Journal of Korean Association
of Cancer Prevention
2001; 6(4): 201-213

Development of Dietary Assessment Tools for an Epidemiological Study with a Small Number of Participants

Hachung Yoon[1], Robert Benamouzig[2] and Keun-Young Yoo[1]

[1]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea
[2]Service de gastro-enterologie, Hôpital Avicenne. 125, rue de Stalingrad, 93009 Bobigny cedex, France

Environmental factors, including diet and chemo-preventive agents such as aspirin, vitamins and antioxidants, seem to be implicated in colorectal carcinogenesis. A randomized intervention trial is underway to test the efficacy of aspirin on colorectal adenoma recurrence. To understand dietary patterns of participants in this trial and to identify dietary factors associated with colorectal adenoma development, a case-control study was designed. The present study aimed at developing and validating dietary assessment instruments for use in a small study population. A food composition database as well as programs for data-entry and for converting to daily dietary intake were developed. Estimated dietary intakes were compared to dietary patterns available on the French population. We observed highly positive correlations with the nutritional intakes estimated by the E3N study's instruments for the same population (r=0.94 for total calorie intake; $p < 10^{-3}$). Our dietary assessment instruments are capable of measuring dietary intakes within the framework of the French population's real dietary pattern and rank them correctly in the study population. The easy use of these dietary assessment instruments shows their potential to be used in other studies with a small number of subjects, aiming at evaluating associations between diet and disease.

Key Words: Colorectal adenomas, Dietary assessment tools, Food frequency questionnaire

INTRODUCTION

Colorectal cancer is the second most frequent cause of death from cancer in Western countries.[1)] Many lines of evidence have shown that dietary factors influence all stages of colorectal carcinogenesis.[2)] Aspirin and non steroidal anti-inflammatory drugs (NSAID) seem to have a chemopreventive effect against colorectal cancer.[3)] A randomized interventional trial is underway to test the efficacy of aspirin on colorectal adenoma recurrence (the APACC study: **A**ssociation pour la **P**revention par **A**spirine du **C**ancer **C**olorectal). To have insight into the dietary patterns of the study population and to evaluate dietary factors related to the development of colorectal adenomas, a nested case-control study was designed.

The Food Frequency Questionnaire (FFQ) is the most commonly used dietary investigation instrument

Corresponding author : Hachung Yoon, Department of Preventive Medicine, Seoul National University College of Medicine
28 Yeongeon-dong, Jongno-gu, Seoul 110-799, Korea
Tel: 02-740-8329, Fax: 02-3673-3540, E-mail: heleney@snu.ac.kr

in epidemiological studies attempting to relate usual dietary intake to disease. The FFQ is easy to administer by the participant alone, inexpensive to process and able to correctly represent the dietary pattern.[4] A self-administered dietary questionnaire for French meal patterns, was developed for use in a large French cohort study aiming at studying cancer risk factors on 100,000 women, called E3N (Etude Epidmiologique auprs de Femmes de la Mutuelle Gnrale de l'Education Nationale) and was validated together with the analyzing instruments.[5] Owing to the relatively small size of our study population, the direct application of these instruments to the APACC study did not seem easy. The 290 subjects needed for the APACC study were calculated, expecting that daily aspirin administration would bring about a 30% reduction in the actual recurrence rate, 43% 4 years after a polypectomy. The present study therefore aimed at developing and validating dietary analyzing instruments, the most adapted to the APACC study plan, for an already validated questionnaire.

STUDY DESIGN

The APACC Study is a prospective, randomized, double-blind, placebo-controlled multicenter clinical trial, designed to determine the efficacy of regular low dose aspirin intake (160 or 300 mg/d *versus* placebo) in reducing the recurrence rate of colorectal adenomas. The study involves 49 gastroenterology centers from various parts of France. After a four-weeks run-in period to test treatment compliance, eligible subjects (aged 18～75 years with at least one histologically confirmed adenoma) were randomized into either the intervention or the control groups. Participants visit the Gastroenterologist Investigator every four months, and information on compliance, tolerance of the treatment and concomitant disease is obtained. The efficacy of aspirin will be evaluated by systematic control colonoscopies in years 1 and 4 after enrollment. The APACC study has been approved by the ethical committee of Saint Germain en Laye Hospital for studies with human subjects (CCPPRB N^{o} 96006) and by the 'Comit National Informatique et Libert (CNIL)' which advocates that all medical information be confidential and anonymous.

