An investigation of the validity and reliability of a food intake questionnaire

B. Johnson,* A. Hackett,† M. Roundfield‡ and A. Coufopoulos§

*Department of Community Nutrition and Dietetics, Abercromby Health Centre, Grove St., Liverpool L7 7HG, †Reader in Community Nutrition, School of Education & Community Studies, Liverpool John Moores University, IM Marsh Campus, Barkhill Road, Aigburth, Liverpool L7 6BD, ‡Department of Mathematics and Computing, Liverpool John Moores University, Byrom St., Liverpool, §Department of Health Studies, University of Central Lancashire, Preston PC1 2HE, UK

Correspondence
Brian Johnson,
Department of Community Nutrition and Dietetics,
Abercromby Health Centre,
Grove St.,
Liverpool L7 7HG, UK

Keywords
dietary questionnaire, reliability, validity.

Accepted
October 2001

Abstract

Aims To evaluate the reliability and criterion validity of a food intake questionnaire (FIQ) designed for use in schoolchildren.

Methods Study of reliability: 98 young people aged 13–14 years attending two schools in deprived areas of Liverpool completed the FIQ on three separate occasions over a 3-month period.

Validity study Ninety-six young people (aged 11–13 years) completed the FIQ and 2 weeks later completed a 3-day food diary (with interview).

Results The FIQ gave consistent response on separate occasions over the 3-month reliability study period. Levels of agreement were consistent between survey combinations. Analysis of variance showed no differences in mean score for food groups between surveys. Pearson correlations for mean scores estimated by separate FIQ ranged from 0.42 for fibre food group to 0.76 for negative marker food group; the majority of the correlations were above 0.5. The data suggested the FIQ should be able to detect a change of ±10% in eating habits. The validity study provided modest but significant Pearson correlations between energy intake, fat intake as a percentage of energy intake and sugars intake derived from 3-day diaries, and mean scores for the fatty, sugary and negative marker food group assessed by the FIQ.

Conclusions The results from both studies provide an indication of the FIQ's reliability, and suggest it has criterion validity for fatty and sugary and negative marker foods.

Introduction

Community dietitians, along with other clinical practitioners, are being asked to provide evidence on the effectiveness of their practice (NHS Executive, 1997; BDA, 1999). Measuring the effectiveness of community nutrition initiatives involving changes in eating habits requires a dietary assessment method that is reliable and cost-effective for use with large numbers of subjects. The food intake questionnaire (FIQ) has been developed for this purpose and has been extensively used over a number of years in different settings (Hackett *et al.*, 1989, 1997; Johnson & Hackett, 1997). The relative or criterion validity of a new dietary assessment method is often assessed against an established method of 'indisputable quality' because there is no universally accepted method of dietary assessment that can be used as an absolute 'gold standard' for comparison (Bingham, 1987; Kemm & Booth, 1992). Reliability is concerned with the consistency or precision of the measurement made when repeated under the 'same conditions' assuming that no real change has taken place (Cameron & Van Staveran, 1988; Willett, 1992). Willett (1992), describing the reproducibility (reliability) of food frequency questionnaires, stressed the need to repeat measurements over a number of intervals of time in order to separate time changes in diet from learning effects. In practice, therefore, reliability has two components: genuine variation from one occasion to another and apparent variation (or difference) due to uncertainty (error) inherent in the measure used.

Materials and methods

Reliability study

Two schools were recruited from two areas of Liverpool representative of social and health deprivation (based on the receipt of Income Support). Each child who agreed to take part, completed the self-administered FIQ on three separate occasions (all weekdays), at roughly equal intervals, over a 3-month period. It was assumed that any real change in eating habit over this time would be minimal (Bingham, 1987).

The FIQ used in the study was a self-administered, adapted 24-h recall method details of which have been published elsewhere (Johnson & Hackett, 1997). The questions related to the consumption of particular foods on the previous day, not the previous 24 h. Children were asked to answer the question: 'Did you at any time yesterday eat any amount of…'; the basic stem question was followed by a list of food related items. The diet records were also analysed by aggregating foods with similar characteristics, e.g. the number of sugary foods claimed to have been eaten, allowing mean scores to be calculated. For the purpose of this study the following aggregated groups were used: fatty ($n = 10$), sugary ($n = 13$), fibrous ($n = 10$), low-sugar foods ($n = 3$) and alternative fats ($n = 5$). These groups were then combined to produce two larger groups of marker foods, positive markers ($n = 20$) and negative markers ($n = 24$) describing foods that dietitians would advise people to eat more of or less of (Johnson *et al.*, 1999).

