow-up Median	Absolute difference Median	Relative change (%) Median	Spearman correlation coefficient (r_s)*	
					Crude	Energy adjusted
Food groups						
Potatoes (g/d)	69·6	67·6	−1·7	−2·6	0·53	0·52
Vegetables (g/d)	110·6	115·1	4·8	4·4	0·45	0·45
Legumes (g/d)	2·4	2·6	0·0	2·3	0·48	0·48
Fruits (g/d)	105·2	110·9	5·7	5·6	0·49	0·49
Dairy products (g/d)	204·6	202·7	−3·0	−1·7	0·55	0·55
Cereals/cereal products (g/d)	166·6	177·6	7·4	4·7	0·44	0·43
Meat/meat products (g/d)	68·5	68·1	0·0	0·0	0·62	0·61
Fish/shellfish (g/d)	15·1	15·6	0·0	0·1	0·52	0·51
Eggs/egg products (g/d)	9·4	9·4	0·0	0·1	0·52	0·49
Added fat (g/d)	19·6	17·4	−1·9	−10·7	0·44	0·40
Sugar/confectionery (g/d)	26·9	24·9	−1·6	−8·5	0·59	0·53
Cakes (g/d)	44·7	37·9	−4·7	−14·2	0·59	0·54
Non-alcoholic beverages (g/d)	1445·4	1657·3	195·8	14·3	0·47	0·48
Alcoholic beverages (g/d)	80·1	76·1	−0·8	−1·8	0·77	0·76
Condiments/sauces (g/d)	11·6	11·0	−0·8	−7·7	0·51	0·48
Soups/bouillon (g/d)	26·6	26·6	0·0	0·2	0·50	0·48
Miscellaneous (g/d)	0·3	0·4	0·0	6·5	0·52	0·52
Nutrients						
Energy (kJ/d)	6955·9	6896·1	−14·2	−0·9	0·53	
Protein (g/d)	59·5	60·1	0·6	1·0	0·50	0·49
Fat (g/d)	64·9	61·7	−2·8	−4·5	0·53	0·41
Carbohydrates (g/d)	182·6	185·6	1·5	0·9	0·51	0·49
Ethanol (g/d)	5·8	5·6	−0·1	−4·6	0·77	0·77
β-Carotene (μg/d)	2250·8	2336·8	82·2	4·0	0·49	0·49
Vitamin C (mg/d)	93·6	99·7	5·9	6·7	0·45	0·45
Calcium (mg/d)	724·9	747·1	18·4	2·7	0·50	0·53

* All correlations reached significance level $P<0·0001$.

for periods from 1 to 10 years (Pietinen *et al.* 1988; Jarvinen *et al.* 1993; Goldbohm *et al.* 1995; Willett, 1998; Masson *et al.* 2003). In a former reproducibility study over 1 year, using a precursor of the current FFQ, the correlation coefficients of food group intake ranged from 0·49 for bread to 0·89 for alcohol with a median coefficient of 0·70 (Bohlscheid-Thomas *et al.* 1997*b*). Goldbohm *et al.* (1995) collected FFQ data annually over 5 years. The results of the regression analysis indicated a slight decline in the median consumption of most nutrients and food groups over time. Concerning macronutrients, the correlation coefficients for the unadjusted FFQ data were mostly between 0·60 and 0·70 (Bohlscheid-Thomas *et al.* 1997*a*). Low correlation can be caused by the method of dietary measurement or real changes of dietary habits. In the validation study, the FFQ-derived measurements were compared with data collected with the 24 h diet recall method used twelve times over 1 year as standard measurement. The correlation coefficients for food group intake between FFQ and 24 h recall data had a median of 0·45 (range 0·19 for legumes and 0·90 for alcoholic beverages) (Bohlscheid-Thomas *et al.* 1997*a*,*b*). The ratio of within- to between-person variations was larger than 1 for almost all food groups, however, the authors noted a lack of precision of the reference measurements (Bohlscheid-Thomas *et al.* 1997*a*). Kipnis *et al.* (2003) proposed as a structure of measurement error besides within-person random error, two biases (intake-related and person-specific). The intake-related bias may lead to the flattening slope phenomenon which was found to be more pronounced in energy-adjusted data. The person-specific bias reflects the difference between total within-person bias and its intake-related component (Kipnis *et al.* 2003). Measurement error of dietary intake often attenuates the disease-related risk estimates and may consequently miss diet–disease associations. The development of refined biological markers such as doubly labelled water for the assessment for the energy expenditure may help to be better able to distinguish between real changes of diet and measurement errors (Kaaks *et al.* 2002; Subar *et al.* 2003).

