University of Southampton

DEVELOPMENT OF DIETARY ASSESSMENT METHODS FOR USE IN THE SOUTH ASIAN COMMUNITY

by

Norimah A. Karim

Doctor of Philosophy

Institute of Human Nutrition

May 1997

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE, HEALTH AND BIOLOGICAL SCIENCES

HUMAN NUTRITION

Doctor of Philosophy

DEVELOPMENT OF DIETARY ASSESSMENT METHODS FOR USE IN THE SOUTH ASIAN COMMUNITY

by Norimah A.Karim

The main objective of this study was to develop and validate methods for assessing dietary intakes in the South Asian community. A 75-food item food frequency questionnaire (FFQ) and a food checklist were developed. Together with a 24 hour recall of intake, the FFQ and food checklist were calibrated against a 4-day weighed record (WR) as the reference method.

The calibration studies were carried out in two parts. The first calibration study was of the FFQ in both men and women while the second calibration study was that of the food checklist and the 24 hour recall in women only. In the FFQ calibration study, a total of 58 subjects (23 men, 35 women) between the ages of 19 and 76 years of Pakistani, Punjabi and Gujerati origin participated. In both calibration studies, demographic information of each subject was gathered. This allowed us to determine any differences between responders and non responders. For the first calibration study, results for men and women combined together suggested reasonable agreement between FFQ and WR estimates of intake, gender specific agreement was, however, poor. The ranking of subjects was poor with gross misclassification in the range of 5% to 14% in men and 15% to 29% in women. This may in part be due to the small sample size and limited statistical power of gender specific comparisons.

In the second calibration study, 44 women between the ages of 17 and 60 years of Pakistani, Punjabi and Gujerati origin participated. There was also poor correlation between nutrient estimates of food checklist vs WR and 24 hour recall vs WR; 13% to 16% and 14% to 27% of the subjects in the food checklist and 24 hour recall respectively were grossly misclassified in their ranking.

In both studies, subject recruitment and completion of the study protocol were major problems, which undermined the confidence with which one can interpret the results. Further development in the FFQ and food checklist are required before these dietary assessment methods can be used to evaluate dietary intake in the South Asian population.

List of Publications

Abstracts

1. N.A. Karim and B.M. Margetts (1995). Comparison of a food frequency questionnaire with a 4-day weighed record in South Asians. The Proceedings of the Nutrition Society 55,75A.

2. N.A. Karim and B.M. Margetts (1997) Macronutrient sources in South Asian foods. The Proceedings of the Nutrition Society 56,27A.

Tables of Contents **Page**

List of tables

List of Figures

List of Appendices

Acknowledgement

I would like to express my appreciation to a number of people who have supported me in the preparation of this thesis. Firstly I would like to thank Dr Barrie Margetts as my supervisor for his guidance and encouragement which helped in the direction of this thesis and for his patient reading of draft chapters. I am also grateful to Prof Alan Jackson for his useful guidance and comments. I wish to thank Universiti Kebangsaan Malaysia for giving me the study leave to do this thesis. I would like to thank Parveen Damani and Kirpal Vedwan for introducing me to the South Asian subjects and without whom I would not have been able to produce this thesis. I am grateful to the Association of Commonwealth Universities for awarding me the Commonwealth Staff Scholarship Award to study at the University of Southampton, United Kingdom. Lastly but not least I wish to thank my colleagues at the Institute of Human Nutrition and Public Health Medicine, my parents and especially my husband, Zainal who always gives me his support and encouragement and my lovely children Hanna, Johan, Rizal and Zain for keeping my faith as a mother.

INTRODUCTION

INTRODUCTION

No dietary assessment methods have been developed and validated for use in the South Asian community in the United Kingdom (UK). The objectives of this thesis are therefore :

Main objective:
Develop dietary assessment methods for use in the South Asian community.

Specific objectives:

1. Establish the main food sources of macronutrients in the different South Asian communities in Southampton
2. Develop a food frequency questionnaire (FFQ) and a food checklist covering 90% of the main food sources in the South Asian diet
3. Assess the relative validity of the FFQ compared with a weighed dietary record
4. Assess the relative validity of the food checklist compared with a weighed dietary record
5. Assess the relative validity of a 24 hour recall compared with a weighed dietary record
6. Describe the likely measurement error introduced by using each method as a measure of the true exposure in an epidemiological study
7. Assess the limitations on the generalizability of the findings because of the potential sampling bias and poor response rate

The thesis is organized into five chapters. Each chapter is outlined as follows. The first chapter discusses the review of the literature relevant to this study; it is divided into seven main sections. Firstly, the description of the South Asian community is presented; secondly is a description of diabetes prevalence studies of Indians in India and South Asians in the UK; thirdly the dietary habits, general meal patterns and nutrient intake studies of South Asians that have been carried out in the UK since the 1970's are reviewed; fourthly the theoretical requirement of dietary assessment methods, measurement errors and the various methods of dietary assessments are discussed; fifthly, previous validation studies are reviewed. The final two sections of chapter one review the comparative dietary methods in epidemiological studies and assumptions of using a biomarker.

The second chapter presents the methododology used in the present study. The objectives of the study are discussed first and followed by a discussion of the development of the FFQ. The main focus of the discussion is on the methodological issues in the calibration studies of the FFQ, food checklist and the 24 hr recall with the 4-day weighed record.

The results of the study are discussed in chapter three. These include the results of the preliminary survey and the calibration studies of the FFQ, food checklist and 24 hr recall.

In chapter four, firstly, the discussion concentrates on the meal pattern and the sources of macronutrients in the South Asian diet. The discussion on the calibration studies follows; and the last two sections focus on the biases in the study design and measurement errors of dietary methods in the calibration studies.

Conclusions of the study are discussed in chapter five. The implications of the result of the calibration studies on the dietary assessment methods which may be used in the South Asian community follows. Some recommendations for future studies are then presented.

CHAPTER ONE

REVIEW LITERATURE

CHAPTER 1 REVIEW OF THE LITERATURE

The review of the literature consists of seven sections. The first section focuses on the origin and population of South Asians in the UK (1.1). This is followed by the prevalence studies of diabetes of Indians in India and South Asians in UK (1.2). The third section discusses the dietary habits of South Asians (1.3) focussing on the general meal patterns (1.3.1) and nutrient intake studies of South Asians in UK (1.3.2)

The fourth section highlights the theoretical requirements of dietary assessment methods and measurement errors in dietary measurements (1.4.1 through 1.4.4). This is followed by the review of dietary assessment methods (1.4.5) which include weighed record and estimated record (1.4.6), recall method (1.4.7), dietary history (1.4.8) and questionnaire (1.4.9).

Validation studies of FFQ are discussed in the fifth section (1.5). This is further broken down into validation studies of FFQ vs weighed record (1.5.1) and validation studies of FFQ vs food record (1.5.2). Section six reviews the only comparative dietary method study among the South Asian community (1.6) as well as the comparative dietary methods in epidemiological studies (1.6.1).

The final section discusses the assumptions of using biomarkers in dietary studies (1.7) and is followed by the discussion on validation studies using biomarkers (1.7.1). The discussion on the predictive equations available for the calculation of basal metabolic rate (BMR) concludes the review of the literature (1.7.2.).

1.1 South Asians in the United Kingdom

In the United Kingdom, there is an estimated one and half million South Asians who originated from the main subcontinent of India or from East Africa (Census, 1991). The South Asians from the subcontinent of India are immigrants mainly from the states of Punjab and Gujerat in India, while the others are from Pakistan and Bangladesh. The East African South Asians originate mainly from Kenya and Uganda. South Asians comprise 2.7% of the total United Kingdom population. Similarly in Southampton where the calibration study is carried out, South Asians make up 2.8% of the city population (Table 1.1).

Table 1.1 Distribution of South Asian ethnic groups in the total United Kingdom (UK) and Southampton population (% of population) (Census,1991)

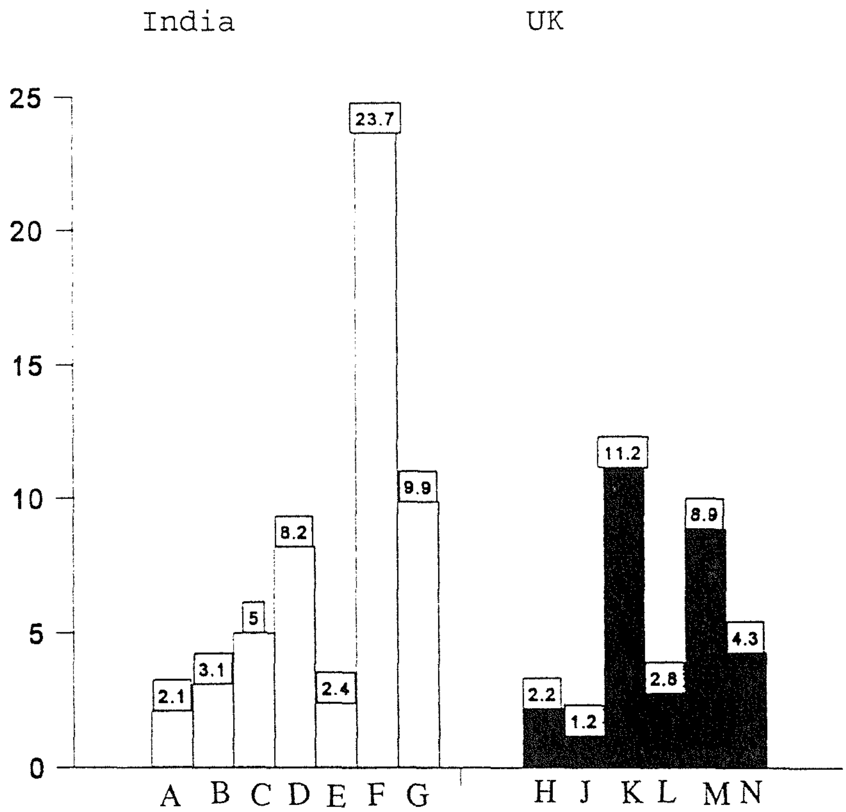
Ethnic groups	UK population	% of UK population	Southampton population	% of Southampton population
Indians	840,255	1.5	3,863	2.0
Pakistanis	476,555	0.9	993	0.5
Bangla-deshis	162,835	0.3	515	0.3
Total	1,479,645	2.7	5,371	2.8

1.2 Diabetes among South Asians in the United Kingdom

One of the major health problems in the South Asian community is diabetes. Since the 1980's, there have been two large diabetes surveys in the UK, both comparing prevalence between South Asians and Europeans (Mather and Keen,1985; Simmons et al,1989).

Figure 1.1 shows the cross sectional prevalence studies that have been carried out in Indians in India and South

Fig 1.1 Cross sectional studies on the prevalence of diabetes among Indians in India and South Asians and Europeans in the UK



A Multicentre (Ahuja,1979) 2.1%

B New Delhi (Verma et al,1986) 3.1%

C Madras (Ramachandran et al,1988) 5.0%

D Madras, urban (Ramachandran et al,1992) 8.2%

E Madras, rural (Ramachandran et al, 1992) 2.4%

F Madras, urban elderly (Ramachandran et al,1994) 23.7%

G Madras, rural elderly (Ramachandran et al,1994) 9.9%

H South Asian, Southall } 2.2%

J European, Southall } (Mather and Keen, 1985) 1.2%

K South Asian men, Coventry } 11.2%

L European men, Coventry } 2.8%

M South Asian women, Coventry } 8.9%

N European women, Coventry } (Simmons et al,1989) 4.3%

Table 1.2 Summary of diabetes survey of Indians in India and South Asians in UK

Diabetes survey	Diagnosis criteria	Ethnic group	Age group	Prevalence
A.Multicentre, India (Ahuja,1979)	FBG >5.6 mmol/l D >6.7 mmol/l	North Indian	>15 years old	2.1%
B.New Delhi (Verma et al,1986)	Household asked whether any resident is a known diabetic	North Indian	>15 years old	3.1%
C.Madras (Ramachandran et al,1988)	IGT $\geq$7.8 mmol/l D $\geq$11.1 mmol/l	South Indian Dravidian and Tamil Naidu	>20 years old	5.0%
D.Madras, urban (Ramachandran et al,1992)	IGT $\geq$7.8 mmol/l D $\geq$11.1 mmol/l	South Indian Dravidian and Tamil Naidu	>20 years old	8.2%
E.Madras,rural (Ramachandran et al,1992)	IGT $\geq$7.8 mmol/l D $\geq$11.1 mmol/l	South Indian Dravidian and Tamil Naidu	>20 years old	2.4%

Table 1.2 continued

Diabetes survey	Diagnosis criteria	Ethnic group	Age group	Prevalence
F.Madras, urban elderly (Ramachandran et al, 1994)	IGT $\geq$7.8 mmol/l D $\geq$11.1 mmol/l	South Indian Dravidian and Tamil Naidu	>60 years old	23.7%
G.Madras, rural elderly (Ramachandran et al, 1994)	IGT $\geq$7.8 mmol/l D $\geq$11.1 mmol/l	South Indian Dravidian and Tamil Naidu	>60 years old	9.9%
H.Southall, UK (Mather and Keen, 1985)	Household asked whether any resident is a known diabetic	mainly Punjabis	>15 years old	2.2%
K.Coventry, UK (Simmons et al, 1989)	IGT $\geq$7.8 mmol/l D $\geq$11.1 mmol/l	mainly Punjabis	>15 years old	11.2% (men) 8.9% (women)

FBG Fasting blood glucose
IGT Impaired glucose tolerance
D Diabetes

Asians and Europeans in the United Kingdom. The summary of the prevalence studies is outlined in Table 1.2. The methods of diagnosis for diabetes and impaired glucose tolerance (IGT) in all these studies, except the multicentre study in India were based on the criteria of the World Health Organization (WHO,1985). Therefore an individual is diagnosed and confirmed as having diabetes if after going through an oral glucose tolerance test (OGTT) (WHO,1985), the blood glucose level (venous plasma) is greater than 11.1 mmol/l, while an individual is considered as having impaired glucose tolerance if his/her blood glucose (venous plasma) is 7.8 mmol/l to 11.1 mmol/l.

In the studies before 1985, for example, the Southall survey in UK and the New Delhi study in India, diabetes was ascertained by asking each household whether any resident was a known diabetic. In the multicentre study in India in the 1970's, the criteria used for diagnosing diabetes were different from the WHO criteria. In the multicentre study, diabetes is diagnosed based on fasting blood glucose level as well as after the OGTT. Here, an individual is diagnosed as diabetic when the fasting blood glucose is 5.6 mmol/l, while after OGTT, the blood glucose level is greater than 6.7 mmol/l. This criterion was even lower than the later diabetes study in the 1980's.

The earlier multicentre study conducted in the 1970's by the Indian Council of Medical Research in six urban centres demonstrated a prevalence ranging from 0.9% (Delhi) to 3.7% (Ahmedabad) and a mean of 2.1% (Ahuja, 1979). In the 1980's a study carried out in Darya Ganj, a suburb of New Delhi showed a higher crude prevalence of 3.1% (Verma et al,1986). Ramachandran et al (1988) conducted a study at a township (Madras) in South India and found that the prevalence of diabetes was much higher (5%) than that in the Indian Council Medical Research study in the 1970's. This implied that the prevalence of diabetes in India in the 1980's was increasing, even though the criteria for the diagnosis of diabetes was more stringent than that used in the multicentre study in the 1970's.

In the 1990's more studies were carried out in India, comparing the prevalence of impaired glucose tolerance and diabetes between the urban and rural population. Ramachandran et al (1992) reported that the prevalence of diabetes was 8.2% in the urban compared with 2.4% in the rural areas. The prevalence was higher in both men (8.4%) and women (7.9%) of the urban compared to the rural areas (2.6% in men and 1.6% in women). This was further confirmed in a more recent study, which showed a greater prevalence of diabetes in the elderly population of urban (23.7%) versus the rural (9.9%) (Ramachandran et al,1994). However both studies reported that despite a higher prevalence of diabetes in the urban population, the occurrence of impaired glucose tolerance was similar in both the urban (12.4%) and rural (14.9%) populations (Ramachandran et al,1992;Ramachandran et al,1994).

In United Kingdom, the Southall Diabetes Survey showed that the prevalence of diabetes in South Asians was at least two times higher than that in the Europeans (2.2% vs 1.2%) (Mather and Keen, 1985). The Coventry diabetes study on the other hand demonstrated a prevalence of four times higher in South Asian men (11.2% vs 2.8%) and two times higher in South Asian women (8.9% vs 4.3%) compared to the Europeans (Simmons et al,1989). A more recent study in London on the prevalence of diabetes in South Asians also revealed a prevalence almost four times higher (McKeigue et al,1991). A common feature in all these surveys was that the South Asian subjects studied were mainly from Punjab; 82% in the Southall Survey, 70% in the Coventry study and about 62% in the London study. In a later report, Simmons et al (1992) showed that the prevalence of diabetes could be different even between the ethnic groups of South Asians. It was reported that the age adjusted prevalence for diabetes was highest among the Gujerati Muslims and followed by Punjabi Hindus, Pakistani Muslims, Punjabi Sikhs and Gujerati Hindus in the men (Table 1.3). A similar trend was also observed in the women.

Table 1.3 Age adjusted prevalence of diabetes (per 1000) of ethnic groups in South Asian population (Simmons,1992)

Ethnic group	Men	Women
Punjabi Sikh	89	75
Pakistani Muslims	91	103
Gujerati Muslims	160	204
Gujerati Hindus	84	88
Punjabi Hindu	113	116

Most of the diabetes surveys carried out in India were carried out in New Delhi or Madras and the surrounding rural areas if comparisons of diabetes prevalence between urban and rural areas were made. There were two major ethnic groups in these Indian studies. The ethnic groups were either South Indian Dravidian or Tamil Naidu (mainly Madras region) or North Indian (New Delhi area). However, these two groups shared a common Hindu religion. This implied that they followed similar religious dietary restrictions. However, their staple foods differed. The South Indians consumed rice as their staple (Ramachandran et al,1988) while the North Indians, chapatti (Gelbier and Taylor,1985;Wharton and Eaton,1983).

Therefore from the prevalence studies discussed above, it was demonstrated that there were ethnic (north and south Indian) differences between the Indians in India and the South Asian ethnic group (Punjabi) in the UK. However, it could be possible that the Indians in India and the immigrants from India in UK have a higher tendency to develop IGT or diabetes than the Europeans in UK. It happened that the diabetes survey in the UK was carried out in an area where there was a higher concentration of Punjabis than the other South Asian ethnic groups. Thus this might not reflect the actual prevalence of diabetes in the various South Asian ethnic groups.

The reasons for the differences in the prevalence of diabetes between South Asians and Europeans still remain unknown. A lot of studies have been carried out to investigate these differences and some have focused on the diet of these populations. These studies will be discussed in Section 1.3.2. In the next section, the dietary habits of South Asians are discussed. Throughout the thesis, terms such as nutrient, diet, dietary, meal, habits and practice are often referred to in the discussion of food. In order to clarify the meaning of these terms, each is defined according to Jackson (1992)(Appendix 3).

1.3 Dietary habits of South Asians

South Asians consist of communities who originate from the various parts of the Indian sub continent namely India, Pakistan and Bangladesh or from East Africa. Various ethnic groups comprise this South Asian population. The Gujeratis come from western India, the Gujerat state, while the Punjabis originate from the north west, Punjab. The other groups namely the Pakistanis are from Pakistan; the Bangladeshi or Sylheti and Bengalis from Bangladesh (Gelbier and Taylor, 1985;Wharton and Eaton,1983). Since these communities are from different parts of the Indian continent, they each have different methods of food preparation depending on their religious beliefs, culture and tradition. The religion of South Asians is closely related to their origin, therefore the Gujeratis are normally Hindus or Jains as well as Muslims, while the Punjabis are usually Sikhism followers. The Bangladeshis and Pakistani are Muslims.

Although the first generation of South Asians have migrated from the Indian main subcontinent and East Africa since the late 1960's to early 1970's, evidence has suggested that they still retain their dietary habits and cultural practices especially those who migrated when they were adults. (Hunt,1977). Some of these South Asian immigrants have even maintained traditional ideas about the role of diet in maintaining health and treating diseases (Bailey et al,1986). The original diet of the South Asians is also associated with the prevailing climate and agricultural conditions (Gelbier and Taylor,1985).

1.3.1 Meal pattern of South Asians

The dietary habits of South Asians are basically similar, though there are variations in the preparation and the amount of seasoning used within families and communities (Attariwala,1977). In general cereals are the main ingredients of the meal in Bangladesh, India and Pakistan. However, the Pakistanis and Punjabis normally consume chapattis while the Gujeratis eat more rice. The Pakistanis usually eat chapatti made with wheat flour while the Gujeratis and Punjabis consume chapattis which are made with wheat or maize flour. The staples of all these ethnic groups also consist of paratha and puri (Wharton and Eaton,1983). The basic diet of the South Asians normally consists of chapattis (made with 85% wheat flour), white rice, vegetables, pulses, beans and fruits. The vegetables eaten by the Gujeratis and Punjabis are like those eaten by the British in the UK. However, the Pakistanis tend to consume familiar tropical vegetables such as aubergine, okra and karela, which are usually more expensive. Fruits eaten are those available in the UK like apples, oranges, bananas, grapes and dates. The fats in the diet are ghee (clarified butter) and cooking oil, which is usually corn oil. The diet also consists of high fried snacks, sweetmeats (made from evaporated milk and sugar) and carbonated drinks. Milk is also consumed regularly (Peterson et al,1986;Gelbier and Taylor,1985;Wharton and Eaton,1983).

The meal times of South Asians are flexible. Breakfast can be consumed with the men before they leave for work, eaten with the children before school or taken with other women and young children in the household before noon. Breakfast foods are the most 'westernized' consisting of bread with butter or jam, eggs either fried or boiled and even cereals with milk. A traditional breakfast would consist of paratha. The lunch meal would contain curry of meat from the previous night, eaten with chapatti. The main meal of the day which is dinner also serves chapatti, curry of meat, fish, vegetables or pulses depending on the ethnic group (Wharton and Eaton,1983). Table 1.4 summarizes the meal pattern of the various ethnic groups in the South Asian community.

Despite the similar meal pattern in all the ethnic groups, the method of food preparation tend to vary. For example, the staple diet in all the groups except the Bangladeshi is chapatti. The staple diet of the Bangladeshis is rice. The preparation of chapatti also differs between ethnic groups. The Pakistanis make large chapattis (8 to 10 inches in diameter) compared with the Gujeratis and Punjabis who prepare smaller chapattis (5 to 8 inches in diameter). The Punjabis also make thicker ones, the Gujeratis thinner ones, while the Pakistanis, intermediate in size. Butter or oil is often spread on the cooked chapatti by the Gujerati, in addition to the little oil already added in the chapatti dough (Wharton and Eaton,1983).

According to Kalka (1988), the dietary habits of the Gujeratis have not changed very much from the native

Table 1.4 Meal patterns of the various ethnic groups in the South Asian community

Meal	Gujerati	Pakistani	Punjabi	Bangladeshi
Breakfast	Paratha or chapatti OR cereals with milk OR bread and butter and jam OR fried or boiled eggs WITH tea	paratha or chapatti OR cereals with milk OR bread and butter and jam OR fried or boiled eggs WITH tea	paratha or chapatti OR cereals with milk OR bread and butter and jam OR fried or boiled eggs WITH tea	paratha and chapatti OR cereals with milk OR bread and butter and jam OR fried or boiled eggs WITH tea
Lunch	Chapatti and rice with curry of meat,vegetable or dal and fresh fruit	chapatti or paratha with curry of meat,vegetable or dal and fresh fruit	chapatti and paratha with curry of meat, vegetable or dal and fresh fruit	rice with curry of fish,meat, vegetable or dal
Dinner	similar to lunch	similar to lunch	similar to lunch	similar to lunch

Gujerat although some changes in the breakfast foods have taken place. The main meal normally comprises two phases; the first one consists of chapatti eaten with cooked vegetables or vegetables with pulses, pickles, chutney served with water, buttermilk, plain yoghurt and sliced vegetables. The second phase would include rice and soup (dal). Since most Gujeratis are strict Hindus, they are usually vegetarians.

Studies investigating the recipes and the nutrient composition of composite dishes consumed by the South Asians have not been carried out until very recently. Kassam-Khamis et al (1995) determined the dishes most commonly eaten by the different South Asian ethnic groups in the UK. The nutrient content and the extent of variation in the nutrient content were then evaluated.

The results of this study among Punjabis (in Southall) and Gujeratis (in Brent and Wembley) demonstrated that there appeared to be a wider range of dishes commonly eaten by the predominantly vegetarian Gujeratis compared to the omnivorous Punjabis. As shown in the earlier dietary habits studies, chapatti was the most commonly eaten food in both groups. At the same time, 72% of the Gujeratis and only 30% of the Punjabis reported consuming rice. Both the Gujeratis and Punjabis consumed milk but the form in which milk was taken differed. Punjabis drank Indian tea (milk and tea boiled together) and ate yoghurt while the Gujeratis consumed lassi (yoghurt drink).

The recipes for the same dishes could vary even within the same ethnic group and between ethnic groups. Thus there was often a wide variation in energy content between two recipes for the same dish with similar ingredients. The variations usually reflected the fat content in the recipes. However, when comparing dishes which were common in both the Punjabis and the Gujeratis which had similar ingredients, the recipes of the Gujeratis appeared more often to have a higher fat content than the recipes of the Punjabis.

Earlier studies in the South Asian communities focused on their food habits and meal patterns whereas more recent studies have looked at the nutrient intakes derived from meals. These studies compared the nutrient intakes of South Asian with the Europeans. These comparative studies are discussed in the next section.

1.3.2 Nutrient intake studies of South Asians: cross sectional studies in the United Kingdom

Most of the dietary intake studies of South Asians have investigated the nutrient intake of the ethnic group which was predominant in the area where the research was carried out. For example, the Gujerati Hindus were the most studied in the London area (Khajuria and Thomas,1992; Peterson et al,1986;Kalka,1988). There were also studies on the vitamin D intake of South Asian Muslims in the Rochdale area; reevaluating the health education campaign to improve vitamin D intake (Stephens et al,1982). Some dietary intake studies have also been carried out on pregnant women (Hindus and Muslims) in Birmingham (Wharton and Eaton,1983) and in Harrow, London (Abraham et al,1987).

Studies comparing the nutrient intakes between South Asians and Europeans are shown in Table 1.5. One of the earlier studies comparing nutrient intakes between South Asian households with the British National Food Survey was carried out in the 1980's (McKeigue et al,1985). The study showed that the energy consumption of the South Asian household was 11% higher than the British household. The fat intake was comparable but striking differences were observed in the lower intakes of saturated fat and cholesterol and much higher intakes of polyunsaturated fat

Table 1.5 Comparative studies of nutrient intakes between South Asians (SA) and Europeans

Author	Study design	Population	Dietary method	Results	Comments
1.McKeigue et al (1985)	Cross sectional study	184 Gujeratis	Household inventory and FFQ	Significantly higher energy,CHO and P/S ratio in SA but comparable fat. Lower SAFA and cholesterol in SA	Distribution of energy intake of SA diet was comparable to the Europeans
2.Silman et al (1985)	Cross sectional study	12 Bangladeshi	Weighed raw ingredient and food intake	Significantly higher,energy,fat, CHO and P/S ratio in SA	Bangladeshis have different food habits than other SA group, thus a different result than other studies
3.Miller et al (1988)	Cross sectional study	20 Gujeratis	5-day weighed record and FFQ	Significantly lower energy but higher P/S ratio in SA. Comparable fat and CHO.	Energy intake contradicts other studies
4.Smith et al (1993)	Cross sectional study	Pakistanis (n=78) and Gujeratis (n=44)	3 day diet diary	Significantly higher P/S ratio and fibre in SA. Comparable energy, fat and CHO	Energy intake contradicts other studies
5.Sevak et al (1994)	Cross sectional study	Mainly Punjabis and Gujeratis. Pakistanis are a minority (n=92)	7-day weighed record	Significantly lower energy,protein and fat but higher P/S ratio and CHO in SA.	Results in agreement with other studies only in P/S ratio

SAFA - saturated fatty acid
CHO - carbohydrate
P/S - polyunsaturated fatty acid/saturated fatty acid

in the South Asian households. Thus a higher polyunsaturated /saturated (P/S) ratio of diet was observed in the South Asian households. The carbohydrate, protein and fat contribution to the total energy intake were 49%, 12% and 39% respectively. This distribution of macronutrients within energy intake was comparable to those determined in the National Food Survey which were 45% carbohydrate, 13% protein and 42% fat (McKeigue et al,1985).

Silman et al (1985) evaluated the nutrient intake of a small group of Bangladeshi men (n=12) in London using the weighed raw ingredient and weighed cooked food record. The results of this study showed that the Bangladeshis have a very high intake of fat and carbohydrate, thus a high energy intake when compared with the estimated UK daily consumption. The P/S ratio was also high (0.46) in the Bangladeshis.

In another comparative nutrient intake study, 20 Gujerati Indian men completed a questionnaire and 5-day weighed records (Miller et al,1988). The nutrient intakes were compared with Europeans. In contrast to the intakes of the Bangladeshis (Silman et al,1985), the Gujeratis in this study consumed significantly lower intake of energy than the Europeans. However the percent total fat and carbohydrate intake was comparable to the Europeans. Again in this study, the P/S ratio was also high (0.52), in agreement with studies by McKeigue et al (1985) and Silman et al (1988).

A recent study on the nutrient intake of South Asian and European men in Bradford has shown that there was significant variations in the types of foods consumed between the two population groups, as well as within the South Asian religious subgroups (Smith et al,1993). With regards to the estimated daily nutrient intakes, significant differences were observed in the protein, total sugars, added sugars, fibres and P/S ratio. The former three nutrients were higher in the Europeans while the latter two were higher in the South Asians. However, no significant difference was shown in the intake of energy, carbohydrate and total fat. When the type of carbohydrate consumed was listed, the Europeans were shown to consume much higher added sugars (61g vs 39g) than the South Asians. The total sugar consumptions was also more in the Europeans than the South Asians (107g vs 79g). Fibre intake, however, was significantly higher in the South Asians than the Europeans. Percent of carbohydrate which contributed to the total energy intake was slightly higher in the South Asians than the Europeans, and this difference was significant.

When comparisons were made in the type of fatty acids consumed, this study showed that the Europeans consumed slightly higher saturated fatty acids but lower polyunsaturated fatty acids in their diet than the South Asians. This was also reflected in the P/S ratio which was higher in the South Asians. The monounsaturated fatty acid intake was similar in both ethnic groups (Smith et al,1993).

The South Asian subjects participating in the Bradford study consisted predominantly of Muslim men with a minority of Hindu men. Between these two religious South Asian subgroups, there was no significant difference in the intake of added sugars and total sugar. The estimated intakes of energy, protein, carbohydrate, fibre, total fat and P/S ratio were significantly different. The Hindus consumed significantly higher intakes in all these nutrients except protein. The components of carbohydrate intake such as added sugars and total sugars were similar in both these religious groups. However, the fibre intake was significantly higher in the Hindus than the Muslims. This difference could be attributed to the vegetarianism that was commonly practised among the Hindus, and this was shown in the higher consumption of vegetables. The total fat intake was significantly higher among the Hindus. This was demonstrated in the higher fat contribution to the total energy intake in the Hindus than the Muslims. Despite these differences, the saturated fatty acid and monounsaturated fatty acid intake were almost the same. The only difference was in the polyunsaturated fatty acid consumption where the Hindus consumed more, thus a higher P/S ratio in the their diet.

In another recent study of South Asians and Europeans men in the west London borough of Ealing, the estimated energy intake, fat and protein were significantly higher in the Europeans than the South Asians. Carbohydrate intake, mainly starches, was consumed in a significantly higher amount in the Asians than Europeans. When the fatty acid components were compared, the Europeans in this study, just as in the Bradford study also consumed significantly higher saturated fatty acid and a lower polyunsaturated fatty acid in the diet. However, in this study the Europeans also consumed a higher monounsaturated fatty acid (Sevak et al, 1994).

1.3.3 Summary critique of S.Asian nutrient intake studies

The studies comparing nutrient intakes between the South Asians and Europeans were all cross sectional studies and the results obtained were inconsistent, especially in the intake estimates of energy and fat (Sevak et al,1994;Smith et al,1993;Miller et al,1988;McKeigue et al,1985;Silman et al,1985). There was consistency, however, in the higher carbohydrate and fibre intake as well as a higher P/S ratio in South Asians than Europeans in all studies. Energy and fat intakes were either higher (Silman et al,1988), comparable (McKeigue et al,1985;Smith et al,1993;Miller et al,1988) or lower (Sevak et al,1994) in the South Asians compared with the Europeans (Table 1.5). The inconsistent results obtained could possibly be due to the different ethnic groups which participated in these studies as well as the different dietary measurement methods used to evaluate the nutrient intake.

The number of subjects who participated in these studies also varied between 12 to 184. In cross sectional studies, having a larger number of subjects would increase the power of the study (Cole,1991). However, in these comparative studies, it would be difficult to suggest which study reflected the true nutrient intakes of the South Asian community. It was possible that there was sampling bias in these studies, especially if the study was a small study. Studies with a small number of subjects would have less power to detect any nutrient differences between South Asians and Europeans. However, those studies with a larger number of subjects would have more power to detect any nutrient intake differences.

The South Asians in the London study were predominantly Punjabis with a smaller number of Gujeratis and Pakistanis (Sevak et al,1994), while the Bradford study subjects were mainly Pakistanis (Smith et al,1993). The earlier studies consisted of predominantly Bangladeshis (Silman et al,1985) and Gujeratis (McKeigue et al,1985;Miller et al,1988). It was realized that the dietary habits of each religious subgroup were basically similar although they might have different regional origin and culture. However, the inconsistent results obtained in these studies might be a reflection of the variations in the method of food preparation which could cause the difference in the nutrient intakes in these ethnic groups. Kassam-Khamis et al (1995) also demonstrated that the recipes in different ethnic groups could be different.

Currently there is only one study of the South Asian community which has compared a weighed inventory with a diet history. However, this study was conducted in South Asian infants (Dugdale and Harbottle,1993). Presently there is no dietary assessment method which has been calibrated for use in the adult South Asian community. It could be possible that the inconsistent results demonstrated by the comparative dietary studies carried out so far are the result of the different dietary methods used. A reliable and validated dietary assessment method is needed to evaluate the dietary intake of the South Asian community. Thus it is the general objective of this study to develop a dietary assessment method, specifically a food frequency questionnaire (FFQ). Before this is discussed further in the methodology (Chapter 2), the theoretical requirement of dietary assessment methods and measurement errors are discussed.

1.4 Theoretical requirement of dietary assessment methods

Before a dietary assessment method is selected to evaluate the dietary intake of the population to be studied, some factors need to be considered. The first question should be what do we want the dietary method to measure. Then, the dietary exposure of interest is defined. The next step is to determine the appopriate dietary method to use to measure the dietary exposure. It is important that the selection of the appropriate method to measure the exposure is considered while still at the design stage of the study. This is because the different types of epidemiological studies will require different measures of dietary exposure, therefore will likely use different types of dietary methods of assessment.

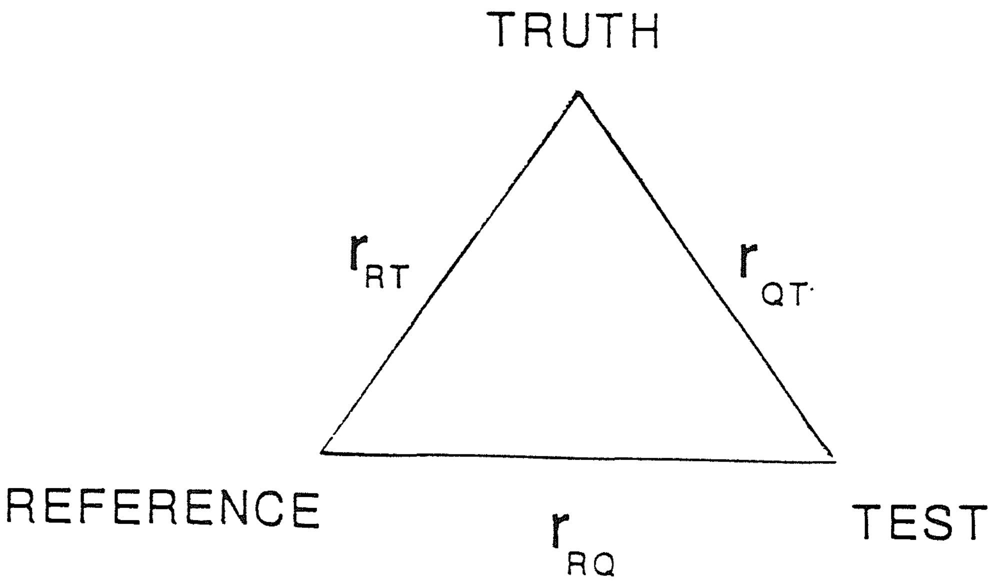
Generally, epidemiological studies can be categorized into experimental and observational studies. In experimental studies, the investigator assigns the exposure to the subjects. In observational studies, the investigator has no control over the subjects' level of exposure. Both experimental and observational studies can be carried out on a population or individual level. A community trial is an example of a population based experimental study, while a clinical trial is an example of individual based experimental study. Examples of individual based observational studies include cross-sectional studies, case-control studies and cohort studies. Population based observational studies include household surveys and ecological studies. Generally, an individual based study has an advantage over a population based study. This is because an individual based study can estimate risk of disease directly in relation to the exposure (Margetts,1991).

