European Journal of Clinical Nutrition (2004) 58, 674–680
 0954-3007/04 $25.00
www.nature.com/ejcn

ORIGINAL COMMUNICATION

Reproducibility and validity of coffee and tea consumption in Italy

M Ferraroni[1], A Tavani[2]*, A Decarli[3], S Franceschi[4], M Parpinel[5], E Negri[2] and C La Vecchia[2,6]

[1]*Dipartimento di Medicina, Chirurgia e Odontoiatria, Università di Milano, Polo San Paolo, Milano, Italy;* [2]*Istituto di Ricerche Farmacologiche 'Mario Negri', Milano, Italy;* [3]*Dipartimento di Scienze Biomediche e Biotecnologia, Università di Brescia, Brescia, Italy;* [4]*International Agency for Research on Cancer, Lyon Cedex, France;* [5]*Centro di Riferimento Oncologico, Aviano (Pordenone), Italy; and* [6]*Istituto di Biometria e Statistica Medica, Università di Milano, Milano, Italy*

Objective: The reproducibility and validity of coffee, decaffeinated coffee and tea intake has not been adequately studied, particularly in Italy, where coffee drinking is peculiar in terms of type and amount of coffee consumed.
Design: We compared coffee and tea consumption, measured by two interviewer-administered food frequency questionnaires (FFQ), with average intake derived from two 7-day dietary (DD) records (the reference method) on 395 volunteers. The Pearson correlation coefficients (*r*) were used to assess both reproducibility and validity of information on coffee intake.
Results: A satisfactory level of reproducibility and validity of the pattern consumption was observed for coffee, decaffeinated coffee and tea. The reproducibility for both sex combined showed *r* of 0.74–0.78 for coffee, 0.57–0.65 for decaffeinated coffee and 0.61–0.67 for tea. The validity was about 0.70 for coffee, around 0.58 for decaffeinated coffee and 0.56–0.60 for tea intake.
Conclusions: The FFQ is a satisfactorily reliable and valid instrument for collecting information on coffee, decaffeinated coffee and tea intake.
Sponsorship: Supported by the Italian Association for Research on Cancer, Milan, Italy, and by the Commission of the European Communities (Contract No.: QLKI-CT-2000-00069).
European Journal of Clinical Nutrition (2004) **58,** 674–680. doi:10.1038/sj.ejcn.1601864

Keywords: Coffee; epidemiological methods; food frequency questionnaire; reproducibility; tea; validity

Introduction

Although coffee consumption has been studied in relation to the risk of several diseases, the reproducibility and validity of the information on its intake, as obtained by a food frequency questionnaire (FFQ), have not been adequately studied in Mediterranean populations, where the amounts and mixtures of ground coffee and brewing methods are different compared to North America and Northern Europe (percolated, filter or boiled) (D'Amicis & Viani, 1993). The information is even more scanty for tea and for decaffeinated coffee, which in Italy are not widely consumed as coffee (Istituto Centrale di Statistica, 1984; Pagano *et al*, 1988). Reproducibility is defined as the ability of an instrument to provide similar values in repeated measurements. Validity is the ability of an instrument to record what it is intended to measure, that is, the instrument is compared with a gold standard, defined in the present study by a 7-day diary (7-DD) record (Decarli *et al*, 1996). Although the FFQs lack a proper gold standard, it is more likely that a 7-DD reports more correctly the real consumption. Usually in this kind of studies, both reproducibility and validity are quantified using the Pearson or Spearman correlation coefficient (*r*), which assesses the agreement between the information obtained with the different methods.

With reference to reproducibility, in the Older Iowa Women Study, the Pearson *r* for caffeine consumption was 0.95, however based on only 44 women (Munger *et al*, 1992). The Spearman *r* for coffee intake among the 201 men and women participating to the reproducibility study of the self-administered questionnaire of the Tromso Study was 0.65 (Jacobsen & Bonaa, 1990), and that of another Norwegian National Dietary Survey was 0.91, evaluated

***Correspondence: A Tavani, Istituto di Ricerche Farmacologiche 'Mario Negri', Via Eritrea 62, 20157 Milan, Italy.**
E-mail: tavani@marionegri.it
***Guarantor*: A Tavani.**
***Contributors*: CLV and SF conceived and coordinated the study. MF, EN and AD managed and analysed data. AT wrote the paper. MP gave valuable contribution on nutrition.**
Received 26 March 2003; revised 4 July 2003; accepted 2 August 2003

among 90 responders to a self-administered questionnaire (Johansson *et al*, 1997). In an Italian study of 294 subjects interviewed in hospital and at home, the reproducibility of coffee consumption was 0.88 (D'Avanzo *et al*, 1996).