On the food group level, correlation coefficients remained similar or were slightly attenuated after adjustment for energy intake. Differing from the observation in food groups and micronutrients, in macronutrients the correlation coefficients for total fat and protein intake in men and fat, protein and carbohydrates intake in women declined after energy adjustment. Adjustment for energy was suggested to correct partially for measurement error (Willett, 1998). Energy adjustment had the strongest impact on the energy-contributing macronutrients such as total fat intake in both genders, whereas in women the correlation coefficient for carbohydrates was worse in the unadjusted data (0·51 and 0·49). The present observation suggests that the macronutrients were differently prone to reporting errors (Voss *et al.* 1998).

The agreement between the same or adjacent ranks of subjects according to food or nutrient intake levels reached generally about 70 %. Extreme misclassification (opposite quintiles) occurred at maximum in 3 % of the participants.

Table 4. Agreement of the test–retest classification to the same, adjacent or opposite quintile and weighted κ statistics for classification to the same quintile for selected food groups and nutrients by gender

	Men				Women			
	Same quintile (%)	Adjacent quintile (%)	Opposite quintile (%)	Weighted κ	Same quintile (%)	Adjacent quintile (%)	Opposite quintile (%)	Weighted κ
Food groups								
Potatoes (g/d)	35·4	37·9	2·3	0·36	36·1	37·8	2·2	0·37
Vegetables (g/d)	31·2	37·0	3·0	0·28	33·3	36·9	2·6	0·31
Legumes (g/d)	36·3	33·3	2·5	0·33	37·4	33·9	2·1	0·35
Fruits (g/d)	35·9	35·5	2·0	0·35	35·4	35·8	2·0	0·34
Dairy products (g/d)	40·7	36·9	1·2	0·44	37·5	37·0	1·8	0·39
Cereals/cereal products (g/d)	32·5	36·7	2·6	0·30	32·2	36·3	2·5	0·29
Meat/meat products (g/d)	39·2	37·2	1·3	0·42	39·9	38·0	1·0	0·44
Fish/shellfish (g/d)	36·4	35·7	2·0	0·36	35·7	37·1	1·8	0·36
Eggs/egg products (g/d)	36·6	37·8	1·9	0·38	36·8	36·5	2·0	0·37
Fat (g/d)	32·8	38·4	2·4	0·32	32·1	37·7	2·6	0·30
Sugar/confectionery (g/d)	38·8	39·2	1·4	0·43	39·1	38·1	1·5	0·42
Cakes (g/d)	38·3	38·9	1·2	0·42	37·8	38·3	1·3	0·41
Non-alcoholic beverages (g/d)	34·0	37·3	2·6	0·33	34·4	36·7	2·5	0·33
Alcoholic beverages (g/d)	50·3	36·3	1·0	0·57	50·3	37·8	0·7	0·60
Condiments/sauces (g/d)	34·6	37·5	2·1	0·35	34·6	37·5	1·8	0·35
Soups/bouillon (g/d)	36·4	36·8	2·0	0·37	35·1	37·2	2·0	0·35
Miscellaneous (g/d)	38·3	35·3	3·2	0·36	39·3	36·3	2·5	0·39
Nutrients								
Energy (kJ/d)	34·9	38·7	1·8	0·36	34·8	39·0	1·8	0·36
Protein (g/d)	35·0	38·1	1·7	0·36	34·4	38·1	2·1	0·35
Carbohydrates (g/d)	34·7	37·9	1·9	0·35	34·8	38·0	1·9	0·36
Fat (g/d)	35·6	37·5	1·8	0·36	34·9	38·1	1·7	0·36
Ethanol (g/d)	50·1	36·3	1·1	0·57	50·5	37·2	0·6	0·59
β-Carotene (ug/d)	32·7	37·4	2·6	0·31	33·5	37·9	2·2	0·33
Vitamin C (mg/d)	31·7	37·4	3·1	0·29	32·4	36·6	2·6	0·30
Calcium (mg/d)	35·5	38·0	1·8	0·37	34·2	37·8	1·9	0·34

The present results indicate a favourable performance of the FFQ and give confidence in the ranking of subjects according to dietary exposure. Comparing the agreement of classification according to food group intake levels obtained by 24 h recalls and FFQ, Bohlscheid-Thomas (*et al.* 1997*a*) found values in a similar range as the present values. Self-selection of subjects whose motivation is above average is likely for the whole cohort. However, the participation rate during follow-up was high, indicating that no further selection process during follow-up may have happened.