Sample

All children in school year nine (aged 13–14 years) of each school were invited to participate; 250 children were eligible to take part.

The method used to administer the FIQ has been reported (Johnson & Hackett, 1997). Each survey was conducted in the same classroom during the morning school session. Each child completed the survey individually; responses were input directly into a computer and stored on disc, and later downloaded into a mainframe computer for analysis.

Validity study

Sample

One hundred and six young people aged 11–13 years (school years 7 and 8) who were actively involved in a school-based nutrition education programme in a secondary school in Rugely, Staffordshire, agreed to take part in the study. Each young person completed the FIQ and 2 weeks later completed a 3-day food diary (with interview). Subjects who failed to complete both a 3-day diary and the FIQ were not included in the final analysis.

3-day food diary and interview

Details of the diary and interview method have been reported (Nathan *et al.*, 1996). Each subject was provided with a pocket-sized food diary to record all foods eaten. On the fourth day children were interviewed to clarify dietary information and to assess portion sizes using a calibrated food atlas (Nelson *et al.*, 1997), which was not used in the previous surveys. Nutritional intake was calculated using Microdiet version 9.1 (University of Salford, 1995) with the fifth edition of *McCance & Widowson's, Composition of Foods* (Holland *et al.*, 1991) and all available supplements.

Statistical analysis

Reliability study

Data were analysed using Statistical Package for the Social Sciences (SPSS Inc., 1999). For the purpose of this study, reliability was defined as the level of agreement represented by a subject giving an identical response to the same item on two occasions (either YY or NN). Agreement between survey combinations were compared (see Tables 5 and 6) for the responses given by each subject for each survey combination: FIQ1 and FIQ2, FIQ1 and FIQ3, and FIQ2 and FIQ3.

The figures indicate the number of subjects answering either 'Yes' or 'No' to the same question on each occasion and as a proportion of the total sample who completed each item. Also, 95% confidence intervals were calculated and expressed as a percentage range (Bowling, 1997).

The width of the 95% confidence intervals for the proportions were calculated using the following formula (Gardner & Altman, 1986):

$$CI = \pm 1.96\sqrt{[p(1-p)/n]}$$

where p = sample proportion n = sample size.

One-way analysis of variance was conducted to assess differences in mean scores; standard error of the mean and ranges are given for the fatty, sugary, fibre, Positive marker foods and Negative marker foods groups.

Reliability of responses over time

To compare the stability of responses over time (interindividual variation), proportions of subjects reporting eating each food were ranked for all surveys. The foods ranked 1–10 and 46–56 for each of the three surveys were compared, i.e. the foods claimed to be eaten by the most and the least numbers of children (see Tables 3 and 4). In addition, Pearson correlation coefficients were calculated for scores between surveys (Table 8).

Validity study

The FIQ was not designed to estimate nutrient intake; the foods included were to reflect current dietary advice only. Nevertheless, if the foods listed are relevant they might be expected to reflect, to some extent, intake of key nutrients, for example the number of sugary foods might be expected to be positively correlated to the intake of sugars. Pearson correlation coefficients were calculated to compare intake of fat, sugars, energy and fibre assessed by the 3-day diary with mean scores from FIQ food groups: fatty, sugary, fibre, Positive markers and Negative markers.

Results

Reliability survey

The numbers of children who completed surveys 1–3 are 153, 159 and 153, respectively; 98 completed all surveys. Of the 250 children eligible to take part, 98 completed all three survey records and so were used for the analysis. Table 1 details the sex and age distribution of those children completing all three FIQ records.