Among studies testing both reproducibility and validity, in a Finnish study using a self-administered questionnaire, 121 men reported a mean intake of 398 g of coffee in the first questionnaire and of 414 g in the second one, and the Pearson *r* for validity ranged between 0.72 and 0.79 (Pietinen *et al*, 1988). In the Nurses' Health Study, the Pearson *r* for validity was about 0.75 for coffee and 0.83 for tea, while those for reproducibility were, respectively, 0.80 and 0.86 (Salvini *et al*, 1989), and in the Health Professional Follow-up Study, the *r* for reproducibility was 0.92 (Feskanich *et al*, 1993). These were among the items having the highest *r*. The Pearson *r* for reproducibility of the questionnaire of the Kuopio Breast Cancer Study was 0.74 for coffee and 0.67 for tea intake, and the corresponding values for testing validity were 0.94 and 0.78 (Männisto *et al*, 1996). The Pearson *r* for reproducibility of the Spanish dietary questionnaire of the EPIC Project was 0.79 in men and 0.51 in women for coffee and tea together and those for validity were 0.80–0.88 in men and 0.57–0.73 in women (EPIC Group of Spain, 1997); the corresponding values for the German part of the project were all above 0.70 (Bohlscheid-Thomas *et al*, 1997).

The pattern of coffee and tea drinking are different in Italy compared to North America and Northern Europe, as in Italy coffee is almost totally espresso and mocha, while tea is not commonly consumed (Istituto Centrale di Statistica, 1986; Pagano *et al*, 1988). In this study, we analysed the reproducibility and validity of information on the intake of coffee, decaffeinated coffee and tea by comparing intake measured by two interviewer-administered FFQs with the average intake derived from two 7-DD records administered to the same volunteers in the overall data set and in strata of selected covariates.

Materials and methods

Study design

The design of the present study has already been described (Franceschi *et al*, 1993, 1995; Tavani *et al*, 1995; Decarli *et al*, 1996). Briefly, we recruited volunteers from three Italian provinces (Pordenone, Forlì and Genoa) by advertisement on the local press, television and nonprofit organisations (Franceschi *et al*, 1993). All volunteers older than 35 y were questioned on their health status, and chronically ill subjects and those on special diets were excluded. The selected 452 volunteers (median age 50 y) completed twice an interviewer-administered FFQ, the first time in autumn–winter 1990–1991 and the second time 3–10 months later (6 months on the average), that is, in spring–summer 1991, which allowed a more accurate estimate of the overall annual consumption of coffee and tea. The same volunteers were instructed to weigh and record day by day in a diary all foods and beverages consumed during two periods of seven consecutive days (7-DD). Among the 452 participants, 395 (87%) (130 males and 265 females, median age 52 y, range 35–69) completed twice the 7-DD record starting on the first Sunday after the FFQ administration on both occasions. Of the 57 volunteers not included in this validation study, 54 had only one 7-DD record and for three the 7-DD record included fewer than 6 days of data collection. Participants were told to be taking part in a dietary survey, but were not aware of the study aims. A nutritionist reviewed all diaries to recover incompletely recorded information.

FFQ

The FFQ referred to the year preceding the interview, included a short section on sociodemographic characteristics, a section investigating a few aspects of the general dietary pattern, several questions to assess fat intake, and questions on the average weekly frequency of consumption of 78 foods or food groups and the most common recipes in the Italian diet, plus some open questions. All questions on diet included the portion size. Occasional consumption (1–3 times per month) was coded 0.5 per week and less than once per month was coded as zero.

The information on coffee, decaffeinated coffee, cappuccino and tea intake included the weekly number of cups consumed. The FFQ did not distinguish between coffee drunk in the coffee shop (espresso) and coffee made at home (mocha), and we assumed half consumption as espresso and half as mocha coffee. As the information on the two 7-DD records was expressed as millilitre of beverage consumed, we transformed the number of cups in millilitre, assuming that a cup of mocha coffee was 67.5 ml, a cup of espresso 30 ml, a cup of cappuccino (espresso coffee and whole milk) 50 ml and a cup of tea 135 ml. We have not analysed the validity of these volumes on the diary; however, these are relatively standard volumes, as indicated in different sources (Venturini, 1984; Nicoli *et al*, 1987), although the amount of caffeine contained in a cup of coffee is rather variable. Coffee intake refers to the sum of espresso, mocha and cappuccino. Small amounts of coffee or tea added to food in the preparation of recipes were not considered, but this is not common.