Therefore, whatever study design is used, it is imperative that an appropriate measure of exposure is used. For example, in the present study, a cross sectional design is used. We want to be able to divide the subjects into thirds of distribution and determine how risk of developing diabetes varies across the thirds of energy and fat intake. It is not important to know the absolute energy and fat intake of each subject, but it is necessary to ensure that the subject is correctly ranked in the distribution. Thus, because this is our objective, we are developing and calibrating a FFQ for the South Asian community, then use the FFQ to measure our dietary exposure.

When choosing a dietary assessment method to evaluate the usual intake of a community, it is essential that the method selected can measure what it is supposed to measure. Measuring dietary intake is a complex process and the measured intake is usually not a perfect proxy of the true intake. Measurement errors are likely whenever the evaluation of intake is made.

In many epidemiological studies, we require a method which is cheap, simple and can measure usual intake. However, it is usually not possible to have an absolute measure of intake. Therefore to determine whether the method used to measure the usual intake is valid, it is neccessary that the method is compared with another method which from previous research has been considered to provide more accurate measures. Theoretically the dietary method selected to measure the usual intake is the test method, while the more accurate method which the test method is compared with is the reference method. The reference method also acts as a proxy of the true intake which we wish to measure. The relationship between test and reference measures and the truth is shown in Figure 1.2. We can also determine whether the reference method is valid itself by establishing the validity of the reference method, which can be done by comparing it with a biomarker (assuming that the biomarker is valid).

Figure 1.2 Schematic diagram of the relationship between test and reference measures and the truth (Nelson,1996)



r_{RT}, correlation between the reference method and the truth

r_{QT}, correlation between the test method and the truth

r_{RQ}, correlation between the reference method and the test method

1.4.1 Measurement errors

An issue that is gaining importance and considerable interest in measuring dietary intake is measurement errors involved in the dietary assessment methods used. Measurement errors are basically the errors made when making dietary measures. There are generally two types of errors, which are random and systematic errors. In epidemiological studies, these errors can occur at two different levels, within person and between persons. In measuring a dietary exposure among a group of indiviuals, measurement errors that occur can either be random or systematic. For random error, the mean of many repeated measures approaches the true value, while for systematic error, the mean of repeated measurements does not approach the true value (Willett,1990).

These measurement errors are always present. Thus, the main objective during the measurement of diet is to minimize the measurement error in the study design. It is also important that these measurement errors are considered in the analysis and in the interpretation of the results.

1.4.2 Types of measurement errors

Normally when evaluating the dietary intake, we would like to measure the usual or habitual diet. In determining the usual diet, there are sources of measurement errors involved. These measurement errors can affect the relationship between the exposure and outcome being investigated. As described earlier, the measurement errors are random error and systematic error (bias). An example of random error in dietary measurement is the true variability in individuals'diets, which are within person variability and between person variability. Random within person error is a day-to-day fluctuation in dietary intake. This is due to changes and variation in daily food intake. Random within person error can be measured by carrying out a reproducibility study, which is making replicate measurements for a subset of subjects.

Random between person error can occur when using one or a few measurements per subject in the presence of random within person error. Random between person error will also result if systematic within person errors are randomly distributed among subjects. This implies that in a random between person error, an overestimation for some subjects is counter balanced by an underestimation for some others. Therefore, the mean for a large group of subjects is the mean intake for the group.

The second type of measurement error is the systematic error. Repeated measurement of diet within a subject may also lead to systematic error. Systematic within person error occurs when subjects consciously or unconsciously underestimate or overestimate their food intake. This systematic within person error is more likely to occur when subjects respond to a standardized questionnaire. When this type of questionnaire is repeated, this systematic error is likely to recur. The result of this repeated error is the mean of the repeated measurements for an individual for example, will not approach the individuals's true mean (Willett,1990).

Systematic between person error results from systematic within person error that affects subjects non randomly. It may occur for a number of reasons. An example of systematic between person error occuring is an ommission of a commonly eaten food on a standardized questionnaire. This will affect all subjects responding to the questionnaire in the same direction but not to the same degree, as each subject does not always use the food equally.

In assessing dietary intake, it is important that the measurement errors and the technical errors of the measuring instrument can be differentiated. Both random errors and systematic errors have to be accounted for, as they can affect the estimate of dietary intake. The effect that random error has is to reduce the power to detect the relationship being investigated in a study. For systematic errors, however, it can either increase or decrease the association of the relationship being investigated. To reduce the random error and increase the power of the study, a large sample size should be recruited into the study (Margetts and Nelson,1995). To reduce the systematic error, a calibration study could be carried out.

Other errors involved in the evaluation of dietary intake also include technical errors, for example, errors due to the properties of the measuring instrument itself, as well as further errors during conversion of food data into nutrient data using food tables (Plummer and Clayton,1993). These technical errors can be minimized by good quality control and following standardized procedure in the study protocol (Margetts and Pietinen,1996).

1.4.3 Effect of measurement errors in a calibration study

Calibration studies are defined as studies which evaluate the relative validity of measures. In calibration studies, the parameters which are required to correct measures of association for a given increase of intake are evaluated (Margetts and Pietinen,1996). In the present study, a FFQ and a food checklist that we developed are calibrated against a 4-day weighed record. The relative validity of these dietary assessment methods are checked against a proxy of a biomarker, the multiples of BMR.

In using a FFQ and a food checklist as test methods, both random and systematic measurement errors are likely to occur. A common error is the systematic error whereby, generally subjects tend to underestimate or overestimate the frequencies of intake, as well as the portion sizes of foods eaten when responding to a standardized questionnaire. Due to these measurement errors, this may result in a loss of power in the study. To correct for the random error, ideally a repeat measurement of the FFQ and food checklist can be carried out. A possible solution to increase the power of the study is to recruit a large number of subjects into the study. However, if no repeated measurements (reproducibility study) are carried out, the effect here is a reduced power in the precision of the dietary intake estimate by the test methods.

The systematic errors in the FFQ and the food checklist include subjects underestimating or overestimating their food intake. Systematic errors are also likely if certain foods commonly eaten are ommited from the FFQ and food checklist. The effect of this type of error is an underestimate of these foods (specificly the nutrients present in these foods) in the same direction in all subjects. To reduce the systematic error, the list of foods included in the FFQ and food checklist should be appropriately selected. These foods should cover 90% of the total energy and macronutrients in the adult diet (Cade and Margetts,1988;Block et al,1986). This reduces the between person systematic error. However, the systematic within person errors cannot be avoided as it is likely that subjects underestimate or overestimate their frequencies of intake and portion sizes.

To reduce the systematic errors of the FFQ and food checklist (test methods), the calibration study is carried out. The relative validity of the test methods are usually compared with a weighed record (WR)(reference method). Nevertheless, even the WR may not be sufficient to estimate systematic error as data collected with WR can also include systematic errors. Therefore to find a reference method which can give an unbiased estimate of food intake, a biomarker is used. One of the most widely used biomarkers is the 24-hour urinary nitrogen excretion for checking the relative validity of protein intake. In some studies, it is not practical to use a 24-hour urinary nitrogen excretion as a measure of protein intake, especially when the subjects are reluctant to provide urine specimen. A proxy of a biomarker which can be used is the multiples of BMR. The advantage of using a biomarker as another reference method is because random errors between the biomarker, the test methods and the first reference method are mutually independent of each other.

In theory, since the weighed record is regarded as a more superior method than the FFQ and the food checklist, it should be able to estimate the systematic error. However, data collected with weighed records may also include systematic errors. Thus, the use of a biomarker helps in identifying errors in the dietary methods.

There are two circumstances that can occur with the presence of measurement errors in a calibration study. Firstly, if the random measurement errors in the reference method and the test method are independent of each other, the random error of the test method will be overestimated, but the validity coefficient will be underestimated. Secondly, if the random measurement errors in the reference method and the test methods are correlated, then we cannot tell whether the random error and the validity coefficient will be under- or overestimated (Kaaks et al,1994). The outcome depends on the size of the independent random error in the reference method (gives underestimation of the validity coefficient) and the strength of the correlation between the random errors of both methods (gives overestimation of the validity coefficient).

To determine the measurement errors present in a calibration study, there are now several measurement error models that can be used. These models are discussed in the next section.

1.4.4 Measurement error models

In the last few years, several authors have introduced measurement error models which seek to clarify the different types of measurement errors in dietary assessment methods. One such model was introduced by Kaaks et al (1994). In this model, the measured dietary intake includes the systematic error (bias) and random error. Several assumptions are made in this model and they are as follows:

1. the systematic error (bias) is linearly related to the true intake.
2. the true intake is normally distributed.
3. the random error is independent of the true intake, normally distributed with mean zero (0) and variance s_E^2.

This measurement error model can be shown as below:

$D_i = \alpha + \beta T_i + E_i$,

whereby

D_i = measured intake of individual i

T_i = true intake of individual i

α = systematic constant measurement error

β = proportional scaling factor

E_i = random measurement error of the individual.

Each term in the equation indicates the measurement errors that can occur when making dietary measurements. For example, the term α , which is the systematic constant error shows the average tendency to over or underestimate intake by a constant amount. This is also called the bias in the measurement. The proportional scaling factor, β , on the other hand reflects the average tendency of the individual to overestimate (if $\beta>1$) or underestimate (if $\beta<1$) intake by an amount which is proportional to the level of true intake. The random measurement error E_i can be divided into within subject random error and random bias or subject specific bias. Normally the random bias is random at group level but becomes systematic at individual level. Thus within subject random error will not be reproduced when repeated measurements are made, while random bias will be reproduced. The effect of this random bias at individual level to a calibration study for example, would be either an overestimation or underestimation of the relationship being investigated.

When we are considering measurement errors at the population level, the systematic constant measurement error, the proportional scaling and the size of the variance of the random error relative to the variance of the true intake are indicators to describe the different dimensions of measurement error. It must also be remembered that the measurement error parameters are determined by the particular dietary assessment method and the population in which it is applied (Beaton,1994).

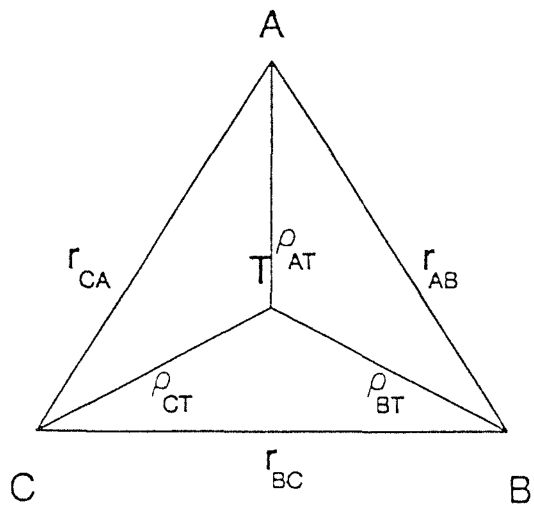
Another model which can be used to investigate the measurement error in dietary assessment is the covariance structure model (Plummer and Clayton,1993). The objective of using this model is to assess the quality of the dietary assessment methods that are being used. This will allow us to separate the study population into high and low intake groups.

The covariance structure model, as the name implies is based on the variance part of the measurement error model. In the calibration and validation of a dietary method for use at individual level, we are interested in looking at the variability of the true average intake and the error variances and covariances. As in the measurement error model of Kaaks et al (1994), in order to use the covariance model, some structure of the error variance matrix has to be imposed. The most appropriate structure is to assume that some of the measurements have uncorrelated errors. But this assumption is difficult to believe, however, we can assume that errors of two different methods on two different occasions are uncorrelated. To estimate the error in the variance and the covariance, several statistical programs can be used. These programs include the LINCS program (Scoenberg and Arminger,1989) and LISREL program available in SPSS (Joreskog and Sorbom,1988).

Recently Ocke and Kaaks (1996) described the method of triads to estimate the level of agreement between three types of measurements. These measurements are usually the test measurement, reference measurement and a biomarker measurement. This relationship is shown in Figure 1.3. The method of triads assumes that the correlation (r) between the three types of measurements are linearly related and the random errors are mutually independent. The validity coefficient, p, is the Pearson correlation between the measured and true intake. It is a measure which incorporates the size of variance of random error relative to the variance of the true intake and the proportional scaling factor. If, for example, the validity coefficient of the three measurements is high, the sample correlations between measurements are expected to be high. On the other hand, if one, two or all the three types of measurements are inaccurate, then the expected sample correlations will be weak.

Looking at the explanation from the opposite approach, if for example, the three sample correlations are high ($r > 0.7$), it could be assumed that the validity coefficient is expected to be close to one. In contrast, if the three sample correlations are low, it will not imply that all the three measurements are inaccurate, but rather the small differences in the low correlation might cause a large differences in the estimated validity coefficient of the three types of measurements.

Figure 1.3 Graphical representation of method of triads. T=true intake variable; A,B, C=measurements of type A,B or C; p=validity coeficient- Pearson correlation between measured and true intake;r=sample correlation between measurements of type A,B and C (Ocke and Kaaks,1996)



$$r_{AB} = \rho_{AT} * \rho_{BT}$$
$$r_{BC} = \rho_{BT} * \rho_{CT}$$
$$r_{CA} = \rho_{AT} * \rho_{CT}$$

$$\rho_{AT} = \sqrt{r_{AB} * r_{CA} / r_{BC}}$$
$$\rho_{BT} = \sqrt{r_{AB} * r_{BC} / r_{CA}}$$
$$\rho_{CT} = \sqrt{r_{CA} * r_{BC} / r_{AB}}$$

When comparing the models to estimate the measurement errors, the method of triads appears to be a practical approach if a biomarker is used as an additional reference measurement to estimate the relative validity of a questionnaire measurement besides a reference method. However, the interpretation of the results obtained by this model can be complicated by the occurrence of Heywood cases. A Heywood case occurs when the estimated validity coefficient is greater than one or when there is a negative correlation in one of the sample correlations (Ocke and Kaaks,1996). Thus when a Heywood case occurs, the method of triads cannot estimate the validity coefficient, as this requires taking a square root of a negative value.

1.4.5 Dietary assessment methods

There are several dietary methods that can be used to assess the dietary intake of individuals or groups of individuals. These methods can either measure the current or past intake. Methods used to evaluate current intake include weighed and estimated records (Bingham,1991). Most other methods aim at assessing the past intake. These methods are the recall method, dietary history and the questionnaire (Nelson,1991). But whatever method is used in a particular study, the challenge is to make sure that the chosen method can correctly measure the dietary exposure with the level of accuracy required for the purpose intended.

In an epidemiological study, the objective of the dietary evaluation is to estimate the usual or habitual intake of individuals or groups and then be able to rank them into distributions, usually thirds or fifths of distribution. Another objective of the dietary evaluation is to reduce the errors of measuring exposure and outcome to within an acceptable level (Margetts,1991). Generally in these studies, the past intakes are assessed. However, in most dietary studies, four dietary assessment methods have been used (Block,1982). These methods are:-

i) weighed record and estimated record
ii) recall method
iii) dietary history
iv) food frequency questionnaire

The discussion of each method follows.

1.4.6 Weighed record and estimated record

The dietary methods which evaluate the current intake are the weighed record and the estimated record (Bingham,1991). In the weighed record, the subjects are required to weigh and record all individual items of food and drink, at the same time anything left over must also be weighed and recorded (Bingham,1981). Before they begin to weigh their foods, the subjects are visited at their own home and taught how to weigh and record their food as accurately as possible. A demonstration is usually carried out on the usage of the weighing scale. The subjects are also encouraged to show their competence in using the scale to the researcher. At the end of the home visit, a notebook or a record book and written instructions are left with the subjects. The number of days required for the weighing of food may range between three to seven days, or sometimes even longer, depending on the nutrient of interest in the study.

In the past decade, a simpler method of keeping weighed record has been introduced. The PETRA (Portable Electronic Tape Recorded Automatic) scales, which are accurate to ± 1g can automatically record verbal descriptions and weight of foods on dual track cassette. This avoids the necessity of subjects having to keep written records. The PETRA scales also do not disclose the weights of the foods eaten to the subjects (Bingham,1987). This thus avoids any intention, if there is any by the subjects, in underreporting or overreporting certain foods.

The weighed record has been regarded as one of the accurate methods of evaluating food intake since more than two decades ago (Marr,1971). This is because with direct weighing of food there is less chance of making mistakes in reporting. However, despite the high accuracy in measuring the food intake, there are several drawbacks in the weighed record. Firstly, the completion of the record requires a high level of numeracy and literacy amongst the subjects. The subjects who participate must also have a high degree of commitment to use the scales at all their meals. Due to these two factors, this method tends to limit the recruitment and the completion of the study (Edington et al,1989;Bingham, 1981). However the usage of the PETRA scales may alleviate these problems. Another disadvantage of the weighed record is the inconvenience of having to weigh all foods eaten. This may result in changes in the frequency of eating and the type of foods eaten (Edington et al,1989). Due to some of the impracticalities involved in the weighing of food, an alternative method like the estimated record has been used. Several large studies have used the estimated record to evaluate current intake of their subjects (Edington et al,1989;Edington et al 1987;Braddon et al,1988).

In the estimated weighed record, all foods and drinks consumed are usually estimated in household measures instead of being weighed. Careful instructions about recording estimated portion sizes are also given. Photographs of small, medium and large portions of many commonly eaten foods can also be used to help the subjects in estimating the food intake (Edington et al,1989;Nelson et al,1994). The estimated weighed record can be carried out for three to seven days or longer.

The advantage in using the estimated record is the lack of inconvenience to the subjects in having to weigh their foods at every meal. Thus it is more appealing and practical to the subjects. Because of this, a higher cooperation rate can be achieved with a 4 to 7 day estimated intake using random subjects (Bingham et al,1987). It has been suggested that if the estimated records are collected satisfactorily, they can be as reliable as weighed record in evaluating food intake (Edington et al,1989;Bingham et al,1994).

The estimated record method has the advantage of recruiting more subjects. Its main problem is in the analysis. Coding each diary involves identifying and determining the weight of a large number of foods. This process can take a long time, for example at least 45 minutes for a seven-day food record (Edington et al,1989). Another problem of the estimated record method is the required training which is usually needed before the subjects involved in any study can be more precise in their estimation (Bingham and Nelson,1991; Bollard et al,1985).

1.4.7 Recall method

The recall method is a dietary method which assesses past intake for the last 24 hours (Nelson,1991). The 24 hour recall method is the simplest of the recall methods, in that the subjects are asked in a systematic way to recall their food intake from the previous day. The quantities of food intake estimated are stated whenever possible in ordinary units such as a slice of bread, a glass of milk or a pat of butter. It will also be helpful if food models of measured quantities are used to assist the subjects in describing their serving size (Bingham,1987).

An advantage of the 24 hour recall is that it can be self administered or administered by a person with little training in a fairly short time. Since it also involves only one day of recording for each subject, a maximum number of subjects can be recruited into the study (Bingham,1981). The 24 hour recall only requires the memory of a recent intake, thus the recall is more precise, therefore the quantities of food estimated may be more accurate. However since it evaluates only a single day's intake, it may not be representative of an individual's intake since his or her diet vary from day to day (Block,1982). However it can be reliable enough to evaluate the current food consumption of a group (Morgan et al,1978).

In order to achieve a greater representativeness of the recall method, a 3-day or 7-day recall can be carried out (Block,1982). However, since the subjects are required to recall intakes of the past three or seven days, the loss in accuracy may exceed gain in representativeness as memories fade quite quickly beyond the most recent day or two (Block,1982;Morgan et al,1978).

The recall method has been used to evaluate the dietary intake in the recent past. Two other methods which have the ability to assess the usual or habitual intake in both the recent and distant past are the dietary history and the questionnaire. These methods are discussed in the next two sections.

1.4.8 Dietary History

The dietary history method was developed by Burke (Burke,1947). It mainly consists of an extensive interview which aims to investigate the usual or habitual diet of an individual (Block,1982). More specificly, the dietary history usually is made up of three parts of evaluation. The first part is a detailed interview to investigate the usual intake of a wide variety of foods. This is then followed by the food frequency list which acts as a cross check list. The final part is a three day record (Nelson,1991).

The starting point of the interview is the 24 hour recall, which finds out the food consumption meal by meal of the individual. This is then followed by the probing of day to day and seasonal variations in the food intake to establish a usual pattern of consumption. The cross check is important in making clarification of the information already gathered about the usual intake. It also improves the accuracy of the dietary history, amount of food intake list which record the frequency and the usual portion size and the type of foods (Bingham,1987; Nelson,1991). The three day record is just an additional way to check the usual intake for a given interval. It has been suggested to be the least valuable part of the dietary history (Bingham,1987).

The dietary history has a few advantages over the 24 hour recall and the questionnaire. This is because it can provide data on the meal patterns, food consumption and nutrient intake of individuals. Since it mainly involves the interviewer's skills, subjects taking part in the dietary history do not require literacy or numeracy skills (Nelson,1991).

However, there are disadvantages of dietary history. Firstly, due to the extensive interview, it is very time consuming and difficult to use in studies involving large number of subjects (Pietinen et al,1988). It is also important that the interviewers involved are competently trained to ensure the dietary history carried out is complete, as well as to reduce any differences between interviewers. Another disadvantage is regarding the possible fading of memory in the subjects if they are required to recall diets and regular habits in the distant past, and also to remember the different items in the diet that are consumed (Block,1982;Bingham,1987;Nelson,1991).

As the dietary history can take a long time to administer, more studies which aim to investigate the relationship between diet and disease have used the questionnaire method to evaluate the dietary intake. The questionnaire method is more practical for epidemiological research because they are easy to administer to a large number of individuals in a standardized format (Flegal et al,1988;Willett et al,1985).

1.4.9 Questionnaire

The questionnaire was first developed in 1960 by Wiehl and Reed (Wiehl and Reed,1960). In developing a questionnaire, two important features of the questions that should be considered are that they are reasonably simple for self administration, and secondly, that the questions should be able to extract information that can reveal any differences in the dietary patterns of interest. The questionnaire should also be able to rank and classify items into categories such that the extremes can be recognized.

The questionnaire that has most often been developed is the frequency and amount type (FAQ). In this questionnaire design, subjects are asked how often they usually eat an item of food or drink and how much they eat when they consume these foods (Nelson,1991). This FAQ can be administered quickly in person or by mail to a large number of subjects. Therefore the questionnaire is useful in large scale epidemiological research (Block,1982).

The number of food items included in the questionnaire should be minimized, but at the same time be sufficient to include major sources of nutrients for the majority of the subjects. The frequency responses should be continuous without any gaps, for example 'never', 'one to three times per month', 'once a week' and followed by a number of times a week to number of times a day. The portion sizes included in the questionnaire should mirror the usual consumption size of the population to be studied (Flegal et al,1988;Nelson ,1991;Bingham et al,1994).

Some of the advantages of the questionnaire are in the easier collection of data, easy coding and it can represent intake over an extended period of time (Flegal et al,1988;Willett et al,1985). It is also cheap and can be used to include subjects in widely spread geographical regions through the use of postal questionnaire (Nelson,1991).

As in the other dietary method evaluations, the questionnaire also has its disadvantages. The main drawback is in the amount of groundwork required in the development of the questionnaire and then its validation (Nelson,1991). As in the dietary history, the involvement of more than one interviewer in the administration of the questionnaire can introduce interviewer bias.

Summary critique of dietary assessment methods

In summary, each of the dietary assessment methods discussed above has its own advantages and disadvantages. These are tabulated in Table 1.6. It is clear that there are several dietary methods to choose from in investigating dietary intakes. However, when selecting the dietary method for a study design, some factors need to be considered. The choice of method in some way depends on the population to be studied, its cost effectiveness, number of subjects involved and the expected compliance, its practicality and more importantly, the level of accuracy required of the method. It is also crucial that the method can appropriately measure the exposure of interest in the study design.

Table 1.6 Summary of advantages and disadvantages of the various dietary methods

Dietary method	Advantage	Disadvantage
Weighed record (WR)	* high degree of accuracy in measuring current intake * direct weighing;reduce reporting error	* high level of numeracy and literacy required in subjects * high degree of commitment by subjects * recruitment number is low * completion rate is low
Estimated record	* convenient for subjects when recording food * recruitment number better than WR	* coding,identifying and determining weight of large number of food * training required before subject can estimate precisely
Recall method	* person administering require little training * recruitment number is large * recall is precise enough as memory of recent intake is required	* not reliable to evaluate current consumption of individual * not a representative intake of an individual's diet due to large day to day variation
Dietary history	* provide data of usual intake * literacy and numeracy skills of subject is not important	* time consuming * skilled interviewer required * fading of memory in subjects to recall long past diet
Questionnaire	* easy to administer * cheap * recruitment number is large * evaluate usual intake	* work required in developing questionnaire and its validation * fading of memory in subjects to recall past diet * possible interview bias

One objective of the present study was to develop a dietary assessment method to evaluate the dietary intake of South Asian population. However, the objective of the dietary assessment was to investigate the usual intake and then rank the subjects into the thirds of distribution (low, medium and high intake). This ranking will then be used to determine differences in intake in relation to the risk of developing diabetes in the extreme group. As mentioned in section 1.2, diabetes is a major health problem in the South Asian community. In the methodology section, the rationale for selecting the food frequency questionnaire (FFQ) as the dietary assessment method is outlined. Since there is no FFQ that has been designed for the South Asian community, time and basic initial work will be put into its development and validation. This is focused in Chapter 2. In the next section, validation studies of the FFQ are reviewed.

1.5 Validation studies of food frequency questionnaire

When using any dietary method to assess the dietary intake, it is recommended that the method selected is validated first before it is used in the study population. The objective of the validation study is to provide information that can help in revising the method if necessary, as well as to check for likely misclassification of an individual based on the responses of the method (Nelson,1991). The term validation (validity) and calibration will be used throughout the thesis. Hence for the purpose of clarification, these terms are defined below.

Validity

Validity can be defined as external and internal validity (Margetts and Nelson,1996). **External validity** can also be defined as generalizability. It can be regarded that any study is externally valid if the results observed are a reasonable representation of the true situation. When assessing the generalizability or the external validity of a study, judgement is required in determining whether the results of the study are applicable to the general population from which the study sample is drawn.

Internal validity is defined as the ability of a method or tool to measure what it is supposed to measure (Block and Hartman,1989; Margetts and Nelson,1995). It has been suggested that in most circumstances, it is impossible to measure the validity of a method absolutely (Margetts,1991). However, it is realistic to measure the relative validity of a measure derived (Margetts,1991). The **relative validity** of the dietary measure can be assessed by comparing it with another method that is considered to be more accurate from previous research (Margetts,1991).

Recently, Margetts and Pietinen (1996) described the important differences between a calibration study and a validation study. Studies which evaluate the relative validity of measures are defined as **calibration studies**. In a calibration study, the parameters which are required to correct measures of association for a given increase of intake are evaluated. In contrast, **validation studies** are studies which take into consideration the total relevant measurement errors involved. This will usually include the relevant time frame, the dietary information used to answer the question being investigated, the study population and how the data are to be presented. The implication here is a calibration study is just a part of the requirement of a validation study. Most validation studies carried out in the past had not considered all the relevant measurement errors. Thus looking at earlier studies which were discussed as validation studies, it would appear now that many of the studies were actually calibration studies.

In the next section, validation studies (really relative validity, but validation will be used as a shorthand to describe these studies) comparing the FFQ with weighed record (section 1.5.1) and FFQ with estimated record (section 1.5.2) in European and American populations are discussed. The only comparative dietary study in the South Asian population is discussed in section 1.6. Generally in the validation studies, nutrient estimates between methods are evaluated for their association (by looking at correlation). The subjects are also checked to see whether they are placed into the same distribution or ranking (correctly classified) or if they are ranked in the opposite section of the distribution (grossly misclassified).

In dietary assessment methods, measurement errors that exist in the measurement should also be analysed. According to Plummer and Clayton (1993), we cannot assume that measurement errors between dietary assessment methods are uncorrelated. This is because the fluctuations of diet over time will induce correlation between errors of different methods at the same time. It is also possible that for certain methods the same errors are repeated, thus there may be correlated errors for the same method on different occasions.

When making comparisons between dietary methods, for example, in a validation study, it is important that the measurement errors in each method are considered. There are now several measurement error models which can be used to assess the quality of the dietary measurements (section 1.4).

However, in earlier validation studies (before 1990), the covariance structure model, the measurement error model and method of triads have not been used to estimate measurement errors. Correlations which determine the strength of associations between methods or assess the goodness of the association (Bellach,1993) as well as the ranking of subjects (agreement) have been used widely instead. In this instance, a minimum acceptable level has to be set up to determine whether the association or the agreement are acceptable.

Most validation studies have used a correlation coefficient to assess the validity and reliability of the test method. Most authors will also conclude that a higher correlation coefficient will indicate a better association between the test and the reference methods. If the correlation is near enough to 1 (mostly > 0.7), then the test method is assumed to be 'good enough' (Bellach,1993). The validation process then stops here.

The term 'good enough' is misleading. Most validation studies have correlations between 0.4 to 0.7. This suggests that less than half of the variability in the test method can be explained by the variability in the reference method (Margetts and Pietinen,1996). Thus, as suggested by Bellach (1993), it is useful if the correlation in a validation study provides low value as this would motivate the researcher to look at the error structure, think about reasons and attempt to improve the test method. If the correlation is close to 1, it would not be helpful as the researcher would assume that this is sufficient to accept the test method as valid.

Thus in general, it is recommended that in investigating the relationship between the test measure with the reference measure, it is more appropriate to examine the results based on the primary objective of the validation study that is whether the aim is to look at the association or the level of agreement between measures. This is because the usefulness of the test measure in determining the usual intake can be evaluated by looking at the association between the test and the reference measures and the extent to which the reference measure approaches the truth (Nelson,1996). Various statistical analyses such as the correlation coefficient, the percentage agreement between categories of intake and the kappa statistics can be used to measure the association or the agreement (ranking) between the test and the reference measures. The best approach is to look at several statistical analyses and observe if the results are stable (Bellach,1993).

1.5.1 Validation studies of FFQ vs weighed record (WR)

There have been a number of studies which have validated the FFQ against the weighed record. Some of these studies are discussed in this section. These studies represent some of the validation studies already carried out in the United States (US) and in Europe. These studies are also summarized in Table 1.7.

The common reason why a weighed record has been used as the standard of comparison for the FFQ is because a weighed record does not rely on memory. It has been regarded as the most accurate and feasible method to evaluate food intake since the quantities of food consumed are based on weighed portions (Willett et al,1985). Since there is day to day variation in intakes of nutrients, some studies have collected the weighed record more than once. The multiple weighed records are also carried out to represent the usual intake of the subjects (Willett et al,1985;Tjonneland et al,1991).

Yarnell et al (1983) validated their FFQ with a 7-day weighed record in 119 Welsh men. Generally the FFQ estimated the nutrients higher than the weighed record, but in contrast in this study the FFQ estimated the nutrients lower than the weighed record. However the correlation coefficient for the nutrients evaluated ranged between 0.27 (total carbohydrate) to 0.75 (alcohol). These correlations were statistically significant. 46% (total carbohydrate) to 70% (alcohol) of the subjects were correctly classified while between 7% (alcohol) and 15% (protein) of the subjects were grossly misclassified. Thus in this study it was concluded that the FFQ could not completely replace a more detailed dietary method in epidemiological studies.

In contrast, studies of Willett et al (1985), Tjonneland et al (1991) as well as Thompson and Margetts (1993) revealed that their FFQs were a useful dietary instrument in categorizing subjects according to nutrient intakes in population based studies. This probably could be explained by the subjects who participated in these studies.

The subjects in the validation study of Willett et al (1985) were nurses. Thus they are more likely to be experienced and trained in keeping quality records. The FFQ was validated against four 7-day weighed record. The correlations between methods ranged between 0.18 (protein) and 0.53 (vitamin C without supplements) which were quite low. When the correlation of the FFQ was checked with two 7-day weighed record and then four 7-day weighed record, there was only a slightly better association. This showed that in this study, a two 7-day weighed record is sufficient to find an association between the methods.

The FFQ and the four 7-day weighed record estimated nutrient intakes comparably, although the distribution as estimated by the FFQ were wider. The degree of misclassification of subjects was small, on average about 3% of subjects were grossly misclassified. 74% and 77% of the subjects were correctly placed in the lowest one or two quintile of the weighed record and the FFQ respectively.

FFQ validation studies have also been carried out among cigarette smokers. One such study was carried out quite recently in 301 (122 men,179 women) smokers (Thompson and

Margetts,1993). In this study the FFQ was validated against a 10-day weighed record. Like the study of Yarnell et al (1983), the FFQ tended to underestimate energy and the macronutrients in women. However, FFQ overestimated the micronutrient in both women and men compared with the weighed record. The mean differences between methods were small for the macronutrients and quite large for the micronutrient. The correlation coefficient between methods for the nutrients were between 0.18 (vitamin A) and 0.83 (alcohol) in men, while 0.31 (vit A) and 0.81 (alcohol) in women.

In the Thompson and Margetts (1993) study , the agreement between methods was also evaluated using Bland-Altman plots (1991). Using this plot, it was revealed that for certain nutrients, the agreement between FFQ and weighed record was not consistent across the whole range of intake for certain nutrients. For example in men, at lower energy intakes the FFQ gave a lower estimate for energy while at higher intakes, it gave a higher estimate than the WR. However, for vitamin C in women, there was a higher estimate by the FFQ at low intakes and lower estimates at higher intakes compared with the WR. Thus the Bland-Altman plot had the ability to assess intakes across the range and hence could compare risk across a range of intakes in different populations of subjects.

The misclassification of subjects was only 2% in both men and women in this study. The FFQ used for smokers in this study compared well with the weighed record and is useful in assessing dietary intakes of smokers.

FFQ has been widely used in epidemiological studies throughout the world. For the last ten years, most FFQs developed have been validated (relatively) before they are finally used in larger epidemiology studies. In Denmark, a 92-food item FFQ was validated against two 7-day weighed

Table 1.7 Summary of validation studies of FFQ against weighed record (WR)

Author	Subjects	Correlation coefficient	Subject classification	Comments
Yarnell et el,1983 7d WR	119 Welsh men	0.27 (total CHO) to 0.75 (alcohol)	*FFQ underestimate nutrients than WR *MC 7%-15% *CC 41%-70% Subjects ranked into tertiles	Study concluded that FFQ could replace a more detailed dietary method
Willett et al,1985 Four 7d WR	173 US nurses 34-59 yrs	0.18(protein) to 0.53(vitamin C)	*FFQ comparable with WR *MC 3% *CC 48% Subjects ranked into quintiles	FFQ would be used to evaluate usual intake in population based study
Tjonneland et al,1991 Two 7d WR	144 Dutch subjects 59 men and 85 women 40-64 yrs	0.17(vitamin A) to 0.64 (vitamin C)	*FFQ underestimate than WR *MC 4-7% *CC 70% Subjects ranked into quintiles	FFQ could replace a more detailed dietary method to evaluate dietary intake
Thompson & Margetts, 1993 10d WR	301 UK subjects 179 women 122 men 40-59 yr	0.18(vitamin A) to 0.83 alcohol) in men); 0.31 vitamin A) to 0.81 (alcohol) in women	*FFQ underestimate macronutrient but overestimate micronutrient than WR *MC 2% Subjects ranked into quintiles	FFQ could be used to evaluate usual intake in smokers

MC-misclassification CC-correctly classified record in 144 subjects aged between 40-64 years old (Tjonneland et al,1991). In this validation study, the weighed record estimated nutrients higher than the FFQ in both men and women. This was in agreement with the study in the Welsh men (Yarnell et al,1983). However the mean difference between methods ranged from 1.8% to 23%. Correlation coefficient between methods ranged between 0.17 (vitamin A) to 0.64 (vitamin C).

The agreement in ranking of the subjects was considered good in that 70% of the subjects were correctly classified within plus or minus one quintile. This was based on the assumptions that if the subjects were randomly distributed, about 36% would be correctly classified. Only 4% and 7% of the subjects were grossly misclassified in the highest and lowest fifth respectively. As in the study of the US nurses (Willett et al,1985) and the smokers (Thompson and Margetts,1993), this study concluded that the FFQ was a reasonable instrument to estimate nutrient intakes of the individuals and to rank them accordingly.

