A computerized diet questionnaire for use in diet health education 1. Development and validation

BY M. G. O'DONNELL*, M. NELSON† P. H. WISE AND D. M. WALKER‡

Department of Endocrinology, Charing Cross Hospital, Fulham Palace Road, London W6 8RF

(*Received 14 August 1990 – Accepted 8 October 1990*)

A diet questionnaire was developed in association with a computer program to provide rapid nutritional feedback to the general public. The questionnaire was validated against 16 d of weighed diet records and biochemical variables in blood and urine. The highest Pearson correlation coefficients obtained between the questionnaire and the weighed records were for alcohol, fibre, iron, riboflavin (*r* 0·74, 0·67, 0·66, 0·66 respectively). Striking sex differences were shown in the results; the trend for higher correlations persisted in females. At least 65% of subjects were classified by questionnaire to within one quintile of the classification by weighed record for the majority of nutrients.

Diet questionnaire: Computer program: Validation: Health education

In Britain, over the last 30 years, major changes have occurred in diet: on average, a downward trend in carbohydrate consumption from 50% of dietary energy in 1950 to 44% in 1982; while energy derived from fat has increased from 35% in 1950 to 42·8% in 1983 (Ministry of Agriculture, Fisheries and Food, 1984). Over a similar time-period there has been an increase in diseases of affluence: obesity, coronary heart disease, cancer, and dental caries. Whilst these diseases are multifactorial in origin (stress, smoking, genetic susceptibility, lack of physical exercise), diet is recognized as an important factor in their development. They also now affect a large proportion of the population.

Consumers are increasingly aware and concerned about food and food safety, but also increasingly bewildered and confused by two phenomena: first, by conflicting messages about which foods are good for health, and second, by ignorance of their own nutrient intakes.

Given the importance of diet in health education, and the demand from consumers for specific appropriate advice, we urgently require a system which accurately and rapidly assesses an individual's current dietary intake, and provides practical dietary feedback and advice tailored to the individual.

The present paper describes such a system: the development and validation of a diet questionnaire linked to a computer program for use in the dietary assessment and education of healthy individuals.

METHODS

Development of the questionnaire

The aims of the self-administered food frequency questionnaire were as follows: (1) to establish individual habitual dietary intake over a 3-month period in individuals aged 18–65 years; (2) to estimate intake of the following nutrients: energy, protein, total fat,

Present addresses: * Department of Biological and Biomedical Sciences, University of Ulster, Coleraine BT52 1SA, Northern Ireland; † Department of Nutrition and Dietetics, King's College London, Campden Hill Road, London W8 7AH; ‡ 31 Bridge Avenue, London W6.

polyunsaturated fat:saturated fat (P:S) ratio, salt, sugar, fibre, alcohol, retinol, cholecalciferol, vitamin E, ascorbic acid, thiamin, riboflavin, nicotinic acid, total folate, calcium, zinc, iron and vitamin B_{12}; (3) to classify individuals into three categories of nutrient intake for health education purposes: unacceptably low, nutritionally acceptable, unacceptably high (based on Department of Health and Social Security (1981), National Advisory Committee on Nutrition Education (1983) and Committee on Medical Aspects of Food Policy (1984) recommendations); (4) to provide individualized practical recommendations for dietary modification.

The questionnaire consisted of two sections. The first contained seventy-nine questions on meal patterns, food frequency and types of foods consumed. A total of 196 separate food items were included. The questionnaire contained six frequency categories each of which was assigned a numerical factor: never, 0; once fortnightly or less 0·25; about once weekly, 1; two to three times weekly, 2·5; four to six times weekly, 5; every day or most days, 7. Quantification of food recalled by the subject was expressed in household measures; for example, number of slices of bread or glasses of milk. Household measures were converted to weights on the basis of information published by Crawley (1988). Illustrations were provided to aid memory recall. Questions on general physical activity levels and exposure to sunlight were also included in this section. The second section contained forty-six questions on types of food eaten, cooking methods and on high-salt foods and alcohol. It concluded with some general questions about use of 'health' foods, slimming products and vitamin and mineral supplements. Finally the height, sex, age and weight of the subject were recorded. Each response on the questionnaire had an individual box and boxes were numbered sequentially.

Questionnaire completion took, on average, 40 min. The individual completed the questionnaire by ticking appropriate boxes, i.e. no writing was required. However, opportunities were provided to record details about food consumption or exercise habits which subjects felt unable to express within the limits of the closed questions. A questionnaire with missing data on the frequency of consumption of main meals, for example, breakfast, lunch or evening meal or with more than three unanswered questions in any one section was unacceptable for analysis.