Two hundred and ninety patients were recruited from February 1996 to February 2000. Only 31% (n=89) of this population consisted of women. The mean age of the subjects was 57.63 (SD 9.29) years. Characteristics of the population are described in Table 1. The control group subjects participating in the

Table 1. Some characteristics of the APACC patients (n=290)

	Male (n=201)	Female (n=89)	Total
Age (years): mean (±SD)	57.75 (±9.35)	57.36 (±9.18)	57.63 (±9.29)
Height (cm): mean (±SD)	172.44 (±6.73)	161.40 (±5.77)	-
Weight (kg): mean (±SD)	78.73 (±11.52)	64.24 (±11.02)	-
Body Mass Index (kg/m): mean (±SD)	26.47 (±3.49)	24.68 (±4.23)	-
Smoking habits			
Never smokers: n (%)	68 (33.83)	72 (80.90)	140 (48.28)
Current smokers: n (%)	61 (30.35)	9 (10.11)	70 (24.14)
Ex-smokers: n (%)	72 (35.82)	8 (8.99)	80 (27.59)
Pack-year: mean (±SD)	18.75±16.12	4.25±9.65	14.62±15.96
Personal history of colorectal adenomas: n (%)	51 (26.15)	18 (20.22)	69 (24.30)
Family history of colorectal adenomas: n (%)	31 (15.74)	11 (12.50)	42 (14.48)
Family history of colorectal cancer: n (%)	65 (33.16)	37 (41.57)	102 (35.17)

nested case-control study were recruited after confirming that they were polyp free by a colonoscopy, matching sex and age (±5 years) from the same clinical center as the patients.

DIETARY ASSESSMENT INSTRUMENTS

1) Questionnaire

An important feature of the E3N questionnaire was that it was separated into a quantification part (the first part of the FFQ) and a qualification part (the second part). The reproducibility and validity of nutrient and food intake measurements from this FFQ have been described in detail elsewhere.[5)]

2) Food composition database

A Food composition database, including 211 food items, was derived from the BALI-3 software,[6)] which contains food composition data provided by the CIQUAL (Centre Informatique sur la Qualit des Aliments), Foch Centre and food industries. The BALI-3 food composition Table was compared to that of the E3N's, for data origin, name of items, and nutritional values. Seven hundred and ninety nine items (61.37% of 1302 items) of the BALI-3 Table and four hundred and twenty three (44.90% of 942 items) of the E3N's nutritional values came from the database of the CIQUAL.

3) Data-entry program

Data collected by the FFQ are transformed into computer form by data-entry programs created with ＜C language＞. The dietary data entry proceeds as followings. Information on the quantification part of the FFQ is entered in three steps: ① food type choice (*e.g.* white coffee or black coffee), ② consumption frequency indication: never or less than once a month, 1～3 times a month, or 1～7 times a week, ③ usual consumption quantity description in referring to five portion sizes (smaller than small, small, medium, large, and larger than large) with an album of 42 food item photos, and common measuring units like different types of cup (glass for juice, glass for wine, tea cup, *etc.*) or bread (baguette, sandwich loaf, whole-wheat bread, milk roll, *etc.*). For data in the qualification part, the entry is carried out as follows: if the respondents reported that they eat a food category, answers were coded as 0, 1, 2, 3 respectively for never or seldom (0), every now and then (+), regularly (++), very often (+++) to be weighted in the analyzing program.

Visible and audible warnings occur in the response to inadmissible values. Certain steps following an answer of "never intake" can be skipped. An automatic coding was programmed for that case.

At the end of each block of questions (one food item for the quantification part and one food category for the qualification part), a question to confirm the entry appears and allows the entered data to be modified. A visual verification under a matrix form gives another chance to correct detected errors at the end of each meal section. Two persons independently perform the data-entry. A program comparing these two entries detects any inconsistencies.

4) Dietary intake analyzing algorithm

Nutritional intake and food consumption data are obtained after combining the FFQ and the food composition Table, by weighting using the consumption frequency and the information of the quantification part. The edible proportion of each food is taken into account for estimating the amount of food eaten. Quantities presented by extreme portion sizes are estimated according to formulas used in the E3N study.[7)]

For the food categories that are detailed in the quantification part (milk, yogurt, fresh cheese, cheese, green salad, hot vegetable, fruit, fish, meat, salad dressing, visible fats), data were also weighted by the relative consumption frequency (0,+,++,+++) of food eaten (*e.g.* apple, orange, *etc.*).

Quantities of sugar eaten with yogurt, fresh cheese and hot beverages are added along with the food consumption frequencies. In the same manner, the

types and the quantities of fat eaten with certain foods, called visible fats, are identified and applied with the same frequency as associated foods. These are: ① dressing for green salad, ② butter or margarine spread on bread and crispbread, ③ butter with cheese, ham, canned fish, sausage, seafood, ④ mayonnaise with boiled egg and seafood.

When respondents answered that they prepare food with fat or add fat after cooking, the reported amounts are added to their daily dietary intake by multiplying by the food consumption coefficient. The coefficient is proportional to the consumption frequencies of some foods eaten for principal and accompanied dishes. For meat and fish, the fats were added according to the cooking method preference indicated in the qualification part.

VALIDATION OF DIETARY ANALYSIS PROGRAM

1) Comparison with the E3N's results

To validate our dietary assessment instruments, we compared nutrient intakes calculated by our method with those calculated by the E3N study group, for 42 female subjects included in the E3N study. The relationship was inferred by the agreement of the two estimates.[8] We began with drawing a plot with the same value every time, all points would lie on this line. Fig. 1 shows the plot for total calorie intake. Then a plot of the difference between the two methods against the mean of the difference was drawn. If most of the differences lay between certain levels of fluctuation, the mean-2SD and the mean+2SD, the agreement would be confirmed (Fig. 2). The same types of plots were drawn for all macro- and micronutrients and the agreement between results from two different instruments were verified. These Figs validated our dietary assessment instruments. We present only the plot with the line of equality and that of the difference for total calorie intake in Fig. 1 and 2.