The frequencies for all FIQ items were ranked in order, revealing a high level of stability in the foods claimed to have been eaten (Table 3). Of the foods reported as being eaten by most young people (ranked 1–10), seven appeared in each ranking for FIQ 1, FIQ 2 and FIQ 3 (Table 2). Of

Table 1 Age and gender distribution of subjects who completed three FIQ records

	Age (years)	Boys	Girls	Total
School A	13	13	19	32
	14	9	12	21
School B	13	7	9	16
	14	16	13	29
Total (*n*)		45	53	98

Table 2 Proportion (%) of subjects eating a particular food 'yesterday'. Ranked 1–10 by survey (foods claimed to have been eaten by most children)

	Survey 1		Survey 2		Survey 3	
Rank	Food	%	Food	%	Food	%
1	Drink milk	82	Fizzy drink	76	Fizzy drink	89
2	Fizzy drink	79	Drink milk	72	Drink milk	76
3	Crisps	74	Crisps	69	White bread	71
4	Boiled sweets	72	Chocolate	65	Crisps	71
5	White bread	69	White bread	65	Boiled sweets	68
6	Chips	65	Chocolate biscuits	62	Regular fizzy drink	67
7	Chocolate	65	Butter/margarine	61	Butter/margarine	65
8	Hot drink	65	Milk on cereal	60	Chocolate biscuits	59
9	Fruit	63	Chips	59	Chocolate	59
10	Chocolate biscuits	59	Sugar in drink	59	Hot drink	58

the 10 foods reported as being eaten by least children (ranked 46–56), eight appeared in all three lists (Table 3).

The proportion (%) agreement was calculated for foods in each food group by calculating the number and proportion of individual subjects who gave an identical response to each FIQ item on each occasion (Tables 4 and 5). For example, Table 4 shows that for chocolate biscuits, 72% of children gave the same response for survey 1 compared to survey 2: the width of the 95% CI for the proportion is ± 0.089 (8.9%), giving a confidence interval of 0.63–0.81 (63–81%).

The proportion of agreement and the width of the 95% confidence limits for all aggregated groups showed the degree of consistency between survey combinations. The widest intervals were 17% from the lowest to the highest estimate of the proportion while the narrowest intervals were 4%. Ninety-five per cent confidence limits showed a consistency in values of ± 7–13% for the sugary food group (Table 4). The majority of CI ranges were between ± 8 and 10%; the degree of stability is also consistent when comparing different survey combinations. The levels of agreement in the fatty food group (Table 4) are comparable to the sugary group; CI ranges expressed as a percentage were within ± 4–13%, the majority of values being ± 7–10%. The Fibre food group, alternative fats and low-sugar groups (Table 5) also revealed similar CI ranges for the proportions of young people answering identically FIQ items for individual surveys. CI ranges were ± 6–17% for foods in the alternative fats group. The low-sugar foods group showed the highest CI ranges; values varied from ± 4 to 17 (Table 5).

Table 3 Proportion (%) of subjects eating a particular food 'yesterday'. Ranked 46–56 by survey (foods claimed to have been eaten by fewest children)

	Survey 1		Survey 2		Survey 3	
Rank	Food	%	Food	%	Food	%
46	Low-fat spread	12	Muesli	14	Baked potato	10
47	Fried fish	11	Artificial sweetener	14	Mashed potato	9
48	Baked potato	11	Mashed potato	14	Roast potato	9
49	Mashed potato	11	Fried fish	13	Artificial sweetener	8
50	Roast potato	9	Brown bread	13	Fried fish	8
51	Tinned fish	9	Roast potato	11	Wholemeal bread	7
52	Boiled potato	7	Wholemeal bread	10	Boiled potato	7
53	Sultana bran	7	Sultana bran	10	Sultana bran	5
54	Artificial sweetener	4	Boiled potato	8	Low-fat burger	6
55	Low-fat sausage	4	Low-fat sausage	7	Low-fat sausage	2
56	Low-fat burger	3	Low-fat burger	4	Tinned fish	2

Table 4 Reliability of FIQ combinations: agreement between surveys for negative marker foods