Development of the computer program

The newly developed DIET program which analysed the questionnaire was written in the high-level language 'C', because of its ability to manipulate large amounts of data rapidly, and was compiled on an IBM-compatible desktop computer. The nutrient data-base was McCance and Widdowson's *The Composition of Foods* (Paul & Southgate, 1978) and Wiles *et al.* (1980) plus unpublished results (M. G. O'Donnell, M. Nelson, P. H. Wise and D. M. Walker). The nutrient content of the diet was calculated by multiplying the frequency of consumption by the quantity of food (based on the number of standard portion sizes consumed) multiplied by the nutrient content per 100 g, and expressed as intake/subject per d. The program prompted the user to enter into the computer (IBM-compatible) the identifying number of each box ticked, thereby eliminating the need for coding questionnaires before analysis. The entered numbers were shown on a screen, facilitating data checking. The program provided a utility to output raw nutritional data to a file on disk. A single questionnaire could be entered and analysed within 12–15 min. Two types of feedback were generated. (1) A personalized print-out with recommendations for desired dietary changes accompanied by a nutrition information pack containing leaflets with practical suggestions to achieve the relevant dietary modifications. (2) A print-out for the health professional, doctor or dietitian. This part of the program is being further developed. The print-out will ultimately contain a list of nutrients and the range of intake

in which the individual falls. Further qualitative information on meal patterns and other dietary habits will also be included.

Validation of the questionnaire

The questionnaire was validated against two methods of assessment: (1) the weighed inventory technique (Marr, 1965); (2) nutritional biochemical variables in blood and urine (Bingham, 1987).

Study detail

Subjects were recruited from a General Practitioner's (D.W.) surgery in Hammersmith, London. All were healthy Caucasian individuals aged between 18 and 65 years who were not suffering from any chronic disorder nor following any special or therapeutic diet. Seventy-five subjects were approached, fifty-eight agreed to take part and a total of fifty-two subjects completed the study, twenty-eight females and twenty-four males.

Recruitment and first completion of questionnaire (*QUEST 1*). An introductory letter over the general practitioner's signature was posted to subjects inviting them to take part in the study. They were then visited at home for recruitment, and if they agreed to participate, were given a questionnaire to complete at home at the beginning of the study.

16-d weighed intake. Following recruitment, an appointment was made for each subject to visit the surgery where they were instructed by a nutritionist (M. O'D.) how to weigh and record food consumption. Each subject was supplied with a notebook and weighing scales, and given the dates of the first 4-d period in which to weigh their food. Records were checked at the end of 4 d.

Subjects recorded diet for a further three 4-d periods, at 1-month intervals, using the weighed inventory method. The 16 d of diet thus recorded were sufficient to classify correctly at least 80% of subjects by thirds of the distribution and grossly misclassify less than 5%, for intakes of all nutrients for which feedback was given (Nelson *et al.* 1989).

Each 16-d period consisted of twelve weekdays and four weekend days, and was completed over four 4-d periods both for convenience and to aid compliance.

Urine and blood sampling. Four fasting blood samples were collected concurrently with the four weighed food records. The blood samples were analysed for the following variables: urea and electrolytes, urate, glucose, Ca, Zn, magnesium, phosphate, alkaline phosphatase (*EC* 3.1.3.1), protein, albumin, bilirubin, cholesterol, triacylglycerols, high-density-lipoprotein-cholesterol, aspartate aminotransferase (*EC* 2.6.1.1), hydroxybutyrate dehydrogenase (*EC* 1.1.1.30, 1.1.1.61) (HBD), 25-hydroxy cholecalciferol, γ-glutamyl transferase (*EC* 2.3.2.2) and creatine kinase (*EC* 2.7.3.2). Routine haematology was also performed which included haemoglobin, packed cell volume, mean corpuscular volume (MCV), leucocyte count, erythrocyte count, serum vitamin B_{12}, serum and erythrocyte folate and serum ferritin.

Four 24-h urine samples were collected concurrently with the weighed records. The urine collections were analysed for sodium, potassium, and nitrogen. The collections were verified for completeness using *p*-aminobenzoic acid (PABA) as the marker (Bingham & Cummings, 1983). The subjects were given instructions on how to collect a complete 24-h urine sample and supplied with the necessary equipment and the PABA tablets on the recruitment visit. They were asked to bring the urine collection and the completed questionnaire when they attended the surgery for their first blood test.

Repeat completion of the questionnaire (*QUEST 2*). The questionnaire was completed a second time between 2 and 4 weeks after the completion of the last weighed record (QUEST 2). This questionnaire was returned by post to the investigator.

Analysis. Blood and urine samples were analysed by the Departments of Chemical

Table 1. *Mean daily nutrient intake derived from 16 d of weighed food records and questionnaire 1 (QUEST 1)*†

(Mean values with their standard errors)