Summary critique of the FFQ validation studies with WR

The validation studies of the FFQ against the weighed record have shown that FFQ can be a useful instrument and especially practical to evaluate dietary intake in studies of large populations. This is because the relative validity of the FFQ has been calibrated with a dietary method which has been regarded as a 'gold' standard. However, the FFQ was inconsistent in estimating nutrients because in some studies, FFQ underestimated compared with weighed record (Yarnell et al,1983;Tjonneland et al,1991). In others, the FFQ underestimated macronutrients and overestimated micronutrient compared with weighed record (Thompson and Margetts,1993), while one showed comparable estimates with the weighed record (Willett et al,1985). The mean differences in some of these studies were quite small.

Furthermore the FFQ was able to rank the subjects well, with misclassification of subjects in most studies ranging between 2% to 7%, with the exception of the Welsh study (15%). Thus it is not surprising that the Welsh study rejected the FFQ as a useful replacement instrument for the weighed record.

It is observed that in some of these studies, there were more women subjects than men. This in a way reflected that women were more willing to keep weighed records than men. This is confirmed by a study of Jorgensen (1992) which showed that women and also people in the older age groups were more willing to participate in food record studies.

The correlation between the FFQ and the weighed record in all these studies ranged between 0.17 to 0.83, however most nutrients have correlations between 0.3 to 0.6 between methods. Thus it would be expected that the correlation of most nutrients between FFQ and other dietary methods would fall around these correlations or even better. The Bland-Altman technique has not been widely used in validation studies. The advantage of the Bland-Altman plots is that they can evaluate the agreement between dietary methods more objectively. Any differential misclassification in the estimate of intake can be detected as the plots can determine the agreement across the whole range of intakes.

1.5.2 Validation studies of FFQ vs food record (FR)

Another dietary assessment method which has been regularly used as a method of comparison with the FFQ is the estimated record or the food record (FR). Again there have been numerous validation studies of this kind. A few of these are discussed here. These studies represent FFQ validation studies with FR that have been carried out in various countries. FR has been selected as the reference method because the errors are assumed to be independent to the errors in FFQ (Willett,1990). Goldbohm et al (1994) also suggested that in weighing their foods, untrained subjects are liable to making mistakes. The summary of these studies are tabulated in Table 1.8.

The number of days food records are kept varied between as little as one day to 16 days or more. Margetts et al,(1989) compared their FFQ with the one day FR in 433 men and women between the ages of 35 to 54 years in three English towns. The correlation between nutrients was between 0.15 (vit A) and 0.36 (energy), which was as good as some of the correlations in the FFQ and WR (Table 1.7). The percent of subjects grossly misclassified ranged between 1% to 5%. Considering that a one-day FR is a simple evaluation of current intake while the FFQ evaluates usual intake, it would be expected that the agreement between methods would be low. However the large number of subjects involved in this study compensated for any extremes high or low reporters in the one day FR, thus not greatly affecting the average.

The study of Margetts et al (1989) also showed that with a reasonably large number of subjects, a simple method as a one-day FR can be the reference method. Further more, the power of detecting dietary difference between subjects increase as the number of subjects gets bigger (Cade et al,1988). More importantly, the FFQ validated in this study was able to rank individuals by nutrient intakes with a fairly low gross misclassification.

One of the more recent FFQ validation study carried out in the Netherlands used 3-three day FR as their reference method (Goldbohm et al,1994). This study was a subgroup of a cohort study of diet and cancer. A total of 59 men and 50 women aged 55 to 69 years participated in this validation study. The objective of the three separate 3-day FR was to reflect the varying consumption pattern due to seasonal

Table 1.8 Summary of validation studies of FFQ against food record (FR)

Author	Subjects	Correlation coefficient	Subject classification	Comments
Margetts et al (1989)	433 UK men and women 35-54 years old	0.15 (vit A) to 0.36 (energy)	MC 1%-5% Subjects ranked into quintiles	FFQ could be used instead of a diet record to estimate intakes in population based study
Goldbohm et al (1994)	59 Dutch men and 50 women 55-69 years old	0.49 (vit B1) to 0.86 (alcohol) correlation mostly between 0.6-0.8	FFQ estimated nutrients lower than FR. FFQ was able to rank subjects according to intake of food groups and nutrients. Subjects ranked into quintiles	FFQ could replace the FR as the FFQ could rank subjects accordingly
Rothernberg (1994)	76 Swedish subjects 42 women and 34 men 70 years old	0.35 (total sugar) to 0.60 (energy and fat)	FFQ estimated energy and nutrient higher than FR CC 49% to 68% MC 4% to 13% Subjects ranked into tertiles	FFQ reflected the habitual intake of elderly in Swedes after confirmation with EI/BMR ratio.
Flegal et al (1988)	228 US subjects 64 white men 43 black men 73 white women 48 black women 24-51 years old		FFQ estimated nutrient intakes higher than FR Subjects ranked into tertiles (Kappa statistics)	Relative validity of FFQ can be increased with improved accuracy of frequency

MC-misclassification
EI-energy intake
BMR-basal metabolic rate

differences. A learning effect of the recording of intake was also investigated.

In this study, it was shown that the FFQ tended to underestimate nutrient intake compared with FR. However the unadjusted correlation coefficient for nutrients ranged between 0.49 (vitamin B1) to 0.86 (alcohol), with a median of 0.69. The Spearman correlation was between 0.38 (vegetables) to 0.89 (alcoholic beverages) for food groups. This study also revealed that despite subjects doing FR for three separate occasions, there was no learning effect observed. The FFQ was also able to rank the subjects according to intake of food groups and nutrients.

In another study in Sweden, the FFQ was validated against a 4-day FR on 76 subjects (42 women,34 men), with an average age of 70 years (Rothenberg,1994). In contrast to the Dutch study, this FFQ estimated higher energy and nutrient than the FR. The mean value and the mean difference between methods had a linear correlation, the mean difference became greater as the mean value became larger. The correlation between methods for nutrients ranged between 0.35 (total sugar) to 0.60 (energy and fat). The ranking of subjects between methods was satisfactory. 49% to 68% of the subjects were correctly classified into the same tertile while between 4% to 13% were grossly misclassified.

This study also validated the FFQ against a biomarker, urinary nitrogen as well as, energy intake/basal metabolic rate ratio (EI/BMR). This made the FFQ validation study more important , with more standards to compare with for its relative validity. Although the FFQ overestimated the intake of nutrients compared with the FR, the EI/BMR ratio of the FFQ seemed to estimate energy intake more reasonably than FR when compared with the doubly labelled water cut off limit. The doubly labelled water method has been used as an objective way of validating reported energy intake, as a measure of habitual intake (Livingstone et al,1990; Black et al,1991). Therefore, if the reported energy intake was lower than the cut off limit, this implied that the reported intake did not reflect the habitual intake. In this study, FR tended to underestimate usual or habitual intake. Thus this study concluded that the FFQ would be useful in estimating the habitual intake of their elderly subjects (Rothenberg,1994).

Flegal et al (1988) in the USA validated their FFQ with 16 days food record as the reference standard. The subjects were 228 black and white men and women aged 24 to 51 years. The food records were taken on four occasions over a period of a year. The reason for this distribution was to include seasonal variation as in the study of Goldbohm et al (1994).

The results showed that the FFQ estimated nutrient intakes higher than the FR in all the four race-sex subgroups. However the mean difference between methods were higher in black men and women compared to their white counterparts. The correlation between methods were also low especially in black subgroups. This lower correlation could possibly be explained by the higher percentage of black men and women who were underreporting their intake in the FR. Thus the agreement in classification by tertile was also poor.

In this study, the authors investigated the effect of certain types of discrepancy on the estimates of energy intake between FFQ and FR. It was revealed that the discrepancies that most affected the correlation between the methods were the frequency and the serving size portion estimates. However the correlation between methods was especially reduced due to the difference between person in estimating frequency.

This study thus showed that despite more days of recording food intake, the correlation between the FFQ and FR was still low and the ranking of subjects remain poor. Greater emphasis in a better estimation of frequency of food and serving size in the FFQ are required to improve the validity between methods. This study also demonstrated that subjects of different races responded differently to the dietary method used in the assessment.

Summary critique of FFQ validation studies with FR

In summary, the FR method in general is still an acceptable reference method to choose for validating FFQ. However, the ability of the subjects to give a more accurate estimate in the frequency and portion size for the FFQ and record their usual food intake will improve the relative validity between methods. The study of Margetts et al (1989) demonstrated that even with a one-day food record, a significant correlation between methods could be achieved as long as this is compensated with a larger recruitment number to increase the power of the study. On the other hand, Flegal et al (1988) was not able to suggest that their FFQ could replace the FR due to the discrepancies between methods despite using a reference method that could reflect a more usual intake.

In the last few years, it has become a common practice in validation studies to include a biomarker to evaluate the relative validity of the test method. The comparison with a biomarker was carried out in the study of Rothenberg (1994). In that study, although the FR and FFQ had a large variation in the nutrient estimates based upon the EI/BMR ratio, they concluded that the FFQ gave a more realistic estimate of energy intake than the FR, thus leading them to suggest that the FFQ could be used to replace the FR.

In validation studies, low correlation coefficients are usually regarded as a lack of association between methods studied. According to Delcourt et al (1994), the type of population studied also has an impact on the correlation coefficient. Generally a heterogeneous population will be a better choice than a homogeneous one because they will have a wider range of energy intakes and a greater variability of diet composition. Very low response rates in a validation study will also cause a low range of intake.

However, according to Bellach (1993), a low correlation between dietary measures could be useful. This is because this will only encourage the researcher to evaluate the reasons behind the poor association with more detail. In contrast, a high correlation although always interpreted as indicating good association, can be misleading as the test measures are often assumed to be valid, when it might not neccesarily be such. Several statistical analyses which could provide consistent results between test and reference measures should be the appropriate way to draw conclusions regarding the relative validity of the test measures.

1.6 Comparative dietary method study in the South Asian community

There have been a lot of validation studies already carried out in Caucasian communities. Since the 1960's, there has been an influx of immigrants from the Indian subcontinent into the United Kingdom (Hunt,1977). They brought their food habits and cultural practices. It has been speculated that the diet of the South Asians might have a role in the higher prevalence of diabetes and coronary heart disease (McKeigue et al, 1991). However, results of dietary evaluation studies in this community have been inconsistent (Section 1.3.2). Since the early 1990's, there has been only one study which compared dietary methods to evaluate dietary intake in the South Asian community. This study is discussed here.

The study was carried out in Sheffield among South Asian children between 4 to 40 months of age. Weighed dietary inventories (using the Petra scale) were compared with a dietary history. One hundred and seventeen children participated in the study (Harbottle and Duggan,1993).

The results showed that the dietary history gave a higher estimate of energy and protein than the weighed inventory. The difference in estimates between methods was significant in the 12 to 18 months age group. Weighed inventory and dietary history estimates between well and unwell children were not significantly different but the weighed inventory estimates of energy intake of well children was significantly higher than unwell children. On the other hand, dietary history estimated energy intake similarly in well and unwell children. This thus suggested that the weighed inventory was a more sensitive indicator of short term intake and the effects of ill health. The dietary history method did not show any difference in energy intake between well and unwell children, maybe because while giving the dietary history the mothers tended to recall healthy intakes rather than when the children were unwell.

Harbottle and Duggan (1993) suggested that both weighed inventory and dietary history could be used to estimate energy intake within the acceptable range according to Prentice et al (1988) in these children. However, dietary history had a better compliance than weighed inventory for this South Asian population. Most South Asian mothers lacked the familiarity with the weighing technique required in the weighed inventory. Demands from other siblings in the household made the weighed inventory less acceptable to these mothers. This thus implied that the weighed inventory method might not be a suitable dietary measure to be used in South Asians especially for women with young children. The dietary history appeared to be a better choice, however the subjects demonstrated a tendency to forget intakes of food in the distant past, thus possibly introducing a recall bias.

Since this was the only study comparing dietary methods used for the South Asian community, no assessment could be made regarding the relative validity of the methods used. This however, highlighted the need for a proper calibration study for dietary assessments methods that will be used to evaluate dietary intakes of South Asians in future dietary studies.

1.6.1 Comparative dietary methods in epidemiological studies

Besides validating a dietary method with another reference method, there have been a few studies which compared various dietary assessment for epidemiological research. The summary of these studies are shown in Table 1.9.

One such study was carried out recently (Bingham et al,1994). The objective of the comparative study was to find a method which was less troublesome for subjects and could replace the weighed record in the assessment of usual diet in a study of diet and cancer. Several dietary assessment methods such as the 24 hour recall, FFQ, food checklist and estimated diet records were compared for their accuracy with the 16-day weighed record. A total of 160 women aged 50-65 years participated in this study.

The results showed that the best dietary method which could estimate nearly as well as the WR was the 7-day estimated record. The correlation coefficient between estimated record and WR was the highest (0.46 to 0.88) and the estimated record was able to classify a greater proportion of individual values into the correct fourth of the distribution (40% to 70%).

Table 1.9 Summary of comparative dietary method in epidemiology studies

Author	Subjects	Correlation coefficient	Subject classification	Comments
Bingham et al (1994)	160 UK women 50-65 years old	WR vs FR (0.46-0.88) WR vs FFQ (0.13-0.90)	CC 40% to 70% CC 30% to 50% MC overall 0% to 10% Subjects ranked into quintiles	FR estimated nearly as well as the WR
Morgan et al (1978)	400 Canadian women	Correlation between methods (DH vs 24 hr recall vs 4d FR) were high	No ranking of subjects	DH gave the highest nutrient estimates, followed by 4d FR and 24 hr recall
Stuff et al (1983)	40 US lactating women	FFF vs 7DR (0.00-0.24) 1DR vs 7DR (0.42-0.63) 3DR vs 7DR (0.74-0.91)	FFF gave the highest estimate. Best agrreement between 3DR and 7DR CC 63% in 1DR vs 7DR CC 78%-90% in 3DR vs 7DR CC <50% in FFF vs	Short term dietary method correlated better with another short term method
Sorenson et al (1985)	50 US subjects 37 women 17 men 25-80 years old	24 hr recall correlated better with diet diary. DH correlated better with FFQ	FFQ gave the highest mean estimates Subjects ranked into tertiles	Dietary methods evaluating the same time span correlated better than if evaluating opposite time span

WR-weighed record
FR-food record
FFQ-food frequency questionnaire
1DR-1d record
7DR-7d record

CC-correctly classified
MC-misclassification
FFF-food frequency form
3DR-3d record
DH-dietary history

The two FFQ used in this study did not agree well with the WR in that the FFQ significantly overestimated some nutrients compared with the WR, with Spearman correlation coefficients ranging between 0.13 to 0.89 (Cambridge FFQ) and 0.39 to 0.90 (Oxford FFQ) respectively. However in general, the percent of subjects correctly classified ranged between 30% to 50% for most nutrients while gross misclassification were between 0% to 10%. Again the problem in the FFQ as in the study of Flegal et al (1988) was the overestimation of frequency of consumption. The choice of portion size to be included into the FFQ of this study was also overestimated.

Studies comparing dietary methods had already been carried out in the late 1970's. Morgan et al (1978) compared 24 hour recall, dietary history and 4-day estimated record. This was carried out in a case control study of diet and breast cancer where 400 healthy subjects participated. In order to compare association between methods, only the correlation coefficient was investigated. There was no evaluation of classifying individuals into distribution of values. Dietary history gave the highest nutrient estimate compared with the 24 hour recall and the 4-day estimated record. The correlation coefficient for various nutrients were similar and highly correlated.

In the 1980's, studies comparing different dietary methods continued to be evaluated. Stuff et al (1983) compared a 1-day record, 3-day record and food frequency form (FFF) against a 7-day record. The 7-day record was selected as the standard because it was thought to be the best compromise between obtaining accurate information with minimal imposition on the subjects lifestyles. In measuring the agreement between methods by classifying the intakes into three categories, the kappa statistics was used.

The results of this comparative study showed that the FFF gave the higher mean energy and nutrient estimate compared with the diet records. The 3-day record indicated the best agreement with the 7-day record and in contrast the FFF showed a poor agreement with the 7-day record. Classification of intakes into categories demonstrated that the 1-day record could classify 63% of the individual correctly and the percent of subjects correctly classified increased to 78% to 90% between 3-day record and 7-day record. The percent correctly classified between the FFF and the 7-day record was less than 50%. This was reflected in the low correlation between methods (0.08-0.20).

The objective of the FFF was to simplify the collection and analysis of nutrient intake data, as well as to be administered quickly to large number of subjects. However, the correlation between FFF and the standard method which was the 7-day record, was poor. It was suggested that the poor association could be the result of an overestimation of food intake by the subjects (lactating mothers) and conversely an underestimation in the 7-day record.

Sorenson et al (1985) compared four dietary methods in evaluating nutrient intake in 50 subjects. Two methods measured long term intake (modified dietary history and FFQ), in contrast the other two methods measured short term intake (2-day diet diary and 24 hour recall). The objective of the study was to compare the methods for determining nutrient intake from estimates of food intake.

The results of this study showed that the FFQ gave higher mean estimates than the other three methods. This result was in agreement with the study of the lactating mothers (Stuff et al,1983) which also had higher estimates by the FFQ. The mean estimates of the dietary history although slightly higher were closer to the 24 hour recall and diet diary. Again in comparing the agreement between methods, the kappa statistics was used to rank the individuals into thirds of intake. The agreement between methods was reasonably good if the methods measured the same duration of intake. The two short term methods and the two long term methods correlated better with each other than comparison between combination of short term and long term method.

Although the dietary history was a method which evaluated usual intake, the result implied that the mean estimates of the dietary history compared better with the short term intake method. This suggested that the three method excluding the FFQ probably estimated actual intake better than the FFQ. Again it was indicated that there might be an overestimation of amount and frequency of usual intake by the subjects when completing the FFQ.

Summary critique of comparative dietary methods in epidemiological studies

In summary, studies comparing various dietary methods in determining nutrient intake have been ongoing for the past 20 years. Although earlier studies only looked at the correlation between methods to evaluate association between them, later studies also determined the ranking (agreement) of individuals into the distribution of intakes. This strengthened the outcome of the result as more than one statistical evaluation which could provide consistent results could confirm any association or agreement between dietary methods.

The objective of the FFQ was to evaluate nutrient intake in a larger group of individuals, at the same time it could be administered quickly and was less imposing on the subjects. However even in the earlier studies (Stuff et al,1983; Sorenson et al,1985) and in the more recent study (Bingham et al,1994), the FFQ tended to give higher mean estimates than other dietary methods. The problem with the higher estimate of the FFQ in all these studies was mostly narrowed down to the probable overestimation of the frequency of intake as well as the overestimation of portion size. However if the aim of the dietary measurement is to be able to rank the individuals into the thirds of distribution of intake, then the FFQ can sufficiently do this. Thus if the correlation and the agreement of the FFQ with another reference method produced a consistent result, FFQ could be used in large epidemiological studies (section 1.5.1 and 1.5.2).

It has been suggested that dietary methods which evaluated the same reference time should agree better than if each determined different duration. This was demonstrated in the agreement between methods in the study of Stuff et al (1983) whereby there was good agreement between the 3-day record and 7-day record (short term intake) while poor agreement between FFF and 7-day record (long term and short term intake). The more recent comparative study also showed that the best agreement was between the weighed record and 7-day estimated record, as both methods evaluated short term intake (Bingham et al,1994). An earlier study also indicated the same trend (Morgan et al,1978).

In future validation studies and any comparative studies of dietary methods, other reference method besides a dietary reference method should be considered. More studies are utilizing biomarkers such as 24 hour urinary nitrogen (Rothenberg,1994; Bingham et al,1995). This could provide a check on reported protein intake. Other studies have used the energy intake/BMR (EI/BMR) ratio as a check on reported energy intake (Black et al,1991;Rothenberg,1994). These studies are discussed in section 1.7.1.

1.7 Assumptions of using a biomarker

The use of biomarkers to validate dietary methods have been carried out for many years now. In comparing the dietary assessment method with a biomarker, it would be possible to evaluate the quality of the dietary measurement used. An important requirement of the biomarker is that it is sensitive to the dietary intake (van't Veer et al,1993). But in using a biomarker, for example, the 24 hour urine nitrogen excretion, some assumptions are made of the biomarker, usually depending on the purpose of the study. Biomarkers can act as an estimate of a dietary intake, as an estimate of the amount present in the body tissues as well as an estimate of the amount available in the tissues (Bates et al,1991).

In the present study, the objective was to develop methods to evaluate the dietary intake of the South Asian community. Furthermore only the ranking of the nutrient intake of the subjects into the thirds of the distribution was of interest, and not the absolute intake. Thus the assumption made of the biomarker here was that it acted as a proxy of the measurements in the diet that were being evaluated. The nutrients of interest in this study included energy, fat, carbohydrate and protein intake.

One of the most widely used biological marker is the 24-hour urine nitrogen (Bates et al,1991;Bingham,1994). It has been suggested that the 24 hour urine nitrogen can estimate the contribution of protein to total energy intake (Isaksson, 1980). However, there are still no quantitative markers for carbohydrate and fat consumption, thus the estimate of total energy intake cannot be obtained using a biomarker. Energy expenditure, however, can be measured with the doubly labelled water technique, but it is too expensive to be used in an epidemiological study. A proxy for energy expenditure is to calculate the EI/BMR ratio of individuals. This acts as a simple check on the relative validity of energy intake (Bingham,1994). We can also use the average daily energy requirement of adults as multiple of BMR to check the relative validity of energy intake

(FAO/WHO/UNU,1985). The BMR of the subjects can be calculated using the BMR equation of Schofield et al (1985) (section 1.7.2 for more details). This equation however requires information on at least the age and weight of the subjects.

Most validation studies of dietary assessment methods using the 24 hour urine nitrogen as a biomarker have used it a proxy of protein intake. These validation studies are discussed in the next section.

1.7.1 Validation studies using biomarkers

Van Staveren et al (1985) validated their dietary history method with the 24-hour urine nitrogen excretion. In their study, it was demonstrated that the mean difference between nitrogen excretion and nitrogen intake was zero for the subjects (22 men,22 women) studied. This led them to suggest that the dietary history estimated the protein intake well, as there was a good agreement between the nitrogen excretion and intake. However, the 95% confidence interval of the mean difference was rather wide for men compared to women.

In another study , this time in Sweden, Rothenberg (1994) validated the FFQ and 4-day record with the 24-hour urinary nitrogen. The FFQ used in this study estimated a higher protein intake compared with urinary nitrogen and the correlation was not statistically significant (r=0.19). The food record however, had a better correlation (r=0.47) in estimating protein intake with the urinary nitrogen. The gross misclassification of FFQ and food record versus urinary nitrogen were 18% and 11% respectively.

Recently, Bingham et al (1995) validated weighed records and other methods of dietary assessments using the 24-hour urine nitrogen. A total of 156 women completed the 16 day weighed records and collected eight 24-hour urine samples. Since the distribution of urinary N: diet N ratio was not normal, the results were sorted and examined as fifths of the distribution of urinary N:diet N ratio.

Other dietary assessment methods validated with the 24-hour urine nitrogen besides the weighed record were the FFQs (Oxford and Cambridge), 24 hour recall, 7 day checklist (with portion and without portion) and the estimated food record. The correlation between weighed record and the 24 hour urine nitrogen was the highest (r=0.78-0.87), followed by the estimated food record (r=0.60-0.70). The correlation between the FFQs, 24 hour recall and food checklist with the urinary nitrogen were lower, which were between 0.01 to 0.5.

In another also recent study, Heitmann and Lissner (1995) used the 24-hour urine nitrogen to validate their dietary history. A total of 323 subjects completed both the dietary history and 24-hour urine collection. The results of this study indicated that the estimates of total energy intake and total protein from the dietary history were lower than those from the urinary analysis. Overall 85% and 72% of the subjects underreported energy and protein respectively.

Summary critique of validation studies using biomarkers

In summary, the studies validating dietary methods with the 24-hour urine nitrogen had correlations in the range of 0.1 to 0.8, which was quite similar to the studies validating dietary methods with another dietary reference method (Table 1.7 and Table 1.8). However, the correlation was high between the weighed record and urinary nitrogen as well as between food record and urinary nitrogen. The urinary nitrogen has been considered to be free from individual bias (Plummer and Clayton,1993). Thus this implied that if a dietary method has a good association with the urinary nitrogen, it would be likely that the quality of the dietary measurements is good. On the other hand, the correlation between the FFQ, checklist and 24 hour recall with the urinary nitrogen is low, indicating a poor association, thus demonstrating a lower quality dietary measurement.

When using urine collection to validate a dietary assessment method, it is important that the 24-hour urine collections are complete. A high degree of commitment is required in the subjects when collecting their urine samples. Thus to ensure that the urine collections are complete, a marker, usually para-amino benzoic acid (PABA) tablets are given to be taken with a meal to the subjects. All the studies above have used the PABA check as a marker for completeness of the 24-hour urine collections.

The results of the validation studies discussed above have shown that the weighed record had the best correlation in estimating protein intake with the urinary nitrogen (Bingham et al,1995). However, the study of van Staveren et al (1985) also demonstrated that the dietary history had a very good agreement with the urinary nitrogen in estimating protein intake. Unfortunately in the van Staveren et al study (1985), the number of subjects who participated was too small to draw firm conclusions.

In contrast Heitmann and Lissner (1995) showed that their dietary history underestimated both the energy and protein intake compared with the urinary nitrogen. Rothenberg (1994) on the other hand indicated that the food record had a better correlation than the FFQ in estimating protein intake with the urinary nitrogen. However, it could be argued that the higher correlation between food record and urinary nitrogen than FFQ with the urinary nitrogen could partly be explained by the same time reference for the collection of urine and the food record. The FFQ in contrast reflected usual intake.

In general the validation studies using the 24 hour urine nitrogen excretion to validate protein intake showed inconsistent results. However, validation studies which did not use biomarkers can pose more problems when interpreting the results. This is because it is more likely that the errors in two dietary methods are correlated (van't Veer et al,1993). Biomarkers on the other hand can provide a fair reference for comparisons of two (or more) other methods of dietary assessment, thus increasing the possibility for a useful comparison.

Despite the 24-hour urinary nitrogen being considered as an accurate biological marker (Isaksson,1980), in certain occasions it is not possible to collect the 24 hour urine samples especially in certain groups of population. Thus the 24-hour urine nitrogen determination might not be practical. In instances like this, an alternative method that can be used to check on the validity of energy intake is the evaluation of energy intake/basal metabolic rate ratio (EI/BMR).

Thus far there is no study yet which has validated a dietary assessment method with a biomarker in the South Asian community. This is not surprising as there was only one study which compared dietary methods in that community (section 1.6). Some problems which were faced while working with the South Asian subjects were also discussed in section 1.6. Other difficulties encountered with these subjects in the present study are highlighted in the methodology (Chapter 2). In the next section, the discussion of the BMR equations that are available are explored.

1.7.2 Predictive equations for basal metabolic rate (BMR)

The report on energy and protein requirement by FAO/WHO/UNU have suggested the use of energy expenditure rather than energy intake as the basis for estimating energy needs in man (WHO,1985). This suggestion has emphasized the importance of using the appropriate BMR equation for the proper population group to estimate the energy intake. Since 1985, BMR equations that have been used to predict energy requirement have been based on the equation developed by Schofield, Schofield and James (1985) which predicted BMR for both sexes between the ages of 0 to 30 years to greater than 60 years. This Schofield equation has also been the basis for the equation used by FAO/WHO/UNU report (WHO,1985). The database was based on 11,000 BMR measurements, but most values were taken from European and North American subjects.

Since the late 1980's, a few studies have been carried out in generating a predictive equation for BMR of Indian population and other population living in the tropics (Soares and Shetty,1988;Henry and Rees,1991;Soares, Francis and Shetty,1993; Hayter and Henry,1994). These studies have shown that the Schofield et al (1985) equations were not valid for Indian population as well as for other populations in the tropics. It was demonstrated that the BMR values of these Indian and tropical population were lower than those predicted (Henry and Rees,1988;Henry and Rees,1991;Soares,Francis and Shetty,1993).

In the review of BMR studies conducted in the tropics, Henry and Rees (1991) demonstrated that the actual BMR of people living in the tropics was significantly lower than that predicted by the Schofield equation, by an average of 8% . The percent of estimation by the Schofield equation was higher in men than in women. It was suggested that the lower BMR in the tropical population could be due to a combination of factors such as climate, diet, ethnic background and body composition.

More recently, Soares et al (1993) generated a predictive equation for BMR of well nourished Indian men. The subjects were well nourished Indian males between 18-60 years old with BMR less than 25, with access to ad libitum energy and protein intake. Two age-specific equations for Indian males that related to BMR and body weight were generated. The age band were 18-30 years and 30-60 years. When the BMR predicted by equations of Schofield et al (1985) and that of Soares et al (1993) were compared, the BMR predicted by the Schofield equation consistently overestimated the BMRs of Asian Indians.

The Soares et al (1993) BMR equation was also compared with those of Henry and Rees (1991) to predict the BMR of Asian Indians and Americans. The results demonstrated that both the equations correctly predicted the BMR for Asians and Americans. This thus suggested that the Indians and other tropical populations and present day Americans had similar BMRs when matched for gender,age and body weight. These suggestions were in contrast to those views which assumed that the Asian Indians and peoples in the tropics have lower BMR than their matched counterparts (Schofield et al, 1985;Henry and Rees,1991).

The implication brought forward by the new predictive equation generated by Henry and Rees (1991) as well as Soares et al (1993) were that the equation of Schofield which was derived over 16 years ago might not be valid for the prediction of BMR of population groups worldwide (Soares et al,1993). This was further supported by the work of Hayter and Henry (1994) which showed that 45% of the database used to derive the Schofield equation comprised of Italians who were primarily young males leading a physically active lifestyle (miners,labourers, military servicemen and students of military college). This apparently increased the BMR of the Italians. Due to the dominance of the total number of Italian subjects involved in predicting the Schofield et al (1985) equation, there appeared to be a bias in the predictive equation.

CHAPTER TWO

METHODOLOGY

CHAPTER 2 METHODOLOGY

This chapter consists of five main sections. The first section (2.1) discusses the objectives of the study. The second covers the development of the food frequency questionnaire (FFQ), (2.2) which explains the sequence of steps undertaken in the design of the FFQ. The third section (2.3) which forms the major discussion in the chapter focuses on the calibration study. This includes the dietary method used (2.3.1), the calibration of test methods with the BMR multiples (2.3.2), the subjects involved (2.3.3), and some problems faced during the calibration study (2.3.4 and 2.3.5). The chapter concludes with the discussion of the role of chance, bias and effects of other variables on the study design (2.4), analysis methods used for the dietary data (2.5), including statistics (2.5.1), adjustments of FFQ (2.5.2), analysis of adjusted FFQ (2.5.3), and analysis of measurement errors with the method of triads (2.5.4).

2.1 Objectives of study

There are general and specific objectives for this study as outlined below.

2.1.1 General objective

The general objective of this study is to develop dietary assessment methods which can be used to evaluate the dietary intake of the South Asian community.

2.1.2 Specific objectives

There are two specific objectives in this study and they are as follows:

1. To develop and calibrate a food frequency questionnaire (FFQ) to be used in the dietary evaluation of the South

Asian community.

2. To develop and calibrate a food checklist as an alternative method of the FFQ in the dietary evaluation of the South Asian community.

2.2 Development of the Food Frequency Questionnaire (FFQ)

There are several dietary assessment methods that can be used to evaluate the dietary intake of a group or individual. These methods have been reviewed in section 1.4.5. The method used in any particular study is selected because it is expected that the method can correctly measure the dietary exposure with the level of accuracy required for the purpose intended. However in this study, we are interested to investigate the usual intake of the South Asian community and be able to rank the macronutrient intakes accordingly into thirds of distribution.

Similarly, several dietary assessment methods can be used to evaluate the dietary intakes of the South Asian community. The weighed record has been regarded as one of the better methods (Marr,1971). On the other hand the estimated record has also been suggested to be as reliable as the weighed record in evaluating food intake (Bingham et al,1994). Other alternative methods include 24 hour recall, dietary history and the questionnaire. Each of these methods have their own advantages and disadvantages (Table 1.6).

In the past, studies evaluating the nutrient intakes of the South Asians have used different dietary assessment methods. The methods used so far were household inventory, weighed record, diet diary , FFQ and dietary history. Thus it was not surprising that the results generated in each study were different, and it was difficult to make any comparisons between studies. The majority of the results shown in these studies were reflecting the absolute intakes of the South Asians studied. The FFQ used in some of the studies were not designed for the South Asian community, thus it might not be appropriate to assume that the nutrient intake actually reflected those of the South Asians. Due to the limited dietary intake studies in the South Asian community, and with different methodology used, it was difficult to suggest which study demonstrated their usual dietary intake.

Dietary assessment methods such as the weighed record, food diary and 24 hour recall can determine the current intake. Since the objective of this study is to determine the usual intake of the community, the two dietary methods that have the ability to evaluate usual intake are the dietary history and the FFQ. However, the disadvantages of using the dietary history are that it is time consuming and requires a skilled interviewer. Thus after considering the relevant factors and the objective of our study, the choice of the dietary assessment method is narrowed down to the FFQ as it is cheap, easy to administer and can still evaluate usual intake. For the development of the FFQ several steps were planned and undertaken. This is shown in the flow chart in Figure 2.1.

In evaluating the dietary intake of a population, it is not possible to measure the intake absolutely (true intake). Thus the dietary method selected to measure the dietary exposure is normally compared with another method that is considered to be more accurate from past research. This is referred to as relative validity (Margetts,1991). In this study, the relative validity of the dietary measure (food frequency questionnaire (FFQ)) which is the test method is compared with the 4-day weighed record which is the reference method). In recent years, it has been suggested that the measurement errors of dietary methods are correlated. Thus in evaluating the relative validity of the test method, we also have to include a proxy for energy

Figure 2.1 Flow chart of the steps considered in the development of the FFQ

Preliminary study
(24 hour recall to assess current meal pattern and investigate the macronutrient sources in the South Asian diet)

↓

develop FFQ
(75 food item)

↓

calibrate FFQ
(4-day weighed record,)
(BMR multiples)

↓

analysis of FFQ with WR

↓

result of the calibrated FFQ

↓

adjustment made on FFQ

↓

analysis of adjusted FFQ

↓

result of adjusted FFQ

intake which is the BMR multiples as one of the reference measures for our calibration study. The external validity or the generalizability, that is whether the findings can be taken as a reasonable representation of the true situation, of this study is considered by imposing selection criteria for the study samples.

2.2.1 Preliminary Dietary Survey

In order to develop the FFQ, the first important consideration was to establish an idea on the meal patterns of the South Asian community. Thus a preliminary dietary survey was carried out. The objectives of the survey were to determine the current meal patterns of the South Asians as well as to evaluate the energy and macronutrient sources in the South Asian diet.

2.2.2 Subjects and Method

This survey was carried out at a General Practice clinic in Southampton. This practice was selected because 95% of the patients registered with the practice were of South Asian origin. Permission to recruit the prospective subjects was obtained from the Practice Manager. The criteria for selection of the prospective subjects were that they be South Asians, were literate and consented to be interviewed. The methodology chosen to evaluate the meal pattern was the 24 hour recall. The recall method was used because the main purpose of the survey was to evaluate the current meal patterns as well as the energy and macro-nutrient sources in South Asian foods. The 24 hour recall method was appropriate for measuring current diet in groups of subjects (Bingham and Nelson,1991). This method also has the advantage of being simple and it allows a larger number of subjects to be recruited with minimum resources. Since the recall method also requires less effort from each subject, excellent compliance can be achieved with this method (Bingham and Nelson,1991). More importantly, at this point of the research process, we were only interested in establishing the general meal pattern and not the absolute intake of the nutrients consumed by the subjects. Thus the recall method was sufficient to evaluate this current meal pattern.

The survey was carried out for three days a week for eight weeks. There was no target number of subjects calculated for this survey. The main aim was to recruit as many subjects as possible during this survey period. There were some limitations here, in that only literate South Asians would be the majority recruited here. However, we were also able to recruit some housewives and unemployed South Asians (Chapter 3). In this survey, 66 South Asian patients attending the clinic agreed to participate, although only 62 subjects completed the recall interview. The results generated in this survey are discussed in the next chapter. However, the general guideline considered in developing the FFQ in this study was taken from the preliminary dietary survey results.

2.2.3 Objective of the FFQ

The food frequency questionnaire (FFQ) is the preferred method to evaluate longterm usual dietary intake in population based epidemiological studies. A FFQ method is simple, easy to administer and requires minimal effort from the subjects (Boeing et al,1989). Previous studies on the evaluation of nutrient intakes in South Asians had been faced with difficulties such as language and cultural barriers (Silman et al,1985). Thus it was not surprising that some South Asian women especially, found the technique of weighing of food very difficult and demanding (Harbottle and Duggan,1993). Participation rates when a weighed record method had been used were shown to be low (Sevak et al,1994). Having considered the various dietary assessment method options that could be used in the study and the objective of our study, we opted for the FFQ method to evaluate nutrient intake in this study because the FFQ data can be collected easily and can represent usual intake of an extended period of time (Flegal et al,1988;Willett et al,1985).