Table 2 shows total calorie and nutrient intakes evaluated by the two methods. Total calorie estimates are very similar, showing 2043.28±579.44 (mean± SD) kcal/d from the E3N instruments and 2047.55± 598.35 kcal/d from the APACC instruments. We observed no significant difference between the two methodologies, using Student test and Wilcoxon signed rank test for all macro-and micro nutrient intake as well as percentages of macronutrient to total calorie, except dietary fiber ($p < 0.05$). Estimated Pearson correlation coefficients and Spearman rank correlation coefficients showed highly positive correlations ($p < 0.001$ for all studied nutrients, Table 2).

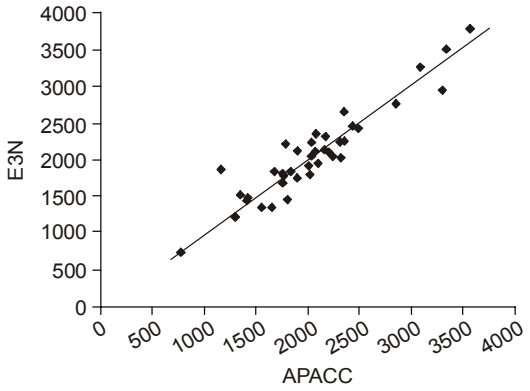

Fig. 1. Daily calorie intake (kcal/day) estimated by the E3N's and the APACC's tools with the line of equality. line of equality. If the two methods give exactly the

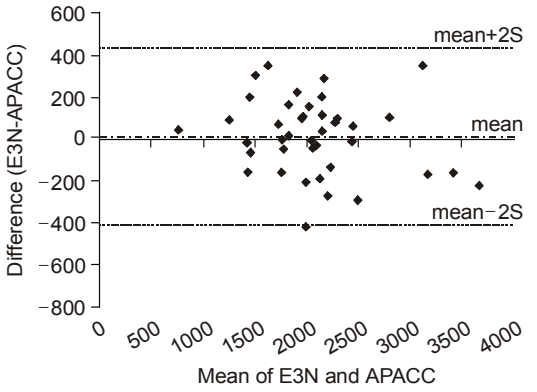

Fig. 2. Difference between the E3N's estimate and the APACC's against the mean of the two estimates with lines of the mean and ±2SD for daily total calorie intake (kcal/day).

Table 2. Daily nutrient intakes estimated by the methods of the APACC and the E3N for 42 female subjects included in the E3N study

	E3Ntools Mean±SD	APACC tools Mean±SD	Correlation$		
			r_p	r_s	r_c
Total calories (Kcal)	2043.28±579.44	2047.55±598.35	0.94	0.90	0.94
Proteins (g)	87.57±20.68	93.45±23.37	0.89	0.85	0.89
% Energy[1)]	18.4±73.53	19.76±3.78	0.81	0.72	0.82
Carbohydrates (g)	220.72±77.2 9	212.53±73.62	0.94	0.90	0.94
% Energy[1)]	44.88±6.60	43.60±6.61	0.78	0.75	0.78
Total Fats (g)	79.50±25.84	79.95±28.46	0.83	0.82	0.83
% Energy[1)]	36.65±5.47	36.64±5.85	0.61	0.60	0.61
Saturated fat (g)	31.87±13.38	34.33±14.32	0.86	0.88	0.86
Monounsaturated fat (g)	24.26±10.53	26.88±10.16	0.75	0.71	0.75
Polyunsaturated fat (g)	11.82±5.62	11.46±4.52	0.43	0.43	0.43
Cholesterol (g)	384.00±191.00	379.59±125.83	0.89	0.86	0.89
Fiber (mg)*	20.54±6.12	23.50±6.55	0.76	0.78	0.77
Alcohol (g)	14.73±16.51	14.86±16.40	0.98	0.99	0.98
Calcium (g)	1002.12±342.51	1129.17±426.70	0.60	0.69	0.60
Iron (mg)	13.49±3.82	13.85±3.88	0.88	0.87	0.88
Vitamin C (mg)	145.96±69.60	120.75±53.12	0.83	0.82	0.84

[1)]percentage in total calories without alcohol
r_p: Pearson correlation coefficient
r_s: Spearman correlation coefficient
r_c: Corrected correlation coefficient
$All correlations are significant at the level $p < 0.001$
*The only significant difference of means between two methods

To take inter-individual variability into account, we also evaluated the deattenuated correlation coefficients (r_c) according to the following formula[9)]:

$$r_c = r_o \sqrt{(1+\gamma/k)}$$

where r_c is the corrected correlation between two analyzing methods, r_o is the observed correlation (we used the Pearson correlation coefficient r_p), γ is the ratio of estimated within-person and between-person variation from the two estimations, and k is the number of repeated estimations (in this study k=2). The three correlation methods showed practically the same coefficient.