7-DD record

In the diary, the amount in millilitres of each hot beverage consumed was recorded. In the few cases where the information on the amount was missing, the same amount used in the questionnaire was used. Coffee or tea added to recipes was not considered.

Since 7-DD records were considered for the analysis when at least 6 consecutive days were recorded, the reference mean daily intake was obtained, for each volunteer, from 13 to 15 different estimates of daily consumption (378 participants with two 7-DD, 10 with one 6-DD and one 7-DD, and seven with one 8-DD and one 7-DD). For each subject, the reference mean daily intake was obtained as the average of

the two 7-DD records, used as reference method in this validation study.

Statistical analysis

Means and s.d.'s of coffee, decaffeinated coffee and tea intake were estimated from each questionnaire and from the 7-DD records.

Since the distributions were skewed towards high values, hot beverages values were $\log_e(x+0.001)$−transformed, in order to improve normality. The Pearson *r* (Snedecor & Cocharn, 1980) were used to assess both the reproducibility of the questionnaire and the validity compared to the reference method. The Spearman *r* were similar and thus not presented. Partial *r* were computed after allowance for study design variables (age, sex, centre) and for total energy (r_a) by regression methods. The correlation was thus evaluated between the hot beverage component residuals (ie, observed hot beverage intake minus intake predicted on the basis of energy intake and the study design variables) (Willett, 1990). Further, we evaluated the deattenuated unadjusted *r* (r_b), in order to take into account the interindividual variability of hot beverage intake as given by the 7-DD (Liu *et al*, 1978). The deattenuated *r* were calculated as: $r_a/\sqrt{1/[1+(s_w^2/n\times s_b^2)]}$ where s_w^2 and s_b^2 represent the within and between variance in the 7-DD, and *n* is the number of replicates for each participant, that is, the number of days of diary. For the reproducibility study, the intraclass *r* (r_c) was considered to measure the proportion of total variation due to between-individual variability (Snedecor & Cochran, 1980). The intraclass *r* was calculated as: $(s_b^2 - s_w^2)/[s_b^2 + (n-1)s_w^2]$ with $n=2$, the number of replicated measurements of FFQ. We do not present *P*-values for *r*, because they have chiefly descriptive purposes; however, with the available number of participants, all *r*-values higher than 0.10 are significant at 0.05 level.

The agreement between the categories based on the distribution values estimated by the FFQs and those estimated by the two 7-DD was assessed by the overall proportion of subjects similarly classified in the same category. The number of categories considered for each hot beverage depends on its distribution of consumption, that is, quintiles for coffee, tertiles for tea and two categories (nondrinkers/drinkers) for decaffeinated coffee (Landis & Koch, 1977).

Results

The distribution of study participants by centre, sex, age and educational level is shown in Table 1. Women were 69%, the median age was 52 y and the median educational level was 10 y. Coffee was consumed by 91.6, 90.4, 93.2 and 92.7% of subjects, as reported, respectively, in the first and second questionnaire and in the first and second 7-DD. The corresponding values for decaffeinated coffee were 11.4, 11.9, 8.1 and 8.4%, and for tea were 58.5, 54.4, 45.1 and 31.1% (not shown). Among consumers, mean consumption in the first and second 7-DD were, respectively, 112.8 and 100.4 ml/day for coffee intake, 50.0 and 52.2 ml/day for decaffeinated coffee and 104.1 and 95.3 ml/day for tea.