Presently, there is no FFQ that has been developed and calibrated for use in the South Asian community. The main objective of developing the FFQ was to design a questionnaire that can be used to evaluate the dietary intake of South Asians. However, the objective of the FFQ was to determine the usual intakes of the individuals and then to rank them into the thirds of distribution. In order to make sure that the questionnaire was valid for use among the South Asian community, a calibration study for the FFQ was carried out. This calibration study is discussed in section 2.3. The next section focuses on the design of the FFQ.

2.2.4 Design of the Food Frequency Questionnaire

In the preliminary survey, the current meal patterns and the energy as well as the macronutrient sources in the South Asian diet were determined. The results from the preliminary survey are discussed in the next chapter. The determination of the energy and macronutrient sources are explained here.

The mean energy and macronutrients which include fat, carbohydrate and protein intake were calculated from the 24 hour recall interview (n=62) carried out in the preliminary survey at the GP clinic. The mean intake of each food item consumed by the subjects was determined. The energy, fat, carbohydrate and protein contributed by each food eaten were then calculated manually using the food composition tables and their supplements. The cumulative percent distribution of energy and the macronutrients were then calculated for each food list until at least 90% of the total energy, fat, carbohydrate and protein respectively had been included. This food list was then included in the FFQ. Thus the designed questionnaire is a food frequency questionnaire consisting of a list of 75 food items. Incidently these foods were also the foods commonly eaten by the majority of South Asians. Some food items (from the recall interview) which were usually consumed by a minority of South Asians were also included in the FFQ.

The food items were listed according to food groups namely:

1. Breads and cereals
2. Meat, poultry, fish and eggs
3. Milk and milk products
4. Fats and oils
5. Fruits and vegetables
6. Beverages
7. Miscellaneous

The miscellaneous group included food items which were frequently eaten by a minority of the South Asians such as achar, chutney and gur. This group also consisted of items like sugar,chocolate and sweets. The frequency responses were spread into seven responses ranging from rarely or never (1), one to three times per month (2), once a week (3), two to three times per week (4), four to six times per week (5), everyday (6) and two or more times per day (7). There was also a column identifying the standard portion for each food item and the number of portions consumed. The standard portion selected for each food item was based on the average portion size suggested by the Ministry of Agriculture, Fisheries and Food (Crawley,1988).

We were also interested in finding out about the method of food preparation of the South Asian community. In order to tease out any subtle differences in the method of food preparation and the food consumption of the South Asians, some questions regarding the matter were included. Questions on the demographic background were also included in the questionnaire. The FFQ is shown in Appendix 1.

2.3 Calibrating the FFQ

The calibration of the FFQ was carried out between December 1993 and February 1995. The sections that follow discuss the process that took place in the calibration study.

2.3.1 Dietary method used for the calibration study

In calibrating the FFQ, the dietary method used to evaluate the FFQ is important as we have to make sure that the method used will measure what it is supposed to measure. Theoretically the 7-day weighed intake method has been considered to be the best possible method for collecting information on nutritional intake (Bingham,1987). However, a high degree of literacy and total commitment of the subjects are required. Furthermore, the inconvenience of weighing all foods eaten may change the frequency of eating and the type of foods eaten.

For the past few years, dietary records with the estimated weights have been shown to be a satisfactory, repeatable and reliable method to use in a large scale epidemiology studies (Edington et al,1989). Mares-Perlman et al (1993) suggested that a dietary history questionnaire could produce nutrient estimates that rank individuals on the basis of intake of most nutrients similar to estimates from multiple food records.

The estimated food intake method has also been suggested to be a reliable method to use in a large scale epidemiological study. Bingham (1987) had also demonstrated that the cooperation rate was high with the 4-7 day estimated intake using random samples. However, the limitation in this method is in the lack of assessments of validity of records using the estimated weights. Bollard et al (1985) further demonstrated that a trained subject would usually estimate portion sizes significantly better than an untrained subject. In general, individual subjects tended to be imprecise at estimating portion sizes of food compared with weighing them directly.

Despite the daily variation in the food intake of an individual, a 3-4 day food record had been suggested to be optimal in evaluating food intake, taking into consideration the seasonal and weekday variations. It has also been shown that to correctly classify 80% of subjects into the extreme thirds of the distribution for evaluating energy and energy yielding nutrients, food records ranging between 3 to 7 days would probably be sufficient (Bingham,1981).

The 24 hour recall could also be used to calibrate the FFQ. Theoretically it would be better to do multiple 24 hour recalls over a period of time, for example four, six or even twelve recalls over a year period as this would represent a more habitual intake. However, the 24 hour recall might not be that representative of an individual's intake since his/her diet vary from day-to-day (Block, 1982). However, if the objective is to evaluate the intake of a group, the 24 hour recall is reliable enough (Morgan et al,1978).

When we are calibrating a dietary method (test method) with another dietary method (reference method), it is expected that both the test and the reference methods would have some common sources of measurement errors. Some of these errors include using the same food composition tables, the error of measuring the dietary intake as well as the error of some subjects overreporting and underreporting their food intake. These measurement errors between the test method and the reference method could also be positively correlated. The implication of this is that it could be possible that the relative validity between the methods is either overestimated or underestimated.

The use of biochemical markers or biomarkers in a calibration study has become increasingly important. This is because the measurement error of the biomarker is likely to be independent of those of the test and reference methods. Thus any correlation observed between the test method and the biomarker would reflect the level of validity between the two measurements (Willett,1990).
In the present study, we used the BMR multiples as the proxy for biomarker, to check the relative validity of the reported energy intake estimated by the test and reference methods.

In the present calibration study of the FFQ, the measurement errors in the dietary methods used will be evaluated and considered. The results obtained will then be interpreted with caution.

The method chosen to calibrate the FFQ in this calibration study is the 4-day weighed record. The 4-day weighed record is considered to be the most suitable method after taking into consideration the subjects that are going to be involved in this study. An earlier study has shown that the South Asian subjects had difficulties in carrying out the weighing of foods (Harbottle and Duggan,1993). Recruitment of subjects in general was rather difficult due to their unwillingness to participate in the study. Thus although in theory the 7-day weighed record would achieve a better classification of energy and energy yielding nutrients into thirds of the distribution, the compliance rate would probably be lower (Bingham,1981). This might affect the accuracy of the results obtained. The compliance rate would probably be higher with a 4-day weighed record method, considering the fewer number of days required to weigh and record the food. A three to 4-day weighed record had been suggested to be a reliable enough method to evaluate the macronutrients intake such as carbohydrate, protein and fat, which were nutrients of interest in this study (Bingham and Nelson,1991).

The results of the FFQ calibration study showed that the FFQ could not rank the macronutrients into the thirds of the distribution. The percent of subjects misclassified was quite high compared with other similar calibration studies (Thompson and Margetts,1993;Willett et al,1985). The correlation between the FFQ and the WR was also low (more results in Chapter 3). Thus it was imperative that other dietary assessment methods were developed and calibrated. The two methods selected were the 24 hr recall and the food checklist. These methods were selected after considering several problems encountered while calibrating the FFQ with the WR. These problems were discussed in section 2.3.5.

After calibrating the FFQ with the WR, we speculated that it was important that the dietary methods used in the second part should be less troublesome and less imposing on the subjects. The 24 hr recall was chosen because it was simple and a larger number of subjects could be recruited into the study with minimum resources (Bingham and Nelson,1991).

The food checklist was the other method that was selected to be our test method. A checklist is quite similar to an estimated food record. An estimated food record has been suggested to be a reliable method to use (Edington et al,1989). Furthermore recently Bingham et al (1995) demonstrated that the correlation between 16-day weighed record and a food checklist was between 0.45 to 0.69 excluding alcohol.

Since there is no food checklist that could be used for the South Asian community, a food checklist was also developed. The checklist consisted mainly of the food items listed in the FFQ. It is in the form of a little booklet whereby a simple instruction was written at the top of the page and then followed by the food item list. The subjects were asked to check off which foods they had eaten at the end of the day. The subjects were also requested to write the number of portions of foods they had consumed for each food item. A space was left at the bottom corner of the page to record foods which were consumed by the subjects but were not present in the printed list. The list was repeated seven times to represent seven days of food intake. The food checklist is shown in Appendix 2.

2.3.2 Calibration of test methods (FFQ, 24 hr recall and food checklist) with BMR multiples

All the test methods which were the FFQ, 24 hr recall and the food checklist were also calibrated against the BMR multiples. Even though the 24-hour urinary nitrogen is regarded as one of the best biomarkers for protein, and the doubly labelled water technique for energy expenditure (Bates et al,1991), we have little confidence in being able to recruit South Asian subjects who are willing to collect their urine specimens. During the recruitment stage of the calibration study, prospective subjects were asked whether they would give consent to the collection of a 24-hour urine specimen, but a majority of them declined. Thus the best possible method to calibrate the energy intake in these South Asian subjects is by calibrating it against the energy requirement calculated from the multiples of BMR.

Schofield et al (1985) developed the predictive equation of BMR which has been used to assess energy requirement of the population worldwide. However, this equation was shown to be invalid for Indian men and women (Soares and Shetty,

1988) as well as for other populations living in the tropics (Henry and Rees,1991). The Schofield et al (1985) equation tended to overestimate BMR of populations living in the tropics by 8% (Henry and Rees,1991). It was suggested that the reason for this overestimation of energy requirement was due to the majority of subjects recruited for the calculation of predictive equation in the database who were very active (military students and personnel, miners and labourers) and therefore might not be representative of the general population. This may bias the predictive BMR equation (Hayter and Henry,1994).

However, in this study, both equations to predict BMR which are Schofield et al (1985) and Henry and Rees (1991) will be used to calculate and estimate the BMR multiples in our subjects. The light physical activity level used to calculate the energy requirement (BMR multiples) is based on the occupational work (FAO/WHO/UNU,1985). We are using the factor 1.55*BMR for men while 1.56*BMR for women. This is in accordance with the subjects in this study who are housewives or retired men or men working in bakeries.

The height and weight of the subjects for the calculation of BMR multiples were taken at the subjects'homes during visits to administer or collect the food checklist. The weight of the subjects was taken with the subjects wearing light clothing and without shoes. The weight was recorded to the nearest ± 0.1 kg using the Soehnle weighing scale (max weight 150 kg). The height of the subjects was measured with the Harpenden pocket stadiometer to the nearest ± 0.1 cm.

2.3.3 Subjects in the calibration study

Initially the subjects involved in the calibration of the questionnaire were patients attending the GP clinic in the earlier dietary survey, but they were from different sets of patients. These subjects were men of South Asian origin and at least 18 years old. The subjects must be literate in the English language. However, due to difficulty in recruiting prospective subjects at the GP clinic, the selection of subjects for the calibration of questionnaire was widened to include women as well. The selection criteria for the women was similar to that used for men. The selection pool for the women included a group of women that attended a South Asian women's group meeting which convened twice a week at the community centre. This selection pool was then widened to include women participating in family groups and men and women attending English classes for South Asians at the various centres in Southampton.

A verbal consent was requested when approaching the subjects for his or her participation in the calibration study. Once the consent was granted, an appointment was set for a questionnaire interview. A second appointment was then set up for a demonstration of usage of dietary scale in their homes, usually two to three weeks after the questionnaire interview. The subjects were then given a demonstration on how to weigh and record their foods for four days, three week days and one week end. To be certain that each subject understood the weighing procedure, he/she was asked to repeat the demonstration to the researcher before a dietary scale (Soehnle) and a food diary were left with them.

For the calibration study, the questionnaire was administered first and then followed by the 4-day weighed record. This sequence has been suggested to be best if the FFQ is to evaluate the usual intake (Nelson,1991). The reason for the sequence of administration is also to reduce attempts by the subjects to recreate responses to the questionnaire like the pattern of the diet they recorded. This sequence also ensures that when the subjects are interviewed with the questionnaire, their responses are likely to be independent of any dietary assessment (Nelson,1991).

In the calibration study for the 24 hr recall and the food checklist, the subjects recruited were restricted to women only. This decision was made because recruitment was very difficult in men. Most young South Asian men were either working or studying, thus were not interested in taking part in the calibration study. In contrast, a lot of the older men (>50 years old) were illiterate in the English language. The illiteracy problem has been demonstrated in the Health and Lifestyle Survey in the ethnic minority groups (Health Education Authority,1994). More significantly, these older men were also reluctant to participate in the study.

Besides reinviting the women who participated in the FFQ calibration study (old), we also recruited new women subjects into the study (new). The same selection criteria applied.

The subjects participating in the 24 hr recall and the food checklist calibration study did not do any weighed food record. This was due to the high rate of refusal in participation if subjects were asked to weigh their food. Therefore for the purpose of comparison with the reference method, we used the weighed record data from the women in the FFQ calibration study. However, the total number of subjects who completed all four dietary assessment methods was 23. The summary of subjects completing the various dietary methods is shown in Table 2.1.

The administration of the food checklist was carried out after a verbal consent was obtained from the subjects. An appointment for the food checklist interview was set up, usually at the subject's home. The subjects were taught on how to complete the food checklist. For subjects who were

Table 2.1 Number of subjects completing the various dietary assessment methods

Methods	Men	Women
Weighed record	23	35
FFQ	48	44
Food checklist	0	48 (27 Old, 21 New)
24 hr recall	0	49 (27 Old, 22 New)
All methods	0	23

illiterate, the instructions for completing the food checklist were explained in the presence of a relative. The food checklist was also translated into the three main ethnic languages of the South Asians, namely Gujerati, Urdu and Punjabi. However , this did not solve the language problem as those who were illiterate in the English language were also illiterate in their native language.

The 24 hr recall was carried out when the food checklist was collected from the subjects. Basically, the subjects were asked systematically to recall the food intake of the past 24 hours. They were also requested to estimate portion sizes of the foods eaten, for example, the size of the chapatti and the amount of curry consumed. Food models were used to help the subjects in estimating the portion size.

2.3.4 Sampling calculation

Since the errors implicit in different levels associated with different correlation coefficients between our test and reference measures, we felt that the minimum size of correlation coefficient we wanted to be able to detect was 0.3 for fat and energy (Bingham et al,1994). We would like the study to have 80% power at 5% significance level. To calculate the sample size required for the study, Table 8.1

from Machin and Campbell (1987) was used. The number of subjects required to detect a significant correlation coefficient of 0.3 for fat and energy was 84 subjects per gender. Table 2.2 shows the sample size required for the study if the size of a statistically significant correlation coefficient we wanted were 0.4 to 0.7 (Machin and Campbell,1987). With the sample size (n=35) of the present study, the power would reach 99% if we could detect a significant correlation coefficient of 0.7. between the test and reference measures.

Table 2.2 Sample sizes for detecting a statistically significant correlation coefficient

correlation coefficient	sample size	power of study
0.7	13	99%
0.6	19	97%
0.5	29	88%
0.4	46	68%
0.3	84	<50%

The number of subjects targeted for recruitment in the FFQ calibration study was 84 subjects for each gender. The same number of subjects was targeted in the second calibration study. However, in the first calibration study, we managed to recruit 100 subjects (54 men, 46 women) but only 58 subjects completed both the weighed record and the FFQ, thus only the results from this 58 subjects were used for analysis, although the demographic characteristics of the non responders were assessed.

In the 24 hr recall and food checklist calibration study, 23 subjects (women) from the FFQ calibration study agreed to reparticipate. We also managed to recruit 36 new subjects. At the end of the study about 52% succeeded in completing the calibration study. Thus the power of the present calibration studies to detect a significant association was about 50%

2.3.5 Problems arising during recruitment of subjects

In the course of recruiting the subjects for the calibration study, several problems arose. The problems encountered in the progress of the calibration study are highlighted here.

The first problem to overcome in penetrating the community was to break down any barriers , one of which is a general suspicion which the community has of the researcher. In order to be part of the community , time was spent talking and explaining to the community about the purpose of the research to be undertaken and in return what advantage the community could attain in terms of health benefits from the study.

The second problem which could be a major problem initially was the language barrier, especially so if the researcher did not speak nor understand the many languages of the South Asian ethnic groups. Most South Asian women and some men were not fluent in English, although some could speak and understand a little. It was also common among them, especially the women, to be illiterate in both their native language and English. However, the language problem was more common among the older men and women, though some younger men and women had the same problem as they came to the United Kingdom only after marriage. This problem was overcome by using a relative as an interpreter when visiting their homes or friends when at community meetings. It would be ideal if a worker from the community could be involved in the calibration study, to help in the language problem that was present in the study. However the use of an interpreter in itself could create another problem, that was whether the questions and answers had been interpreted correctly. Communication problems using interpreters had been confirmed in two studies involving South Asian patients (Leatherdale et al,1978; Ebden et al,1988).

In the initial stage of this study, recruitment of subjects for the calibration of the questionnaire was restricted to men only, after considering that most of the men can speak, write and understand English better than women. The questionnaire interview went very well. However, when it came to weighing and recording the food intake for the calibration study, another problem rose. It was the custom of the South Asian community that the men were the leader in the household and the women were the ones preparing and serving the food. Thus, although initially most of the men agreed to be interviewed with the questionnaire and then weighed and recorded their food, at the end of the study only about 45% of the men who agreed to participate in the calibration study completed the weighed food record.

A meeting with the ethnic liaison officer and the health visitors eventually eased the recruitment problems. The suggestion put forward was to include women into the calibration study. Various women's group meetings and their meeting locations were suggested as the sampling pool. After an initial introduction, the recruitment of subjects proved to be better, though reluctance to participate in some of the prospective subjects still remained.

2.3.6 Problems arising during the food weighing and completing the food checklist

The problems in the calibration study did not end at the recruitment stage. Even in the weighing of the food, other difficulties also appeared. The most common problem here was the much longer time required to complete the 4-day weighed record. It happened quite regularly that the subjects took at least a month and sometimes even longer to complete the food weighing. Most of the time these subjects ended up not completing the food weighing at all. Because the subjects took a longer time to complete their 4-day weighed record, we would expect bias on the days that food was recorded. It would be likely that the days which the subjects chose to record their food intake, were the days they thought they had desirable foods.

Other difficulties included the non realistic weight recorded for the portion of food consumed. From these records it was possible that the subjects did not fully understand the procedure involved in the food weighing and recording or had forgotten how to do it. This error in recording was mostly made by the illiterate South Asians. Although these subjects were sincere in wanting to participate in the calibration study, their lack of education and minimal understanding of English made it difficult for them to contribute effectively.

Although the food checklist was as readily accepted by the subjects as the FFQ, there were some prospective subjects who totally refused to do any extra task which would change their daily routine. There was also a similarity in the time taken to complete the food checklist as in the weighed record, in that a longer time was required to complete the food checklist. However this time, most of the subjects succeeded in returning the finished food checklist. Thus the food checklist appeared to be a more well received (higher participation rate) method to evalute dietary intakes of South Asians in this study.

2.4 Role of chance and effects of other variables in the study design

In a study, various issues need to be considered when interpreting the results. These issues are whether the results have occurred by chance, whether bias has occurred and whether the results have been influenced by other variables.

2.4.1 Role of chance

The role of chance in a study design can be evaluated by hypothesis testing and confidence interval. To test that the association between the FFQ (test method) and the dietary assessment methods (reference method), a hypothesis is set up. In any study, we usually hypothesize that there is no relationship between the exposure of interest and the outcome (null hypothesis) while alternatively there is a relationship between the exposure and the outcome (alternative hypothesis). We can then conduct a test of significance to allow us to conclude that the association found in the study occurred by chance alone. The P-value is a probability statement, which indicates whether the probability or likelihood of obtaining a result at least as large as that observed in a study occurred by chance alone. The level of significance (P-value) is usually set at 0.05, 0.01 or 0.1. If the P-value is higher than the predetermined cut-off-point, it suggests that there is no association between the exposure and the outcome being considered.

If the P-value is less than or equal to 0.05 ($P<0.05$), this indicates that there is a 1 in 20 chance of observing an association as large or larger than found in the study by chance alone, given that there is no association between the exposure and the outcome. This means that chance is an unlikely explanation of the findings, and we reject the null hypothesis and conclude that there is a statistically significant association between exposure and outcome under study. On the other hand, if P-value is >0.05, we would consider that chance cannot be excluded as a likely explanation for the findings and we accept the null hypothesis, and conclude that the findings are not statistically significant at that level.

In this calibration study, to ensure that the results have not occurred by chance alone, the results are presented as means and 95% confidence interval. Confidence interval is another measure which could be used to evaluate the role of chance. Confidence interval is defined as the range within which true magnitude of effect lies within a certain degree of assurance (Hennekens and Buring,1987). The width of the confidence interval can indicate the amount of variability inherent in the estimate and thus the effect of the sample size. Thus the wider the confidence interval, the greater the variability in the estimate of the effect and usually due to the smaller study sample size. Similarly, the narrower the confidence interval, the more stable the estimate (usually because of a larger study sample). As in the P-value, the confidence interval can correspond to the significance level of interest (for example, 0.05 approximately equals to 95% confidence interval).

2.4.2 Bias

Besides the bias in the dietary measurements, in general, there are two other types of bias, selection bias and observer or information bias. The former is referred to as any error that arises in the process of identifying the study population while the latter is any systematic error in the measurement of information on exposure and outcome (Henneken and Buring,1987). There are three common sources of information bias. They are social desirability bias, recall bias and interview bias (Margetts,1996).

Initially the criteria set for the subject inclusion were South Asians, at least 18 years old and could read and write in English. To reduce the selection bias of only recruiting the educated South Asians, the selection criteria were expanded to include any South Asian who was at least 18 years old and was willing to participate despite speaking and understanding minimal English. Subjects who could not write were also included. In the second part of the calibration study, we only recruited women subjects. This would introduce gender selection bias, however this bias would be considered in the interpretation of the results. We have however, considered any bias such as age and education because otherwise only younger educated South Asians would fulfill the selection criteria.

The second type of bias which is the information bias, usually occur in case-control studies and retrospective studies. This calibration study has a cross sectional design, thus the information bias in the form of recall bias might be minimal.

The social desirability bias, is a type of bias in which the individual intends to show a desirable image. This sort of bias may occur in women on diets as well as overweight men and women (Margetts,1996). In this study, our comparison of energy intake estimated from our test methods with the EI/BMR ratio demonstrated that many of the overweight subjects underreported their energy intake while some of the underweight subjects overreported their energy consumption. Thus it could be assumed that some form of desirability bias was present in this study.

In the FFQ, the demographic information of each subject was gathered. This allowed us to determine whether those who participated were different from those who dropped out of the study.

Besides the biases discussed above, it is possible that some other variables could be associated with the exposure under study. However, in this study we tried to control these other variables by restricting the inclusion criteria and at the same time reducing selection bias as much as possible.

2.5. Analysis of dietary data

The dietary data were converted to nutrient intake using the commercially available software package (Compeat version 4 for WR) and a software package for FFQ developed at the Wessex Institute of Public Health Medicine, University of Southampton. Each computer software package used a separate but similar nutrient database developed from food composition tables and their supplements (Holland et al,1988; Holland et al,1989; Holland et al,1991a and Holland et al,1991b). The databases used to calculate nutrient intakes were compared for the FFQ, WR, food checklist and the 24 hr recall.

2.5.1 Statistical Analysis

Means and 95% confidence interval (CI) of the mean daily nutrient intakes were computed. Mean nutrient difference (MD) between the test methods were calculated from reference methods (FFQ-WR, food checklist (CL)-WR, 24 hr recall (RCL)-WR) together with 95% CI. In addition, comparisons were also made using energy adjusted nutrient intakes (Willett,1990). Energy adjusted nutrients were computed as residuals from the regression model with the total energy intake as the independent variable, and absolute nutrient intake as the dependent variable. The ability of the FFQ, food checklist and the 24 hr recall method to correctly rank individuals was assessed by the Spearman rank correlation coefficient and the kappa statistics. Agreement between methods for energy, protein, fat and carbohydrate was also examined using the Bland-Altman technique in which the difference in the nutrient value between dietary methods was plotted against their mean for each subject (Bland and Altman,1986).

2.5.2 Adjustments and correction of the FFQ

The result of the first analysis of the comparison for the nutrient intakes for the FFQ and WR showed a large mean difference for various nutrients. Due to these large mean differences, some adjustments were made to the FFQ specifically to the chicken, eggs and meat groups (now referred to as the meat group) as well as the fats and oils group (now referred to as the fat group).

In the meat group, adjustments were made on the portion sizes of the meat dishes. For example, the chicken, lamb and beef portion sizes were reduced from 300g to 200g per portion. The fish (not fried) portion was decreased to 75g from the original 300g portion. These reductions in the portion sizes were based on the average portion sizes of the meats consumed and recorded in the weighed records. In the unadjusted FFQ, the meat group were consistently overestimated.

Adjustments were also made on the frequency of the fat group. In the unadjusted FFQ, the frequency of the fat consumption was everyday, as curry dishes were eaten everyday, and oils or ghee were used to fry the curry spices. However, the nutrient content of the curry meat group dishes had already taken fats into account. Thus by considering the fat group frequency as an everyday consumption, the estimated fat intake has been accounted for twice. The frequency of the fat group was thus adjusted accordingly in the adjusted FFQ.

2.5.3 Analysis of adjusted FFQ dietary data

The data from the adjusted FFQ were converted to nutrient intake using the same software package as mentioned in section 2.5. The calculated nutrient intakes were then compared with the adjusted FFQ and WR. A similar procedure was also repeated for the statistical analysis.

2.5.4 Analysis of measurement errors using the method of triads

There were three test methods used in the study namely the FFQ, 24 hr recall and the food checklist. In order to calibrate the relative validity of the test methods, the 4-day weighed record was used as the reference method. In this study, we have also used the multiples of BMR as the proxy of the biomarker to check the relative validity of the energy estimated by the dietary methods. According to Clayton and Plummer (1993) and van't Veer et al (1993), it was very likely that the measurement errors between these dietary methods were correlated. However, using the BMR multiples could provide a fair comparison of the other dietary assessment methods because the errors were likely to be uncorrelated.

We evaluated the measurement errors in the FFQ calibration study and the food checklist as well as the 24 hr recall calibration study using the method of triads (Kaaks,1996). This method allows the validity coefficient of the test measurement to be estimated from a triangular comparison between FFQ, food checklist and 24 hr recall (test methods), weighed record (reference method) and BMR multiples (proxy for biomarker). This method assumes that the measurements are linearly related to the true intake and have independent random errors. The graphical representation is shown in Figure 1.3 (p 33).

CHAPTER THREE

RESULTS

CHAPTER 3 RESULTS

This chapter is organized into four main sections. The first focuses on the results of the preliminary survey (3.1) This is broken down into two subsections which are the sources of macronutrients in the South Asian diet (3.1.1) and results on the nutrient intakes of South Asians in the preliminary survey (3.1.2). This is followed by the results of the FFQ calibration study (3.2) which describes the demographic characteristics of subjects (3.2.1) and discusses the comparison of FFQ with the 4-day weighed record (3.2.2). The third section discusses the results of the food checklist and 24 hr recall calibration study (3.3). Comparison of food checklist, 24 hr recall with 4-day weighed record are discussed in section 3.3.1. The final section focuses on the measurement errors in the calibration study (3.4).

3.1 Results of preliminary survey

The preliminary survey was carried out with two main objectives. Firstly we wanted to establish the meal pattern of the South Asian community and secondly we wanted to determine the food sources of the main macronutrients (fat, carbohydrate and protein) in their South Asian diet. The meal pattern shown in Table 3.1 reflected those of four major ethnic groups in Southampton which were the Gujeratis, Punjabis, Pakistanis and Bangladeshis. The demographic data of the subjects were also gathered and represented in Table 3.2.

The staples of South Asians mainly consisted of chapatti eaten with meat or vegetable curry. The type of curry consumed, whether meat or vegetable depended on whether they were vegetarians or omnivores. This type of diet was influenced by their religion. Generally the Hindus (Gujeratis) and some Sikhs (Punjabis) were vegetarians

Table 3.1 Basic meal pattern of South Asians in Southampton

Breakfast	bread spread with butter or margarine cereal with milk sometimes egg (fried or boiled) tea always with milk and sugar
Lunch	always chapatti and sometimes rice* eaten with meat/chicken/fish/vegetable/dal curry (usually from the previous day's leftover) sometimes English meal
Dinner	as lunch meal (freshly cooked for dinner)
Before bed	a glass of milk

* Only the Bangladeshis and Gujeratis ate rice everyday

while the Muslims (Pakistanis) were omnivores. It was also a common practice among the omnivores to occasionally eat vegetable curry with their chapatti. The Bangladeshis were the only group that consumed rice with meat or vegetable curry and ate chapatti infrequently. The Gujeratis, in contrast, always consumed chapatti and rice with vegetable curry and dal (lentil) soup. Although the meal pattern generally appeared very similar between ethnic groups, subtle differences were teased out during the recall interview (Table 3.3). It was found that the method of preparation of chapatti and the curry differed between these ethnic groups. The Gujerati households prepared the chapatti with wheat flour and ghee or vegetable oil as the main ingredients while the Pakistanis and Punjabis normally used only wheat flour and water when making the chapatti dough. However, a minority of Pakistani and Punjabi households also added ghee and vegetable oil in the chapatti dough.

The way chapatti was consumed after cooking was also different. The Gujeratis (Hindus and Muslims) would spread either ghee or butter on the chapatti before consumption while the Pakistanis and Punjabis would consume the chapatti as it was (dry). Chapatti size also varied between ethnic groups. The Gujeratis made the smallest chapatti, followed by the Punjabis while the Pakistanis made the biggest chapatti. The major difference in the curry preparation was in the type of fat used. Some households (majority of the Gujeratis) used ghee in food preparation which required frying spices. Other households (majority of Pakistanis and Punjabis), however, used only vegetable oil, with only a minority of these two ethnic households also using ghee for frying spices.

Tea was widely consumed by the South Asians. The number of cups of tea drank per day ranged between two and ten. The tea was normally prepared with milk and two teaspoons of

Table 3.2 Demographic data of subjects in the preliminary survey (n=62)

Sex	male	26 (42%)
	female	36 (58%)
Marital status	married	43 (70%)
	single	16 (26%)
	divorced	2 (3%)
	widowed	1 (1%)
Ethnic	Pakistani	12 (18%)
	Punjabi	38 (57%)
	Gujerati	9 (14%)
	Bangladeshi	7 (11%)
Occupation	Unemployed	15 (23%)
	Housewife	17 (25%)
	Businessman	5 (8%)
	Bakery employee	5 (8%)
	Student	10 (15%)
	Others	15 (21%)
Lived in UK	<5 years	12 (18%)
	6-10 years	4 (6%)
	11-20 years	23 (35%)
	>20 years	27 (41%)

Table 3.3 Subtle differences in the method of food preparation and food consumption of the South Asian community

	Gujerati	Pakistani	Punjabi
Use vegetable oil or ghee in chapatti dough	Yes	No/Yes	No/Yes
Spread ghee or butter on chapatti	Yes	No/Yes	No/Yes
Use ghee to fry spices	Yes	No/Yes	No/Yes

sugar. Fizzy drinks were also consumed frequently but usually taken with the meal. The amount taken varied between 250 ml and 2 litres per day. Fizzy drink consumption was more common in the younger subjects. Beer was consumed by a small number of subjects, usually Gujerati and Punjabi men in amounts between 4 and 7 pints per week.

Chutney or pickles and achar were consumed by some of the subjects depending on their preference for these foods. There were three types of chutney which were commonly eaten. These included the spicy, salty and sweet chutney. Those who enjoyed chutney consumed it at least once a day.

The meal pattern of the South Asians in this preliminary survey resembled closely to those South Asians in Birmingham and in London (Kalka,1988;Wharton and Eaton,1983). The preparation of chapatti and curry of the South Asians were also similar. The habit of spreading butter and ghee on chapatti was practised by the Gujeratis in this study and also in Birmingham.

The meal pattern of the Punjabi and Gujerati groups in Southampton mirrored those of their counterparts in London (Kassam-Khamis et al,1995). The consistency of results across centres in the UK suggested that meal patterns were likely to be set before migration and maintained while in the UK. This implied that regardless of where the South Asians settled down when they first migrated to the UK, the meal pattern of these groups was very much influenced by their origin in the Indian subcontinent.

3.1.1 Macronutrient sources in the South Asian diet

To date no study has determined the food sources of energy and macronutrients in the South Asian diet. A total of 62 men and women of South Asian origin were interviewed and macronutrient intakes derived using the 24 hour recall method. The energy, fat, carbohydrate and protein sources in the South Asian diet respectively are tabulated in Table 3.4 through Table 3.7. The percentage distribution of energy and fat in South Asian foods are also shown graphically in Figure 3.1 and Figure 3.2. The sources of energy and the macronutrients were listed in the order of the size of their contribution to the total energy intake. The cumulative percent distribution of energy, fat, carbohydrate and protein was calculated for each food list until at least 90% of the total energy, fat, carbohydrate and protein intake had been included.

The preliminary study demonstrated that the main sources of energy in the South Asian diet were chapatti, meat curry and milk and dairy products. These foods were also important sources of fat and protein. Chapatti, breads and rice were the main sources of carbohydrate in the diet. In general, it can be summarized that important sources of energy, fat, carbohydrate and protein in the diet were mainly concentrated in the chapatti, meat curry and breads as these foods contributed between 30% to 50% of the energy and macronutrient intake. The food sources of energy and macronutrients in the South Asian diet were different from those of the English diet (Cade and Margetts,1988). There

Table 3.4 Sources of energy in the South Asian diet

Food	% of total energy intake	Cumulative %
1.Chapatti	17.6	17.6
2.Meat curry	16.4	34.0
3.Milk and dairy	14.6	48.6
4.Breads	6.9	55.5
5.Vegetable curry	4.1	59.6
6.Rice	3.8	63.4
7.Paratha/naan	3.5	66.9
8.Buns/pastries	3.3	70.2
9.Fruits/juices	2.8	73.0
10.Sugar	2.7	75.7
11.Breakfast cereals	2.3	78.0
12.Eggs	2.2	80.2
13.Chips	2.2	82.4
14.Crisps	1.9	84.3
15.Savouries	1.6	85.9
16.Lentils	1.5	87.4
17.Soft drink	1.2	88.6
18.Meat,not curry	1.0	89.6
19.Puddings	0.4	90.0

Table 3.5 Sources of fat in the South Asian diet

Food	% of total fat intake	Cumulative %
1.Meat curry	32.0	32.0
2.Dairy products (butter etc)	12.5	44.5
3.Milk	11.8	56.3
4.Paratha/naan	9.5	65.8
5.Vegetable curry	5.0	70.8
6.Eggs	3.9	74.7
7.Buns/partries	3.7	78.4
8.Chapatti	3.1	81.5
9.Savouries	2.8	84.3
10.Crisps	2.5	86.8
11.Meat,not curry	1.9	88.7
12.Chips	1.7	90.4

Table 3.6 Sources of carbohydrate in the South Asian diet

Food	% of total CHO intake	Cumulative %
1.Chapatti	30.1	30.1
2.Breads	11.5	41.6
3.Rice	6.8	48.4
4.Milk	5.8	54.2
5.Sugar	5.7	59.9
6.Fruits/juices	5.6	65.5
7.Breakfast cereals	4.1	69.6
8.Paratha/naan	3.9	73.5
9.Buns/pastries	3.7	77.2
10.Vegetable curry	3.4	80.6
11.Chips	3.2	83.8
12.Soft drink	2.6	86.4
13.Meat curry	1.8	88.2
14.Lentils	1.7	89.9
15.Crisps	1.2	91.1

Table 3.7 Sources of protein in the South Asian diet

Food	% of total protein intake	Cumulative %
1.Meat curry	31.2	31.2
2.Chapatti	18.5	49.7
3.Milk /dairy products	16.6	66.3
4.Breads	7.8	74.1
5.Eggs	4.7	78.8
6.Vegetable curry	3.8	82.6
7.Meat,not curry	3.8	86.4
8.Lentils	3.3	89.7
9.Paratha/naan	2.6	92.3

Figure 3.1 Percentage distribution of sources of energy in the South Asian foods

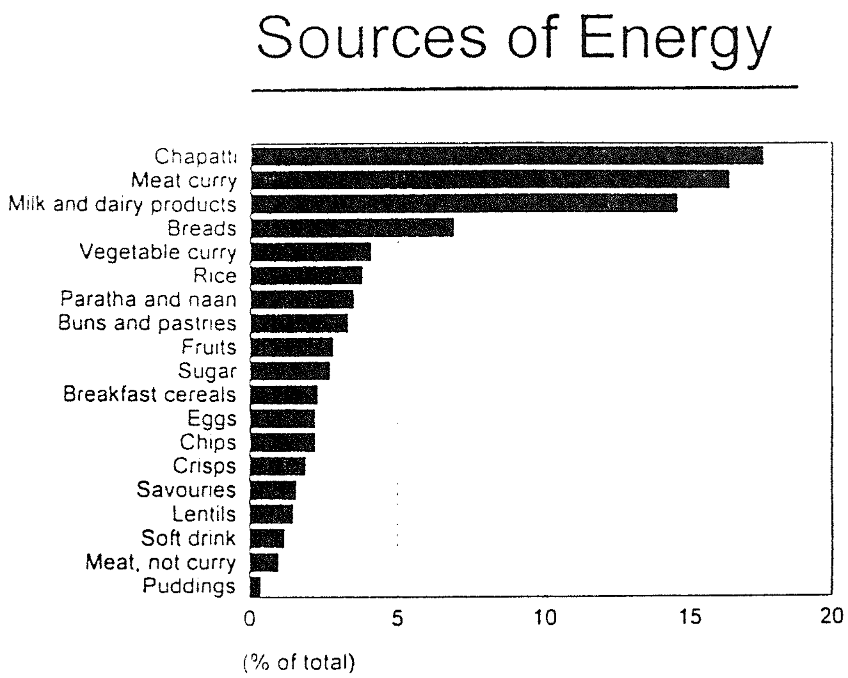
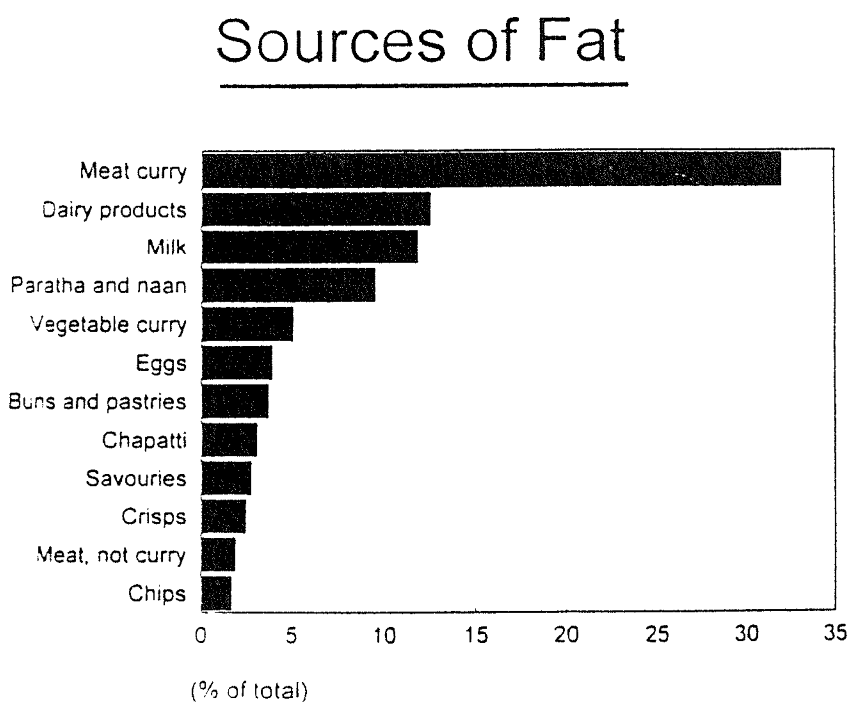


Figure 3.2 Percentage distribution of sources of fat in the South Asian foods



was less variety in the foods eaten in the South Asian diet compared with the English diet. This was reflected in the much shorter list of food items required to contribute to at least 90% of the total energy and macronutrient intake. This also demonstrated that the diet consumed by the South Asians seemed to be quite monotonous and homogeneous.