Nutrient	Weighed record		Quest 1		Statistical significance of difference
	Mean	SE	Mean	SE	
			Males (*n* 24)		
Energy (MJ (kcal))	10·6 (2581·1)	0·41 (99·4)	6·3 (1535·2)	0·38 (91·7)	***
Protein (g)	92·0	3·9	62·8	3·6	***
Carbohydrate (g)	273·1	13·7	172·2	12·8	***
Total fat (g)	106·3	5·9	60·4	4·3	***
Total sugars (g)	111·6	7·4	60·9	6·4	***
Alcohol (g)	32·7	7·1	10·3	0·3	***
Dietary fibre (g)	22·0	1·6	17·6	1·3	**
Calcium (mg)	1051·2	63·2	638·5	42·9	***
Iron (mg)	17·3	1·1	12·1	1·1	***
Zinc (mg)	11·7	0·6	7·9	0·5	***
Sodium (mg)	3640·1	164·7	1597·2	108·8	***
Retinol (μg)	703·9	90·9	623·9	65·7	
Carotene (μg)	2561·2	282·8	2642·5	321·7	
Thiamin (mg)	1·4	0·1	1·1	0·1	***
Riboflavin (mg)	2·1	0·28	1·5	0·1	***
Nicotinic acid (mg)	23·9	0·9	16·6	1·4	***
Vitamin B_{12} (μg)	6·4	0·7	5·5	0·5	
Total folate (μg)	226·2	11·1	155·7	9·6	***
Ascorbic acid (mg)	71·8	7·4	56·3	5·8	**
Vitamin D (μg)	3·1	0·3	3·5	0·6	
Vitamin E (mg)	6·1	0·3	4·3	0·4	***
P:S ratio	0·36	0·13	0·54	0·2	***
Percentage energy from total fat	36·7	5·3	35·3	5·7	
Percentage energy from alcohol	8·9	9·3	6·1	5·9	***

		Females (*n* 28)			
Energy (MJ (kcal))	7·5 (1831·1)	0·33 (80·8)	5·9 (1428·1)	0·43 (106·5)	**
Protein (g)	63·9	2·7	56·5	4·1	*
Carbohydrate (g)	193·1	9·9	161·5	13·9	*
Total fat (g)	81·4	4·4	57·1	5·3	***
Total sugars (g)	87·1	6·7	65·7	7·3	*
Alcohol (g)	16·8	2·9	10·2	0·2	***
Dietary fibre (g)	15·5	1·2	15·5	1·3	
Calcium (mg)	798·6	56·9	607·3	47·6	***
Iron (mg)	12·3	1·0	10·5	0·8	*
Zinc (mg)	8·0	0·4	7·4	0·6	
Sodium (mg)	2231·8	117·5	1622·4	81·1	***
Retinol (μg)	933·6	147·0	532·3	60·2	**
Carotene (μg)	2445·1	194·3	2965·1	337·1	
Thiamin (mg)	0·9	0·1	0·9	0·1	
Riboflavin (mg)	1·5	0·1	1·3	0·1	***
Nicotinic acid (mg)	16·2	0·7	13·6	1·2	*
Vitamin B_{12} (μg)	5·2	0·7	4·2	0·5	
Total folate (μg)	150·2	8·3	149·4	9·8	
Ascorbic acid (mg)	57·8	5·8	76·6	9·0	
Vitamin D (μg)	2·6	0·3	2·9	0·5	
Vitamin E (mg)	5·0	0·5	4·2	0·4	
P:S ratio	0·33	0·16	0·41	0·2	*
Percentage energy from total fat	39·3	4·4	35·1	7·3	**
Percentage energy from alcohol	6·7	6·3	6·6	8·2	***

P:S, polyunsaturated fat: saturated fat.
* $P < 0{\cdot}05$, ** $P < 0{\cdot}01$, *** $P < 0{\cdot}001$.
† For details, see p. 4.

Pathology and Haematology at Charing Cross Hospital, London. Questionnaire data were analysed by the DIET program as described previously.

Weighed records were coded and sent for analysis of nutrient intake to the Dunn Nutrition Centre, Cambridge. Statistical analysis was carried out using SPSS (SPSS.SPSSX, 1988).

RESULTS

Table 1 shows mean nutrient intakes derived from 16 d of weighed food records and from QUEST 1 for males and for females. Energy intakes, from the weighed intake values, were on average 87% of the recommended daily allowance (RDA) (Department of Health and Social Security, 1981) in men and 83% of the RDA in women. Intakes of vitamins and minerals were on average equal to or greater than the RDA. Table 1 shows that in males the questionnaire, when compared with the weighed records, underestimated the intake of all nutrients with the exception of carotene and vitamin D. The differences between the means were all highly statistically significant (paired *t* test) except for retinol, carotene and vitamin D. Table 1 also shows that the questionnaire underestimated nutrient intake in females, but to a lesser extent than in males. The values for fibre (15 g) and thiamin (0·9 mg) were equal. The questionnaire estimates of mean intake of ascorbic acid, vitamin D and carotene were higher than those established by weighed record. The women showed fewer statistically significant differences between questionnaire and weighed record means compared with the men; indeed, for fibre and most vitamins there were no differences.

Table 2 presents Pearson's correlation coefficients for QUEST 1 *v.* the means of 16 d weighed records by sex and for sexes combined. The highest values obtained for sexes combined were for alcohol, fibre, Fe and riboflavin (*r* 0·74, 0·67, 0·66 and 0·66 respectively). Lowest correlations were obtained for energy and the macronutrients, protein, fat and total sugar, with *r* values ranging from 0·3 to 0·4. The most striking feature is the clear sex difference for many nutrients. Males achieved much higher correlation coefficients than females for total sugar (0·60 *v.* 0·19) and for ascorbic acid (0·75 *v.* 0·23). Conversely, females achieved substantially higher values than males for protein (0·56 *v.* 0·33), Ca (0·67 *v.* 0·38), Zn (0·69 *v.* 0·33) and total folate (0·54 *v.* 0·29). Correlation coefficients for Na were not significantly different from zero.