	FIQ 1 vs. FIQ 2			FIQ 1 vs. FIQ 3			FIQ 2 vs. FIQ 3		
Negative markers	*n*	%	95%CI range ±	*n*	%	95%CI range ±	*n*	%	95%CI range ±
Sausages	97	80	8	97	77	8	97	96	4
Pies & pasties	96	78	8	97	62	10	70	94	6
Burgers	97	81	8	98	81	8	97	88	7
Crisps	97	74	9	98	73	9	97	7	9
Full fat margarine	57	75	9	56	77	11	56	61	13
Chips	88	80	8	97	62	10	97	64	10
Fried vegetables	98	80	8	98	88	6	98	78	8
Butter	94	69	9	97	73	9	84	67	10
Roast potato	98	88	6	98	82	8	92	87	7
Fried fish	98	82	8	98	85	7	98	85	7
Whole milk	74	78	9	72	78	10	91	66	10
Sugar in drink	97	79	8	96	67	9	54	61	13
Puddings	90	68	10	96	75	9	94	73	9
Ice cream	98	69	9	98	55	10	98	69	9
Sugar on food	98	8	8	94	71	9	94	73	9
Boiled sweets	95	77	8	97	74	9	96	67	9
Chocolate	98	78	8	97	74	9	67	68	9
Milkshake	97	74	9	97	65	9	91	75	9
Chocolate biscuits	96	72	9	96	7	9	98	62	10
Plain biscuit	97	64	10	96	67	9	96	71	10
Frosties	98	76	8	98	78	8	98	76	8
Cakes & Tarts	98	59	10	97	58	10	96	59	10
Fizzy drink	97	78	8	98	8	8	86	86	7
Still drink	97	69	9	97	72	9	97	69	9

n = Total number of subjects who answered the question. % = The proportion of the total sample who gave an identical response to the same FIQ question on separate occasions, i.e. YY or NN. CI Range = ± 95% CI expressed as a percentage.

Analysis of variance

One-way analysis of variance revealed no significant differences in responses for the aggregated food groups by survey (Table 6).

The reliability of the FIQ was also assessed by comparing mean scores for each food group by survey using Pearson correlation coefficients (Table 6). All correlations were above 0.5 and the values ranged from $r = 0.41$ for the fibre food group to $r = 0.76$ for the negative marker group.

Validity study

One hundred and seven children agreed to take part. Eleven subjects failed to complete both a 3-day food diary and the FIQ and were eliminated from the analysis (Table 8).

Table 9 shows the Pearson correlation coefficients between energy intake, total fat intake as a percentage of energy intake, total sugars intake and fibre intake derived from the 3-day diary and the mean score for foods in FIQ aggregated groups. There were low but significant correlations for the FIQ fatty food group with energy intake and fat intake as a percentage of energy intake. The sugary food group was significantly correlated to total sugars intake and energy intake. There were also significant correlations between the negative marker group for energy and total fat intake as a percentage of energy intake. The fibre, alternative fats, low-sugar and positive marker groups were not significantly correlated with the estimates of nutrient intake.

Discussion

Food frequency questionnaires have been developed (Willet, 1992) to measure the intake of macro- and micronutrients as an alternative to more elaborate dietary assessment methods that are not practical for routine community. Dietary

Table 5 Reliability of FIQ combinations: agreement between surveys for positive marker foods

	FIQ 1 vs. FIQ 2			FIQ 1 vs. FIQ 3			FIQ 2 vs. FIQ 3		
Negative markers	*n*	%	95%CI range ±	*n*	%	95%CI range ±	*n*	%	95%CI range ±
Low fat spread	56	82	15	48	75	7	91	69	11
Semi-skim milk	75	71	10	72	74	10	62	81	11
Low fat burger	97	96	4	96	93	7	97	92	5
Low fat sausage	98	95	4	98	94	5	98	91	6
PUFA	56	66	12	56	61	12	42	64	15
Diet still drink	30	73	16	35	77	14	29	79	17
Diet fizzy rink	70	79	10	78	68	10	68	75	10
Artificial sweetener	96	89	6	92	91	6	94	85	7
Baked beans	97	85	7	95	85	7	96	79	4
Salad	98	74	8	97	67	9	97	64	10
Peas	97	72	9	96	78	8	95	73	9
Baked potato	95	85	7	95	88	7	96	80	8
Fruit	98	76	8	88	80	8	87	78	9
Brown bread	98	78	8	98	83	7	98	83	7
Sultana bran	97	87	7	97	92	5	98	87	7
Wholemeal bread	97	86	7	97	89	6	98	89	6
Muesli	98	82	8	98	84	7	98	84	7
Branflakes	97	82	8	98	82	8	97	81	8
Mashed potato	97	87	7	98	86	7	97	85	7
Boiled potato	93	90	6	97	89	6	95	85	7

n = Total number of subjects who answered the question. % = The proportion of the total sample who gave an identical response to the same FIQ question on separate occasions, i.e. YY or NN. CI Range = ± 95% CI expressed as a percentage.