2) Validation in using a part of the APACC population

The dietary data-collection file of the APACC consists of four elements: ① an information form, ② a green covered FFQ, ③ a white covered questionnaire to obtain information about age, sex, occupation, personal and family medical histories, and reproductive history for women, ④ a subject's consent form. The Gastroenterologist Investigator distributes the file to the patients (case and control) at the recruitment visit. Study participants are asked to give their present dietary habits. After filling in the two questionnaires, each subject sends them back in the stamped envelope to the APACC study coordination

Table 3. Percent of missing data and imputation methods on each variable in the APACC patients (n=77)

	Male (n=49)		Female (n=28)		Replacement by
	total missing[1] n (%)	subject[2] n (%)	total missing[1] n (%)	subject[2] n (%)	
Consumption frequency	81 (0.72)	28 (57.14)	54 (0.84)	18 (64.29)	○ *X* for month → 2 times/month ○ *X* for week → 3 times/week ○ Else → average of 2/month and 3/week
Consumption quantity	75 (0.67)	30 (61.22)	83 (1.29)	21 (75.00)	○ Quantity of the same food item from other meal of the same individual ○ Else → Medium portion size or unity of the food (croissant, egg, *etc.*)
Quantities of visible fats					
Eaten with hors d'uvre or cheese	17 (34.69)		8 (28.57)		Median of observed data according to sex
Used for cooking	11 (22.45)		9 (32.14)		Iteration of a multiple linear regression[3]
Added after cooking	10 (20.41)		6 (21.43)		Iteration of a multiple linear regression[3]
Salad dressing	3 (6.12)		1 (3.57)		Mean adjusted for age, BMI, and quantity of salad, by sex
Food type identification at 2nd part	24 (5.44)	13 (26.53)	19 (7.54)	6 (21.43)	Standard (not low-fat) type of the food

[1] total missing: number of missing data concerning the variable
[2] subject: number of subjects having at least one missing datum concerning the variable
[3] variables entered in the regression model: age, BMI, eaten quantity of principal and accompanied dishes

Table 4. Mean±SD daily total calorie and nutrient intake in 77 patients of the APACC study

	Male (n=49)	Female (n=28)
Total calories (kcal/day)	2978.90±757.12	2323.99±631.48
Proteins (g/day)	122.72±32.78	101.02±24.68
% energy[1]	18.16±2.77	18.62±3.12
Carbohydrates (g/day)	296.97±103.42	240.28±90.05
% energy[1]	42.94±8.15	42.52±7.03
Total fats (g/day)	117.18±35.91	96.10±33.46
% energy[1]	38.96±7.22	38.86±6.36
Saturated fat (g/day)	41.05±14.90	32.45±12.61
Monounsaturated fat (g/day)	32.23±9.73	24.85±8.17
Polyunsaturated fat (g/day)	12.52±3.47	8.99±3.08
Cholesterol (g/day)	574.21±285.40	445.65±194.45
Fiber (mg/day)	26.14±9.33	22.16±8.53
Alcohol (g/day)	35.43±32.51	13.41±15.24
Calcium (mg/day)	1227.60±506.29	1282.61±443.07
Iron (mg/day)	20.39±6.08	16.71±4.40
Vitamin C (mg/day)	89.42±48.66	111.79±66.04

[1] percentage in total calories without alcohol

center.

Three hundred files were sent and one hundred and thirty (43.33%: 103 cases and 27 control) were returned by December 1999. A total of 76% (n=99) seemed able to be exploited. The mean (±SD) age of the patients was 56.10 (±11.04) years for 49 men and 57.71 (±10.12) years for 28 women. The Body Mass Index (BMI; weight/height^2) was 26.08±3.43 (meanSD) for males and 24.8±62.80 for females. Gender proportion, age and BMI were not different from the 290 recruited subjects. For the moment, we will not exploit the dietary data for the 27 controls because of the small number of responded questionnaires.

We could not estimate the dietary habits for 34% (n=31) of respondents who had marked "X" instead of writing a number for consumption frequency. These questionnaires were sent back to the participants for correction. When there was only one missing value for a set of questions (food type, frequency and quantity of consumption), we tried to apply an imputation. Table 3 shows the per cent missing on each variable for 77 cases and imputation methods used. We found a tendency that women generated slightly more missing values than did men, but no significant difference was observed between the two sexes. The amounts of sugar added to coffee or tea and those of butter or margarine spread on bread are not presented in Table 3. The missing response proportion was very small (0.4% for butter and 0.3% for sugar), we therefore replaced by the mean of the observed data. Rules used for the missing data imputations of the quantify of visible fat consumption, are the following: ① verify the normality of variable distributions. Transformation into natural logarithm was carried out if necessary, ② choose the predictive variables[10] using "stepwise" option in "PROC REG" procedure of SAS, from the suggested set of predictor variables given by Y=[X,U] where,

Y : fat quantity
X : covariates (for present study they are age, BMI)

U : quantities of food eaten with fats

③ then choose an imputation method, considering the distribution normality and the selected covariates (Table 3). All imputation models were evaluated separately for men and women.