Thus in the development of an FFQ for the South Asian community, it was essential that a food list different from the one used for the English diet was included. This new list would cover at least 90% of the total energy, fat, carbohydrate and protein intake in the adult diet. This was clearly shown in the findings in our study which demonstrated that a different set of foods and much less variety of food items were consumed by the South Asians.

3.1.2 Nutrient intakes of South Asians in the preliminary survey

Although the objective of the preliminary survey was to evaluate the current meal pattern of the South Asians, we decided to analyse the nutrient intake of the subjects. The food intake obtained from the 24 hour recall was analysed for nutrient content using the Compeat Version 4 software package. While analysing for the nutrient composition of the foods consumed in the South Asian diet, a few problems pertaining to the Compeat software package were identified. There were a lot of South Asian foods missing in the package. Some nutrients, (although not analysed in this study) such as fatty acids (saturated, monounsaturated and polyunsaturated) were also incomplete in most of the South Asian foods available in the package. The nutrient content of the foods not present in the package were estimated by substituting the food with a food that was as similar as possible to that in the package.

Since the analysis of the nutrient content in the preliminary study, further work had been carried out in updating the nutrient composition of the composite dishes consumed by the South Asians. Thus far, only the nutrient composition of dishes commonly eaten by the Gujeratis and Punjabis have been analysed and reported. Further work updating the nutrient composition of dishes of the Pakistanis and the Bangladeshis are currently being undertaken (Kassam-Khamis et al,1995).

The implication of the incomplete nutrient composition of foods in the Food Composition Table for the South Asian foods would be an underestimation of some of the nutrients for certain foods. The completeness of the nutrients for some foods of certain ethnic groups (Gujerati and Punjabi), while incompleteness in the other groups (Pakistani and Bangladeshi) could further widen the nutrient intake differences between these groups, if there were any differences in the first place.

The nutrient intakes of the subjects are broken down by gender in Table 3.8. The mean energy intakes of the men and women were 2237 kcal per day and 1727 kcal per day respectively. As expected, the intakes of energy and macronutrients namely fat, carbohydrate and protein were higher in men than women. However, the macronutrient contribution to the total energy intake were very similar in both genders.

Among the ethnic groups of the South Asian population studied, the macronutrient intake showed a large range in the energy intake of both the men and women (1893-2336 kcal/day in men;1628-2493 kcal/day in women)(Table 3.9). Even the fat, carbohydrate and protein contribution to the total energy varied widely in all the ethnic groups and in both genders. The percent contribution of fat, carbohydrate and protein intake in this preliminary study compared very well with those of the National Food Survey (44% fat, 44%

Table 3.8 Mean (SD) for energy and nutrient intakes of the subjects in the preliminary survey

Nutrient	Male (n=26)	Female (n=36)	All (n=62)
Energy,kcal	2237 (876)	1727 (615)	1942 (773)
Fat,g	108.9 (52.3)	85.6 (37.4)	95.3 (45.4)
% Fat	43.8	44.6	44.1
CHO,g	247.2 (105.9)	186.5 (74.9)	214.1 (93.4)
% CHO	44.2	43.1	44.0
Protein,g	68.9 (29.4)	52.7 (20.8)	59.6 (25.8)
% Protein	12.3	12.3	12.3

CHO-carbohydrate

Table 3.9 Mean (SD) for energy and nutrient intakes between ethnic groups of South Asians

Ethnic	Indian		Pakistani		Bangladeshi		Punjabi	
	men (n=4)	women (n=5)	men (n=4)	women (n=7)	men (n=2)	women (n=3)	men (n=16)	women (n=21)
Energy, kcal	1893 (488)	1834 (428)	2263 (155)	1634 (509)	2028 (99)	2493 (405)	2336 (1073)	1628 (651)
Fat,g	95.6 (43.5)	94.6 (15.3)	119.1 (18.5)	88.2 (28.2)	88.9 (10.2)	134.6 (31.8)	112.0 (62.3)	76.1 (39.6)
% Fat	43.9	47.3	47.3	48.4	39.4	48.4	41.9	40.6
CHO,g	212.6 (48.2)	194.9 (72.9)	222.7 (57.7)	172.0 (77.2)	263.7 (31.8)	266.0 (26.7)	276.8 (126.4)	198.6 (77.2)
% CHO	42.9	38.9	36.9	39.5	48.4	40.3	45.6	47.2
Protein,g	46.4 (18.8)	62.2 (16.3)	87.9 (19.4)	48.8 (15.6)	59.8 (1.2)	70.4 (5.1)	70.8 (32.2)	49.4 (23.2)
% Protein	9.7	13.7	15.6	12.2	11.8	11.4	12.1	12.1

cabohydrate, 12% protein vs 42% fat, 45% carbohydrate, 13% protein). However, these results could not be interpreted as a true reflection of nutrient intakes of the population as the number of subjects involved here was too small.

3.2 Results of the FFQ calibration study

The calibration of dietary methods in this study was carried out in two parts. The first part was the calibration of the FFQ as the test method with the 4-day weighed record as the reference method. This section discusses the results of the FFQ calibration study. The second calibration study calibrated the food checklist and the 24 hr recall against the 4-day weighed record.

The FFQ calibration study was carried out over a period of 15 months between December 1993 and February 1995. A total of 100 subjects agreed to participate in this study. However, at the end of the study, 58 subjects completed both the FFQ interview and the 4-day weighed record (responders), 34 subjects completed the FFQ interview but did not complete the weighed record (non responders) and 8 subjects did not complete either the FFQ interview or the weighed record. The breakdown of responders and non responders according to ethnic groups and gender are outlined in Table 3.10 (described in more detail on p 116). Demographic data of the subjects are represented in Table 3.11.

Table 3.10 Responders and non responders according to gender and ethnic groups in the first calibration study

	Responders (n=58)		Non responders (n=34)	
Ethnic	Men	Women	Men	Women
Pakistani	6	20	5	4
Gujerati	9	8	5	2
Punjabi	8	7	13	3
Bangladeshi	0	0	2	0
Total	23	35	25	9

Table 3.11 Demographic characteristics of responders and non responders in the FFQ calibration study

		Responders	Non responders
Sex	Male	23 (40%)	25 (74%)
	Female	35 (60%)	9 (26%)
Marital status	Married	53 (91%)	31 (91%)
	Single	3 (5%)	2 (6%)
	Divorced	2 (4%)	1 (3%)
No.of children	0-5 children	52 (90%)	32 (94%)
	6-10 children	6 (10%)	2 (6%)
No.in household	0-5 person	40 (69%)	18 (53%)
	6-10 person	18 (31%)	16 (47%)
Ethnic	Gujerati	17 (29%)	7 (21%)
	Pakistani	26 (45%)	9 (26%)
	Punjabi	15 (26%)	16 (47%)
Live in UK	< 5 years	7 (12%)	2 (6%)
	6-10 years	5 (9%)	3 (9%)
	11-15 years	2 (3%)	4 (12%)
	> 15 years	44 (76%)	25 (73%)
Education	0-6 years	10 (17%)	6 (18%)
	> 6 years	48 (83%)	28 (82%)
Age	Male	46 years (19 to 69 years)	44 years (18 to 46 years)
	Female	45 years (19 to 70 years)	31 years (21 to 46 years)

3.2.1 Demographic characteristics of subjects in the FFQ calibration study

In the FFQ calibration study, the Pakistanis represented the majority of the responders, followed by the Punjabis and Gujeratis while the non responders were mostly Punjabis. Both the responders and non responders in the calibration study were mostly married and had five or fewer children. For the responders, 69% had at least five people living in the household, 76% had been living in the UK for at least 15 years and 83% received more than six years of education either in India or both in India and the UK. The trend was similar in the non responders in that the majority (94%, 53% and 73% respectively) had five children or less, at least five people in the household and received more than six years of formal education (Table 3.11).

The demographic data showed that the majority of responders and non responders in this survey consisted of both men and women who had few children and did not have extended family living with them. This was demonstrated in the smaller percentage of the subjects with 6-10 person in the household. Most of the subjects attended at least six years of formal schooling, mostly in India. The younger South Asians had more than six years of formal education, both in India and UK.

The demographic information collected showed that there were some differences in the demographic characteristics between the responders and non responders. This was evident in the gender and ethnic group criteria. There were more women who were responders (60%) compared to men. A possible reason for this higher response in women was the required weighing of food in the calibration study. As discussed in section 2.3.5 regarding problems encountered during the recruitment of subjects, South Asian men were not involved in the food preparation in their homes, thus a higher drop out rate occurred among men. In terms of ethnicity, there was no obvious reason why there was a difference in participation.

The gender difference between the responders and non responders could have made a difference in the calibration study. These differences included how well men and women weighed and recorded their food intake as well as how they could remember and estimate their usual intake when responding to the FFQ. There was a higher percentage of women (66%) than men (28%) who were overweight and obese in the calibration study (Table 3.12). Thus it could be possible that the overweight and obese individuals changed their food intake during the weighing of food so as to show a more acceptable food intake to the researcher. If the obese individuals intentionally reduced their food intake during the weighing of food, this would underreport the usual intake. The implication on the calibration study would be a lower absolute mean nutrient intake in the weighed record, thus possibly making the mean differences between the two dietary measures which were the FFQ and the WR larger than it should be. This was assuming that the subjects did not underreport the FFQ.

Table 3.12 Body Mass Index (BMI) of subjects (%) in the calibration study

BMI	Classification	Men	Women
< 20	Underweight	11%	10%
20-25	Normal	61%	24%
25-30	Overweight	28%	42%
> 30	Obese	0%	24%

3.2.2 Comparison of food frequency questionnaire (FFQ) with the 4-day weighed record (WR)

When comparing the FFQ (test method) with the WR (reference method) in estimating the nutrient intake, the mean nutrient differences (MD) between the two methods were calculated from mean of FFQ minus mean of WR (FFQ-WR). The percent mean difference, which was ((mean FFQ-mean WR)/mean WR) * 100 was also calculated. The results of the FFQ shown have been adjusted from the original FFQ, and only the adjusted FFQ (now referred as FFQ) are shown.

Table 3.13 and Table 3.14 represent the comparison of the mean daily nutrient intakes between the FFQ and WR for men and women respectively. Even though the number of subjects in this calibration study was small, it was important that the analysis were carried out separately for each gender. This was because it was likely that men and women did not respond to dietary assessment similarly (Nelson,1996). The FFQ tended to give higher estimates for energy, fat, carbohydrate and protein as well as other nutrients such as sugar, starch and dietary fibre in both men and women. The percent mean difference between the two measures was in general within 10% for energy (10%), fat (7%) protein (4%) and starch (-6%), while in the range of 13% to 74% for carbohydrate, dietary fibre and sugar in women. The percent mean difference for dietary fibre in men was 9%. However, for the other nutrients, the mean differences between measures ranged between 11% to 99%.

The results presented in Table 3.13 and Table 3.14 only showed the results of the FFQ with the WR. Some modifications were made to the original FFQ and these modifications were regarding portion sizes and frequencies to specific food groups. Adjustment to the portion size was made to the meat group. Changes were made to the frequencies of fats and oil group. These adjustments were

Table 3.13 Comparison of mean daily energy and nutrient intakes between WR and FFQ in men (n=23)

Nutrient	Mean WR (95% CI)	Mean FFQ (95% CI)	MD	%MD
Energy, kcal	2202 (1901,2502)	2604 (2280,2928)	402	18%
Fat,g	98.98 (82.46,115.51)	109.93 (89.89,129.96)	11.0	11%
CHO,g	272.19 (234.29,310.08)	339.02 (305.93,373.24)	66.8	25%
Protein, g	67.32 (57.38,77.25)	78.36 (67.73,88.99)	11.0	16%
Sugar,g	68.79 (53.16,84.41)	137.14 (114.46,159.81)	68.4	99%
Starch,g	182.96 (155.67,210.25)	199.07 (180.42,217.72)	16.1	9%
Dietary fibre, g	25.52 (21.49,29.56)	33.47 (30.11,36.84)	7.95	31%

MD, mean difference, (FFQ-WR)
%MD,percent mean difference, (mean FFQ-mean WR/mean WR) * 100
CI, confidence interval

Table 3.14 Comparison of mean daily energy and nutrient intakes between WR and FFQ in women (n=35)

Nutrient	Mean WR (95% CI)	Mean FFQ (95% CI)	MD	%MD
Energy, kcal	1647 (1487,1808)	1803 (1639,1967)	156	10%
Fat,g	70.95 (59.24,82.66)	75.81 (61.17,84.44)	4.86	7%
CHO,g	210.82 (191.9,229.8)	239.02 (215.5,262.5)	28.2	13%
Protein, g	54.77 (48.09,61.44)	56.86 (51.98,61.02)	2.09	4%
Sugar,g	59.28 (47.23,71.34)	103.08 (88.50,117.67)	44.6	74%
Starch,g	140.80 (130.1,151.5)	133.10 (119.4,146.8)	-6.8	-6%
Dietary fibre, g	20.60 (18.33,22.86)	24.42 (22.38,26.45)	4.13	18%

MD, mean difference, (FFQ-WR)
%MD, percent mean difference, (mean FFQ-mean WR/mean WR) * 100
CI, confidence interval

explained in the methodology (section 2.3.8). The adjustments made to the portion size of the meat group and the frequency to the fat group contributed to a better agreement between the WR and the FFQ.

The agreement between the two measures (FFQ vs WR) for energy, fat, carbohydrate and protein was also examined using the Bland -Altman technique (Bland and Altman,1986). The difference in nutrient value (FFQ-WR) between the dietary measures was plotted against their mean for each subject. Initially the scatter plot was plotted using logged data. However since there was no difference in the plots shown (probably due to a small sample size) regardless of the data used, the scatter plots using the raw data are presented. The reason for using the raw data is for an easier interpretation of the results later.

Figure 3.3a through Figure 3.3d show the Bland-Altman plots for energy, fat, carbohydrate and protein intakes in men. The plots for energy and fat revealed that the FFQ gave a higher estimate for energy and fat throughout the whole range of intake. The plots for carbohydrate and protein demonstrated a similar trend.

A similar Bland-Altman plot was carried out to examine the agreement between FFQ and WR in women. Figure 3.4a through Figure 3.4d show the plots for energy, fat, carbohydrate and protein intakes in women. The Bland-Altman plots showed that the FFQ gave a higher estimate for energy, fat and protein than WR at lower intakes but the FFQ gave a lower estimate than WR at higher intakes, thus there appeared to be a differential bias. The plot for carbohydrate demonstrated that throughout the whole range of intake, the FFQ gave a higher estimate.

Figure 3.3a   Bland-Altman plot for energy in men

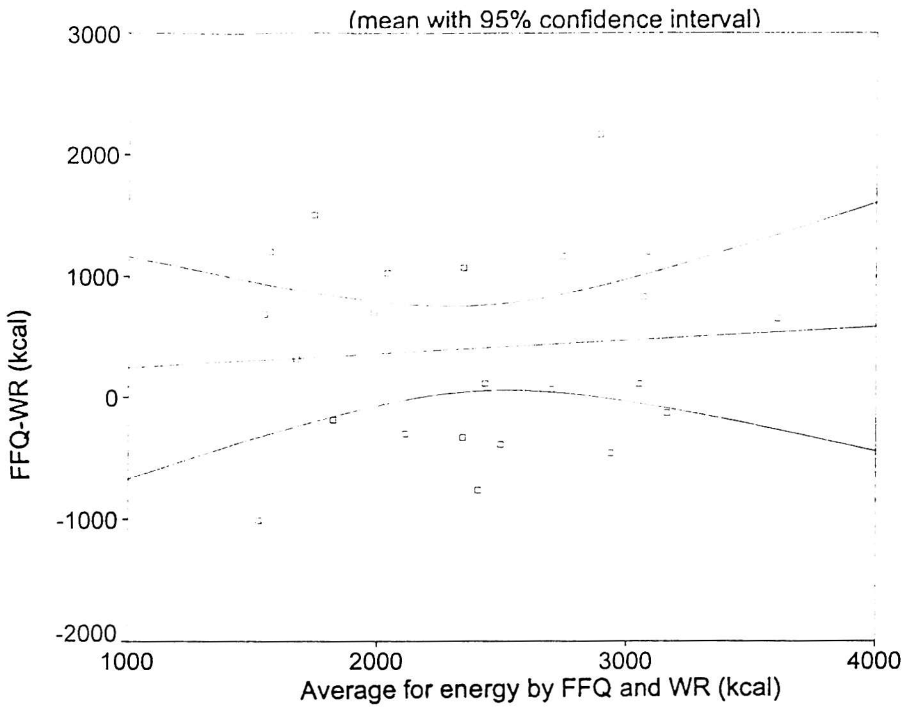


Figure 3.3b   Bland-Altman plot for fat in men

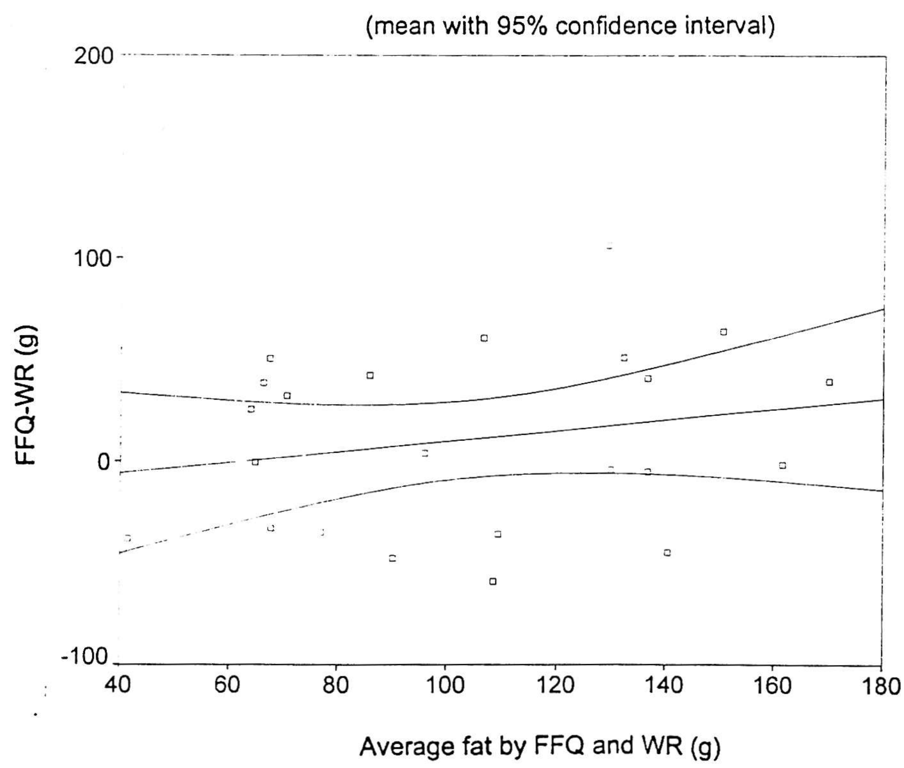


Figure 3.3c Bland-Altman plot for carbohydrate in men

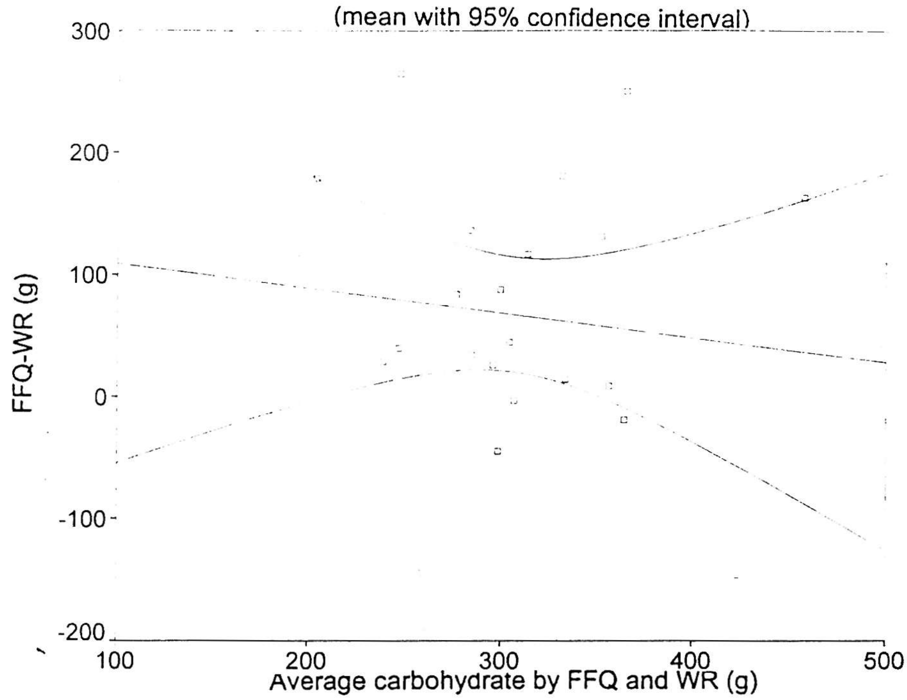


Figure 3.3d Bland-Altman plot for protein in men

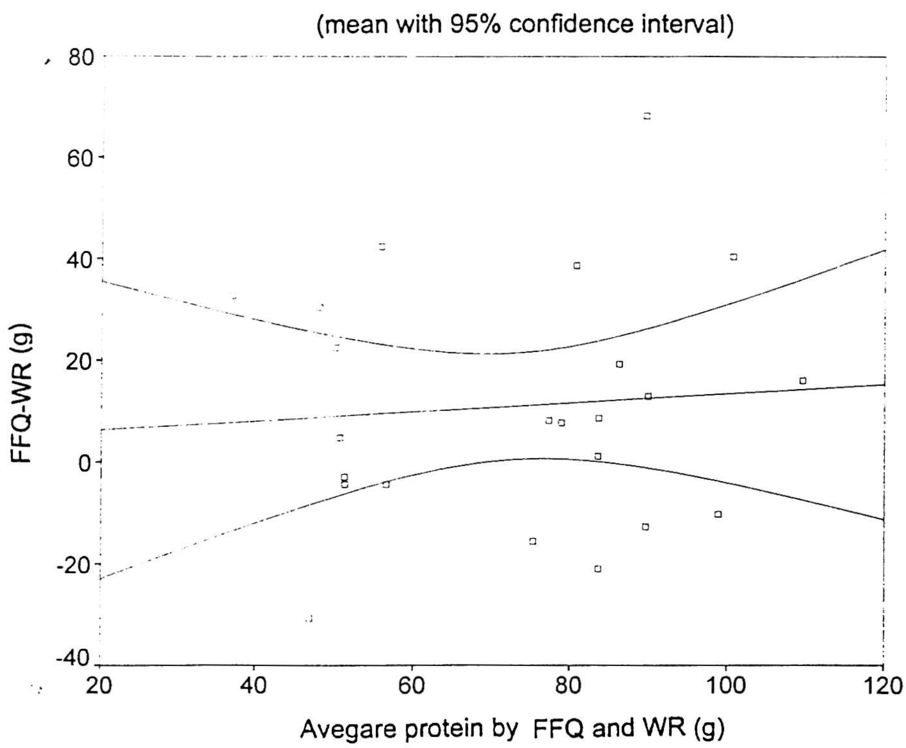


Figure 3.4a Bland-Altman plot for energy in women

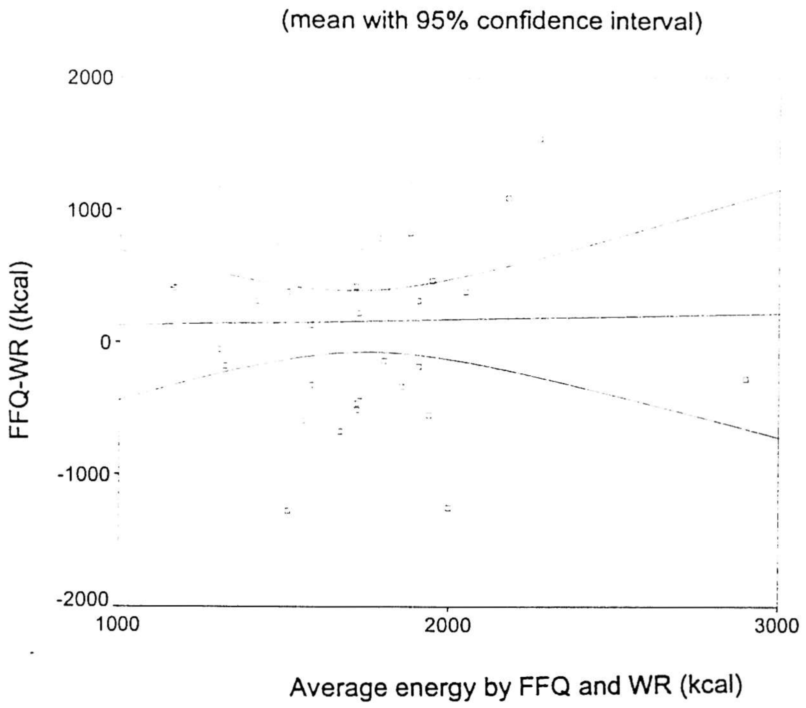


Figure 3.4b Bland-Altman plot for fat in women

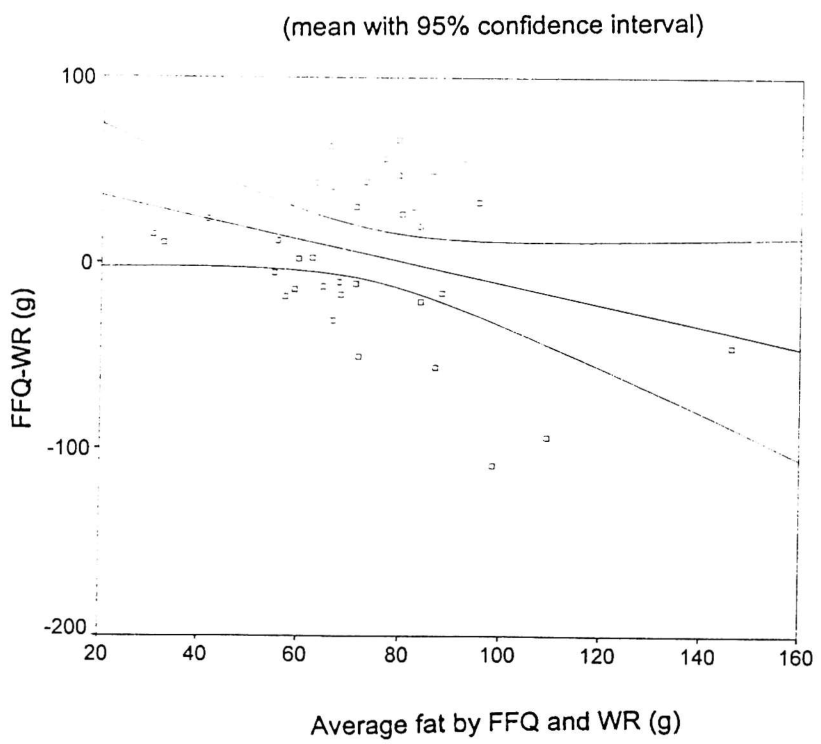


Figure 3.4c Bland-Altman plot for carbohydrate in women

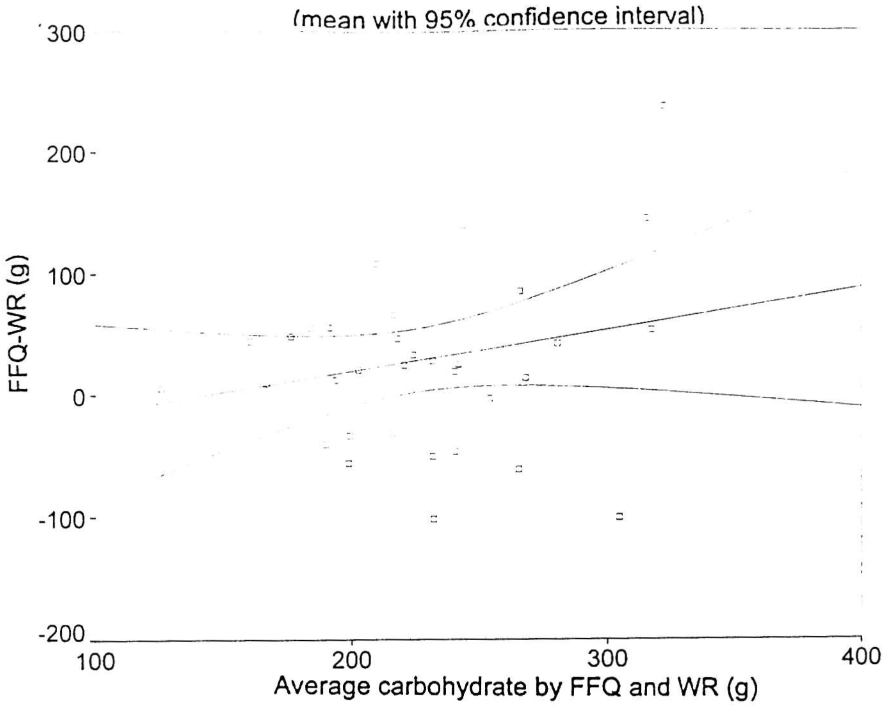
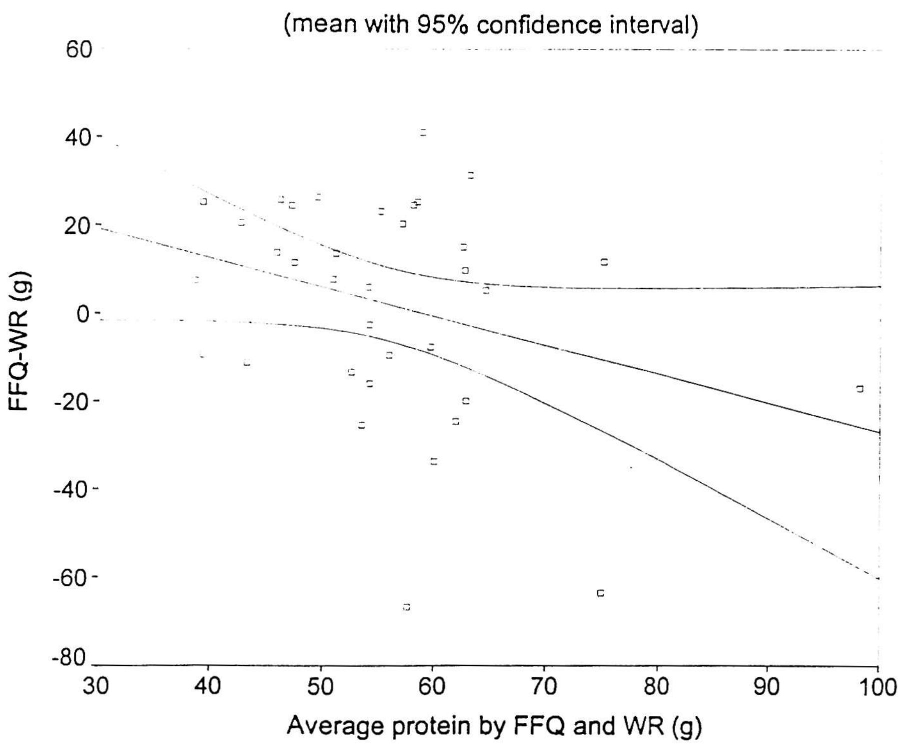


Figure 3.4d Bland-Altman plot for protein in women



The Spearman rank correlation coefficient was calculated for the FFQ and WR, both energy adjusted and unadjusted (Table 3.15). Rank correlation coefficient for energy adjusted and unadjusted were statistically significant for all the nutrients which were energy, fat, carbohydrate and protein only when the data for men and women were pooled together (most likely due to a larger number of subjects). The correlations ranged between 0.26 to 0.38 for the energy unadjusted while were between 0.25 and 0.36 for energy adjusted. The correlations of the selected nutrients analysed, except carbohydrate, were negative in women, but the correlations for all nutrients were positive in men. However, none of the correlations were significant in either men or women separately. For the energy adjusted, the correlations of the macronutrients were not ststistically significant except for protein in men and carbohydrate in women. Generally the correlations were higher for the adjusted energy nutrients than the unadjusted in women.

Table 3.15 Spearman rank correlation coefficients[a] for selected nutrients in the food frequency questionnaire and weighed record

Nutrient	All (n=58)		Men (n=23)		Women (n=35)	
	unadj[1]	adj[2]	unadj	adj	unadj	adj
Energy	0.29[a]		0.39		-0.16	
Fat	0.26[a]	0.25	0.46	0.10	-0.05	0.29
CHO	0.38[a]	0.36[a]	0.20	0.23	0.28	0.36[a]
Protein	0.27[a]	0.36[a]	0.55	0.44[a]	-0.17	0.23

[a] Two tailed tests for significance P<0.05.

[1] energy unadjusted nutrients

[2] energy adjusted nutrients (by method of Willett,1990)

The ranking of the subjects into the thirds of distribution was also evaluated. Table 3.16 shows the percent of subjects who were correctly classified and grossly misclassified between the two measures. Subjects are categorized as correctly classified if they are ranked in the same third of the distribution by both measures, while they are defined as grossly misclassified if they are ranked in the opposite third of the distribution by both measures. The level of agreement (%) in ranking of subjects revealed that between 15% to 29% of the women were grossly misclassified, while the percentage was lower in men (5% to 14%, mean 9%). The level of agreement between measures was also evaluated using the kappa statistics (Table 3.17). Kappa value would indicate the proportion of subjects similarly classified at each ranking of FFQ and WR correctly by chance agreement. The correlation between measures were poor ($k < 0.2$) for most nutrients in both men and women . Only carbohydrate ($k=0.23$) in women and protein ($k=0.29$) in men were weakly correlated (Altman,1994).