Table 1 Distribution of 395 participants in a validation study according to selected characteristics and centre, Italy 1990–1991

	Centre (Province)			
	Pordenone	*Genoa*	*Forlì*	*Total no. (%)*
Sex				
Males	83	26	13	122 (30.9)
Females	147	85	41	273 (69.1)
Age (y)				
<40	44	5	8	57 (14.3)
40–49	81	19	11	111 (28.1)
50–59	65	41	17	123 (31.1)
≥60	40	46	18	104 (26.3)
Education (y)				
≤10	129	42	31	202 (51.1)
>10	101	69	23	193 (48.9)

Table 2 shows the mean daily consumption of coffee, decaffeinated coffee and tea estimated by the 7-DD and the two FFQs; it also shows the Pearson, deattenuated and intraclass *r* and the overall proportion categorised in the same category of beverage record. Compared to the estimates from the 7-DD, the two FFQs tended to overestimate coffee, decaffeinated coffee and tea intake both in men and women. The estimated coffee intake in men was 90.98 ml in the 7-DD, 98.53 in FFQ1 and 103.60 ml in FFQ2; corresponding values in women were 103.03, 110.73 and 112.11 ml. The comparison of the 7-DD with FFQ1 yielded a Pearson *r* (r_a) of 0.77 in men and 0.64 in women, and deattenuated *r* (r_b) were comparable. The comparison of the estimates of the two FFQs (reproducibility of the questionnaire) yielded a Pearson *r* (r_a) of 0.85 in men and 0.68 in women, and the values were similar when taking into account the intravariability of consumption (r_c). The overall proportion of subjects falling in the same category of coffee intake in the 7-DD and in the two questionnaire administrations was 37.7% for men and 40.0% for women. In FFQ1 and FFQ2, 61.5% of men and 60.8% of women were classified in the same quintile. For decaffeinated coffee, the estimated intake with the 7-DD were 4.45 and 4.05 ml, respectively, in men and women, and 9.87 and 7.34 ml in men, and 6.60 and 9.62 ml in women in FFQ1 and FFQ2, respectively. The Pearson *r* for validity were 0.54 (r_a) in men and 0.59 (r_a) in women for FFQ1. The corresponding index for reproducibility of the questionnaire were 0.64 (r_a) in men and 0.65 (r_a) in women. The overall proportion categorised in the same category of decaffeinated coffee intake was 90.0% between the 7-DD and FFQ1 in men and 90.6% in women. The proportion of those classified in

Table 2 Comparison of coffee, decaffeinated coffee and tea consumption (ml/day) estimated according to the reference method (mean of the two 7-day diary, 7-DD) and two food frequency questionnaires (FFQ1 and FFQ2) in all subjects and in strata of sex (The data are from 395 participants in a study of reproducibility and validity, Italy 1990–1991)

	7-DD		*FFQ1*		*FFQ2*		*Pearson's correlation coefficient*[a]						*Percent of subjects classified in the same category*[b]		
	Mean	*s.d.*	*Mean*	*s.d.*	*Mean*	*s.d.*	*7-DD vs FFQ1*		*7-DD vs FFQ2*		*FFQ1 vs FFQ2*		*7-DD vs FFQ1*	*7-DD vs FFQ2*	*FFQ1 vs FFQ2*
Beverage							r_a^c	r_b^d	r_a^c	r_b^d	r_a^c	r_c^e			
Coffee															
All	99.06	62.44	106.71	73.49	109.31	70.26	0.69	0.71	0.74	0.76	0.74	0.78	39.2	42.2	61.0
Men (*n*=130)	90.98	56.75	98.53	69.31	103.60	66.78	0.77	0.79	0.81	0.83	0.85	0.74	37.7	37.7	61.5
Women (*n*=265)	103.03	64.78	110.73	75.25	112.11	71.87	0.64	0.66	0.70	0.72	0.68	0.80	40.0	44.5	60.8
Decaffeinated coffee															
All	4.18	19.11	7.68	31.29	8.87	31.37	0.58	0.58	0.60	0.58	0.65	0.57	91.1	89.9	91.9
Men (*n*=130)	4.45	22.32	9.87	41.85	7.34	27.84	0.54	0.54	0.62	0.62	0.64	0.49	90.0	93.1	92.3
Women (*n*=265)	4.05	17.36	6.60	24.53	9.62	32.99	0.59	0.61	0.55	0.55	0.65	0.63	90.6	88.3	91.7
Tea															
All	38.29	76.48	52.24	87.90	49.58	89.69	0.56	0.60	0.50	0.54	0.61	0.67	60.8	61.5	69.9
Men (*n*=130)	43.29	108.89	56.00	102.04	56.52	109.85	0.67	0.71	0.46	0.51	0.60	0.81	64.6	61.5	71.5
Women (*n*=265)	35.84	54.00	50.40	80.20	46.18	77.93	0.51	0.54	0.53	0.55	0.61	0.55	58.9	61.5	69.1