Table 4 shows the average daily intakes of total calories and macro- and micro-nutrients in 77 patients, after the treatment of missing data. Total calorie intake was 2978.90±757.12 kcal/d (mean±SD) in males and 2323.99±631.48 kcal/d in females. Contributions of proteins, carbohydrates and total fats to total calories without alcohol were no different in males or in females. Table 5 shows the energy and macronutrient assignment to each meal.

DISCUSSION

A study to evaluate validity and reliability is valid for a specific population with a specific questionnaire.[11] In the validation study for the FFQ used in our study, all participants were women.[5] The question therefore was whether the same validation results would be obtained from a male population, when the E3N's FFQ was applied to our study. In the present study, we observed highly positive correlations among foods eaten in a food category between the two sexes (not reported). Moreover, the missing data generating pattern did not differ in males or females. We can thus infer that the FFQ is also applicable to men.

More than three-quarters of the subjects had at least one missing variable. Since we could not guess the true value, it was impossible to distinguish the hypothesis of "Missing Not At Random (M.N.A.R.)" from the "Missing At Random (M.A.R.)" for a given explicative variable. However, none of the questions -on type, frequency or quantity of food eaten- in the FFQ did not seem to make the subject not to answer. The choice of missing data management methods depends upon the size of the bias likely to occur. Generally, it is not necessary to adopt refined methods to manage data in which the problematic data are below 5%.[12] Since the ranking of consumption in the study population is of greatest interest, we chose the method presenting the highest correlation coefficient with the observed data. Without the missing data treatment, we considered information concerning the problematic answers as no consumption. We evaluated the difference in intake of total calories or certain nutrients brought by missing data treatment. We found no statistically significant difference in the dietary intakes before and after the imputation. The draws for imputation in the present study seem thus reasonable in view of the covariates and the observed data.

In the present validation study, our analysis of nutrient intakes compares favorably with those of the E3N in forty-two females included in the E3N study. We selected the correlation measurement method, because this method reflects well the rank of the amount eaten.[13] Spearman rank correlation coefficients were positive and significant for all nutrients including specific ones suggested to be related to colorectal adenomas such as total fats, saturated fat, monounsaturated fat, polyunsaturated fat and dietary calcium. Even for dietary fiber, for which we observed a significant difference between the two estimations, Spearman rank correlation coefficient was high enough (r_s=0.78, $p < 0.001$) to guarantee a good concordance.

We observed somewhat higher dietary calcium and iron intakes compared to the intakes estimated by the E3N's instruments. In addition, intakes of these nutrients in the population included in the APACC study also seem to be higher than RDA for the French population[14] and other studies performed on the French population.[15,16] This can be attributed to a large water consumption. Men drank tap and bottled water (mean±SD) 1046.17±858.87 ml/d and women drank 1231.84±1039.17 ml/d, excluding water in foods. Water consumption provided an average of 12.09% (147.35 mg/d) of calcium and 7.69% (1.67 mg/d) of iron for men and 15.10% (197.28 mg/d) and 13.41% (2.45 mg/d) for women, respectively. Many food composition Tables do not include the item

Table 6. Total calorie intakes observed from studies conducted in France

	Method	Age (Yrs)	Male				Female				First author, publication year [reference]
			Energ (kcal/d) mean±SD	% Energy without alcohol (mean)			Energy (kcal/d) mean±SD	% Energy without alcohol (mean)			
				Protein	Carbo-hydrate	Fats		Protein	Carbo-hydrate	Fats	
On healthy population											
ASPCC	7 days DR	55~64	2097[1]	18.0	41.1	40.8	1708[1]	18.7	39.7	41.4	Rigaud, 1997 [20]
Burgundy	DH	30~79	2278±521[2]	15.9[6]	43.9[6]	40.3[6]					Boutron, 1989 [21]
ESVITAF	7 days DR	18~45	3112±598	13.8	40.8	45.4	2418±484	14.2	39.8	46.0	ESVITAF, 1986 [22]
INCA	7 days DR	≥15	2513±588	17.7	44.2	38.1	1944±421	17.3	43.9	38.8	CREDOC, 2000 [18]
PETRA	4 days DR		2121±588[3]	17.6[6]	37.0[6]	45.4[6]	-	-	-	-	Bonifaci, 1997 [23]
SU.VI.MAX	24h DR	45~60	2428±623	18[7]	38[7]	37[7]	1811±482	18[7]	41[7]	38[7]	Galan, 1998[16]
Val de Marne	FH	50~65	2540±678	15.1	40.4	33.1	1660±454	18.0	42.4	35.2	Hercberg, 1991 [15]
On patients											
E3N	FFQ	40~65	-	-	-	-	2350±766	15.0	41.8	43.2	Clavel-Chapelon, 1997 [17]
Marseille	FFQ	M 59[4] F 62[4]	2243[5]	15.8[6]	45.5[6]	38.7[6]	1662[5]	16.6[6]	39.6[6]	43.9[6]	Marcquart-Moulin, 1987 [24]
APACC	FFQ	31~75	2979±757	18.2	42.9	41.1	2324±631	18.6	42.5	38.9	-