Table 3 shows correlations for QUEST 1 *v.* QUEST 2. The highest correlations for sexes combined were obtained for alcohol (0·88), ascorbic acid (0·81), riboflavin (0·76) and sugar (0·78). Values for these nutrients were similar in both males and females. Females achieved higher values than males for the majority of nutrients, notably for energy (0·78 *v.* 0·47), fat (0·74 *v.* 0·51) and fibre (0·82 *v.* 0·59). Ascorbic acid and total sugar showed good repeatability between questionnaires despite weak correlations with the weighed records.

Table 4 shows correlations between biochemical variables and dietary intake by weighed record and questionnaire. Correlations were examined where some degree of association was expected. Significant associations were demonstrated between weighed records and selected biochemistry for sexes combined. These associations were much weaker or non-existent for the questionnaire for sexes combined. For sexes separately, the associations between biochemistry and weighed records were all weakened or not significant. The significant values for the questionnaire *v.* biochemistry for sexes combined remained significant for sexes separately: γ-glutamyltransferase and alcohol, urinary N and protein in men, serum and dietary Zn in women. Curiously, the correlation of urinary N with dietary protein by sexes separately was stronger for the questionnaire than the weighed record.

Slopes are presented to illustrate the questionnaire's ability to predict nutrient intake. They also illustrate systematic bias in the questionnaire. Fig. 1 shows the slope for dietary Fe intake estimated by 16 d of weighed record *v.* QUEST 1. The correlation was 0·66 (sexes combined), the slope was 0·8, and it can be seen that the questionnaire is a reasonably good

Table 2. *Correlation coefficients (Pearson's) for questionnaire 1 (QUEST 1) v. means of nutrient intake derived from 16 d weighed food records*†

Nutrient	Males (*n* 24)			Females (*n* 28)			Sexes combined (*n* 52)
	r	95% CI		*r*	95% CI		
Energy	0·42*	0·70	0·01	0·34*	0·63	−0·04	0·35**
Protein	0·33	0·65	−0·08	0·56**	0·77	0·24	0·43**
Carbohydrate	0·52**	0·76	0·15	0·26	0·58	−0·13	0·35**
Total fat	0·44*	0·72	0·04	0·40*	0·68	0·04	0·39**
Total sugars	0·60***	0·81	0·26	0·19	0·53	−0·20	0·31*
Alcohol	0·79***	0·90	0·57	0·82***	0·91	0·64	0·74***
Dietary fibre	0·71***	0·87	0·43	0·66***	0·83	0·38	0·67***
Calcium	0·38*	0·68	−0·03	0·67***	0·83	0·40	0·53***
Iron	0·62**	0·82	0·29	0·72***	0·86	0·47	0·66***
Zinc	0·33	0·65	−0·08	0·69***	0·85	0·43	0·43**
Sodium	−0·02	0·39	−0·41	0·39*	0·67	0·02	0·09
Retinol	0·13	0·51	−0·29	0·46**	0·71	0·10	0·31*
Carotene	0·60**	0·81	0·26	0·05	0·42	−0·33	0·31*
Thiamin	0·66***	0·84	0·35	0·47**	0·72	0·12	0·52***
Riboflavin	0·59**	0·80	0·24	0·78***	0·89	0·57	0·66***
Nicotinic acid	0·55**	0·78	0·19	0·51**	0·74	0·17	0·54***
Vitamin B_{12}	0·44*	0·72	0·04	0·49**	0·73	0·14	0·49***
Total folate	0·29	0·62	−0·13	0·54**	0·76	0·21	0·36**
Ascorbic acid	0·75***	0·89	0·50	0·23	0·56	−0·16	0·33**
Vitamin D	0·14	0·51	−0·28	0·39*	0·67	0·02	0·27*
Vitamin E	0·54**	0·77	0·17	0·33*	0·63	−0·05	0·39**

CI, confidence interval.
* $P < 0·05$, ** $P < 0·01$, *** $P < 0·001$.
† For details, see pp. 4, 5 and 8.

predictor of Fe intake despite significant differences between questionnaire and weighed record mean values for intake. Fig. 2 shows the slope for dietary ascorbic acid intake by 16 d of weighed record *v.* QUEST 1. The correlation (sexes combined) was 0·33 and the slope was 0·27. There is a great deal of scatter and the questionnaire's ability to predict ascorbic acid intake is poor.

The analysis of PABA in the 24-h urine collections produced a very wide range of percentage recoveries. Percentage recovery ranged from 3 to 113 in men and from 5 to 131 in women. Of the subjects, 65% had recoveries below 85%, the accepted value for a complete collection (Bingham & Cummings, 1983). It is very unlikely that this number of collections was incomplete. The analytical technique may have underestimated PABA, losses may have occurred during storage or subjects may have failed to take the PABA tablets. The very high values may have been due to contamination of urine with paracetamol preparations or B-vitamins. The correlations reported, therefore, include all urines, and this may have weakened the results, as it is probable that some of the urine collections were incomplete.