Table 3.16 Level of agreement (%) in ranking between WR and FFQ for selected nutrients

Nutrients	Men		Women	
	% GM*	%CC**	%GM*	%CC**
Energy	14	38	27	21
Fat	9	41	29	26
CHO	9	41	15	49
Protein	5	52	29	27

* GM Grossly misclassified **CC Correctly classified

Table 3.17 Kappa statistics (k) for selected nutrients between FFQ and WR

Nutrients	Men	Women
Energy	0.07	-0.19
Fat	0.12	-0.12
CHO	0.00	0.23
Protein	0.29	-0.10

The statistical analysis carried out (Spearman rank correlation, kappa statistics and ranking into thirds of distribution) to determine the association between the two measures in the calibration study demonstrated a consistent result. If one statistical analysis revealed a weak association or a poor agreement between the two measures, the other analysis showed a similar result.

According to Bellach (1993), if several statistical analyses could produce stable results, this could mean that the association or the agreement between measures observed was reliable. This was because no one analysis could determine the validity of a measure, as each analysis had its limitations.

Table 3.18 and Table 3.19 show the underreporters and overreporters of energy intake based on comparison with EI/BMR ratio. We used the cut off point of < 1.2*BMR as the indication of underreporters while > 1.8*BMR as the indication for overreporters. This is because a BMR of < 1.2*BMR and > 1.8*BMR are considered as an erroneous estimate of habitual intake (Bingham,1994). In women, on average 52% underreported their energy intake when reporting using the WR, while only 7% underreported energy intake when reporting using FFQ. The percentage of women overreporting using the WR was only 5% (mean) while the percentage was 21% (mean) when using the FFQ. Black (1996) in her review also reported that there was bias in underreporting energy intake among women, however, there was no bias in men.

Table 3.18 Underreporters and overreporters of energy intake by dietary methods in women

	Schofield BMR			Henry and Rees BMR		
Method	<1.2	1.2-1.8	>1.8	<1.2	1.2-1.8	>1.8
FFQ	3(7%)	32(74%)	8(19%)	3(7%)	30(70%)	10(23%)
WR	19(56%)	13(39%)	1(5%)	16(48%)	16(48%)	1(5%)

Table 3.19 Underreporters and overreporters of energy intake by dietary methods in men

	Schofield BMR			Henry and Rees BMR		
Method	<1.2	1.2-2.1	>2.1	<1.2	1.2-2.1	>2.1
FFQ	3(18%)	10(58%)	4(24%)	1(6%)	12(70%)	4(24%)
WR	7(41%)	9(53%)	1(6%)	7(41%)	9(53%)	1(6%)

In the present study, the trend in men was similar to that seen in women. There was a higher percentage (mean 41%) of men who were underreporters when reporting energy intake using WR compared with FFQ (mean 12%). The percentage of overreporters in men was smaller in the WR (mean 6%, with only one subject) than in the FFQ (mean 24%).

The comparison of energy intake of the dietary methods with EI/BMR ratio showed that the measures using the FFQ were more reliable and realistic, as a higher percentage of subjects were within the normal EI/BMR ratio. Measures using the WR tended to place more subjects as underreporters. Comparison with the energy requirement indicated that many subjects reported energy intake even below their energy requirement for both methods of assessment.

Generally, in the South Asian community, the estimation of energy intake using the WR demonstrated that there was a tendency for the subjects to underreport their food intake in both men and women. However, the underreporting was more serious in women. The FFQ method on the other hand was able to estimate energy intake rather well as an average of 72% (women) and 64% (men) reported an energy intake within an acceptable EI/BMR ratio.

Table 3.20 and Table 3.21 show the underreporters and overreporters of energy intake calculated from the mean of FFQ and WR estimates by body mass index (BMI). In women, as their BMI increased, the percentage of underreporters also increased. As the BMI decreased, the percentage of overreporters increased. There was no specific trend in men, although the number in each group was very small.

Table 3.20 Underreporters and overreporters of energy intake by body mass index (BMI) in women

BMI	Classification	n	<1.2 BMR	1.2-1.8 BMR	>1.8BMR
<20	underweight	5	0(0%)	0(0%)	5(100%)
20-25	normal	17	8(47%)	6(35)	3(18%)
25-30	overweight	13	6(46%)	3(23%)	4(31%)
>30	obese	8	5(63%)	2(25%)	1(12%)

Table 3.21 Underreporters and overreporters of energy intake by body mass index in men

BMI	Classification	n	<1.2BMR	1.2-2.1 BMR	>2.1 BMR
<20	underweight	2	2(100%)	0(0%)	0(0%)
20-25	normal	11	4(36%)	6(55%)	1(9%)
25-30	overweight	4	1(25%)	3(75%)	0(0%)
>30	obese	0	0(0%)	0(0%)	0(0)

In this study, it would appear that underweight women overreported their energy intake while overweight and obese women underreported their energy intake. Several other studies have reported similar findings (Heitmann and Lissner,1995; Livingstone et al,1990). However it might not be appropriate to generalize this to the general South Asian population as the sample size was small.

3.3 Results of the food checklist and 24 hour recall calibration study

The second part of the calibration study was carried out between August 1995 and February 1996. In this calibration study, only women subjects were recruited. The reason for this has been explained in the methodology (section 2.3.3). Fifty nine subjects were recruited including 23 women subjects from the first calibration study. However, 15 (25%) subjects dropped out by the end of the study, leaving 44 women who completed the food checklist and the 24 hr recall.

In the second calibration study, the newly recruited subjects (n=21) did not complete any weighed record. These subjects only completed the food checklist and the 24 hr recall. The subjects who were recruited into the second calibration study (n=27) from the first calibration study however, completed the weighed record, food checklist and the 24 hr recall. Therefore in the comparisons of energy and nutrient intakes, the results of the completed weighed record (n=35) from the first calibration study was used (Table 3.23).

Table 3.22 shows the demographic characteristics of responders and non responders. Generally both the responders and non responders were married (>50%) and had less than six children. However, 82% of the responders had at least six people living in the household. In contrast, all the non responders had at the most, five people living in the household.

In terms of ethnicity, as in the first calibration study, Pakistani women (66%) made up the majority of responders, followed by the Gujeratis (25%) and the Punjabis (9%). As in the first calibration study, there was a higher percentage of Punjabi women non responders. A majority of the responders (84%) and non responders (100%) had been living in the UK for more than 15 years and most had 6 years of education (86% responders vs 87% non responders).

The demographic characteristics demonstrated that a majority of the women who participated in this calibration study were married, had five or less children but probably living with extended family. In contrast there were as many married and single women who were non responders. All of the non responders lived in a smaller household compared to the responders.

Table 3.22 Demographic characteristics of responders and non responders in the second calibration study

		Responders	Non responders
Sex	Female	44 (75%)	15 (25%)
Marital status	Married	34 (77%)	8 (53%)
	Single	7 (16%)	7 (47%)
	Divorced	3 (7%)	0 (0%)
No of children	0-5 children	44 (100%)	15 (100%)
	6-10 children	0 (0%)	0 (0%)
No in household	0-5 person	8 (18%)	15 (100%)
	6-10 person	36 (82%)	0 (0%)
Ethnic	Gujerati	11 (25%)	3 (20%)
	Pakistani	29 (66%)	2 (13%)
	Punjabi	4 (9%)	10 (67%)
Live in UK	< 5 years	0 (0%)	0 (0%)
	6-10 years	5 (11%)	0 (0%)
	11-15 years	2 (5%)	0 (0%)
	> 15 years	37 (84%)	15 (100%)
Education	0-6 years	6 (14%)	2 (13%)
	> 6 years	38 (86%)	13 (87%)
Age	Female	40 years (22-65 years)	19 years (16-28 years)

The women subjects who participated in the calibration studies (both the FFQ and the food checklist) were recruited mainly from women's group meetings held at the South Asian community centres in Southampton. There were two different groups of women who attended the meeting. One group consisted mainly of elderly and a smaller number of younger Gujerati women. The second group consisted of women from the three main ethnic groups, Pakistani, Punjabi and Gujerati, however these women were younger and predominantly Pakistanis. Thus it was not surprising that there were more Pakistani women being recruited into the study. However, there was no obvious reason for the higher drop outs among the Punjabis.

Among the non responders, almost half (47%) were single women. These women were mostly rather young (< 20 years old), were born in the UK and therefore literate in the English language. However, through personal interaction with them, it appeared that they were not committed and not interested to participate in the study in the first place, thus many dropped out eventually. Although these young women consented to take part in the study, this was likely to be due to a feeling of obligation as well as not wanting to disappoint the researcher. Married women who dropped out gave family commitment (domestic duties and caring for children) as their reason for not completing the study.

3.3.1 Comparison of food checklist (CL), 24 hr recall (RCL) with the 4-day weighed record (WR)

Table 3.23 represents the comparison of the mean daily nutrient intakes between food checklist and WR as well as 24 hr recall and WR respectively. The food checklist tended to give higher estimates for energy, fat, carbohydrate and sugar than WR, while lower estimates for starch. The 24 hr recall however gave lower estimates for energy, fat and starch but higher estimates for sugar than WR. Estimates of protein and fibre were comparable between the three measures.

The mean differences between measures as well as the percent mean difference were calculated. These results are also shown in Table 3.23. Generally the percent mean difference between measures were within 10% for energy (8% CL, 3% RCL), fat (8% RCL), protein (4% CL, 1% RCL), starch (10% CL, 8% RCL) and fibre (2% CL, 9% RCL). However the percent mean difference between measures for sugar were 51% (CL vs WR) and 35% (RCL vs WR) repectively. When we compared the percent mean difference between measures, the test measures (FFQ,CL,RCL) vs the reference measure (WR), it would appear that all the three test measures gave similar estimates of the nutrient intakes as more than half of the nutrients analysed were within 10% mean difference.

To examine further the agreement between food checklist and WR as well as 24 hr recall and WR, Bland-Altman plots were carried out. Figures 3.5a through 3.5d show the plots for energy, fat, carbohydrate and protein intakes of food checklist vs WR. The plots for energy , fat and protein showed that the food checklist gave higher estimates for energy, fat and protein than WR, but at higher intakes the mean difference between measures was greater. The plot for carbohydrate revealed that at lower intakes, WR estimated carbohydrate higher than food checklist while at higher intakes, the WR estimated lower than the food checklist.

In the Bland-Altman plots for 24 hr recall vs WR (Figure 3.6a through Figure 3.6d), the plots for energy, fat and carbohydrate showed that at the lower end of intakes, the WR estimated higher energy, fat and carbohydrate than 24 hr recall, while at the higher end, the 24 hr recall estimated higher than the WR. However, the plot for protein did not show any trend. Throughout the whole range of intakes, there was almost equal higher estimates of protein by both

Table 3.23 Comparison of mean daily intake by weighed record (WR) and 24 hr recall (RCL) and food checklist (CL) with the 95% confidence interval (CI) in the second calibration study

Nutrient	Mean WR (95%CI)	Mean RCL (95%CI)	Mean CL (95%CI)	MD RCL	%MD RCL	MD CL	%MD CL
Energy, kcal	1647 (1487,1808)	1605 (1466,1744)	1782 (1628,1937)	-42	-3	135	8
Fat,g	70.95 (59.24,82.88)	65.45 (57.60,73.29)	83.42 (75.79,91.05)	-5.5	-8	12.5	18
CHO,g	210.82 (191.8,229.8)	212.91 (192.8,233.0)	219.22 (197.7,240.8)	2.1	1	8.4	4
Protein, g	54.77 (48.09,61.44)	55.29 (49.13,61.45)	52.61 (48.26,56.96)	0.5	1	-2.2	-4
Sugar,g	59.28 (47.23,71.34)	80.00 (67.17,92.28)	89.33 (75.67,102.99)	20.7	35	30.1	51
Starch,g	140.80 (130.1,151.5)	129.61 (117.6,141.6)	127.91 (115.7,138.7)	-12.9	-9	-13.6	-10
Fibre,g	20.60 (18.33,22.86)	18.76 (16.11,21.41)	20.29 (18.24,22.34)	-1.8	-9	-0.3	-2

MD, mean difference, (CL-WR);(RCL-WR)

%MD, percent mean difference, mean CL-mean WR/mean WR*100;

mean RCL-mean WR/mean WR*100

Figure 3.5a Bland-Altman plot for energy in women

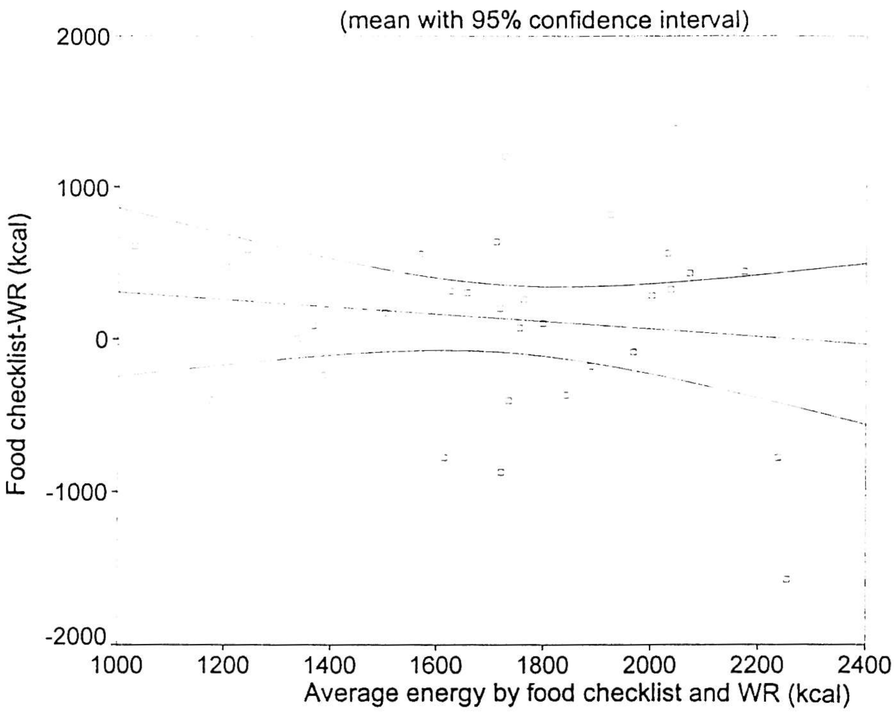


Figure 3.5b Bland-Altman plot for fat in women

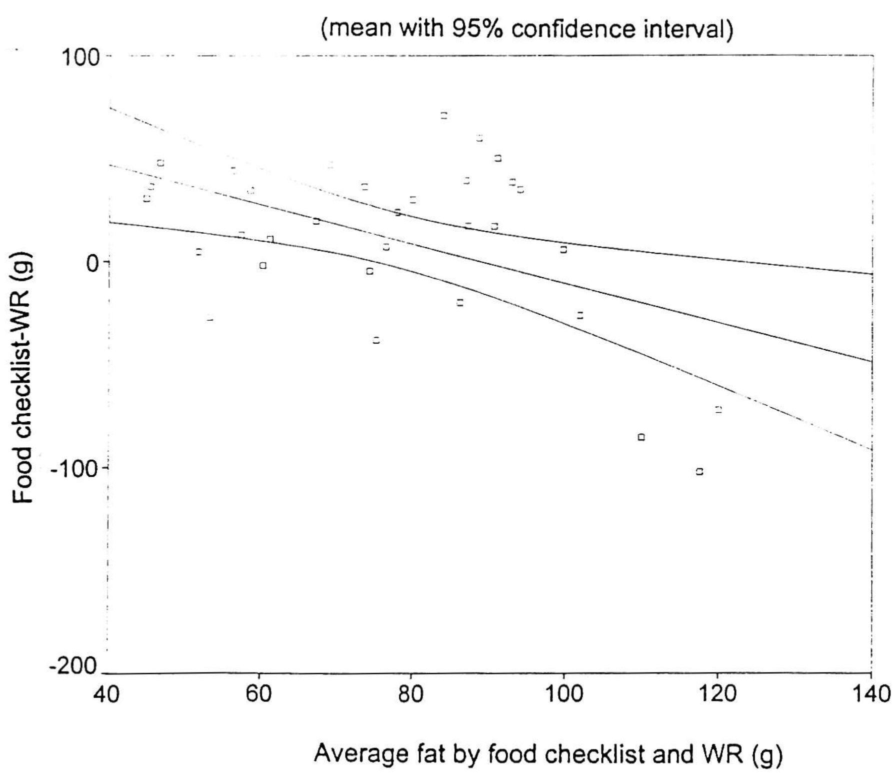


Figure 3.5c Bland-Altman plot for carbohydrate in women

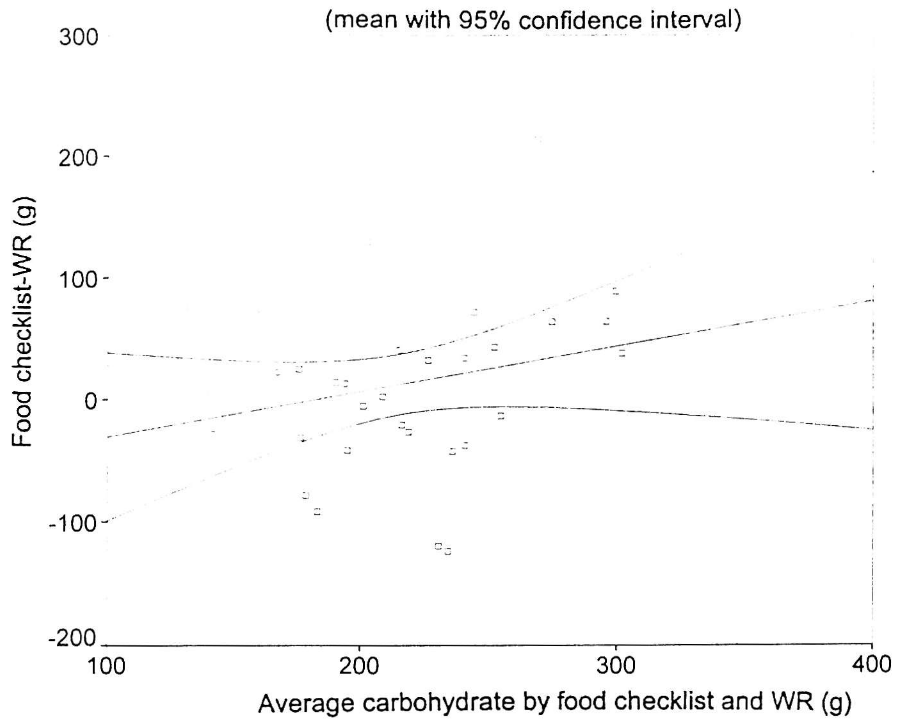


Figure 3.5d Bland-Altman plot for protein in women

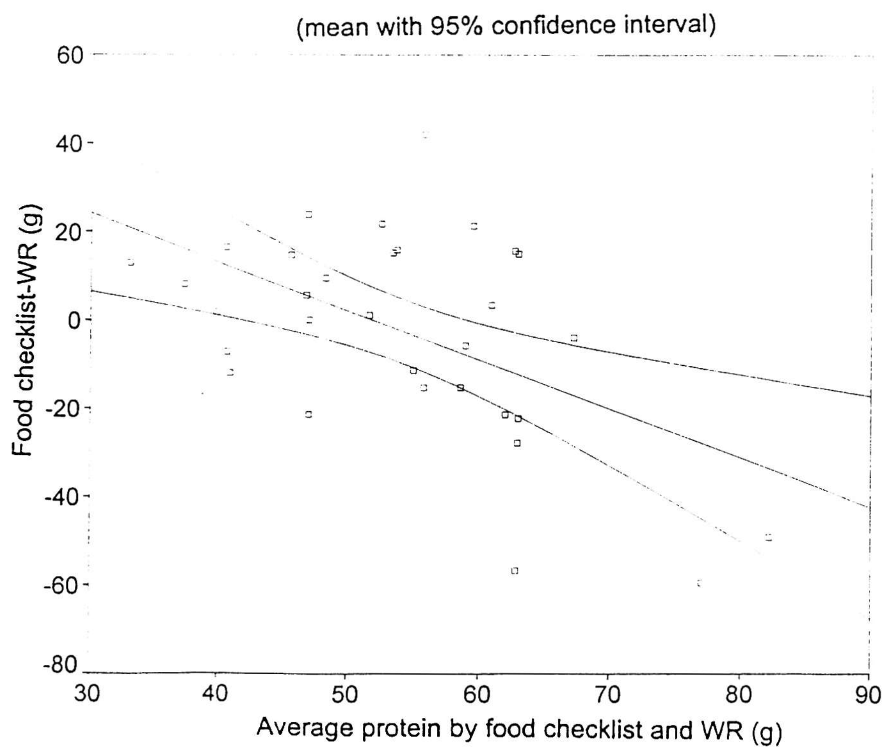


Figure 3.6a Bland-Altman plot for energy in women

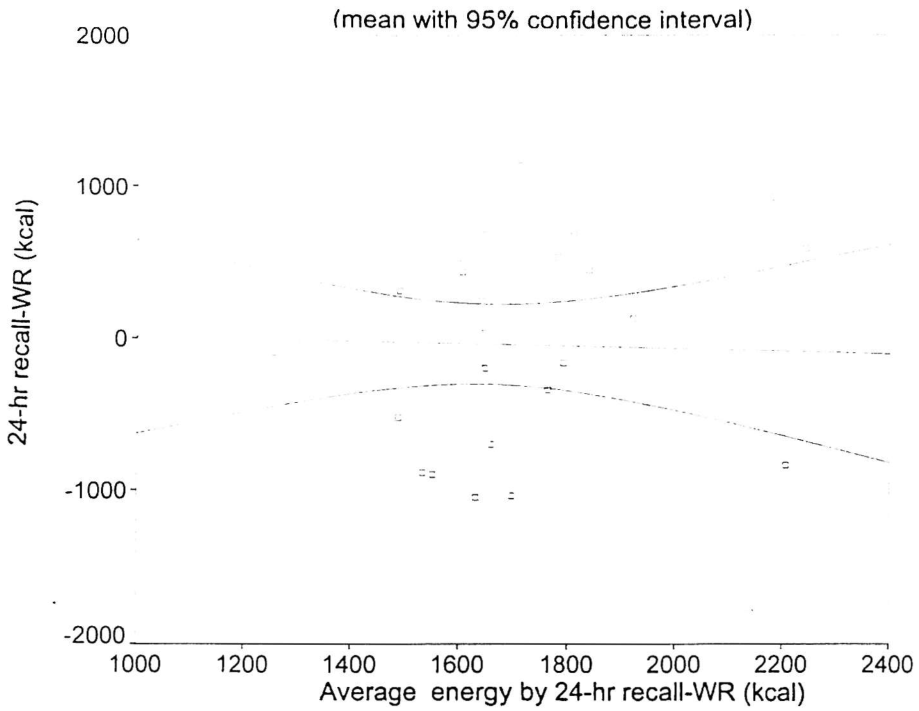


Figure 3.6b Bland-Altman plot for fat in women

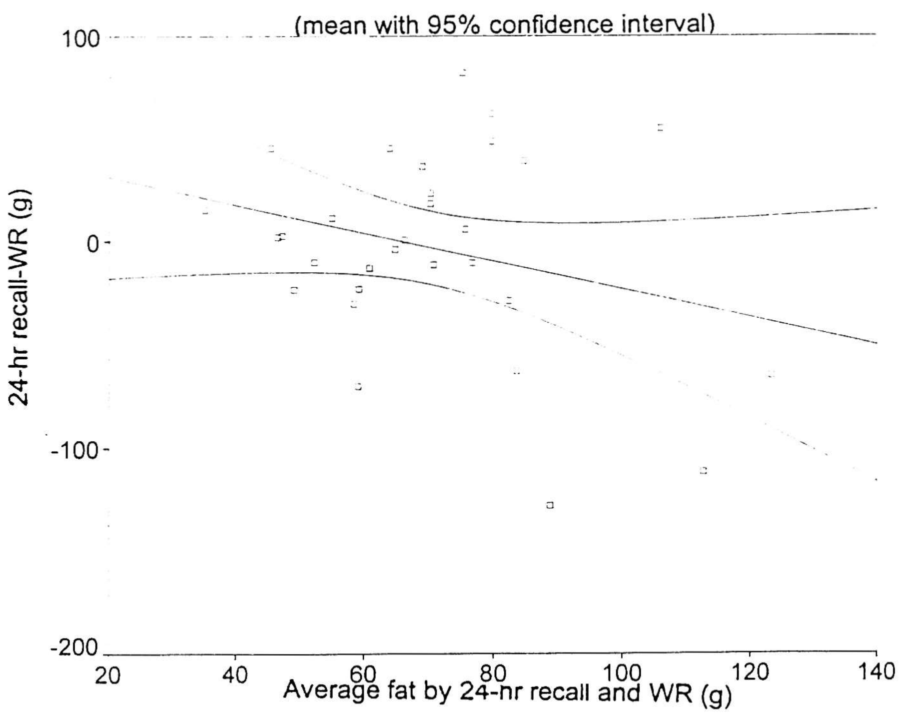


Figure 3.6c Bland-Altman plot for carbohydrate in women

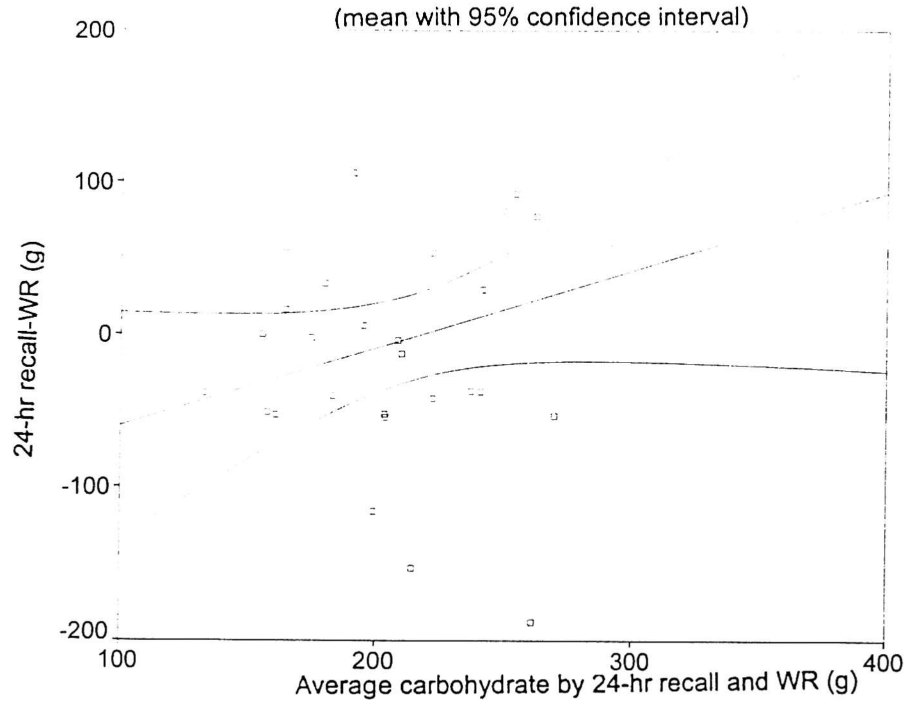
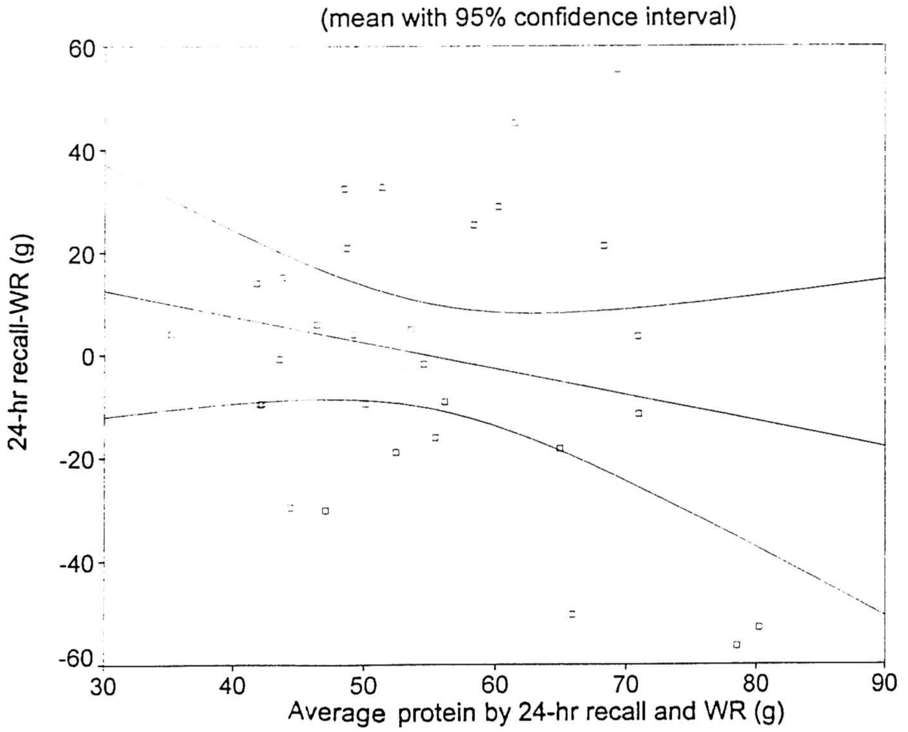


Figure 3.6d Bland-Altman plot for protein in women



the 24 hr recall and WR.

The Spearman rank correlation coefficient for food checklist and WR as well as 24 hr recall and WR are tabulated in Table 3.24 for both energy adjusted and energy unadjusted nutrients. The correlations between measures ranged between 0.05 to 0.26 for energy adjusted and -0.02 to 0.32 for energy unadjusted. However, none of the correlations were significant.

Table 3.24 Spearman rank correlation coefficients for selected nutrients in the food checklist and 24 hr recall with weighed record

Nutrient	Food checklist		24 hr recall	
	unadj[1]	adj[2]	unadj	adj
Energy	0.10		-0.15	
Fat	0.16	0.16	-0.17	0.22
Carbohydrate	0.32	0.06	0.26	0.26
Protein	-0.02	0.05	-0.15	0.22

No correlations were statistically significant.
[1] energy unadjusted nutrients
[2] energy adjusted nutrients

Table 3.25 shows the level of agreement in the ranking of subjects between measures. In general the food checklist method had the highest percentage of subjects correctly classified for all nutrients, except for protein. The food checklist method also had the lowest percentage of subjects grossly misclassified for all nutrients. Both the FFQ and 24 hr recall methods could only classify between 21% to 49% of the subjects into the thirds of the distribution, while grossly misclassified between 14% to 29%. The kappa statistics between WR and 24 hr recall and WR and food checklist are tabulated in Table 3.26. The correlations between WR and 24 hr recall were poor ($k<0.2$) for all nutrients, while the correlations between WR and food checklist were poor for all nutrients, except energy

(k=0.24)

Table 3.25 Level of agreement (%) in ranking between food checklist (CL) and 24 hr recall (RCL) with weighed record

Nutrients	CL		RCL	
	%GM[a]	%CC[b]	%GM	%CC
Energy	13	56	23	27
Fat	13	44	27	23
CHO	13	45	14	35
Protein	16	19	25	28

[a] %GM, Grossly misclassified, subjects are ranked in the opposite thirds of the distribution

[b] %CC, Correctly classified, subjects are ranked in the same thirds of the distribution

Table 3.26 Kappa statistics (k) for selected nutrients between 24 hr recall and WR and food checklist and WR

Nutrients	Recall	Food Checklist
Energy	-0.10	0.24
Fat	-0.15	0.16
Carbohydrate	0.01	0.17
Protein	-0.08	-0.20

The energy intake estimated by the 24 hr recall method and the food checklist were also compared against the EI/BMR ratio. Table 3.27 show the underreporters and the over-reporters of energy intake by dietary methods based on the comparison with the EI/BMR ratio. About half of the subjects underreported their energy intake when reporting with the WR (mean,52%) and 24 hr recall (mean,52%). Only 7% of subjects underreported with the FFQ. On the other end, an average of between 21% and 20% of the subjects overreported the energy intake with the FFQ and food checklist methods respectively. However among the dietary methods used to estimate energy intake, the FFQ appeared to have the most number of subjects which could be categorized as having a normal EI/BMR ratio .

Table 3.27 Underreporters and overreporters of energy intake by dietary methods

Method	Schofield BMR			Henry and Rees BMR		
	<1.2	1.2-1.8	>1.8	<1.2	1.2-1.8	>1.8
FFQ	3(7%)	32(74%)	8(19%)	3(7%)	30(70%)	10(23%)
CL	14(32%)	21(49%)	8(19%)	12(28%)	22(51%)	9(21%)
RCL	23(53%)	18(42%)	2(5%)	22(51%)	17(40%)	4(9%)
WR	19(56%)	13(39%)	1(5%)	16(48%)	16(48%)	1(5%)

3.4 Results of measurement errors using the method of triads

The measurement errors in both calibration studies (FFQ and food checklist) were estimated using the validity coefficient with the method of triads (Kaaks,1996). The measurements to be evaluated were FFQ, food checklist and 24 hr recall as the test measurements, 4-day weighed record as the reference measurement, while the marker was the energy requirement calculated from BMR multiples. We evaluated the validity coefficient for both men and women. The results are shown in Figure 3.7a (for men) and Figure 3.7b (for women) for the FFQ calibration study. Figure 3.7c and Figure 3.7d show the results of the food checklist and 24 hr recall calibration study in women.

In men, the sample correlations were moderate (0.39) for the association between FFQ and WR. However, the correlation was low and negative for BMR multiples and FFQ as well as for BMR multiples and WR respectively (-0.19 vs -0.14). The estimated validity coefficient were 0.73 for the FFQ and 0.54 for the WR. The validity coefficient for BMR multiples was low (0.26).

In women, the sample correlations were low and negative between FFQ and WR. However, the correlation for FFQ and BMR multiples was 0.22, while WR and BMR multiples was 0.08. Due to the negative correlation in one of the sample correlations in women, the method of triads could not estimate the validity coefficient as this would require taking the square root of a negative value.

In the 24 hr recall calibration study, there was also a negative sample correlation observed. This was between the 24 hr recall and WR (-0.14). The other two sample correlations were 0.24 for the association between 24 hr recall and BMR multiples, and 0.08 for WR and BMR multiples.

For the food checklist calibration study, sample correlations were very low for food checklist and WR (0.05) and WR and BMR multiples (0.27). Therefore the estimated validity coefficient were 0.41 for food checklist, 0.12 for WR and 0.66 for BMR multiples. When using the method of triads, it might still be possible to get a high validity coefficient, although all three sample correlations were low. However, small differences in sample correlations that were low might result in large differences in the estimated validity coefficient of the three types of measurements.

CHAPTER FOUR

DISCUSSIONS

CHAPTER 4 DISCUSSION

There are two main sections in this chapter. The first section discusses the general introduction of the chapter (4.1). This is followed by discussions on the meal patterns of South Asians (4.1.1) and energy and macronutrient sources in the South Asian diet (4.1.2). Section 4.1.3 focuses on the development of the FFQ and followed by the discussion on the FFQ calibration study (4.1.4). In the fifth section, the calibration study of the food checklist and 24 hr recall are discussed (4.1.5). The final section discusses the bias in the study design (4.2) and measurement errors in the calibration studies (4.2.1).

4.1 General introduction

The general objective of the study was to develop a dietary assessment method which could be used to evaluate dietary intake of the South Asian community. Specifically, we wanted to design a method which could determine the longterm usual intake of this South Asian community. This method then could be used to determine the relationship between diet and disease (specifically diabetes) in the South Asian population. This is because diabetes is one of the major health problems among the South Asians. According to two large diabetes surveys in the UK, South Asians seemed to have a prevalence of diabetes, which was two to four times higher than the Europeans (Mather and Keen,1985; Simmons et al,1989). Although the actual cause(s) of diabetes remains unknown, several factors such as intrauterine growth retardation, body fat distribution, physical activity and diet might be involved in the aetiology (Hales et al,1991;Phipps et al,1993;McKeigue et al,1991;Randrianjohany et al,1993).

The role of diet in the aetiology of diabetes remains unclear. However, macronutrients especially carbohydrate and fat have been suggested to influence the development of diabetes. Dietary fibre was suggested to delay gastric emptying and caused a smaller rise of glucose in the blood during oral glucose tolerance test (Ray et al,1983). This would therefore have a protective effect. The type of fat in the diet might affect the development of insulin resistance. Diets which were high in saturated, monounsaturated or polyunsaturated (omega-9) fatty acids could lead to severe insulin resistance (Waldhausl et al,1989). Insulin resistance could lead to hyperglycaemia, which is a symptom of diabetes.