DR : Dietary Record FH: Food History FFQ: Food Frequency Questionnaire

[1]median: 5th and 95th percentile 1186~3188 for males and 952~2343 for females, respectively

2)pilot test for a case-control study, adjusted for age and sex

3)transformed from kJ into kcal; original value=8866±2458 Kj/day

4)mean age

5)median in patients

[6]% energy estimated from mean value of each macronutrient

7)% energy with alcohol (% alcohol of total calories are 7% in males and 3% in females)

bottled water or Mineral water. Since we analyzed the mineral composition of bottled water (mineral and spring), it is not surprising to find higher calcium and iron intakes than other studies which used a food composition Table without “Mineral water” item.

Table 6 presents the total calorie intakes and energetic assignment by macronutrients, observed on the French population. Globally, the absolute value of total calorie intake in our population seems to be higher than in other studies. The dietary intakes obtained from our population were similar to the results obtained by the same FFQ from 7505 females, a representative sample of participants in the E3N study.[17] Energy consumption in the APACC’s female population (mean±SD 2428±765 kcal/d) was comparable to the E3N’s (2350±766 kcal/d). We must consider that the population included in the APACC study was an average ten years older and 5 kg heavier than the E3N’s subjects (mean±SD 57.3±9.2 *versus* 48.9±6.7 years; 64.2±10.7 *versus* 59.1±9.1 kg). In nutritional epidemiology, energy-related Figs of nutrient intakes are important and we found similar proportions of macronutrient’s contribution to total calorie consumption to those in other studies (Table 6). Breakfast and snacks had clearly more carbohydrates than lunch or dinner for the population of the

Table 7. Number of food items in FFQ and observed total calories by FFQ and dietary record

First author, publication year	Country	Age	Number of items	Total calories mean±SD (kcal/d)				Reference study
				FFQ		DR*		
				Men	Women	Men	Women	
Pisani, 1997	Italy	M 50.5[1] F 49.4[1]	47	2453±697	2094±608	2757±557	2164±532	The EPIC study [25]
Caan, 1998	USA	60～79	105	1711±619	1467±518	2455±766	1894±606	Block’s HHHQ[2] [26]
Martinez, 1999	USA	40～80	113	1825±583		1916±564		AFFQ for WBF Trial[3] [27]
Caan, 1998	USA	60～79	126	2057±660	1855±513	2357±739	1921±611	Willet’s FFQ [26]
Bonifacj, 1997	France	M 42.4[1] F 40.8[1]	134	2121±588[4]		2144±485[5]		To develop dietary assessment instruments for the French in Mediterranean region [23]
Bohlscheid-Thomas, 1997	Germany	35-64	158	2305±733		2193±520		The EPIC study [28]
Ocké, 1997	Netherlands	M 42.6[1] F 49.0[1]	178	2772	1912	2701	1798	The EPIC study [29]
Katsouyanni, 1997	Greece	M 25～67	190	2476±993	-	2385±607	-	The EPIC study [30]
Van Liere, 1997	France	F 40～65	209	-	2034±643[6]	-	1673±348[7]	The EPIC study [5]

[1]mean age
[2]Block Health Habits and History Questionnaire
[3]Arizona Food Frequency Questionnaire for Wheat Bran Fiber Trail
[4]Transformed from KJ into kcal; original value=8866±2458 KJ/day
[5]Transformed from KJ into kcal; original value=8692±2027 KJ/day
[6]Transformed from KJ into kcal; original value=8502±2688 KJ/day
[7]Transformed from KJ into kcal; original value=6993±1455 KJ/day
*Dietary Record (DR)
for the EPIC study : 24-hour recall
for the report of Caan *et al.* [26]: DARCCD DH (Diet, Activity, and Reproductive Risks for Colon Cancer Diet History)
for WBFT: 4 day diet record

APACC study. Lunch is the most caloric meal of the day and is rather proteic (Table 5). These findings are comparable to results of the INCA study. 27.1% and 9.7% of carbohydrates came from the breakfast and snacks, respectively. 38.9% of daily total calories and 47.4% of daily protein intake of the French population were assigned to lunch.[18] The present study assures us that our dietary assessment instruments are capable of measuring the dietary consumption within the framework of the French population's real dietary pattern.