Examination of correlation coefficients is one aspect of questionnaire validation. An alternative approach is to consider the questionnaire's ability to classify individuals correctly according to levels of intake, as this is the aim in diet health education when seeking to provide appropriate dietary advice. The ability of the questionnaire to classify individuals correctly into fifths of the distribution was examined for all nutrients for which feedback was provided.

Table 5 shows classification of subjects by fifths of nutrient intake for weighed records *v.* QUEST 1. Between 20 and 50% of individuals were classified in the same fifth by both

Table 3. *Correlation coefficients (Pearson's) for questionnaire 1 (QUEST 1)‡ v. questionnaire 2 (QUEST 2)‡*

Nutrient	Males (*n* 22)			Females (*n* 27)			Sexes combined (*n* 49)
	r	95% CI		*r*	95% CI		
Energy	0·47*	0·73	0·08	0·78	0·89	0·57	0·66
Protein	0·48*	0·74	0·09	0·61	0·80	0·31	0·58
Carbohydrate	0·62**	0·82	0·29	0·81	0·91	0·63	0·74
Total fat	0·51**	0·76	0·13	0·74	0·87	0·51	0·64
Total sugars	0·74	0·88	0·48	0·81	0·91	0·63	0·78
Alcohol	0·93	0·97	0·84	0·90	0·95	0·75	0·88
Dietary fibre	0·59**	0·80	0·24	0·82	0·91	0·64	0·71
Calcium	0·46*	0·73	0·07	0·69	0·85	0·43	0·60
Iron	0·50**	0·75	0·12	0·77	0·89	0·56	0·64
Zinc	0·40*	0·69	−0·00	0·63	0·81	0·34	0·57
Sodium	0·64**	0·83	0·32	0·77	0·88	0·54	0·70
Retinol	0·64**	0·83	0·32	0·84	0·92	0·68	0·75
Carotene	0·49*	0·75	0·11	0·54**	0·76	0·21	0·51
Thiamin	0·60**	0·81	0·26	0·69	0·85	0·43	0·65
Riboflavin	0·74	0·88	0·48	0·78	0·89	0·57	0·76
Nicotinic acid	0·54**	0·77	0·17	0·73	0·87	0·49	0·66
Vitamin B_{12}	0·59**	0·80	0·24	0·83	0·92	0·66	0·71
Total folate	0·54**	0·77	0·17	0·78	0·89	0·57	0·67
Ascorbic acid	0·85	0·93	0·68	0·80	0·90	0·61	0·81
Vitamin D	−0·09†	0·33	−0·48	0·70	0·85	0·44	0·26*
Vitamin E	0·50**	0·75	0·12	0·67	0·83	0·40	0·60

CI, confidence interval.
All values were statistically significant ($P < 0·001$) except as indicated.
* $P < 0·05$, ** $P < 0·01$.
† $P > 0·05$.
‡ For details, see pp. 4, 5 and 8.

techniques (correctly classified). The percentage correctly classified was similar for protein, fat and carbohydrate (25–27), but slightly higher for dietary fibre (38). The classification of alcohol was particularly high (53·8%) while that of carotene was poor (19%).

Less than 4% of subjects were grossly misclassified for all nutrients, with the exception of carbohydrate, Na and retinol. At least 65% of subjects were classified by questionnaire to within one quintile of the classification by weighed record for all nutrients apart from total fat, total sugar, sodium and carotene (63·5, 61·5, 53·8 and 61·4% respectively).

DISCUSSION

The newly-developed DIET program is one of the first validated systems designed to provide rapid nutritional assessment and feedback for the general public. The questionnaire can be completed without any professional help or supervision. The analysis program is 'user friendly' and can be operated by non-nutritionist personnel. A similar approach has been developed in Australia, for use with diet diaries or a food-frequency questionnaire, but data entry is more complex, and the authors do not give the time taken to process data. There is limited information on the validation of the system (Baghurst & Record, 1984). Wise *et al.* (1986) in Aberdeen have expanded an existing program to provide interactive nutrition education with dietary input from food records and diaries.

The present study was limited to Caucasian adults. Other ethnic groups and adolescents were excluded due to different food habits which would require different emphases in the questionnaire.

The results of the 16 d weighed food intake (Table 1) are very similar to other recent

Table 4. *Correlation coefficients (Pearson's) between blood and urine variables and dietary intake assessed by 16 d of weighed records and questionnaire 1 (QUEST 1)*†

Serum biochemistry	Diet	Sexes combined: Weighed record *r*	Sexes combined: QUEST 1 *r*
Urea	Protein	0·41**	0·11
Creatinine	Protein	0·47***	0·08
Zinc	Zn	0·30*	0·27*
γ-GT	Alcohol	0·39**	0·26*
Urine (/24 h)			
Sodium	Na	0·40**	0·18
Potassium	K	0·50***	0·08
Nitrogen	Protein	0·56***	0·34**

		Males: Weighed record			Males: QUEST 1		
		r	95% CI		*r*	95% CI	
Urea	Protein	0·21	0·57	−0·21	−0·004	0·40	−0·41
Creatinine	Protein	0·12	0·49	−0·31	−0·07	0·34	−0·46
Zn	Zn	−0·08	0·34	−0·46	0·16	0·53	−0·26
γ-GT	Alcohol	0·36*	0·66	−0·06	0·35*	0·65	−0·07
Urine (/24 h)							
Na	Na	0·008	0·41	−0·40	0·25	0·59	−0·17
K	K	0·19	0·55	−0·23	0·08	0·47	−0·33
N	Protein	0·18	0·54	−0·24	0·36*	0·67	−0·05