[a]Based on $\log_e(x+0.001)$ transformed values.
[b]Quintiles for coffee, nondrinkers/drinkers for decaffeinated coffee, tertiles for tea.
[c]r_a: estimated from multiple regression analysis including terms for age, centre, total energy and sex.
[d]r_b: deattenuated correlation coefficient.
[e]r_c: intraclass correlation coefficient.

the same category of consumption in FFQ1 and FFQ2 was 92.3 in men and 91.7 in women. For tea, the estimated intake with the 7-DD was 43.29 and 35.84 ml, respectively, in men and women, and in the two FFQs was 56.00 and 56.52 ml in men, and 50.40 and 46.18 ml in women. The Pearson *r* for validity were 0.67 (r_a) in men, and 0.51 (r_a) in women for the FFQ1. The Pearson *r* for reproducibility of the questionnaire were 0.60 (r_a) in men and 0.61 (r_a) in women. The overall proportion categorised in the same category of tea intake between 7-DD and FFQ1 in men was 64.6%, and in women 58.9%. The proportion of those classified in the same category between FFQ1 and FFQ2 was 71.5% in men and 69.1% in women.

The influence of age, educational level and body mass index between the 7-DD and the two interviews on the degree of validity and reproducibility of the FFQ in relation to the consumption of coffee, decaffeinated coffee and tea is shown in Table 3. People aged 60 y or more tended to drink less coffee (82.24 ml compared to 98.99 and 113.38 ml in the other two strata of age according to the 7-DD and about 30%

Table 3 Comparison of coffee consumption (ml/day) estimated according to the reference method (mean of the two 7-day diary, 7-DD) and two food frequency questionnaires (FFQ1 and FFQ2) by age, education and body mass index (The data are from 395 participants in a study of reproducibility and validity, Italy 1990–1991)

	7-DD		*FFQ1*		*FFQ2*		*Pearson's correlation coefficient*[a]					
	Mean	*s.d.*	*Mean*	*s.d.*	*Mean*	*s.d.*	*7-DD vs FFQ1*		*7-DD vs FFQ2*		*FFQ1 vs FFQ2*	
							r_a^b	r_b^c	r_a^b	r_b^c	r_a^b	r_c^d
Coffee												
Age (y)												
<50 (*n*=168)	98.99	64.82	117.50	79.90	115.77	74.21	0.79	0.81	0.80	0.81	0.77	0.79
50–59 (*n*=123)	113.38	66.66	113.88	73.52	122.94	71.96	0.55	0.56	0.73	0.75	0.78	0.82
≥60 (*n*=104)	82.24	48.01	80.81	54.54	82.74	53.09	0.62	0.63	0.63	0.64	0.65	0.60
Education (y)												
≤10 (*n*=204)	102.61	62.48	105.40	72.91	111.45	69.72	0.66	0.69	0.67	0.70	0.68	0.78
>10 (*n*=191)	95.27	62.33	108.12	74.26	107.02	70.96	0.70	0.72	0.78	0.80	0.77	0.78
Body mass index (kg/m^2)												
≤23 (*n*=141)	94.19	58.22	108.40	70.31	111.14	70.44	0.65	0.66	0.71	0.73	0.70	0.79
<23–>25 (*n*=104)	100.83	65.96	100.00	70.56	104.74	67.69	0.67	0.69	0.73	0.76	0.88	0.83
≥25 (*n*=150)	102.41	63.87	109.78	78.42	110.75	72.15	0.76	0.80	0.77	0.79	0.65	0.74
Decaffeinated coffee												
Age (y)												
<50 (*n*=168)	2.05	9.75	4.25	16.97	6.92	24.81	0.55	0.56	0.51	0.53	0.73	0.72
50–59 (*n*=123)	2.11	13.08	4.98	23.26	6.23	23.39	0.48	0.52	0.47	0.51	0.51	0.41
≥60 (*n*=104)	10.07	31.50	16.41	50.32	15.16	45.43	0.64	0.65	0.66	0.68	0.65	0.56
Education (y)												
≤10 (*n*=204)	4.39	20.90	7.30	34.22	8.79	34.67	0.55	0.55	0.53	0.54	0.62	0.59
>10 (*n*=191)	3.95	17.03	8.08	27.91	8.96	27.50	0.60	0.61	0.60	0.61	0.67	0.54
Body mass index (kg/m^2)												
≤23 (*n*=141)	3.35	10.89	7.25	24.04	7.63	24.97	0.61	0.63	0.57	0.60	0.58	0.49
<23–>25 (*n*=104)	1.91	11.53	4.73	19.41	7.70	25.08	0.40	0.40	0.34	0.35	0.66	0.69
≥25 (*n*=150)	6.54	27.43	10.13	42.11	10.86	39.69	0.62	0.63	0.70	0.73	0.71	0.58
Tea												
Age (y)												
<50 (*n*=168)	34.80	50.25	51.77	95.21	56.88	97.39	0.58	0.66	0.37	0.42	0.55	0.54
50–59 (*n*=123)	43.07	100.86	55.35	90.25	40.69	82.50	0.49	0.50	0.55	0.57	0.67	0.77
≥60 (*n*=104)	38.29	78.62	49.33	72.03	48.31	84.55	0.63	0.65	0.66	0.70	0.66	0.84
Education (y)												
≤10 (*n*=204)	36.32	84.41	45.66	76.81	36.87	76.40	0.59	0.62	0.53	0.57	0.59	0.72
>10 (*n*=191)	40.40	67.13	59.27	98.09	63.16	100.43	0.53	0.56	0.47	0.49	0.63	0.63
Body mass index (kg/m^2)												
≤23 (*n*=141)	43.28	59.32	57.24	76.85	63.19	95.56	0.57	0.62	0.54	0.57	0.66	0.55
<23–>25 (*n*=104)	31.22	47.03	37.92	52.68	36.81	68.08	0.58	0.63	0.39	0.45	0.60	0.39
≥25 (*n*=150)	38.52	102.81	57.47	112.99	45.64	95.83	0.55	0.57	0.53	0.54	0.58	0.81