In general, longer food frequency lists overestimate intake, whereas shorter lists underestimate intake as compared to quantitative methods of 24-hour recall and the food record.[19] Table 7 shows total calorie intakes estimated by the food frequency questionnaire and the dietary records with the number of food items in the FFQs, published between 1997 and 2000 in adult populations. In agreement with Krebe-Smith et al.,[19] we found higher total calorie intakes estimated by FFQ in which the number of food items is above 130, than those estimated by a dietary record. The FFQ, with 209 items, used in our study also showed a higher total calorie estimation than the 24 hour recall dietary record.[5]

The present study demonstrated that our dietary assessment instruments successfully convert the FFQ to daily nutrient intakes. It was ascertained that our dietary assessment instruments can correctly rank the subjects in their distribution of dietary intakes. Our food composition database remains incomplete (17～24%) for certain micronutrients such as retinol, carotene, vitamins D, E, C, B6 & B12, thiamin, riboflavin, niacin, pantheoleic acid and folic acid. They are all of interest in epidemiologic studies on diet and cancer. Therefore a second food composition database is under construction to analyze the micronutrients as fully as possible. The dietary assessment instruments provide a global view of the dietary pattern in subjects included in the APACC study as well as guaranteeing the equality between intervention arms. They also identify dietary factors related to the development and recurrence of colorectal adenomas. Further, the dietary assessment instruments can be exploited for other studies, with small a number of subjects, aiming to evaluate the relationship between dietary pattern and a disease.

ACKNOWLEDGEMENTS

The authors would like to thank all Gastroenterologists Investigators and pathologists participating in the APACC study, EY Cha, MTT Nguyen, M Sauve and Ph Dufreche for their assistance. The APACC study is financially supported by the Association pour la Recherche sur le Cancer (ARC), the Socit Nationale Franaise de Gastro-Entrologie (SNFGE), the Sanofi-Synthelabo Company, and the Programme Hospitalier de Recherche Clinique 96 (PHRC96).

REFERENCES

1) Ghadirian P, Maisonneuve P, Perret C, Lacroix A, Boyle P. Epidemiology of sociodemographic characteristics, lifestyle, medical history, and colon cancer: A case-control study among French Canadians in Montreal. *Cancer Detec Prev* 1998; 22: 396-404.
2) Shike M. Diet and lifestyle in the prevention of colorectal cancer: An overview. *Am J Med* 1999; 106: 11S-15.
3) Benamouzig R. Do aspirin or non steroidal anti-inflammatory drugs decrease the risk of colorectal cancer? *Gastroenterol Clin Biol* 212: S22-27.
4) Margetts BM, Cade JE, Osmond C. Comparison of a food frequency questionnaire with a diet record. *Int J Epidemiol* 1989; 18: 868-873.
5) van Liere MJ, Lucas F, Clavel F, Slimani N, Villeminot S. Relative validity and reproducibility of a French dietary history questionnaire. *Int J Epidemiol* 1997; 26: S128-136.
6) Leblanc JCh, Yoon H, Kombadjian A, Verger Ph. Nutritional intakes of vegetarian population in France. *Eur J Clin Nutr* 2000; 54: 443-449.
7) Lucas F, Niravong M, Villeminot S, Kaaks R, Clavel-Chapelon F. Estimation of food portion size using photohraphs: validity, strenghts, weaknesses and recommendations. *J Hum Nutr Diet* 1995; 8: 65-74.
8) Bland JM, Altman DG. Statistical methods for assess-