Table 6 Analysis of variance for aggregated food group by survey

	FIQ 1		FIQ 2		FIQ 3		
Food group	Mean	SEM	Mean	SEM	Mean	SEM	F prob.
Fatty foods	3.2	0.16	3.2	0.24	3.0	0.2	0.36
Sugary foods	7.5	0.29	8.3	0.43	7.8	0.42	0.43
Fibre foods	2.4	0.16	2.4	0.17	2.9	0.17	0.42
Alternative fats	1.6	0.12	1.5	0.12	1.4	0.12	0.20
Low sugar foods	1.3	0.08	1.3	0.08	1.4	0.07	0.69
Negative markers	12.8	0.78	15.9	1.57	12.2	0.76	0.83
Positive markers	6.1	0.3	5.9	0.4	5.6	0.35	0.39

F probability (between groups D.F. 2).

Table 7 Spearman correlation coefficients (*r*) between agregated Food group by FIQ

Food group	FIQ 1 and FIQ 2	FIQ 1 and FIQ 3	FIQ 2 and FIQ 3
Fatty	0.59	0.55	0.59
Sugary	0.58	0.62	0.69
Fibre	0.45	0.42	0.44
Posmark	0.57	0.55	0.56
Negmark	0.68	0.76	0.71

All correlations significant $P = < 0.005$.

Table 8 Age and sex distribution of subjects

Age (years)	*n*	Boys (*n*)	% of total	Girls (*n*)	% of total
11	29	13	14	16	17
12	48	21	22	28	29
13	15	7	7	8	8
Missing	3	0	0	3	3
Total	96	41	43	55	57

Table 9 Pearson correlation coefficients (*r*) between 3-day diary and FIQ food groups

	FIQ				
3-day diary	Fatty	Sugary	Fibre	Negative markers	Positive markers
Energy (kJ)	0.20*	0.28*	0.03	0.23*	0.01
Fat (%)	0.36*	0.27*	−0.17	0.34*	−0.06
Sugars (g)	0.09	0.23*	0.12	0.12	0.07
Fibre (g)	−0.057	0.12	0.04	−0.03	0.03

*All correlations significant $P = < 0.05$.

recommendations ultimately have to be given in terms of foods offering suggestions for a healthy diet based on changes in food intake (HEA, 1997; MAFF, 1997). This is also consistent with the way health professionals formulate dietary advice into practical food advice for the public. The FIQ was developed to reveal broad changes in food intake over time and has been used to measure the eating habits of schoolchildren serially in city- and nationwide surveys. In this context the FIQ may be a practical and cost-effective method for collecting food intake data and in evaluating the effectiveness of nutrition interventions. It could be expected that the correlations measuring agreement between the FIQ and the 3-day diary at best would be modest. The significant positive correlations found are encouraging and suggest that the FIQ has criterion validity. The Pearson correlation coefficients for energy, fat and sugars intake seen in this study are similar to those found by Arnold *et al.* (1995), who evaluated a questionnaire specifically designed to estimate nutrient intake. The lowest correlations between the 3-day diary and FIQ were found with the fibre food group; this finding was disappointing and could relate to a lack of precision within the FIQ in estimating fibre foods. The FIQ asks only if a food was consumed rather than how much of the food was consumed. For example, 71% of young people reported that they consumed fruit; it could be expected that this proportion would be closer to 100% if the five-a-day target is to be reached.

The levels of agreement between surveys and the significant correlations between the FIQ and the 3-day diary give confidence in the FIQ's ability to monitor changes in food intake over time, and to detect differences between groups. Given the simple nature of the FIQ stem-question the correlation coefficients support the contention that asking young people food-based questions about diet, using a convenient tool, is nevertheless valid.

There were no significant differences in mean scores for separate food groups between surveys, and if it is assumed that intake did not change significantly during the study period, these results show the extent of changes which the FIQ should be able to detect (± 4–17%). Pearson correlations for mean scores estimated by separate FIQ ranged from 0.42 for fibre to 0.71 for the negative marker group. The majority of the *r*-values were above 0.5, the lowest being associated with the fibre food group. Since some change in eating habits almost certainly did occur over the 3-month study period, which would lower correlations, and that correlations with discrete variables are attenuated (Appleton *et al.*, 1986), this indicates the FIQ has reliability.