[a]Based on $\log_e(x+0.001)$ transformed values.
[b]r_a: estimated from multiple regression analysis including terms for age, centre, total energy and sex.
[c]r_b: deattenuated correlation coefficient.
[d]r_c: intraclass correlation coefficient.

less coffee according to the two FFQs) and more decaffeinated coffee compared to younger people (10.07 ml compared to 2.05 and 2.11 ml in the other two strata of age according to the 7-DD and about 2–4 times more decaffeinated coffee according to the two FFQs). Reproducibility and validity for coffee were higher in younger and more educated subjects, but no other major differences in the mean consumption of coffee, decaffeinated coffee and tea nor in the Pearson *r* for validity or reproducibility were found across strata of covariates.

Discussion

This study indicates and further quantifies satisfactory levels of reproducibility and validity of the information on coffee, decaffeinated coffee and tea consumption obtained from the FFQ, as reflected by *r*-values all above 0.50 and around 0.70 for coffee intake, the beverage most commonly drunk. The consumption pattern of coffee should be reasonably stable during time, as people's opinion on coffee drinking in Italy has not substantially changed over the last 10 y (Istituto Centrale di Statistica, 1986; Pagano *et al*, 1988).

For the same questionnaire, we have previously shown that the Spearman *r* of reproducibility for alcohol drinking was around 0.70–0.75, with values comparable for wine, but lower for beer and spirits, that is, for less frequently drunk beverages (Ferraroni *et al*, 1996). Thus, the FFQ is a reliable and valid instrument for collecting information on major beverages in Italy. The difficulties in validating dietary variables include that FFQ lacks a true gold standard of reference. We used the usual solution to calibrate the FFQ results against another dietary instrument, the 7-DD, which is more likely accurate (Salvini *et al*, 1989; Männistö *et al*, 1996). It is possible, although not quantified in this study, that some inaccuracy is present also in the 7-DD and that any such inaccuracy is more relevant for beverages that may be underreported more than main foods.

Most published reports on reproducibility and validity of information on coffee and tea intake gave similar results, while there is little information for decaffeinated coffee.

As expected, reproducibility and validity tended to be higher among younger, more educated subjects; apparently it was also better for men and heavier participants (body mass index over 25 kg/m^2). The FFQ slightly overestimated coffee intake in either sex; however, the difference is smaller than 10% and the overall proportion of subjects correctly classified into the same category of consumption is satisfactory. The percentage of subjects classified in the same category of intake was the lowest for coffee and the highest for decaffeinated coffee, which mostly depends on the use of quintiles for coffee, tertiles for tea and two categories for decaffeinated coffee, due to the different frequency of consumption of the three items. Moreover, there is also some exposure misclassification when coffee is used as a surrogate measure of caffeine exposure (Brown *et al*, 2001). The data presented on validity mostly refer to the comparison between intake reported in FFQ1 and 7-DD, as FFQ1 is more similar than FFQ2 to the conditions in which the questionnaires are used in epidemiological studies.