ing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307-310.
9) Hu FB, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, Sampson L, Willett WC. Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* 1999; 69: 243-249.
10) van Buurn S, Boshuizen HC, Knook L. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999; 18: 681-694.
11) Thompson FE, Byers T. Dietary assessment resource manual. *J Nutr* 1994; 124: 2245S-2317.
12) Garcia-Acosta S, Clavel-Chapelon F. Dealing with missing, abnormal and incoherent data in E3N cohort study. *Rev Epidm et Sant Publ* 1999; 47: 515-523.
13) Stram DO, Hankin JH, Wilkens LR, Pike MC, Monroe KR, Park S, Henderson BE, Nomura AM, Earle ME, Nagamine FS, Kolonel LN. Calibration of the dietary questionnaire for a multiethnic cohort in Hawaii and Los Angeles. *Am J Epidemiol* 2000; 151: 358-370.
14) Dupin H, Abraham J, Giachettii I, eds. *Apports nutritionnels conseills pour la population franaise.* Paris: Tec&Doc Lavoisier. 1992.
15) Hercberg S, Preziosi P, Galan P, Deheenger M, Dupin H. Dietary intakes of a representative population sample in a district of Paris area (Val de Marne). II. Energetic nutrient intakes. *Rev Epidmiol Sant Publ* 1991; 39(3): 233-244.
16) Galan P, Yoon HC, Preziosi P, Viteri F, Valeix P, Fieux B, Brianon S, Malvy D, Roussel AM, Favier A, Hercberg S. Determining factors in the iron status of adult women in the SU.VI.MAX study. *Eur J Clin Nutr* 1998; 52: 383-388.
17) Clavel-Chapelon F and the E3N group. E3N, a French cohort study on cancer risk factors. *Eur J Cancer Prev* 1997; 6: 473-478.
18) CREDOC-AFSSA-DGAL. *Enqute nationale sur les consommations alimentaires.* Paris: Tec&doc Lavoisier. 2000.
19) Krebs-Smith SM, Heimendingers J, Subar AF, Patterson BH, Pivonka E. Estimating fruit and vegetable intake using food frequency questionnaires: a comparison of instruments. *Am J Clin Nutr* 1994; 59: 283S.
20) Rigaud D, Giachetti I, Deheeger M, Borys JM, Volatier JL, Lemoine A, Cassuto DA. Enqute franaise de consommation alimentaire I. Energie et Macronutriments. *Cah Nutr Dit* 1997; 32: 379-389.
21) Boutron MC, Faivre J, Milan C, Lorcerie B, Estve J. A comparison of two diet history questionnaires that measure usual food intake. *Nutr Cancer* 1989; 12: 83-91.
22) ESVITAF. Vitamin status in three groups of French adults. *Ann Nutr Metabol* 1986; 30: 28-50.
23) Bonifacj C, Gerber M, Scali J, Daures JP. Comparison of dietary assessment methods in a Southern French population: use of weighted record, estimated-diet records and a food-frequency questionnaire. *Eur J Clin Nutr* 1997; 51: 217-31.
24) Macquart-Moulin G, Riboli E, Cornee J, Kaaks R, Berthezne P. Colorectal polyps and diet: A case-control study in Marseilles. *Int J Cancer* 1987; 40: 179-188.
25) Pisani P, Faggiano F, Krogh V, Palli D, Vineis P, Berrino F. Relative validity and reproducibility of a food frequency dietary questionnaire for use in the Italian EPIC Centers. *Int J Epidemiol* 1997; 26: S152-160.
26) Caan BJ, Slattery ML, Potter J, Quesenberry CP, Coates AO, Schaffer DM. Comparisons of the Block and the Willett self-administered semiquantitative food frequency questionnaires with an interview-administered dietary history. *Am J Epidemiol* 1998; 148: 1137-1147.
27) Martinez ME, Marshall JR, Graver E, Whitacre RC, Woolf K, Ritenbaugh C, Alberts DS. Reliability and validity of a self-administered food frequency questionnaire in a chemoprevention trial of adenoma recurrence. *Cancer Epidemiol Biomarkers Prev* 1999; 8: 941-946.
28) 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. *Int J Epidemiol* 1997; 26: S71-S81.
29) Ock MC, Bueno-de-Mesquita H, Goddijn HE, Jansoen A, Pols MA, Staveren WAV, Kromhout D. The Dutch EPIC food frequency questionnaire II. Relative validity and reproducibility for nutrients. *Int J Epidemiol* 1997; 26: S49-S58.
30) Katsouyanni K, Rimm EB, Gnardellis C, Trichopoulos D, Polychronopoulos E, Trichopoulou A. Reproducibility and relative validity of an extensive semi-quantitative food frequency questionnaire using dietary records and biochemical markers among Greek school teachers. *Int J Epidemiol* 1997; 26: S118-127.

적은 수의 대상자를 가진 역학적 연구에서 식이습관 측정도구의 개발

[1]서울대학교 의과대학 예방의학교실, [2]Service de gastro-enterologie, Hpital Avicenne. 125, rue de Stalingrad, 93009 Bobigny cedex, France

윤 하 정[1], Robert Benamouzig[2], 유 근 영[1]

식생활습관을 비롯하여 아스피린, 비타민과 항산화제 등을 포함한 화학적 암예방 물질 등은 대장-직장의 발암과정에서 중요한 역할을 하는 것으로 알려져 있다. 대장 선종 제거술 후에 재발을 억제하는 아스피린의 효과를 검증하기 위한 한 개입연구가 프랑스인들을 대상으로 진행 중에 있다. 이 연구에 참여하는 대상자들의 식생활습관을 탐구하고, 선종의 발생 위험에 연관된 식이요인을 규명하기 위하여 환자-대조군 연구가 계획되었다. 본 연구는 이 환자-대조군 연구에서 대상자들의 식생활 습관을 추정하는 도구를 개발하고 실용성을 확증하는 것을 목적으로 하였다. EPIC study의 프랑스 부분인 E3N study에 맞게 개발된 식이섭취빈도 설문지를 적은 수의 대상자 집단에서 분석하기 위하여 식품성분표, 설문자료를 입력하는 프로그램, 전산 입력된 자료를 일일 식품 및 영양소 섭취량으로 전환하는 프로그램 등이 개발되었다. 추정된 일일 영양 섭취량을 E3N 자체의 분석도구를 이용하여 분석한 수치와 비교해 본 결과, 매우 높은 상관계수를 관찰하였다 (총열량 섭취량에 대한 상관계수 $r=0.94$; $p<10^{-3}$). 본 연구에서 개발된 식이 습관 측정도구는 대상자들의 식생활 습관을 프랑스 국민들의 실제 식습관의 범위내에서 반영하며, 집단내에서는 대상자들의 섭취량을 순위별로 층화하여 분류하는 데에 적합한 것으로 판단된다. 개발된 도구들의 용이한 사용법은 질병과 식생활의 연관성을 규명하는 것을 목적으로 하는 다른 연구에도 쉽게 이용될 수 있을 것으로 사료된다.

Key Words: 대장 선종, 식생활 측정도구, 식이섭취빈도 조사법