In addition, foods ranked according to reports by most and least subjects also confirmed the stability of responses. As expected from other surveys (DHSS, 1989; Adamson *et al.*, 1993; Hackett *et al.*, 1997), popular foods included: boiled sweets, chocolate, fizzy drinks and crisps (Table 3). The foods reported least often included: low-fat sausages, high-fibre breakfast cereals, baked potato, and mashed potato (Table 4) – unfortunately these are often desirable foods.

The widest confidence intervals were related to foods with the lowest proportion of subjects who claimed to have consumed the item, as a result of branched questions, e.g. drinking milk: subjects can select semi-skimmed, skimmed or whole milk as appropriate. The 95% confidence limits showed a consistency in values of ± 7–13% for the sugary food group. When comparing FIQ1 vs. FIQ2, and FIQ1 vs. FIQ3 for sugary foods, the confidence limits could be expected to deteriorate as the length of time between each survey increases; however, this was not found, and the CI ranges were similar in each combination; FIQ1 vs. FIQ2 were similar to FIQ1 vs. FIQ3 (CI range ± 8–10%). The levels of agreement in the fatty food group were comparable to the sugary group. Confidence limits were within ± 4–13% with the majority of values being ± 7–10%. This suggests that the FIQ should be able to detect a change of ± 10% or greater which we would suggest is acceptable to most dietitians.

The mean scores and standard deviations for food groups were used to estimate the sample size required for an intervention study, or for effective monitoring of dietetic intervention. This information is extremely useful to prevent the possibility of a false negative (type II error) or a false positive (type I error) result. The formula suggested by Hall (1983) assumes an acceptability level for type I and type II error probabilities of 0.05 (appendix 1). Appendix 2 shows the sample size required to detect a specified difference in the mean scores for the negative marker group.

These studies have provided an indication of the FIQ's reliability and validity. It must be stressed that the FIQ was not designed to assess nutrient intakes. It is also assumed that there was stability in eating habits over the three surveys in order to test reliability. The modest but significant correlations give added confirmation that the choice of foods in the list is appropriate. The results presented suggest that the FIQ has criterion validity for sugary and fatty and negative marker food groups. The results also indicate the sample size required to use the FIQ in relation to its power to detect differences.

Acknowledgments

Both studies were assisted by a grant from the NHS Executive (North-west) Research and Development Directorate's PhD Bursary scheme. We would also like to thank all of the young people and teachers in all of the schools for their co-operation.

References

Adamson, A.J., Rugg-Gunn, A.J., Appleton, D.R., Butler, T.J. & Hackett, A.F. (1993) Sugars consumption by 379, 11–12-year-old English schoolchildren in 1990 compared to results in 1980. *J. Hum. Nutr. Dietet.* **6**, 419–431.

Appleton, D.R., Rugg-Gunn, A.G. & Hackett, A.F. (1986) Interpreting the correlation coefficient when one of the variables is discrete. *J. Dent. Res.* **65**, 1346–1348.

Arnold, J.E., Rohan, T., Howe, G. & Leblanc, M. (1995) Reproducibility and Validity of a Food-Frequency Questionnaire designed for use in girls aged 7–12 years. *Ann. Epidemiol.* **5** (5), 369–377.

Bingham, S.A. (1987) The Dietary assessment of individuals: methods, accuracy, new techniques and recommendations. *Nutr. Abstracts Rev. (Series a)* **57** (10), 705–740.

Bingham, S.A. & Nelson, M. (1991) *Assessment of food consumption and nutrient intake*. In *Design Concepts in Nutritional Epidemiology.* eds Margetts, B.M., Nelson, M. Oxford: Oxford University Press.

Bowling, A. (1997) *Research Methods in Health. Investigating Health and Health Services*. Buckingham: Open University Press.

British Dietetic Association. (1999) *Clinical Effectiveness Bulletin. Issue 5. March, 1999*. Birmingham: BDA.

Cameron, M. & Van Staveran, W., eds. (1988) *Manual on Methodology for Food Consumption Surveys.* Oxford: OUP.

Department of Health and Social Security. (1989). *The Diets of British Schoolchildren*. Committee on Medical Aspects of Food Policy. Report on Health and Social Subjects 36. London: HMSO.

Gardner, M.J. & Altman, D.G. (1986) Confidence intervals rather than P values: estimation rather than hypothesis testing. *Br. Med. J.* **292**, 746–750.