		Females: Weighed record			Females: QUEST 1		
		r	95% CI		*r*	95% CI	
Urea	Protein	0·07	0·43	−0·31	0·09	0·44	−0·30
Creatinine	Protein	0·39*	0·66	0·01	0·12	0·47	−0·26
Zn	Zn	−0·02	0·36	−0·39	0·32*	0·62	−0·06
γ-GT	Alcohol	0·31	0·61	−0·07	0·19	0·52	−0·21
Urine (/24 h)							
Na	Na	0·09	0·45	−0·29	0·24	0·56	−0·16
K	K	0·44*	0·70	−0·08	0·04	0·41	−0·34
N	Protein	0·19	0·53	−0·20	0·28	0·59	−0·10

CI, confidence interval; γ-GT, γ-glutamyltransferase (*EC* 2.3.2.2).
Values were statistically non-significant except as indicated.
* $P < 0·05$, ** $P < 0·01$, *** $P < 0·001$.
† For details, see pp. 4, 5 and 8.

studies of weighed diet in British adults (Bingham, 1981; Yarnell *et al.* 1983; Nelson *et al.* 1989) suggesting that the sample was representative of the adult population. Moreover, the 16 d of record were spread over 3 months, making the individual intake more likely to be representative of 'usual diet', providing an appropriate standard against which to validate the questionnaire. It is also clear from Table 1 that the questionnaire underestimated nutrient intake in males, and to a lesser extent in females. This is one of the errors associated with questionnaires (Bingham, 1987) which can be due to a variety of reasons. In our case, under-reporting by subjects of food frequency and portion size, rather than missing food items, is more likely to be the cause of subject underestimation on the questionnaire and constitutes systematic error within the questionnaire.

There are two aspects to the validation of the questionnaire. The first is a comparison of

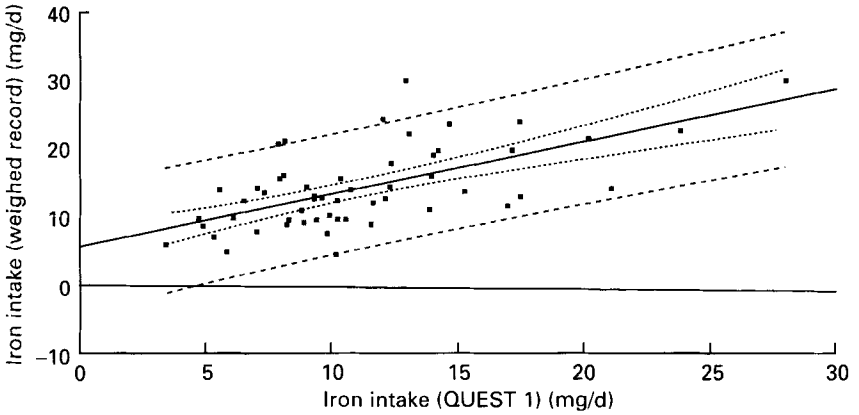

Fig. 1. Iron intakes derived from 16 d weighed record *v.* questionnaire 1 (QUEST 1). Y = Q1 × 0·8 + 5·7 (——); 95% CI Expected Y (.....); 95% CI Predicted Y (– – –). For details of procedures, see pp. 4, 5 and 8.

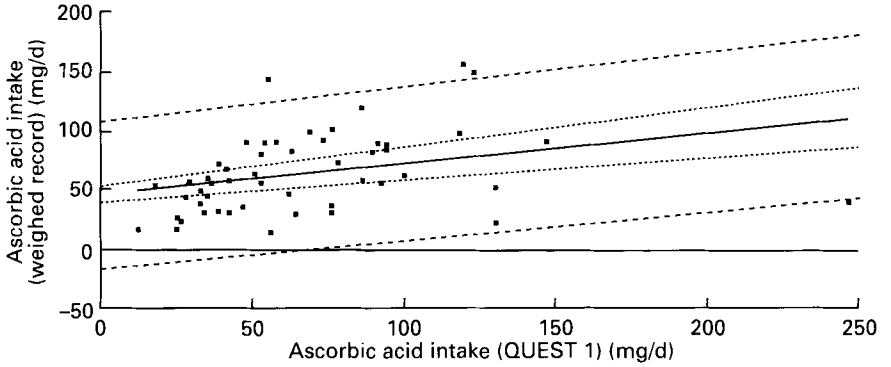

Fig. 2. Ascorbic acid intakes derived from 16 d weighed record *v.* questionnaire 1 (QUEST 1). Y = Q1 × 0·27 + 46 (——); 95% CI Expected Y (.....); 95% CI Predicted Y (– – –). For details of procedures, see pp. 4, 5 and 8.

aggregate measures (means) in groups. It is clear that the present questionnaire substantially underestimates weighed intakes for most nutrients. If, however, the aim is to rank individuals rather than estimate 'true' intake, then the second aspect of validation, correlation and regression, may indicate that the questionnaire is satisfactory for ranking purposes. If the correlation of weighed intake and questionnaire is good, then regression and calculation of predicted intake may in part overcome the problems associated with underestimation of intake by the questionnaire, when the purpose is to classify subjects according to the need for nutritional advice.