When reporting on changes of food group and nutrient intake, several limitations need to be considered. Though identical dietary questionnaires were used, some time-dependent differences in over- or underreporting related to the conceptions of healthy diet may have occurred. In the baseline assessment this systematic error was explored by comparison with a standard method. About 8 % of the initial study sample provided additionally 24 h dietary recall data for calibration purpose of the EPIC subcohorts (Slimani *et al.* 1999), which were analysed on the food group level.

Measuring the long-term reproducibility may be influenced by the decreasing ability to complete the FFQ due to increasing age. However, the strength of the present study is the high participation rate, the application of an identical FFQ, and the calculation of food groups and nutrients according to an identical procedure.

In conclusion, the present results suggest a moderate to good long-term reproducibility of the FFQ used in the EPIC-Heidelberg cohort. After a mean follow-up time of 68·8 months, we observed indications for some modifications in the consumption level of food groups and selected nutrients. The observed deviations may reflect true changes in dietary behaviour, but may also be related to measurement error. For several nutrients the dietary changes were consistent with the results of independent national surveys over time, suggesting that the FFQ results may partly reflect real dietary changes over time. In addition, the observed agreement of classification according to food and nutrient intake levels after several years gives confidence that the dietary data can be used as long-term exposure variables to investigate diet and disease associations. Future analyses on diet and disease risk can additionally benefit from more accurate classification of long-term habitual diet using individual exposure levels based on two dietary measurements, especially for those participants with instable FFQ results.

Acknowledgements

The authors thank the volunteers who participated in EPIC-Heidelberg. EPIC-Heidelberg was supported by the German Cancer Research Centre, grants from the European Community (SI2.326938) and Deutsche Krebshilfe e.V. (70-2488).

References

Altman DG (1991) *Practical Statistics for Medical Research*. London: Chapman & Hall.

Bergmann MM, Bussas U & Boeing H (1999) Follow-up procedures in EPIC-Germany – data quality aspects. European Prospective Investigation into Cancer and Nutrition. *Ann Nutr Metab* **43**(4), 225–234.

Bingham SA, Luben R, Welch A, Wareham N, Khaw KT & Day N (2003) Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet* **362**(9379), 212–214.

Boeing H, Bohlscheid-Thomas S, Voss S, Schneeweiss S & Wahrendorf J (1997) The relative validity of vitamin intakes derived from a food frequency questionnaire compared to 24-hour recalls and biological measurements: results from the EPIC pilot study in Germany. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26**, Suppl. 1, S82–S90.

Boeing H, Korfmann A & Bergmann MM (1999*a*) Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. *Ann Nutr Metab* **43**, 205–215.

Boeing H, Wahrendorf J & Becker N (1999*b*) EPIC-Germany – a source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. *Ann Nutr Metab* **43**, 195–204.

Bohlscheid-Thomas S, Hoting I, Boeing H & Wahrendorf J (1997*a*) Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26**, Suppl. 1, S59–S70.

Bohlscheid-Thomas S, Hoting I, Boeing H & Wahrendorf J (1997*b*) 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* **26**, Suppl. 1, S71–S81.

Bundesanstalt für gesundheitlichen Verbraucherschutz und Veterinärmedizin (1999) *Der Bundeslebensmittelschlüssel (BSL II.3): Konzeption Aufbau und Dokumentation der Datenbank blsdat. BgVV-Hefte* 08/1999. Berlin: BgVV.

Colditz GA, Willett WC, Stampfer MJ, Sampson L, Rosner B, Hennekens CH & Speizer FE (1987) The influence of age, relative weight, smoking, and alcohol intake on the reproducibility of a dietary questionnaire. *Int J Epidemiol* **16**(3), 392–398.

Deutsche Gesellschaft für Ernährung (2004) *Ernährungsbericht 2004*. Rheinbreitenbach: Medien Haus Plump.