Gregory, J.R., Lowe, S., Bates, C., Prentice, A., Jackson, L.V., Smithers, G., Wenlock, R. & Farron, M. (2000) *The National Diet and Nutritional Survey: Young People Aged 4–18 years*. London: Stationary Office.

Hackett, A.F., Jarvis, S.N., Flanaghan, G.J. (1989) The feasibility of using school-based microcomputers computers to collect information on the health-related behaviour of children. *Health Education J.* **48** (1), 39–42.

Hackett, A.F., Kirby, S. & Howie, M. (1997) A national survey of the diet of children aged 13–14 years living in urban areas of the United Kingdom. *J. Hum. Nutr. Dietet.* **10**, 37–51.

Hall, J.C. (1983) A method for the rapid assessment of sample size in dietary studies. *Am. J. Clin. Nutr.* **37**, 473–477.

Health Education Authority. (1997) *Eight Guidelines for a Healthy Diet.* Oxford: HEA in association with Ministry of Agriculture Fisheries and Food and the Department of Health.

Holland, B., Welch, A.A., Urwin, I.D., Buss, D.H., Paul, A.A. & Southgate, D.A.T. (1991) *McCance and Widdowson's the Composition of Foods*, 5th edn. London: HMSO.

Johnson, B. & Hackett, A.F. (1997) Eating habits of 11–14-year-old schoolchildren living in less affluent areas of Liverpool, UK. *J. Hum. Nutr. Dietet.* **10**, 135–144.

Johnson, B., Hackett, A.F., Bibby, A. & Cross, J. (1999) An investigation of the face validity of a food intake questionnaire. Lessons for dietary advice? *J. Hum. Nutr. Dietet.* **12**, 307–316.

Kemm, J.R. & Booth, D. (1992) *Promotion of Healthier Eating. How to Collect and Use Information for Planning, Monitoring and Evaluation.* London: HMSO.

Margetts, B.M. & Nelson, M. (1991) *Design Concepts in Nutritional Epidemiology*. Oxford: Oxford University Press.

Ministry of Agriculture Fisheries and Food (MAFF). (1997) *Healthy Eating. Foodsense Number 5*. London: MAFF.

Nathan, I., Hackett, A.F. & Kirby, S. (1996) The dietary intake of a group of vegetarian children aged 7–11 years compared with matched omnivores. *Br. J. Nutr.* **75**, 533–544.

Nelson, M., Atkinson, A. & Meyer, J. (1997) *Food Portion Sizes. A Photographic Atlas*. London: MAFF.

NHS Executive. (1997) *The New NHS. Department of Health.* London: HMSO.

SPSS Inc. (1999) *SPSS for Windows. V 9.1.* Chicago: SPSS Incorporated.

Willett, W. (1992) *Nutritional Epidemiology*. Oxford: Oxford University Press.

Appendix 1

Assessment of sample size in dietary studies (Hall, 1983)

$$n = 2 \cdot \left[\frac{(z\alpha - z\beta)\sigma}{\delta}\right]^2$$

where, n = the sample size in each group, $Z\alpha$ = the upper α per cent point of a normal distribution (α is the probability of a false positive error), $Z\beta$ = the lower β per cent point of a normal distribution (β is the probability of a false negative error), σ = the SD of the variable under study (mean score of food group): based on the assumption that it is approximately the same for both groups, and δ = the difference in population means that is thought to be of interest.

The formula reduces to:

$$n = 2 \cdot \left[\frac{3.3 \cdot \sigma}{\delta}\right]^2$$

For example, to estimate the sample size required to detect a change in mean score of two foods in the negative foods group:

$$n = 2 \cdot \left[\frac{3.3 \cdot 4.9}{2}\right]^2$$
$$n = 132$$

where: 4.9 = SD of mean score for negative food group, 2 = the difference in mean scores between surveys.

Appendix 2

Estimate of the sample size required to detect a difference in mean score for Negative food group

Difference between group means*	Sample size for each group (*n*)	Total sample size (2*n*)
1	261	522
1.5	116	232
2	65	132
2.5	42	84
3	29	58
4	16	33
5	10	21
6	7	14

*Number of foods. Number of negative foods in group = 24. Mean intake of negative foods 15.9 (mean of three surveys, Table 6).