The highest correlation values were obtained for alcohol (0·74), fibre (0·67), Fe (0·66) and riboflavin (0·66). For alcohol, fibre and riboflavin these findings agree with previous observations that questionnaires are best at estimating nutrients which occur in concentrated amounts in relatively few foods which are readily quantified (Byers *et al.* 1983). For Fe, the good agreement was probably due to careful questioning on meat and offal consumption. Illustrations were also provided to aid recall of portion sizes.

Yarnell *et al.* (1983), in the validation of a self-administered questionnaire designed for use in epidemiological studies of heart disease, obtained a similar value for alcohol (*r* 0·75) but a lower value for fibre (*r* 0·37). The latter may have been due to fewer questions on fibre intake.

In the present study, the correlations obtained for energy and the macronutrients were in the range 0·26–0·56. The correlation coefficient for energy for sexes combined was 0·35,

Table 5. *Classification of subjects by fifths of nutrient intake, weighed record* v. *questionnaire 1 (QUEST 1)**

	Percentage correctly classified	Percentage classified to within one quintile	Percentage grossly misclassified
Energy	32·6	69·2	3·8
Protein	24·9	71·0	1·9
Carbohydrate	26·9	65·2	5·7
Total fat	26·9	63·5	0
Total sugars	23·0	61·5	3·8
Alcohol	53·8	92·3	0
Dietary fibre	38·4	76·9	0
Calcium	36·5	73·0	1·9
Iron	36·4	72·9	0
Zinc	32·6	65·1	0
Sodium	25·0	53·8	7·7
Retinol	28·7	65·1	5·7
Carotene	19·1	61·4	1·9
Thiamin	36·3	72·7	2·3
Riboflavin	34·5	72·9	0
Nicotinic acid	28·7	70·9	0
Vitamin B_{12}	28·7	74·7	0
Total folate	44·2	69·0	1·9
Ascorbic acid	32·7	71·1	1·9
Vitamin D	32·7	67·2	1·9
Vitamin E	36·6	67·1	0

* For details, see p. 4.

similar to the value of 0·30 obtained by Yarnell *et al.* (1983). Low values for protein, fat and total sugar were again similar to those reported by other workers (Yarnell *et al.* 1983; Willett *et al.* 1985). These low values are probably due to the difficulty when using questionnaires, in obtaining accurate estimates of the frequency and amount of consumption of a wide range of foods, and also to the general problems of under-reporting and questionnaire bias. That the error lies in this type of questionnaire rather than the standard for assessing 'usual intake' is indicated by the similarity of the correlations in the different studies regardless of the length of time for which the diet records were kept: 7 d (Yarnell *et al.* 1983), 16 d (present study), or 28 d (Willett *et al.* 1985).

Striking sex differences were shown in the results. Females tended to achieve higher correlations for reproducibility of questionnaire responses (Table 3). This may be due to women spending more time than men in the purchasing and cooking of meals, and they may, therefore, be more aware of food portion sizes and frequency of consumption (Craig & Truswell, 1988). However, for the comparison of questionnaire results with weighed records and biochemistry (Tables 2 and 4) the differences between the sexes was not consistent. The extent of the sex differences indicates the need to validate questionnaires separately for males and females. The extent to which the differences between strength of correlation can be obtained is illustrated by comparing the correlation coefficients for sexes combined with those for sexes separately (Tables 2, 3, 4). Few workers have presented results for males and females separately, hence the difficulty in comparing these findings with previous findings. Epstein *et al.* (1970) compared a short diet questionnaire with a diet history and presented findings for males and females separately. In men and women aged 40 years and over, for energy, total carbohydrate and protein men had higher correlation

coefficients than women (0·71, 0·71 and 0·73 compared with 0·66, 0·53 and 0·53 respectively), whereas for sugar, total fat and fat components women had correlation coefficients greater than men (0·5–0·71 compared with 0·14–0·51). In the present study, men had higher correlations for energy, carbohydrate, fat and sugar, but not for protein (Table 2).

The higher correlations for QUEST 1 *v.* QUEST 2 in women indicate that women are more reliable at reporting their food intake than men, even when validity is poor. For example, vitamin C showed poor validity (*r* 0·23, QUEST 1 *v.* weighed record) but good repeatability (*r* 0·80). It should be borne in mind that this is not a true repeatability study, as the subjects weighed their food between the completion of the questionnaires, and the act of recording may have influenced responses on QUEST 2. Indeed, the correlations for QUEST 2 *v.* weighed record were generally greater than for QUEST 1 *v.* weighed record.

Appreciable sex differences in the validity of reported intake based on questionnaires may, therefore, obscure true sex differences in the relationships between diet and disease.

The 95% confidence intervals shown in Tables 2, 3 and 4 show, however, that the difference in correlations for males and females is not statistically significant (with the exception of vitamin D, Table 3) and, therefore, conclusions from the study are based on results for sexes combined.