The very low prevalence of consumption for decaffeinated coffee and tea is responsible for the very high s.d. values in Tables 2 and 3. For FFQ1 when nondrinkers were excluded, the Pearson *r* were not appreciably modified for decaffeinated coffee (0.61 compared to 0.57) and tea (0.55 compared to 0.56), while decreased for coffee (0.57 compared to 0.69). The information was missing for one subject for coffee, one for decaffeinated coffee and two subjects for tea in the FFQ1; the corresponding values in the FFQ2 were 19, 32 and 33. There were no missing in the diaries, where no information corresponded to no consumption. We also cannot exclude that the differences noted in reproducibility and validity among different demographic groups may be due more to the prevalence of consumption than the agreement of number of cups consumed.

Although coffee and tea give a very low calorie intake, allowance for energy intake was performed to avoid a potential bias of systematic under- or over-reporting. Total energy intake deserves special consideration in nutritional epidemiology because individual differences in total energy intake produce variation in intake of specific nutrients unrelated to dietary composition, because the consumption of most nutrients is positively correlated with total energy intake. This added variation may be extraneous, and thus a source of error in many analyses (Willett, 1998).

There are several important sources of variation and bias in beverage consumption (Beaton, 1994). Variations in reproducibility may reflect not only problems of recall, but also true changes in diet, which is an important source of variation contributing to misclassification of dietary intakes (Willett, 1998), like seasonal patterns in the usage of beverages, which might influence dietary recall (Subar *et al*, 1994). The FFQ did not distinguish between espresso and mocha, and we assumed half consumption as espresso and half as mocha coffee, which may lead to lower coefficients for validity. The effect of order and day of week may be confounding factors in the diaries, and some order effect (earlier days in the 7-day period compared to later days in the 7-day period) cannot be excluded. However, most of the 7-DD started on Sunday, and dietary records were kept for a number of days sufficient to estimate the average intake and cover the period of time corresponding to the questions in the questionnaire (1 week). We have not tested analytically the seasonal variation of coffee intake, but the mean consumption obtained from the two diaries recorded 3–10 months apart was not different. Another possible source of bias in reproducibility studies is sampling of subjects (Ferraroni *et al*, 1996). Volunteers are not representative of the general population, as they are generally more concerned with health and diet than other people, so the reproducibility might be overestimated. This is largely unavoidable in validation studies; however, the involvement of a large number of subjects of both sexes over a wide range

of ages and cultural status gives strength to the present results. Potential bias in the selection of volunteers should hardly affect the comparisons as shown by the similar results across strata of age, sex, education and body mass index.

In conclusion, the present FFQ was developed for a network of case–control studies in different Italian regions, and its validity compared to the reference method (two 7-DD records) was satisfactory and compatible with the best results published so far for coffee. Although neither of the two methods to obtain dietary information is free from errors, we tried to reduce the correlation between errors in the FFQ and 7-DD records (eg definition of portion size). Thus, information on coffee, decaffeinated coffee and tea provided by this FFQ is a satisfactory, reliable and valid measure of intake of these beverages for the purpose of epidemiological inference in the Italian population

References

Beaton GH (1994): Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. *Am. J. Clin. Nutr.* **59**, 253S–261S.

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

Brown J, Kreiger N, Darlington GA & Sloan M (2001): Misclassification of exposure: coffee as a surrogate for caffeine intake. *Am. J. Epidemiol*. **153**, 815–820.

D'Amicis A & Viani R (1993) The consumption of coffee. In: *Caffeine, Coffee and Health*. Monographs of the Mario Negri Institute for Pharmacological Research, ed. S Garattini, pp 1–16. Milan.

D'Avanzo B, La Vecchia C, Katsouyanni K, Negri E & Trichopoulos D (1996): Reliability of information on cigarette smoking and beverage consumption provided by hospital controls. *Epidemiology* **7**, 312–315.