Thus we expected that the dietary method we developed could rank intake of macronutrients that might have a role in diabetes. This ranking should be able to differentiate those who have a high or low intake with sufficient accuracy to be confident that the observed effect or lack of effect is not simply due to chance alone. This was the ideal. The extent to which we achieved this ideal is discussed below.

In the present study we developed two dietary assessment methods which were the FFQ and the food checklist. Together with the 24 hr recall, these methods were calibrated against a 4-day weighed record (reference method). We have also used a proxy of a biomarker which was the BMR multiples to check the relative validity of the energy intake estimated by the test methods and the reference method.

Ideally, to achieve 80% power at 5% significance level in detecting a 0.3 correlation (energy) between the test and the reference methods, a total of 168 subjects (84 subjects per gender) was required in the calibration study. However, in the FFQ calibration study, only 58 subjects (23 men, 35 women) completed both the FFQ and the weighed record. This gave an overall response rate of only 35% (27% in men and

42% in women). This low recruitment number thus reduced the power of the study from 80% to less than 50% (Machin and Campbell,1987).

In the food checklist and 24 hr recall calibration study, only 44 subjects (all women) completed the study, giving a response rate of 52%. Even though the response rate was higher compared with the FFQ calibration study, the correlation between the test and the reference method did not reach statistical significance. This was because the power in the checklist calibration study was also reduced.

The problems that were encountered with recruitment and with the weighing of foods were highlighted and discussed in the methodology. Briefly these problems included the language problems as well as the high rate of illiteracy in the South Asian community. The subjects also tended to take a much longer time to complete their weighed record. These could have resulted in selection bias in the subjects who participated, as literate individuals were more likely to be recruited into the study. It was also possible that there was bias in the days subjects recorded their weighed record. Other reasons for the poor outcome could be contributed to the reference method (WR) that we used. A simpler method like the multiple 24 hr recall could likely increase the recruitment number of the study and possibly would improve the outcome of the study.
The Spearman rank correlation between FFQ and WR showed that the correlations in the selected nutrients except carbohydrate were higher in men than in women for the energy unadjusted, even though the number of subjects was lower in men compared with women. This might imply that in general men recorded their WR and responded to the FFQ better than women. However, in the energy adjusted nutrients, the correlations were higher in the women than men, except protein. The level of agreement in ranking between WR and FFQ for selected nutrients in men was also

better than women. The gross misclassification of subjects was much lower in men compared to women. Therefore when the data for both men and women were analysed together, the Spearman correlation indicated a significant relationship. The significant correlations observed could in part be due to the larger sample size and increased statistical power.

The WR as the reference method did not perform as assumed (the assumption was that WR has a high accuracy of measuring food intake). When the energy intake estimates from the test and reference methods were compared with BMR multiples, more subjects were shown to underreport their energy intake with the WR compared with the FFQ, food checklist and the 24 hr recall. This suggested that although WR has been regarded as an accurate method to estimate nutrient intake, however, in the South Asian community it did not appear so. This could have been partly explained by the general difficulty that the South Asian community had when completing the WR.

The low response rates in the calibration studies indicated that those subjects who responded were likely to be more motivated and more interested in the study than the population as a whole. It could be possible that the results of this study could be an overestimation of the association between the test and the reference method. The FFQ and the food checklist that we developed had the potential to be a useful dietary assessment method to evaluate usual dietary intake in the South Asian community. However, further development of the FFQ and the food checklist in this population is needed. Therefore some caution is required in interpreting the results in the present study.

4.1.1 Meal pattern of South Asians

Few studies have evaluated the dietary habits (Hunt,1977;

Attariwala,1977; Wharton and Eaton,1983) and nutrient intake of South Asians (Sevak et al,1994; Smith et al,1993). However, the methodology used to determine nutrient intake in these studies estimated the absolute intakes. In the present study, the concern was more in the ability of the method to evaluate the relative intakes and particularly be able to rank the subjects into thirds of the distribution. For this reason, we selected the FFQ as our test method. Furthermore to the best of our knowledge, there is no FFQ which has been developed for dietary evaluation of the South Asian community.

Since the present study was the first study to develop a FFQ for South Asians, several steps were required for the development of the FFQ. Firstly, there was no database on foods or food list which could reflect the food sources of the main energy and macronutrients in the South Asian diet. Although there have been studies which have evaluated the meal pattern of South Asians (Kalka,1988; Wharton and Eaton,1983), these studies did not provide any information on the energy and macronutrient sources in the South Asian diet. Thus the preliminary survey carried out was useful. We were able to determine the meal pattern and evaluate the food sources of energy and macronutrients in the South Asian diet.

Our results demonstrated that the meal pattern of the South Asians in Southampton was similar to other South Asians who have settled elsewhere in the UK. This would suggest that the meal pattern of South Asians was influenced by their origin in the Indian subcontinent. The meal pattern also showed that the foods consumed were still very traditional, with the exception of breakfast which was most westernized. This further implied that the meal pattern had been set before migration and still maintained regardless of where they settled in UK. Earlier studies in the 70's confirmed that South Asians still retained their meal patterns after migration especially when they migrated as adults (Hunt,1977; Attariwala,1977). However, there is a need to stress that there were subtle differences in the method of food preparation and food consumption between ethnic groups of the South Asians. For example, the use of ghee (clarified butter), usually for frying spices was most common among the Gujeratis, while not so common among the Punjabis and Pakistanis. However, for deep frying, all ethnic groups used either corn oil or vegetable oil. The difference in the food consumption between these ethnic groups was whether ghee or butter was spread on the chapatti or not before being eaten. The Gujeratis were more likely to consume their chapatti with a little butter or ghee on it while the Punjabis and Pakistanis consumed their chapatti dry.

Recently, Kassam-Khamis et al (1995) showed that the recipes (method of food preparation) of the South Asian ethnic groups differed. This variation in food preparation also occured within the same ethnic group (Attariwala, 1977). The vegetarian Gujeratis consumed a greater variety of dishes than the Punjabis. More significantly, despite common dishes with similar ingredients of both Gujeratis and Punjabis, the recipes of the Gujeratis appeared more often to have a higher fat content (weight in gram) than the recipes of the Punjabis.

The implication of these subtle differences in the method of food preparation and consumption on the nutrient intakes of South Asians is still not known. The consumption of ghee by the Gujeratis would suggest that they would be likely to consume more saturated fats in the diet compared with the Punjabis and Pakistanis. Although we did not analyse the type of fatty acids present in the diet of our subjects in the preliminary survey, Smith et al (1993) have shown that there were significant differences in the polyunsaturated fatty acid, total fat intake and P/S ratio between two

South Asian religious subgroups, which were the Pakistani Muslims and Gujerati Hindus. However, no differences were demonstrated in the saturated fatty acid and monounsaturated fatty acid intakes between the two groups.

As demonstrated in this study as well as various other dietary habit studies, there were differences in the food preparation of the different South Asian ethnic groups. Whether these differences could result in very different macronutrient intakes remains to be further investigated. However, when more studies assessing these differences have been carried out, this may explain the inconsistencies in the results of South Asian dietary studies.

4.1.2 Energy and macronutrient sources in the South Asian diet

As mentioned in the last section, the lack of a database on the food sources of energy and macronutrients in the South Asian diet made the determination of this list imperative in the present study. The three food items mostly consumed by the South Asians were chapatti, meat curry and milk and milk products. These food items contributed at least 30% of the total energy and fat intake in the South Asian diet. Sources of energy and fat in the English diet were contributed by white bread, milk, potatoes, beer and butter (Cade and Margetts,1988)

In the development of an FFQ, it is necessary that the foods which represent the intake of the population to be studied are included in the FFQ. Our results showed that the food list in the South Asian diet was different from the list in the English diet (Cade and Margetts,1988). This suggested that when designing an FFQ for ethnic groups with different dietary patterns or even for a population in different geographical areas, it was important that the foods which contributed to at least 90% of the total energy intake was determined for that group. These foods should then be included in the FFQ. We included foods which contributed to at least 90% of the total energy intake because an adequate capture of total calories would enhance the ability to cover a wider range of nutrients. In order to achieve that, this food list should contain foods representing 90% of the South Asians total caloric intake (Block et al,1986). The inclusion of inappropriate foods would distort the actual food or nutrient intakes of the study population. Since the subjects in this preliminary study consisted of Gujeratis, Punjabis and Pakistanis, the foods identified in our food list could be used to develop FFQ for these three groups.

The shorter list of foods in the South Asian diet reflected that they generally consumed a narrower range of food items. The meal pattern of the South Asians showed that chapatti and meat curry were eaten at least twice a day, at lunch and dinner (sometimes even for breakfast). This further suggested that the foods consumed were still traditional in nature and were homogeneous and rather monotonous. This meal patterns had also been shown by many other studies of South Asians in the UK since the 1970's (Hunt,1977; Attariwala,1977;Wharton and Eaton,1983;Kalka, 1988). Although the sample size in our preliminary study (n=62) was too small to generalize to the South Asian community, our findings were consistent with those of others. Collectively, our study and studies of others could represent the general meal pattern of South Asians. However, ideally a study with a large number of subjects would be likely to be more representative.

4.1.3 Development of the FFQ

Food frequency questionnaires have been used in many epidemiological studies because it is simple, easy to administer and requires minimal effort from the subjects

(Boeing et al, 1989). Furthermore, it can be administered quickly in person or self administered by mail to a large number of subjects at a lower cost (Nelson,1996). However, it is critical that any FFQ used in the sampling population has been validated first in that population.

The FFQ developed in the present study consisted of seventy five (75) food items. Besides including the food items which cumulatively contributed up to 90% of the total energy intake in the FFQ, food items which were regularly eaten by a minority of South Asians have also been added to the list. This was necessary as excluding the food items which were less frequently eaten would mean underestimating some specific nutrients that might be contributed by these foods. According to Block et al (1986), food items selected should be based on their contribution to intake at group level, and this technique was applied for the development of our FFQ.

We outlined the food items according to food groups. The FFQ list started with the staple foods (chapatti, paratha and bread), was followed by the meat group, milk group, fats and oils, fruits and vegetables, beverages and miscellaneous food items. This order of food items followed the South Asian meal pattern rather closely. Thus theoretically we hoped that this would help the subjects to remember average intake more easily. Haraldsdottir (1993) has suggested asking global categories of foods first then followed by specific foods. This could prevent the subjects from having the tendency to overestimate consumption frequency of single food items.

The order of frequency response was in increasing order (never to twice or more per day). Based on an earlier study which compared the effect of various FFQ design on nutrient estimates, a decreasing order of frequency response was suggested to result in higher estimates of nutrient intake

(Wolk et al,1994). We could not determine whether our FFQ would estimate the nutrients accordingly. However, since the design of the frequency response was in increasing order, this in theory should not bias to higher estimates.

It has been reported that an open ended questionnaire layout (exact frequency of intake) could reduce misclassification of ranking (Tylavsky and Sharp,1995). Our FFQ adopted the close ended response categories. The responses ranged from never to twice or more per day, totalling seven responses. This probably could affect the estimate of energy and nutrients for our subjects and in addition the misclassification of ranking as well. However, since our FFQ was administered by interview, clarification for responses which had a wider interval (eg.1-3 times per month) could be made. The ranking of our subjects was poor as there was a high percentage of gross misclassification. The impact of this misreporting of the exact frequency of consumption would be likely to result in a greater measurement error.

The portion size used in the study was based on the standard portion suggested by MAFF (Crawley, 1988). The standard portion was chosen as there were incomplete references regarding portion sizes specific to most South Asian foods. In the course of administering the FFQ to the subjects, we discovered that some of the portion sizes used in the FFQ did not reflect the portions usually eaten by them. The portions suggested by MAFF were adapted from the portion sizes served in the South Asian restaurants which catered for mainly the European consumers. From the weighed record, we found that the portion size of the main food items such as chapatti and meat curry was only half the standard portion of the FFQ. It should be noted that the size of chapatti consumed by South Asians varied depending on their ethnic group. Generally the Gujeratis consumed the smallest chapatti, followed by an intermediate size eaten by the Punjabis. The Pakistanis ate the biggest chapatti compared to the other two groups. The chapatti portion size suggested by MAFF was in agreement with those consumed by the Gujeratis.

Problems regarding the portion size in this FFQ also applied to the chicken, eggs and meat group (now referred to as the meat group). The 300g per portion initially used for the meat group was reduced to 200g per portion when the nutrient content was analysed. The fish portion which was very much over portioned was reduced to 75g per portion from the original 300g per portion. These new portion sizes were based on the average portion sizes of the meat group consumed and recorded in the weighed record.

During the FFQ interview, many subjects found it difficult to estimate the consumption of their meat curry. This was because the meat in the curry were usually cut into small pieces. Different households would very likely cut the meat pieces differently. Thus it would be likely that the subjects had the tendency to wrongly estimate the portions of meat eaten.

The problem of getting the appropriate average or standard portion of the South Asian foods needs to be rectified. This is especially important as the portions suggested by MAFF does not apply to the actual portion size consumed at home. According to Wolk et al (1994), using inappropriate portion sizes could result in differences in the nutrient intake estimates. Thus it would be helpful if a standard portion size for the meat group of South Asian dishes is defined.

The effect of using an inappropriate portion size for the original FFQ was an to overestimate the energy and fat intakes by the FFQ in most of our subjects. This was because the portion size which was wrongly defined was for the food items which were consumed daily by most of the South Asians, namely the chapatti and meat curry. As a result of this, the portion size was adjusted accordingly for the analysis of the nutrient estimates. After adjustment, the estimate of total energy intake was close to that predicted from the BMR data.

The FFQ that we developed was administered in person in the English language. Interviewer administered FFQs provided several advantages. Firstly, the presence of an interviewer was more likely to maintain the respondent's interest, and secondly it could help clarify any potential misunderstanding. Furthermore, clarification in the quantification of portion size could be carried out (Jain et al,1996). Although self administered FFQs might in theory achieve a large recruitment of subjects into the study, we realized that this would be unlikely in this community as many people were still functionally illiterate especially among women and the elderly, even in their own native language (HEA,1994). About 60% of the subjects in our FFQ calibration study were women, while all subjects in the food checklist and 24 hr recall calibration study were women. This could result in selection bias in our women subjects as mainly literate women would be recruited into the study. For this population, a self administered FFQ might be likely to achieve a lower response rate than an interviewer administered FFQ. The low recruitment number might in part result in a limited statistical power to achieve significance in the relationship investigated. The small number of subjects in this study, due to low compliance might result in a biased outcome. It would also be likely that a biased sample was recruited into our study. The effect of this sample bias is discussed in section 4.1.6.

4.1.4 FFQ calibration study

In our effort to develop a suitable method to evaluate the usual dietary intake of South Asians, we have selected the FFQ as the method of assessment. In order to check the relative validity of the FFQ, a calibration study was carried out, using a 4-day weighed record as the reference method. In theory, the 7-day weighed record has been considered as the best method to obtain information on nutritional intake (Bingham,1987). However, this does not imply that the 7-day weighed record would work best in all calibration studies. A high degree of committment and literacy are required of the subjects for the weighed record to work well. Furthermore the weighing of food is quite intrusive and might be likely to make an individual change her or his diet (Jain et al,1996). Due to all these factors, a lower response rate might be expected when using the weighed record in a calibration study.

Many calibration studies have used the 7-day weighed record as the reference method, because theoretically the weighed record has been regarded to be the best method to collect nutrient intake information (Willett et al,1985; Pietinen et al,1988;Bingham et al,1994). However, since our aim was to rank the intake of macronutrients and not determine the absolute intakes estimated by the FFQ, we used a 4-day weighed record instead as this was reported to be optimum enough to evaluate macronutrients (Block and Hartman,1989; Bingham and Nelson,1991). It has also been shown that to correctly classify 80% of the subjects into the extreme third of the distribution for energy and energy yielding nutrients, food records between 3 to 7 days would be sufficient (Bingham,1981). Furthermore we had to consider our study samples. An earlier study of the South Asian community which evaluated nutrient intakes using dietary history and weighed inventory revealed that their subjects had difficulties in completing the weighed inventory

(Harbottle and Duggan,1993). Therefore in order to achieve a large recruitment, we have chosen the 4-day weighed record because this was the most feasible way to carry out the calibration study and in doing so satisfy both the recruitment number and nutrient intake evaluation.

The response rate for our FFQ calibration study was 35%. This was in agreement with the response rate of a Canadian study (Jain et al,1996) but lower than others (Morgan et al,1978;Pietinen et al,1988;Thompson and Margetts,1993). Reasons for the non participation and drop out in our study could be attributed to refusal and unwillingness to cooperate, incomplete FFQ and unreliable weighed records. The obvious difference between those who completed the study and those who dropped out were gender and ethnicity. There were 60% female responders compared with 40% male responders. As reported by the first comparative dietary studies in the South Asian community, there were also problems pertaining to the food weighing in our study. We experienced a higher drop out rate among the male subjects (52%) compared to the female subjects (20%). The reason for this higher drop out rate in men could be attributed to the South Asian culture itself. It was in their culture that the women prepared and served meals to the men. Thus, having to weigh their food was a tremendous effort for these men and many eventually dropped out. Some men relied on their wives or daughter-in-laws to do the food weighing. In this instance, family commitments (domestic duties and caring for the children) were the obstacle for them to complete the weighed food record. It could be possible that some men found the technical aspect of weighing difficult, thus although initially they were interested and commited to the calibration study, towards the end they failed to keep up with the weighing of their food. The subjects (both men and women) also took a much longer time than expected to complete the weighed records. Quite a number of subjects took at least a month or even longer to finish the 4-day record. In the end, they did not complete the food weighing at all. In terms of ethnicity, it was not clear why more Pakistanis and less Gujeratis and Punjabis took part in the study.

As in other calibration studies, we expected the responders to be different from non responders. Responders would likely be more educated, motivated and generally more interested in the study. However, in the FFQ calibration study, the responders and non responders appeared to be rather similar in a lot of their demographic characteristics. The duration of stay in the UK since migration from the Indian subcontinent were more than 15 years in a majority of responders and non responders. However, we could not assume that they were very similar. We would expect that responders were very likely to have acculturised themselves more than the non responders. We could also anticipate that the non responders were more very traditional in their lifestyle. The same assumption could be made about the education. A majority of the responders and non responders had at least six years of education. Since many of them migrated to the UK when they were adults, it would be very likely that they had their education in India.

Through personal interaction and communication with the subjects and prospective subjects, it could be generalised that the responders had a better understanding and communication skills towards the English language than the non responders. The responders were also more likely to be younger, thus more likely to interact and socialise with the host population.

Subjects recruited were approached in centres where social group meetings or English classes were held. Some subjects were also recruited from their surgeries. It would be expected that those who attended the social gathering and the English classes were different from those who did not. Thus this might introduce selection bias into the study.

Nelson (1996) suggested that men and women responded differently to dietary assessment. Therefore it was neccesary that men and women be analysed separately by gender. We used several statistical techniques to evaluate the relationship between our FFQ (test method) and WR (reference method). It has been suggested that the validity of a calibration study could not be determined using correlation coefficient alone, but would require other measures of agreement to describe the relationship between the test and reference measures (Nelson,1996). Bellach (1993) have also recommended that it would be more meaningful if several statistical analyses which produced stable results were used to conclude the association between measures. Thus we have chosen five analyses:

a. Percent mean difference
b. Spearman rank correlation (energy adjusted and unadjusted)
c. Bland Altman technique
d. Distribution of mean nutrient intakes to tertiles (ranking)
e. Kappa statistics

The results of our calibration study showed that there was poor agreement between estimates of nutrient intake derived from the FFQ and the WR for both men and women. Spearman rank correlation were not significant for energy and the macronutrients for both men and women in the energy unadjusted nutrient intake. However, the correlations were significant for protein in men and for carbohydrate in women. The correlations were statistically significant when the data for both men and women were analysed together. This could be attributed to the larger number of subjects, thus giving greater power to detect the difference. The correlations in men and women ranged between 0.20 to 0.55

and -0.05 to 0.28 respectively for energy unadjusted, while ranged between 0.10 to 0.44 and 0.23 to 0.36 for the energy adjusted. For both groups together the correlations were between 0.26 to 0.38 (energy unadjusted) and 0.25 to 0.36 (energy adjusted. The correlations broken down by gender were lower than those reported by others (Willett et al,1985; Pietinen et al,1988), while comparable to the comparative dietary method study of Bingham et al (1994). However, comparison with other studies must be made with caution because of large differences between questionnaires and the populations in which they are applied. The lower correlations in this study could be attributed to the subjects who participated. Our subjects consisted of South Asians who could be regarded as having lower than average level of education and totally unfamiliar with weighing of foods. In contrast, the studies which demonstrated high correlations consisted of volunteers and nurses who were more motivated.

Closer analysis of the differences between measures using the Bland-Altman technique (Bland and Altman,1986) for estimates of energy, fat, carbohydrate and protein was not consistent across the range of intake for energy, fat and protein in women. There appeared to be a differential bias in that at lower intakes, the FFQ gave higher estimates than WR, while at higher intake, the FFQ gave lower estimates than the WR. The Bland-Altman plots for men revealed that the FFQ gave a higher estimate throughout the whole range of intake. The use of Bland-Altman plot here enabled the agreement between measures to be assessed across the whole range of intake. This is an advantage as most epidemiological studies seek to compare risk of disease across the whole range of intake in different group of subjects. This type of comparison could be helpful.

The FFQ generally gave higher estimates than the WR for almost all nutrients analysed. This was expected as comparison of energy estimates with BMR multiples indicated a high percentage of underreporters when subjects reported their diet with WR. However, our findings were in agreement with various other studies which compared the FFQ with the WR (Bingham et al,1994;Rothenberg,1994). Despite the high underreporting by the WR, the higher nutrient estimates derived from the FFQ could possibly be due to the overestimation of frequency of intake as well as the overestimation of portion size. Overestimation of frequency of intake and portion of intake was also shown by others (Flegal et al,1988; Bingham et al,1994). We overestimated the frequency of the fat group for our FFQ. The South Asian foods (dishes) are usually cooked with the fat or oil already incorporated into the cooking. For example, spices were always fried in most South Asian dishes. Thus in administering the FFQ, the fat intake of South Asians were accounted for twice, once in response to the meat dishes and a second time specificly in answering to the fat or oil group. This directly overestimated fat intake. This was adjusted accordingly during the nutrient analysis.

In this study, we also experienced overestimation of portion sizes for foods which contributed significantly to the energy and macronutrient intake. For example, we overestimated the portion size for the meat group by as much as twice the portion size normally consumed by the South Asians. As for chapatti, which was the staple food and eaten daily, the portion size used in the FFQ was more applicable to the Gujeratis. Thus it was necessary to adjust the number of portions consumed accordingly when conducting the FFQ interview.

The ability to accurately estimate usual food intake using an FFQ depended on the food list relevant to the repondent's typical diet as well as the ability to accurately recall and quantify food intake (Wheeler et al,1994). In this FFQ calibration study, difficulty regarding the portion estimation was experienced by the subjects. Unlike the English foods, there is still no standard portion size for a lot of South Asian foods. Thus it would be likely that many subjects either overestimated or underestimated their portion size when responding to the FFQ.

Some of our subjects especially women (66%) were overweight. According to various authors (Nelson et al,1994;Prentice et al,1985;Black et al,1991) overweight subjects tend to underestimate portion size while elderly subjects tend to overestimate portion size. However, generally there was a tendency for subjects to underestimate large portion sizes while overestimating small portion sizes (Nelson et al,1996). Furthermore in a recent study, it was reported that individuals who were embarassed about their weight were more likely to give records which would not be a valid reflection of their usual intake, even those with normal weight (Mela and Aaron,1996).

In contrast to the FFQ, the lower estimates derived from the WR could be attributed to the underreporting by some of the subjects. As weighing food could be considered as intrusive, this might cause subjects to change their diet, most likely towards reducing their usual intake. The percentage of subjects underreporting their energy intake by WR was high (52% in women vs 41% in men). The percentage of underreporting was much lower by the FFQ method (7% in women vs 12% in men). It could also be that the high prevalence of overweight among women increased the tendency to underreport their intake using a WR, while at the same time underestimating portion sizes in the FFQ.

Besides comparing the FFQ with the WR and EI/BMR ratio, we also compared the energy intake estimates with the energy requirement (BMR multiples). It was demonstrated that 72%

men and 80% women did not meet their energy requirement. This further suggested that our subjects underestimated their actual energy intake.

The high percentage of subjects (both men and women) underreporting energy intake with the WR posed an uncertainty regarding the creditability of the WR as a reasonable reference method in this study. Although in theory WR has been regarded as the best possible method to collect nutritional information, it might not appear so in the present study because WR underestimated food intake quite seriously when compared with the EI/BMR ratio. Instead the FFQ seemed to be more realistic in estimating energy intake. Therefore it could be assumed that WR did not estimate actual intake accurately in our calibration studies.

The effect of underreporting in WR and overestimating frequency in FFQ would widen the discrepancy in the estimated mean nutrient intakes. The underreporting in the WR would also imply that the portion sizes of the food items might not reflect the actual (usual) portion size usually consumed by the South Asians. Thus the portion sizes adapted to the FFQ from WR might not be appopriate either.

The percent mean differences between the FFQ and WR were greater than 10% for all nutrients except starch in men. However, in women, the percent mean differences were within 10% for energy, fat, protein and starch. They were in the range of 13% and 74% for carbohydrate, fibre and sugar. Although the cut off point of less than or equal to 10% mean difference had been used as an indicator of a good agreement between measures, it did not appear such when ranking between measures was evaluated. The level of agreement in ranking between WR and FFQ revealed that there was a high percentage of subjects who were grossly misclassified (9% in men vs 25% in women) while only 31% women and 43% men were correctly classified. This implied that many subjects were incorrectly placed in the proper distribution by both measures, thus indicating poor ranking. Kappa statistics carried out revealed that the agreement between measures was also poor (k<0.3)

4.1.5 Food checklist and 24 hour recall calibration study

In the FFQ calibration study, we experienced some technical and recruitment problems. Thus in trying to develop another dietary assessment method as an alternative to the FFQ, we had to select a method which was simple to administer and would not burden the subjects. This method had to be simple enough. It was decided that only South Asian women would be recruited into this calibration study.

Although the food checklist has not often been used as the test method, it could be regarded as similar to an estimated record. The food checklist and 24 hr recall were selected as the test methods. The food checklist used in this calibration study was also developed for the first time in this study. We used the same food list as the FFQ to develop the food checklist. However, the design was based on that of Bingham et al (1987). The food checklist was self administered, whereby subjects had to check off daily which foods they had eaten, for seven consecutive days. This limited our recruitment to literate subjects only. We also translated the food checklist to the various South Asian languages such as Gujerati, Urdu and Punjabi. However, the translated food checklist was not of benefit, as we did not manage to recruit more prospective subjects as women who were illiterate in the English language were also illiterate in their native language.

The 24 hr recall was simple, thus we hoped we could recruit a maximum number of subjects. For both test methods, we used the 4-day weighed record as the reference method.

The response rate in the food checklist and the 24 hr recall calibration study was 52%. The difference between responders and non responders were in the marital status and ethnicity. We would expect more of the married women to drop out of the study, as usually they have more family commitments. However, in this calibration study, about half (47%) of the non respondres were single. These single women were mostly young (< 20 years old). Closer interaction with them showed that they were not interested in the study in the first place, but consented to participate as they felt obliged. In terms of ethnicity, it was difficult to reason out why the participation was as such. A possible reason could be that the Pakistani women and the researcher had a common bonding regarding their religion.

A similar statistical analysis (percent mean difference, Spearman rank correlation, Bland-Altman technique, ranking and Kappa statistics) were carried out to determine the association between the test methods and the reference method. The results showed that there was also poor agreement between estimates of nutrient intake derived from food checklist and WR, as well as 24 hr recall and WR. The Spearman correlation coefficient were not significant for all nutrients for food checklist and WR and 24 hr recall and WR, both for energy adjusted and unadjusted. These correlations were lower than those of Bingham et al (1994).

The food checklist generally gave higher estimates than the WR, while the 24 hr recall on the other hand gave lower estimates than WR. However we found that underreporting was high when estimating energy using WR. When we compared the estimates of energy intake from the food checklist and 24 hr recall with the EI/BMR ratio, an average of 30% of subjects by food checklist and 52% of subjects by 24 hr recall respectively were underreporters. This suggested that using the food checklist and 24 hr recall method would likely result in underestimating energy intake.

Percent mean difference between measures for both food checklist and 24 hr recall with the WR were generally lower than 10% for most nutrients indicating a good agreement between measures. However, other statistical analysis showed otherwise. The ranking of subjects into thirds of distribution was also poor. Food checklist method grossly misclassified an average of 14% of the subjects. This misclassification was however, lower than the FFQ (25%) and the 24 hr recall method (22%). The percentage correctly classified by food checklist, 24 hr recall and the FFQ in this study were lower than another comparative study (Bingham et al,1994).

Food checklist and 24 hr recall methods are dietary methods which assess actual intake. Since the WR is also a method which evaluates actual intake, it would be expected that these methods would have a better association with the WR than FFQ. However this did not appear so.

A possible reason for the low correlation between our test measures and reference measures could be attributed to the homogeneity of our population. This led to a narrower range of energy intake and a lesser variability in their diet composition. The low response rate and the small number of subjects recruited contributed further to the lower association between measures.

Generally it was rather difficult to get prospective subjects to participate in the calibration study. This was due to the language problems that existed between the subjects and the researcher. The reference method (WR) which we used to calibrate the test methods (FFQ, FCL and 24 hr recall) was not appropriate for the South Asian subjects as many of them found the weighing of food difficult. It would be likely that a simpler method,like the multiple 24 hr recall would increase the number of subjects recruited into the study. Personal communication with a social worker in the community provided a better understanding towards the attitude of the community. Subjects were reluctant to disclose their eating habits. The response rate in both of the calibration studies were quite discouraging. A lot more efort and time need to be put into the recruitment strategy to guarantee that future studies would face with less problems.

4.1.6 Bias in the study design

The subjects in our calibration studies consisted of men and women who were literate in the English language. Thus this made sampling bias likely to be present in the study. We would expect that the responders in both calibration studies to be different from those who did not. This was due to the possibility that responders were likely to be more aware of health and more educated than the non responders. Therefore we could assume that the estimated relative validity of this study might be an overestimate of the true relative validity.

Besides the sampling bias, it was not clear whether social desirabilty bias had a role in the way the subjects responded to the FFQ, food checklist and 24 hr recall as well as in the way they recorded their weighed record. Our subjects consisted of 60% women who were overweight or obese. Furthermore, more than two thirds of the subjects did not meet the energy requirement calculated from BMR multiples. Therefore it was not surprising if these subjects underreported their food intake. At the same time they were likely to overreport their consumption of foods viewed as good and underreported consumption of foods regarded as bad. This misreporting of dietary consumption was likely to result in a greater measurement error.

4.1.7 Measurement error analysis with method of triads

Approaches to determine measurement errors in dietary measurements have increased over the past decade. We cannot assume that dietary measurements are error free. Generally the measurement errors in dietary studies can be categorized into two types. They are systematic error and random error. Systematic error or bias is defined as under or overestimation of intake by an individual or group of individuals. Random error is considered as within subject day-to-day variance. Usually random error can be minimized by taking repeated measurements from the same individuals.

In the present study, in using the FFQ, food checklist and 24 hr recall as test methods while the WR as the reference method, both random and systematic measurement errors were likely to occur. The comparison of energy estimates from FFQ, food checklist, 24 hr recall and WR with EI/BMR ratio showed that a high percentage of subjects were underreporting energy intake when recording diet using WR, food checklist and 24 hr recall. On the other hand, subjects tended to overreport energy intake when responding to the FFQ. The presence of the systematic errors whereby subjects underestimated or overestimated food intake had to be accounted for. Systematic errors were also likely in our FFQ, where the portion sizes for some commonly eaten foods (eg. chapatti and meat curry) were inappropriately portioned. The effect of this type of error would be to underestimate or overestimate these foods in the same direction in all subjects.

To correct for the random error, ideally repeat measurements (reproducibilty study) of the test methods could be carried out. A large number of subjects could be recruited into the study to minimize the random errors. In the present study, no reproducibilty study was carried out. The effect to our study would be a reduced power in the precision of the dietary estimates by the test methods.

To reduce the systematic errors of the test methods, the calibration studies were carried out. The relative validity of the test methods were compared with the 4-day weighed record. Nevertheless, even the WR might not be sufficient to estimate systematic errors as data collected could also include systematic errors. We used a proxy of biomarker, the multiples of BMR as our second reference method. The BMR multiples could give an unbiased estimate of food intake. Another advantage of using the biomarker is because random errors between the biomarker, the test methods and the first reference methods were likely to be mutually independent of each other.

We estimated the measurement errors in the FFQ calibration study as well as the food checklist and 24 hr recall calibration study with the method of triads (Kaaks,1996). This method was simple and used an elementary factor approach. However, the method of triads as in other measurement error models assumed that the correlation between the three measurements involved (test measurement, reference measurement and biomarker measurement) were linearly related to the true intake and the random errors were mutually independent.

The results of our evaluation of measurement error with the method of triads showed that in men, the validity coefficient of FFQ and WR were 0.73 and 0.54 respectively (FFQ calibration study). This could be regarded as an overestimation of the validity coefficient and thus could be interpreted as the upper limit for the true validity coefficient. It was likely that the FFQ and WR had positively correlated random errors, as both were dietary methods which relied on information given by subjects about food consumption. These common errors between dietary methods could be contributed by using the same database during conversion of food data into nutrient data, the error due to the properties of the instrument itself, as well as due to short term fluctuations in the diet (Plummer and Clayton,1993). The validity coefficient of the biomarker was low (0.26). It was likely that the random measurement error of the biomarker and the FFQ and WR were truly independent. We could therefore suggest that the estimated validity coefficient of the FFQ using the triad method with the 4-day WR and BMR multiples value was the upper limit. The observed correlations between FFQ and BMR multiples, as well as the WR and BMR multiples was the estimated lower limits of the validity coefficient.

For the measurement errors in the food checklist calibration study, the validity coefficient of the food checklist and WR were 0.41 and 0.12 respectively. The validity coefficient of the biomarker was however, rather high (0.66). In this instance, we could suggest that the estimated validity coefficient of food checklist using the triad method with the WR and BMR multiples value was the upper limit. The observed correlation between food checklist and BMR multiples as well as WR and BMR multiples was the estimated lower limits of validity coefficient.

In interpreting this measurement error model for the FFQ calibration study for men, it would appear that the FFQ measurement was less accurate, as the validity coefficient were in the range of -0.19 and 0.73. There was an underestimation in the FFQ as the validity coefficient for the biomarker was low. It could be assumed that the true validity coefficient would probably lie in a range in which the correlation between FFQ and BMR multiples values and the maximum validity obtained by the method of triad. For the food checklist calibration study on the other hand, the food checklist measurements appeared reasonably accurate with the lower limit for the validity coefficient of 0.27 and upper limit of 0.41.

In the FFQ calibration study, we could not calculate the validity coefficient of the measurement for the women as one of the sample correlations was negative (r between FFQ and WR, -0.04). A negative correlation was also observed in the 24 hr recall calibration study. These negative correlations could be attributed to two possible explanations. Firstly, it could be that the random sampling fluctuations were present in the observed correlations between measurement. The second explanation was that one or more of the underlying model assumptions (which was linear relationship with the truth and independent random errors between measurement) was violated.

In our context, the negative correlation observed in one of the sample correlations could be attributed to a random sampling fluctuation. At the same time we could not dismiss that one or more of the underlying assumptions has been violated. To decrease the size of the sampling fluctuation, a larger number of subjects should be recruited into the calibration study. To ensure a positive sample correlation, an accurate reference and biomarker should be used. In our calibration study, the reference method which was the 4-day weighed record did not produce a more accurate measurement than the FFQ and the food checklist as there were more underreporting with the WR than our test measures. This thus suggested probably with a more accurate reference measurement, the validity coefficient of both calibration studies could have been better.

CHAPTER FIVE

CONCLUSIONS, IMPLICATIONS AND FUTURE RECOMMENDATIONS OF THE STUDY

CHAPTER 5 CONCLUSIONS, FUTURE RECOMMENDATIONS AND IMPLICATIONS OF THE STUDY

5.1 Conclusions of the study

The calibration studies have produced several useful findings. We also learned about some drawbacks in the progress of the study. The conclusions from the study are listed below.

1. The South Asian diet was different from the English diet in that it was still very traditional in nature, monotonous and homogeneous.

2. Three food items, namely chapatti, meat curry and bread contributed 30% to 50% of the major nutrients in the South Asian diet.

3. A South Asian foodlist which we identified had to be used for development of our food frequency questionnaire (FFQ) and the food checklist.

4. The design of the FFQ was in general acceptable, except for some adjustments with regards to standard portion size of some food items (eg. chapatti and meat group).

5. The weighed record which was used as the reference method did not accurately measure food intakes of this community. There was a high percentage of underreporters when subjects reported their dietary intake with WR.