Döring A, Honig-Blum K, Winkler G, Kammerlohr R, Fischer B & Keil U (1998) *MONICA Project Region Augsburg. Data Book. Dietary Surveys 1984/85 and 1994/95 in Middle-aged Men from the City of Augsburg*. Neuherberg: GSF-Forschungszentrum.

Fleiss JL & Cohen J (1973) The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Measure* **33**, 613–619.

Goldbohm RA, van't Veer P, van den Brandt PA, van't Hof MA, Brants HA, Sturmans F & Hermus RJ (1995) Reproducibility of a food frequency questionnaire and stability of dietary habits determined from five annually repeated measurements. *Eur J Clin Nutr* **49**(6), 420–429.

Hu FB & Willett WC (2002) Optimal diets for prevention of coronary heart disease. *JAMA* **288**(20), 2569–2578.

Jarvinen R, Seppanen R & Knekt P (1993) Short-term and long-term reproducibility of dietary history interview data. *Int J Epidemiol* **22**(3), 520–527.

Jungjohann SM, Luhrmann PM, Bender R, Blettner M & Neuhauser-Berthold M (2005) Eight-year trends in food, energy and macronutrient intake in a sample of elderly German subjects. *Br J Nutr* **93**(3), 361–368.

Kaaks R, Ferrari P, Ciampi A, Plummer M & Riboli E (2002) Uses and limitations of statistical accounting for random error correlations, in the validation of dietary questionnaire assessments. *Public Health Nutr* **5**(6A), 969–976.

Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, Ferrari P & Carroll RJ (2002) Bias in dietary-report instruments and its implications for nutritional epidemiology (Review). *Public Health Nutr* **5**(6A), 915–923.

Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, Bingham S, Schoeller DA, Schatzkin A & Carroll RJ (2003) Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* **158**(1), 14–21.

Klipstein-Grobusch K, Georg T & Boeing H (1997) Interviewer variability in anthropometric measurements and estimates of body composition. *Int J Epidemiol* **26**, 174–180.

Kroke A, Klipstein-Grobusch K, Voss S, Möseneder J, Thielecke F, Noack R & Boeing H (1999) 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* **70**, 439–447.

Masson LF, McNeill G, Tomany JO, Simpson JA, Peace HS, Wei L, Grubb DA & Bolton-Smith C (2003) Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. *Public Health Nutr* **6**(3), 313–321.

Mensink G (2002) *Was essen wir heute? Ernährungsverhalten in Deutschland. Beiträge zur Gesundheitsberichterstattung des Bundes*. Berlin: Robert-Koch-Institut.

Pietinen P, Hartman AM, Haapa E, Rasanen L, Haapakoski J, Palmgren J, Albanes D, Virtamo J & Huttunen JK (1988) Reproducibility and validity of dietary assessment instruments. II. A qualitative food frequency questionnaire. *Am J Epidemiol* **128**(3), 667–676.

Riboli E & Kaaks R (1997) The EPIC project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26**, Suppl., S6–S14.

Roddam AW, Spencer E, Banks E, Beral V, Reeves G, Appleby P, Barnes I, Whiteman DC & Key TJ (2005) Reproducibility of a short semi-quantitative food group questionnaire and its performance in estimating nutrient intake compared with a 7-day diet diary in the Million Women Study. *Public Health Nutr* **8**(2), 201–213.

Rohrmann S & Becker N (2002) Development of a short questionnaire to assess the dietary intake of heterocyclic aromatic amines. *Public Health Nutr* **5**(5), 699–705.

Schatzkin A & Kipnis V (2004) Could exposure assessment problems give us wrong answers to nutrition and cancer questions? *J Natl Cancer Inst* **96**(21), 1564–1565.

Slimani N, Deharveng G, Charrondiere RU, *et al.* (1999) Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Comput Methods Programs Biomed* **58**(3), 251–266.

Subar AF, Kipnis V, Troiano RP, *et al.* (2003) Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study. *Am J Epidemiol* **158**(1), 1–13.

Voss S, Kroke A, Klipstein-Grobusch K & Boeing H (1998) Is macronutrient composition of dietary intake data affected by underreporting? Results from the EPIC-Potsdam Study. European Prospective Investigation into Cancer and Nutrition. *Eur J Clin Nutr* **52**(2), 119–126.

Willett W (1998) *Nutritional Epidemiology*. New York: Oxford University Press.

Willett W & Stampfer MJ (1986) Total energy intake: implications for epidemiologic analyses. *Am J Epidemiol* **124**, 17–27.