Decarli A, Franceschi S, Ferraroni M, Gnagnarella P, Parpinel MT, La Vecchia C, Negri E, Salvini S, Falcini F & Giacosa A (1996): Validation of a food-frequency questionnaire to assess dietary intakes in cancer studies in Italy. Results for specific nutrients. *Ann. Epidemiol.* **6**, 110–118.

EPIC Group of Spain (1997): Relative validity and reproducibility of a diet history questionnaire in Spain. I. Foods. *Int. J. Epidemiol*. **26**(Suppl 1), S91–S99.

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

Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB & Willett WC (1993): Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J. Am. Diet. Assoc*. **93**, 790–796.

Franceschi S, Negri E, Salvini S, Decarli A, Ferraroni M, Filiberti R, Giacosa A, Talamini R, Nanni O, Panarello G & La Vecchia C (1993): Reproducibility of an Italian food frequency questionnaire for cancer studies: results for specific food items. *Eur. J. Cancer* **29A**, 2298–2305.

Franceschi S, Barbone F, Negri E, Decarli A, Ferraroni M, Filiberti R, Giacosa A, Gnagnarella P, Nanni O, Salvini S & La Vecchia C (1995): Reproducibility of an Italian food frequency questionnaire for cancer studies. Results for specific nutrients. *Ann. Epidemiol.* **5**, 69–75.

Istituto Centrale di Statistica (ISTAT) (1986): Indagine statistica sulle condizioni di salute della popolazione e sul ricorso ai servizi sanitari—novembre 1983. Note e Relazioni. *ISTAT* (Roma) 1.

Jacobsen BK & Bonaa KH (1990): The reproducibility of dietary data from a self-administered questionnaire. The Tromso Study. *Int. J. Epidemiol*. **19**, 349–353.

Johansson L, Solvoll K, Opdahl S, Bjorneboe GE & Drevon CA (1997): Response rates with different distribution methods and reward, and reproducibility of a quantitative food frequency questionnaire. *Eur. J. Clin. Nutr.* **51**, 346–353.

Landis JR & Koch GG (1977): The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174.

Liu K, Stamler J, Dyer A, McKeever J & McKeever P (1978): Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *J. Chronic Dis.* **31**, 399–418.

Männistö S, Virtanen M, Mikkonen T & Pietinen P (1996): Reproducibility and validity of a food frequency questionnaire in a case–control study on breast cancer. *J. Clin. Epidemiol.* **49**, 401–409.

Munger RG, Folsom AR, Kushi LH, Kaye SA & Sellers TA (1992): Dietary assessment of older Iowa women with a food frequency questionnaire: nutrient intake, reproducibility, and comparison with 24-hour dietary recall interviews. *Am. J. Epidemiol.* **136**, 192–200.

Nicoli MC, Dalla Rosa M, Lerici CR & Severini C (1987): Caratteristiche chimiche dell'estatto di caffè. Nota I. Cinetica di estrazione della caffeina e delle sostanze solubili. *Ind. Alim.* **26**, 467–471.

Pagano R, Negri E, Decarli A & La Vecchia C (1988): Coffee drinking and prevalence of bronchial asthma. *Chest* **94**, 386–389.

Pietinen P, Hartman AM, Haapa E, Räsänen L, Haapakoski J, Palmgren J, Albanes D, Virtamo J & Huttunen JK (1988): Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. *Am. J. Epidemiol.* **128**, 655–666.

Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B & Willett WC (1989): Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. *Int. J. Epidemiol*. **18**, 858–867.

Snedecor GW & Cochran WG (1980): *Statistical Methods*, Seventh Edition. Ames, Iowa: The Iowa State University Press.

Subar AF, Frey CM, Harlan LC & Kahle L (1994): Differences in reported food frequency by season of questionnaire administration: the 1987 National Health Interview Survey. *Epidemiology* **5**, 226–233.

Tavani A, Negri E, Ferraroni M, D'Avanzo B, Decarli A, Giacosa A, La Vecchia C, Nanni O, Parpinel MT, Salvini S, Talamini R & Franceschi S (1995): Influence of some covariates on the reproducibility of an Italian semi-quantitative food frequency questionnaire. *Eur. J. Cancer Prev.* **4**: 319–327.

Venturini M (1984): *Dietologia e Dietetica*. Roma: Editoriale Italiana.

Willett W (1998): *Nutritional Epidemiology*, Vol. **30**. Monographs in epidemiology and biostatistics. New York: Oxford University Press.