Questionnaires have been developed and used extensively in the past to investigate possible links between diet and disease, often with disappointing results. The poor-to-moderate associations between the questionnaire and biochemical variables in Table 4 indicate that questionnaires may not be powerful enough to detect possible relationships between diet and disease mediated by blood biochemistry. Urinary N appears to be the best biochemical measure for dietary protein, based on the correlation of N excretion and protein intake. It is curious, however, that the correlations for men and women fall and become non-significant when analysed separately by sex. The reason for this is not clear. It may relate in part to differences in the completeness of urine collections between subjects. The difficulties in obtaining complete collections from individuals are well known and we were disappointed in our use of PABA to verify the completeness of collection. The need for a more precise tool for the estimation of diet in epidemiological studies remains pressing. There is potential use, however, for the present questionnaire in epidemiological studies of alcohol, fibre and Fe intake.

The criteria for the use of the questionnaire in health education are less strict than those imposed for its use in epidemiological studies as, in health education, one is concerned primarily with the correct classification of individuals into broad groups of intake. The ability of the questionnaire to achieve such classification and to provide appropriate feedback will be reported in a future paper. A study is also currently underway to evaluate the effects of personalized dietary feedback from the questionnaire on diet modification in a group of healthy adults, and to test the use of the questionnaire for purposes of dietary health education in a general practice setting.

The authors wish to acknowledge the help of Miss L. Sankaran in coding the diet records and the Department of Chemical Pathology and Haematology, Charing Cross Hospital, London for analysing the blood and urine samples.

REFERENCES

Baghurst, K. I. & Record, S. J. (1984). A computerised dietary analysis system for use with diet diaries or food frequency questionnaires. *Community Health Studies* **8**, 11–18.

Bingham, S. (1981). The diet of individuals: a study of a randomly chosen cross-section of British adults in a Cambridgeshire village. *British Journal of Nutrition* **45**, 23–35.

Bingham, S. A. (1987). The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutrition Abstracts and Reviews* **57**, 722–726.
Bingham, S. & Cummings, J. H. (1983). The use of 4-aminobenzoic acid as a marker to validate the completeness of 24 hour urine collections in man. *Clinical Science* **64**, 629–635.
Byers, T. E., Rosenthal, R. I., Marshall, J. R., Rzepka, T. F., Cummings, K. M. & Graham, S. (1983). Dietary history from the distant past: a methodological study. *Nutrition and Cancer* **5**, 69–77.
Committee on Medical Aspects of Food Policy. (1984). Diet and cardiovascular disease. *Report on Health and Social Subjects*. no. 28. London: H.M. Stationery Office.
Craig, P. I. & Truswell, A. S. (1988). Dynamics of food habits of newly married couples: food-related activities and attitudes towards food. *Journal of Human Nutrition and Dietetics* **1**, 409–419.
Crawley, H. (1988). *Food Portion Sizes*. London: H.M. Stationery Office.
Department of Health and Social Security (1981). Recommended intakes of nutrients for the United Kingdom. *Report on Public Health and Medical Subjects* no 120. London: H.M. Stationery Office.
Epstein, L. M., Reshef, A., Abramson, J. H. & Bialik, O. (1970). Validity of a short dietary questionnaire. *Israel Journal of Medical Science* **6**, 589–597.
Marr, J. W. (1965). Individual weighed dietary surveys. *Nutrition* **19**, 18–24.
Ministry of Agriculture, Fisheries and Food (1984). Household food consumption and expenditure: 1983. *Annual Report of the National Food Survey Committee*. London: H.M. Stationery Office.
National Advisory Committee on Nutrition Education. (1983). *Proposals for Nutritional Guidelines for Health Education in Britain*. London: Health Education Council.
Nelson, M., Black, A. E. & Cole, T. J. (1989). Between- and within-subject variation in nutrient intake from infancy to old age: estimating the number of days required to rank dietary intakes with required precision. *American Journal of Clinical Nutrition* **50**, 155–167.
Paul, A. A. & Southgate, D. A. T. (1978). *McCance and Widdowson's The Composition of Foods*. London: H.M. Stationery Office.
SPSS.SPSSX (1988). *Statistical Package for the Social Sciences*, 3rd ed. Michigan: SPSS.
Wiles, S. J., Nettleton, P. A. & Black, A. E. (1980). The nutrient composition of some cooked dishes eaten in Britain: a supplementary food composition table. *Journal of Human Nutrition* **34**, 189–223.
Willett, W. C., Sampson, L., Stampfer, M. J., Rosner, B., Bain, C., Witschi, J., Hennekens, C. H. & Speizer, F. E. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *American Journal of Epidemiology* **122**, 51–65.
Wise, A., Lockie, G. M. & Liddell, J. A. (1986). Effective ways of reducing saturated fat intake. *Health Education Journal* **45**, 210–214.
Yarnell, J. W. G., Fehily, A. M., Millbank, J. E., Sweetnam, P. M. & Walker, C. L. (1983). A short dietary questionnaire for use in an epidemiological survey: comparison with weighed dietary records. *Human Nutrition: Applied Nutrition* **37**A, 103–112.

Printed in Great Britain