6. Generally it was difficult to work with the South Asian community because of language problems and the high rate of illiteracy in the community.

7. By using the method of triads, the measurement errors in the FFQ and 24 hr recall among women could not be evaluated

due to the presence of a negative sample correlation. However, the measurement errors in FFQ among men showed that the FFQ underestimated dietary intake. In women, the food checklist method appeared to be able to estimate dietary intake reasonably accurately, with lower limit for validity coefficient of 0.27 and upper limit of 0.41.

8. The test methods had a weak association with the reference method. The ranking of subjects for all test methods were also poor. Gross misclassification of subjects ranged between 13% and 29%. However, these results have to be interpreted with caution as the relative validity of the reference method was poor.

9. Further development of the FFQ and food checklist in this population are required before they can be used for evaluation of dietary intake in the South Asian population.

5.2 Implications of the calibration studies on the dietary assessment in the South Asian community

In this study, we have developed and calibrated an FFQ and a food checklist against a 4-day WR to be used for the evaluation of the dietary intake in the South Asian community. We have also used the 24 hr recall and again calibrated it against the 4-day WR. The objective of developing the dietary assessment methods (FFQ and food checklist) was to rank groups of individuals according to low, medium and high intakes of energy and macronurients. The FFQ and food checklist were the first FFQ and food checklist to be designed and calibrated for the South Asian community.

The subjects involved in our calibration studies consisted of three main South Asian ethnic groups namely the Pakistanis, Punjabis and Gujeratis. The only group which was not included into the calibration study was the

Bagladeshis. The reason for the exclusion of the Bangladeshis was because their meal patterns and food habits were different from the other three main groups. In designing the FFQ and the food checklist, we determined the food sources of energy and macronutrients in the South Asian diet. This information was then used to develop a food list for the FFQ and the food checklist. Therefore we hoped that the FFQ and the food checklist could be used to evaluate nutrient intakes of Pakistanis, Punjabis and Gujeratis.

The weakness of this study was the small number of subjects that we managed to recruit into the study. This might limit the statistical power to achieve significance in the relationship that we were investigating. It would be likely that with a larger number of subjects, a significant relationship could be attained. A large recruitment number would also reduce the random errors in the measurement. Problems pertaining to recruitment have already been discussed. We were only able to recruit South Asian men and women who were literate in the FFQ calibration study. We recruited only women in the food checklist calibration study. Therefore there was selection bias in these studies. The effect of this selection bias would likely be an overestimate of the association between the test and reference measures. This was because we would expect those who responded to the study to be more educated, had a less traditional (South Asian) lifestyle and more aware of health. Caution should be used in generalising the results of this study to the South Asian community at large. However, the results would be useful in aiding future development of dietary assessment methods in this community.

The Spearman correlations of energy and macronutrients were not significant when we analysed the data according to gender. However, the correlations were significant for energy unadjusted and energy adjusted macronutrients when the data were pooled and analysed together. In general the correlations between test and reference measures were poor. The ranking of FFQ, food checklist and 24 hr recall with the weighed record were also poor. Gross misclassification of subjects ranged between 13% and 29% for all three test measures. These results indicated that the relative validity of our test measures were poor.

The correlation between FFQ and WR was 0.25 for the energy adjusted fat. We were only able to rank on average 33% of the subjects into the proper thirds of the distribution (correctly classified). However, using Table 3.1 (Clayton and Gill, 1991), in theory with a correlation of 0.25, about 48% of the subjects can be correctly classified. And using Figure 3.5 (Clayton and Gill,1991), the effect of exposure misclassification (for eg. fat) upon observed relative risk (of developing diabetes) is the attenuation of the relative risk by half. If it is assumed that the true relative risk of diabetes in the South Asian population is four times that in the whole UK population, the results of the present calibration study suggested that the measured effect of diet on risk of diabetes would be half.

Looking at diabetes prevalence studies of South Asians and Europeans in UK (Mather and Keen,1985;Simmons et al 1989;McKeigue et al,1991), these studies showed that the prevalence was two to four times higher in the South Asians compared with Europeans. Although dietary intake studies have been carried out to investigate whether there were differences in dietary intake between these two populations, the results have been inconsistent. There were discrepancies in estimating energy and fat intakes in both populations. Our present preliminary study demonstrated that the diet of South Asians was homogeneous and monotonous. This could result in a large within to between subject variance, thus could lead to gross misclassification when subjects were split into groups (eg. thirds or fifths) based on dietary intake.

The number of subjects involved in previous studies of the South Asian diet were between 12 and 184. The dietary methods used were 7-day weighed record, FFQ or household inventory. Since in some of these studies, the number of South Asian subjects was small, while the variance of dietary estimate could be large, this could result in a greater measurement error in the dietary measure. While the measured dietary intake of the Europeans might be closer to the true measure, the measured dietary intake of the South Asians could be underestimated or overestimated. Thus the comparative studies could not detect any significant difference between the diet of South Asians and Europeans. Therefore the results of the comparative studies would suggest that there was no measured effect of diet on risk of diabetes, when there could have been an effect if the measurement errors in the dietary measures were considered in the interpretation of the results. This was demonstrated in our FFQ, whereby the relative risk of diabetes would be underestimated by half.

The strength of this study was being the first study to develop dietary assessment methods and then calibrate them for use in the South Asian community. The FFQ and food checklist appeared to be methods which were well received by the community. The WR as the reference method was not preferred by the South Asians as it was found to be too technical, complicated and demanding to do. Because of these reasons, recruitment number and response rate were low.

The FFQ, food checklist and the 24 hr recall failed to rank the subjects into low, medium and high intakes according to the WR. On the other hand, we found that the WR was not a reasonable reference method, as a large number of subjects (>50%) underreported their energy intake (based on comparison with EI/BMR ratio). This implied that the test measures would likely have a better correlations with the reference measure, if there was no underreporting bias in the reference measure.

In summary, the objective of our test measures namely the FFQ, food checklist and 24 hr recall was to evaluate the usual dietary intake of South Asians and to rank them to low, medium and high intake. However, the results of the calibration studies had to be interpreted with caution. Our test measures were not able to rank groups of individuals correctly as the reference measure. However, this did not imply that our test methods were not reasonable. This was because comparison of energy estimates with a biomarker suggested that the test methods could estimate better than the reference method. Further development of the FFQ and food checklist are needed before they could be used to evaluate dietary intake of the South Asian community. A more reasonable and feasible reference method should be selected to increase the recruitment number as well as the response rate. It is also important to ensure that the measures are reliable.

5.3 Future recommendations

The FFQ and food checklist would be ideal methods to evaluate usual intake in the South Asian community. However, during the FFQ calibration study, we were faced with a number of problems. These problems would likely have a role for the poor outcome of the calibration studies. Therefore some recommendations are suggested below for future calibration studies.

1. To increase the number of subjects and response rate.
A lot of effort and time are needed to overcome problems

with recruitment. Time should be spent talking and explaining to the community regarding the objective of the research to be undertaken and in return what health benefits the community can attain from the study. The researcher should be able to gain the community's confidence and trust. Cooperation with the community leaders as well as the ethnic liaison officers and health visitors can open up ways to get an introduction to the community. However, a lot of hard work is required beyond this initial introduction as this does not guarantee a good response rate from the community. A possibility in increasing the number of subjects and response rate is to use a dietary assessment method which is simple enough for the South Asian community.

2. To get a reasonable reference method.
Although the weighed record is regarded as an accurate dietary assessment method to get nutritional information, in the South Asian community this does not appear so. There are a few problems with using WR in this community. The South Asian community finds it very difficult to complete the WR because there is a possibility that they could not understand the technical aspect of the weighing even though demonstration on the usage of the scale has been carried out. Therefore an unrealistic weight of foods for portions consumed has been recorded. A longer time is needed by the subjects to complete the weighed record (sometimes as long as a month or longer). A lot of subjects tend to underreport their food records, probably the weighing causes them to change their usual intake. A dietary method which is very simple and can get a large number of subjects into the study will be ideal as the reference method.

3. To include all adult South Asians into the study.
Our calibration study is restricted to include only literate subjects. This is because only literate individuals respond and agree to participate in the

calibration study. There are some subjects who could understand and spoke minimum English, but their weighed record are unreliable. Therefore our subjects are biased in their selection. However, to be able to include illiterate South Asians into the study, an interpreter or a community worker need to be involved in the study.

APPENDIX 1 FOOD FREQUENCY QUESTIONNAIRE

Questionaire on Food Intake of South Asians

Family Background

1 Name:

2 Date of birth:

3 Age:

4 Place of birth:

5 Sex:

6 Marital status: Married/ single/ divorced/ others:

7 No. of children:

8 No. in household:

9 Race/Ethnic:

10 Religion:

11 Home address:

Tel:

12 Duration living in Southampton:

13 Duration living in United Kingdom:

14 Occupation:

15 Education:

16 Do you have diabetes? Yes/No

If yes , for how long? ____ years

17 Does anyone in the family has diabetes?

If yes, please specify ______

DIET AND DIABETES

1. When you eat bread, how many slices would you usually have in one day?

    ____ slices/day

2. When you eat rice, how many plates would you usually have in one day?

    ____ plates/day

3. When you drink milk (including in tea and coffee and on breakfast cereals), how much would you usually have in one day?

    ____ pints/day

4. When you drink tea or coffee, how much sugar would you usually put in the drink?

    ____ teaspoon/cup

5. When you consume fizzy drinks, how much would you usually consume per day?

    ____ can/day

6. When you make chapatti, do you <u>usually</u> add ghee or vegetable oil with the flour during the preparation?

   Yes/No

   If yes, how much ghee or vegetable oil would you <u>usually</u> use in the preparation?

   ______ tablespoons/lb flour

7. When you eat chapatti, how many pieces would you <u>usually</u> have in one day?

   (Please circle number in the appropriate column for the chapatti size and the number of chapattis eaten of this size on the line)

Type	Number of chapattis eaten	Chapatti size small	medium	large
Wholewheat	______	1	2	3
Millet	______	1	2	3
Others	______	1	2	3

8. When you spread ghee or butter on your chapatti, how much would you <u>usually</u> spread on each chapatti?

_______ tablespoons/chapatti

9. When you fry your food, how much fat/ghee or vegetable oil do you <u>usually</u> use for frying the food?

butter	___	tablespoon
ghee	___	tablespoon
margarine	___	tablespoon
cooking oil	___	tablespoon

Type of vegetable/cooking oil used __________

10. Over the last three years, have you changed your use of the following foods?

Food	No change	Eat less now	Eat more now
Cooking oil	1	2	3
Ghee	1	2	3
Butter	1	2	3
Any margarine	1	2	3
Whole milk	1	2	3
Skimmed milk	1	2	3
Eggs	1	2	3

Food	No change	Eat less now	Eat more now
Meat	1	2	3
Fish	1	2	3
Fruits	1	2	3
Vegetables	1	2	3
Chapatti	1	2	3
Rice	1	2	3
Gur	1	2	3
Sugar	1	2	3
Asian fried snacks	1	2	3
Asian sweets	1	2	3
Pickles	1	2	3
Achar	1	2	3
Chutney	1	2	3
Alcohol	1	2	3
Tea	1	2	3
Fizzy drink	1	2	3

FOOD	rare or never (1)	1-3 times /mth (2)	once a week (3)	2-3 times /week (4)	4-6 times /week (5)	every day (6)	2 or more times /day (7)	Standard portion	No of portion	Cooking method
Breads and cereals										
White bread								1 medium slice		
Brown bread								1 medium slice		
Wholemeal bread								1 medium slice		
Poori								1 medium piece		
Paratha								1 medium piece		
Chapatti								1 medium piece		
Cakes								1 slice		
Sweet biscuits								1 biscuit		
Indian sweets (eg gulab jaman)								1 piece		

FOOD	rare or never (1)	1-3 times /mth (2)	once a week (3)	2-3 times /week (4)	4-6 times /week (5)	every day (6)	2 or more times /day (7)	Standard portion	No of portion	Cooking method
Fried snacks (eg samosa)								1 piece		
Breakfast cereal high fibre (eg Wheetabix)								1 medium bowl		
Breakfast cereal ordinary								1 medium bowl		
Rice								1 plate		
Sweet rice								1 plate		
Chicken, eggs and other meats										
Eggs-fried								1 egg		
Eggs-not fried								1 egg		
Fish-fried								1 medium fillet		
Fish-not fried								1 medium fillet		

FOOD	rare or never (1)	1-3 times /mth (2)	once a week (3)	2-3 times /week (4)	4-6 times /week (5)	every day (6)	2 or more times /day (7)	Standard portion	No of portion	Cooking method
Beef								1 slice		
Lamb/Sheep								1 slice		
Pork								1 slice		
Chicken								1 medium portion		
Sausages								1 piece		
Milk and products										
Whole milk includes in tea/coffee								$\frac{1}{2}$ pint		
Semi skimmed milk includes in tea/coffee								$\frac{1}{2}$ pint		
Skimmed milk includes in tea/coffee								$\frac{1}{2}$ pint		
Yoghurt								1 pot		
Ice cream								1 **average** scoop		
Cheese								1 slice		

FOOD	rare or never (1)	1-3 times /mth (2)	once a week (3)	2-3 times /week (4)	4-6 times /week (5)	every day (6)	2 or more times /day (7)	Standard portion	No of portion	Cooking Method
Cheese and egg dishes (eg pizza)								1 slice		
Fats and oils										
Butter								1 tbs		
Margarine								1 tbs		
Ghee								1 tbs		
Cooking/vegetable oil								1 tbs		
Fruits and vegetables										
Oranges								1 medium		
Apples								1 medium		
Bananas								1 medium		
Pears								1 medium		
Strawberries								1/2 pennet		
Mangoes								1 medium		
Nectarines								1 medium		
Peas								1/2 cup		
Long beans								1/2 cup		

FOOD	rare or never (1)	1-3 times /mth (2)	once a week (3)	2-3 times /week (4)	4-6 times /week (5)	every day (6)	2 or more times /day (7)	Standard portion	No of portion	Cooking Method
French beans								1/2 cup		
Dhal								1/2 cup		
Carrots								1/2 cup		
Cabbage								1/2 cup		
Bean sprout								1/2 cup		
Aubergine/ brinjal								1 medium		
Bitter gourd/karela								1 medium		
Okra								1/2 cup		
Cucumber								1/2 cup		
Cauliflower								1/2 cup		
Lettuce								2 leaves		
Onion								1 medium onion		
Garlic								1 clove garlic		
Ginger								1/2 inch		
Tomato								1 medium		
Sweet corn								1 medium		

FOOD	rare or never (1)	1-3 times /mth (2)	once a week (3)	2-3 times /week (4)	4-6 times /week (5)	every day (6)	2 or more times /day (7)	Standard portion	No of portion	Cooking Method
Potatoes-chips-fried								1 medium portion		
Potatoes-not fried								1 egg size		
Miscelaneous /beverages										
Achar								1 tbs		
Chutney/ pickles								1 tbs		
Sweet								1 sweet		
Chocolates								1 chocolate		
Marmalade/ jam								medium covering per slice bread		
Sugar								1 tsp		
Gur								1 tsp		
Sweetener								1 tablet		
Beverages										
Beer								1/2 pint		
Wine or sherry								1 average glass		

FOOD	rare or never (1)	1-3 times /mth (2)	once a week (3)	2-3 times /week (4)	4-6 times /week (5)	every day (6)	2 or more times /day (7)	Standard portion	No of portion	Cooking method
Tea/Coffee								1 cup		
Orange/Other juice								1 glass		
Squash								1 glass		
Fizzy drink								1 can		

APPENDIX 2 DAILY FOOD CHECKLIST

PLEASE FILL IN THE DATE AND PUT A NUMBER ON THE LINE FOR THE FOODS YOU HAVE EATEN TODAY.

Date ___/___/_____ Day __________

<u>BREADS AND CEREALS</u>

white bread ....slices ____
brown bread..........slices ____
wholemeal bread......slices ____
poori................pieces ____
paratha..............pieces ____
chapatti.............pieces ____
cakes................slices ____
sweet biscuits.......slices ____
Indian sweets........numbers ____
Fried Indian snacks..numbers ____
weetabix................... ____
cornflakes................. ____
rice....................... ____
sweet rice/rice pudding..... ____

<u>CHICKEN, EGGS, LAMB, FISH, AND OTHER MEATS</u>

fried eggs................. ____
boiled eggs................ ____
omelette................... ____
scrambled eggs............. ____
fried fish................. ____
fish curry................. ____
beef curry................. ____
lamb curry................. ____
pork....................... ____
chicken curry.............. ____
sausage.................... ____

<u>MILK AND PRODUCTS</u>

whole milk, full cream....cup ____
semi skimmed.............cup ____
skimmed (low fat)........cup ____
yoghurt..................pot ____
ice cream..............scoop ____
cheese, cottage.............. ____
cheese, processed.......slice ____
pizza...................slice ____

<u>BEVERAGE</u>

beer....................... ____
wine/sherry/spirit......... ____
coffee..................... ____
tea........................ ____
orange/fruit juice......... ____
squash/cordial............. ____
coke/other fizzy drinks.... ____

<u>FATS AND OILS</u>

butter............ ____
margarine......... ____
ghee.............. ____
cooking oil....... ____

<u>FRUITS AND VEGETABLES</u>

oranges........... ____
apples............ ____
bananas........... ____
pears............. ____
strawberries...... ____
peaches........... ____
grapes............ ____
peas.............. ____
long beans........ ____
runner beans...... ____
dal............... ____
carrots........... ____
cabbage........... ____
aubergine......... ____
karela............ ____
okra/bendi........ ____
cucumber.......... ____
cauliflower....... ____
lettuce........... ____
onions............ ____
garlic............ ____
ginger............ ____
tomatoes.......... ____
sweet corn........ ____
chips............. ____
baked/boiled potatoes.......... ____

<u>MISCELLANEOUS</u>

achar............. ____
chutney........... ____
sweet/lollipop.... ____
chocolate......... ____
jam/marmalade..... ____
sugar... .... ____
gur............... ____

Other foods eaten but not on list:

APPENDIX 3

Terms relating to discussions of food are defined below

Food

The individual items which are ingested, together they contain varying amounts of energy and nutrients, and together they make up a diet.

Diet

Diet is the sum total of all foods ingested by an individual; the total to sustain life.

Dietary

Dietary is a prescribed course of diet.

Nutrient

Nutrients are chemically defined compounds which required by the body, either individually or together to sustain life.

Meal

Meal is foods at one of the customary times of eating.

Habits/Practice

It is a permanent tendency to perform certain actions. It is also a custom acquired by frequent repetition.

REFERENCES

REFERENCES

Abraham R, Campbell Brown M, North WRS and McFadyen R (1987). Diets of Asian pregnant women in Harrow: iron and vitamins. *Hum. Nutr. Appl. Nutr.* **41A:**164-173.

Ahuja MMS (1979). Epidemiological studies on diabetes mellitus in India. *Epidemiology of diabetes in developing countries*. MMS Ahuja.Interprint. New Delhi. pp 29-38.

Altman DG (1994). *Practical statistics for medical research*. ed. DG Altman. Chapman and Hall. London. pp 396-435.

Attariwala R (1977). Asian children and school meals in the United Kingdom - a question of choice. *J. Hum. Nutr.* **31:**251-255.

Bailey CJ, Day C and Leatherdale BA (1986). Traditional treatments for diabetes from Asia and West Indies. *Prac. Diab.* **3:**190-193.

Bates CJ, Thurnham DI, Bingham SI, Margetts BM and Nelson M (1991). Biochemical markers of nutrient intake. In *Design concepts in nutritional epidemiology*. eds B.M Margetts and M.Nelson. Oxford University Press. New York. pp 192-265.

Beaton GH, Milner J, Gorey P, McGuire V, Cousins M, Steward E, de Ramos M, Hewitt D, Grambsch PV, Kassim NA and Little JA (1979). Sources of variance in 24 hr dietary recall data:implication for nutrition study design and interpretation. *Am. J. Clin. Nutr.* **32**:2456-2559.

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

Bellach B (1993). Remarks on the use of Pearson's correlation coefficient and other associations measures in assessing validity and realibilty of dietary assessement methods. *Eur. J. Clin. Nutr.* **47 (Suppl 2):**542-545.

Bingham SA, McNeil NI and Cummings (1981). A study of a randomly chosen cross section of British adults in a Cambridgeshire village. *Br. J. Nutr.* **45:**23-35.

Bingham SA (1987). The dietary assessment of individuals; methods, accuracy, new techniques and recommendations. *Nutr. Abstr. Rev.* **57:**705-742.

Bingham SA and Nelson M (1991). Amount of food consumption and nutrient intake. In: *Design concepts in nutritional epidemiology.* eds BM Margetts and M Nelson. Oxford University Press, New York. pp 153-192.

Bingham SA (1994). The use of 24hr urine samples and energy expenditure to validate dietary assessments. *Am. J. Clin. Nutr* **59 (Suppl):**227S-231S.

Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, Sneyd MJ, Key TTA, Roe L and Day NE (1994). Comparison of dietary assessment methods in nutritional epidemiology: weighed record vs 24 hr recall, FFQ and estimated diet record. *Br. J. Nutr.* **72:**619-643.

Bingham SA, Cassidy A, Cole TJ, Welch A, Runswick SA, Black AE, Thurnham D, Bates C, Khaw KT, Key TJA and Day NE (1995). Validation of weighed record and other methods of dietary assessement using 24 hr urine nitrogen technique and other biological markers. *Br. J. Nutr.* **73:**531-550.

Black AE, Goldberg GR, Jebb SA, Livingstone MBE, Cole TJ and Prentice AM (1991). Critical evaluation of energy intake using fundamental principles of energy physiology

:2. Evaluating results of published surveys. *Eur. J. Clin. Nutr.* **45:**583-599.

Bland JM and Altman GA (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **i:**307-310.

Block G (1982). A review of validation of dietary assessment methods. *Am. J. Epid.* **124:**492-505.

Block G, Hartman AM, Dresser CM, Gannon J and Gardner L (1986). A data-based approach to diet questionnaire design and testing. *Am. J. Epid.* **124:**453-469.

Block G and Hartman AM (1989). Issues in reproducibility and validity of dietary studies. *Am. J. Clin. Nutr.* **50:**1133-1138.

Boeing H, Wahrendorf J, Heinemann L, Kulesza W, Rywik SL, Sznajd JA and Thiel C (1989). Results from a comparative dietary assessment in Europe: 1. Comparison of dietary information from concurrently applied frequency questionnaires and qualitative measurement instrument. *Eur. J. Clin. Nutr.* **431:**367-377.

Bollard JE, Yuhas JA, and Bollard TW (1988). Estimation of food portion sizes: effectiveness of training. *J. Am. Diet. Assoc.* **88:**817-820.

Braddon FEM, Wadsworth MEJ, Davies JMC, Cripps AA (1988). Social and regional differences in food and alcohol consumption and their measurement in a national birth cohort. *J. Epid. Comm. Health.* **42:**41-349.

Burr ML (1991). Cohort studies. In: *Design concepts in nutritional epidemiology.* eds BM Margetts and M Nelson. Oxford University Press. New York. pp 369-384.

Burke BS (1947). The dietary history as a tool in research. *J. Am. Diet. Assoc* **23**:1041-1046.

C.S.O. (1990). Social Trends 1992. London: Central Statistical Office. H.M.S.O.

Cade JE (1988). Are diet records using household measures comparable to weighed intakes? *J. Hum. Nutr. Diet.* **1**:171-178.

Cade JE and Margetts BM (1988). Nutrient sources in English diet: quantitative data from three English towns. *Int. J. Epid.* **17**:844-848.

Clayon D and Gill C (1991). Covariate measurement errors in nutritional epidemiology. In: *Design concepts in nutritional epidemiology*. eds BM Margetts and M Nelson. Oxford University Press, New York.pp 79-96.

Clayton D (1994). Measurement error: effects and remedies in nutritional epidemiology. *Proc. Nutr. Soc.* **53**:37-42.

Coggon D (1991). Case control and cross sectional studies. In: *Design concepts in nutritional epidemiology*. eds BM Margetts and M Nelson. Oxford University Press. New York. pp 354-368.

Cole TJ (1991). Sampling, study size and power. In: *Design concepts in nutritional epidemiology*. eds BM Margetts and M Nelson. Oxford University Press, New York. pp 53-78.

Crawley H (1988). Food portion sizes. Ministry of Agriculture, Fisheries and Food. London:H.M Stationery Office.

Delcourt C, Cubeau J, Balkau B, Papoz L and CODIAB-INSERM-ZENECA Pharma study group (1994). Limitation of correlation coefficient in the validation of diet assessment method. *Epidemiology*. **5**:518-524.

Ebden P, Bhatt A, Carey JC and Harrison B (1988). The bilingual consultation. *Lancet* **i**:347.

Edington J, Thorogood M, Geekie M, Ball M and Mann J (1989). Assessment of nutritional intake using dietary records with estimated weights. *J. Hum. Nutr. Diet.* **2**:407-415.

Edington J, Geekie M, Carter R, Benfield L, Fisher K, Ball M and Mann JI (1987). Effect of dietary cholesterol on plasma cholesterol concentration in subjects following reduced fat, high fibre diet. *Br. Med. J.* **294**:333-336.

FAO/WHO/UNU (1985). Energy and protein requirements. Report of a joint FAO/WHO/UNU consultation. Technical Report Series 724. Geneva: World Health Organization.

Flegal KM, Larkin FA, Mertzner HL, Thompson FE and Guire KE (1988). Counting calories: Partitioning energy intake estimate from a food frequency questionnaire. *Am. J. Epid.* **128**:749-760.

Gelbier S and Taylor SGBWS (1985). Some Asian communities in the UK and their culture. *Br. Den. J.* **June 8**:417-418.

Goldbohm RA, van den Brandt PA, Brants HAM, van't Veer P, Al M, Sturmans F and Hermus RJJ (1994). Validation of a dietary questionnaire used in a a large scale prospective cohort study on diet and cancer. *Eur. J. Clin. Nutr.* **48**:253-265.

Gregory J, Foster K, Tyler H and Wiseman M. (1990). The dietary and nutritional survey of British adults. London:HMSO

Hales CN, Barker DJP, Clark PMS, Cox LJ, Osmond C and Winter PD (1991). Fetal and infant growth and impaired glucose tolerance at age 64 years. *Br. Med. J.* **303**:1019-1022.

Haraldsdottr J (1993). Minimizing error in the field: quality control in dietary surveys. *Eur. J. Clin. Nutr.* **(Suppl 2)**:S19-S24.

Harbottle L and Duggan MB (1993). Dietary assessment in Asian children - a comparison of the weighed inventory and dietary history methods. *Eur. J. Clin. Nutr.* **47**:666-672.

Harbottle L and Duggan MB (1994). Daily variation in food and nutrition intake of Asian children in Sheffield. *Eur. J. Clin. Nutr.* **48**:66-70.

Hawthorne K, Mello M and Tomlinson S (1993). Cultural and religious influences in diabetes care in Great Britain. *Diab. Med.* **10**:8-12.

Hayter JE and Henry CJK (1994). A reexamination of BMR predictive equations: the importance of geographic origin of subjects in sample selection. *Eur. J. Clin. Nutr.* **48**:702-707.

Health Education Authority (1994). Black and minority ethnic groups in England. MORI Health Research Unit. HEA.

Heitmann BL and Lissner L (1995). Dietary underreporting by obese individual -is it specific or non specific? *Br. Med. J.* **311**:986-989.

Hennekens CH and Buring JE (1987). *Epidemiology in medicine.* ed SL Mayrent. Little,Brown and Company. Boston/Toronto. pp 243-304.

Henry CJK, Piggott S and Emery B (1987). Basal metabolic rate and diet induced thermogenesis in Asians living in Britain. *Hum. Nutr. Clin. Nutr.* **41C:**397-402.

Henry CJK and Rees DG (1991). New predictive equation for the estimation of basal metabolic rate in tropical peoples. *Eur. J. Clin. Nutr.* **45:**177-185.

Hunt S (1977) Traditional Asian food customs. *J. Hum. Nutr.* **31:**245-248.

Holland B, Unwin I and Buss DH (1988). Cereal and cereal products. Third supplement to Mc Cance and Widdowson's composition of foods. H.M. Stationery Office.

Holland B, Unwin I and Buss DH (1989). Milk and milk products. Fourth supplement to Mc Cance and Widdowson's composition of foods. H.M. Stationery Office.

Holland B, Unwin I and Buss DH (1991a). Vegetables, herbs and spices. Fifth supplement to Mc Cance and Widdowson's composition of foods. H.M. Stationery Office.

Holland B, Welch AA, Unwin I, Buss DH, Paul AA and Southgate DAT (1991b). Mc Cance and Widdowson's The Composition of Foods. 5th ed. Royal Society of Chemistry, Ministry of Agriculture, Fisheries and Foods.

Isaksson B (1980). Urinary nitrogen as a validity test in dietary surveys. *Am. J. Clin. Nutr.* **33:**4-12.

Jackson AA (1992). Balance and turnover. In: Lecture notes from a training course for clinicians in practical nutritional support. ed B Moran and AA Jackson. Dept.Human Nutrition. pp 4-7.

Jain MG, Harrison L, Howe GR and Miller AB (1982).

Evaluation of self administered dietary questionnaire for use in a cohort study. *Am. J. Clin. Nutr.* **36**:931-935.

Jain M, Howe GR and Rohan T (1996). Dietary assessment in epidemiology: Comparison of a food frequency and dietary history questionnaire with a 7d food record. *Am. J. Epid.* **143**:953-960.

Jorgensen LM (1992). Who completes seven day food records? *Eur. J. Clin. Nutr.* **46**:735-741.

Joreskog KG and Sorbom D (1988). Lisrel 7. A guide to the program and applications, SPSS. Chicago.

Kaaks R, Riboli E, Esteve J, van Kappel AL and van Staveren, WA (1994). Estimating the accuracy of dietary questionnaire assessments: validation in terms of structural equation model. *Stats. Med.* **13**:127-142.

Kaaks R, Riboli E and van Staveren WA (1995). Calibration of dietary intake measurements in prospective cohort studies. *Am. J. Epid.* **142**:548-556.

Kaaks R, Plummer M, Riboli E, Esteve J, van Staveren WA (1994). Adjustment for bias due to errors in exposure assessments in multicentre cohort studies on diet and cancer: a calibration approach. *Am. J. Clin. Nutr.* **59 (Suppl)**:245S-250S.

Kaaks R (1996). Biochemical markers as an additional measurement in studies on the accuracy of dietary questionnaire meaurements. *Am. J. Clin. Nutr.* (in press).

Kalka I (1988). The changing food habits of Gujeratis in Britain. *J. Hum. Nutr. Diet.* **1**:329-335

Kassam-Khamis T, Judd P, Thomas JE, Sevak L, Reddy S and

Ganatra S (1995). Frequency of consumption and nutrient composition of composite dishes commonly consumed by the South Asians originating from Gujerat and the Punjab. *J. Hum. Nutr. Diet.* **8**:265-277.

Khajuria S and Thomas J (1992). Traditional Indian beliefs about the dietary management of diabetes- an exploratory study of the implications for the management of Gujerati diabetics in Britain. *J. Hum. Nutr. Diet.* **5**:311-321.

Knight TM, Smith Z, Whtteleo A, Sahota P, Lockton JA, Hogg G, Bodfords A, Toop M, Kernohan EEM and Baker MR (1992). Insulin resistance, diabetes and risk markers for ischaemic heart disease in Asian men and non Asian men in Bedford. *Br. Heart. J.* **67**:343-350.

Leatherdale BA, Jhittay P and Gill A (1978). Communication with Asian diabetics. *Br. Med. J.* **2**:1197-1198.

Livingstone MBE, Prentice AM, Strain JJ, Coward WA, Black AE, Barker ME, McKenna PG and Whitehead RG (1990). Accuracy of weighed dietary records in studies of diet and health. *Br. Med. J.* **300**:708-712.

Machin D and Campbell MJ (1987). Significance of the correlation coefficients. In: *Statistical tables for the design of clinical trials.* Blackwell Scientific Publications. Oxford. pp 89-93.

Mares-Perlman JA, Klein BEK, Klein R, Ritter LL, Fisher MR and Freudenheim JL (1993). A diet history questionnaire ranks nutrient intakes in middle aged and older men and women similarly to multiple food records. *J. Nutr.* **123**:489-501.

Margetts BM (1991). Basic issues in designing and interpreting epidemiological research. In: *Design concepts* *in nutritional epidemiology.* eds BM Margetts and M Nelson. Oxford University Press. New York. pp 13-45.

Margetts BM, Cade JE and Osmond C (1989). Comparison of a FFQ with a diet record. *Int. J. Epid.* **18:**868-873.

Margetts BM and Nelson M (1995). Measuring dietary exposure in nutritional epidemiology studies. *Nutr. Research. Rev.* **11:**165-178.

Margetts BM and Pietinen P (1996). Editorial. *Int. J. Epid.* (in press).

Marr JW (1971). Individual dietary surveys: purpose and methods. *World Rev. Nutr. Dietet.***13:**105-164.

Mather HM and Keen H (1985). The Southall Diabetes Survey: prevalence of known diabetes in Asians and Europeans. *Br. Med. J.* **291:**1081-1084.

Mather HM, Verma NPS, Mehta SP, Madhus L and Keen H (1987). The prevalence of known diabetes in Indians in New Delhi and London. *J. Med. Assoc. Thailand.* **70:**54-58.

McKeigue PM, Adelstein AM, Shipley MJ, Riemersma RA, Marmot MG, Hunt SP, Butler SM and Turner PR (1985). Diet and risk factors for coronary heart disease in Asians in Northwest London. *Lancet* **ii:**1086-1090.

McKeigue PM, Marmot MG, Syndercombe Court YD, Cottier DE, Rahman S and Riemersma RA (1988). Diabetes, hyperinsulinaemia and coronary risk factors in Bangladeshis in East London. *Br. Heart. J.* **60:**390-396.

McKeigue P(1993). Coronary heart disease and diabetes in South Asians. Proc. Nat. Conf. of the Health of the Nations. pp 43-53.

McKeigue PM, Shah B and Marmot MG (1991). Relation of central obesity and insulin resistance with high diabetes prevalence and cardiovascular risk in South Asians. *Lancet*. **337**:382-386.

Mela DJ and Aaron JI (1996). Honest but invalid: consumer views on recording their food intake. Nutrition Society Summer Meeting, University of Ulster at Coleraine. Abstract C54.

Morgan RW, Jain M, Miller AB, Choi NW, Matthews V, Munan L, Burch JD, Feather J, Howe GR and Kelly A (1978). A comparison of dietary methods in epidemiologic studies. *Am. J. Epid.* **107**:488-498.

Nelson M (1991). The validation of dietary questionnaire. In: *Design concepts in nutritional epidemiology*. eds BM Margetts and M Nelson. Oxford University Press. New York. pp 266-297.

Nelson M, Atkinson M and Darbyshire S (1994). Food photography 1. The perception of food portion size from photographs. *Br. J. Nutr* **72**:649-663.

Nelson M, Atkinson M and Darbyshire S (1996). Food Photograph 2. Use of food photographs for estimating size and the nutrient content of meals. *Br. J. Nutr.* (in press).

Nelson M (1996) The validation of dietary assessment. In: *Design concepts in nutritional epidemiology*. eds BM Margetts and M Nelson. Oxford University Press. New York. 2nd edition. (in press)

Ocke M and Kaaks R (1996). Biochemical markers as an additional measurement in dietary validity studies: Application of the method of triads with examples from European Perspective Investigation into Cancer and

Nutrition. *Am. J. Clin. Nutr.* (in press).

Office of Population and Census Survey (OPCS) Southampton. 1991

Peterson DB, Dattani JT, Baylis JM and Jepson EM (1986). Dietary practices of Asian diabetics. *Br. Med. J.* **292:**170-171.

Pietinen P, Hartman M, Haapa E, Rasanen L, Haapakosku J, Palmgren J, Albanes D, Virtomo J and Huttunen JK (1988). Reproducibility and validity of dietary assessment instrument. I. A self administered food use questionnaire with a portion size picture. *Am. J. Epid.* **128:**655-666.

Pietinen P, Hartman M, Haapa E, Rasanen L, Haapakosku J, Palmgren J, Albanes D, Virtomo J and Huttunen JK (1988). Reproducibility and validity of dietary assessment instrument. II. A qualitative FFQ. *Am. J. Epid.* **128:**667-676.

Plummer M and Clayton D (1993). Measurement error in dietary assessment: an investigation using covariance structural models. Part 1. *Stats. Med.* **12:**925-935.

Plummer M, Clayton D and Kaaks R (1994). Calibration in multicentre cohort studies. *Int. J. Epid.* **23:**419-426.

Plummer M and Clayton D (1993). Measurement error in dietary studies: an investigation using covariance structural models Part II. *Stat. Med.* **12:**937-948.

Prentice AM, Coward WA, Davies HL, Mugatroyd PR, Black AE, Goldberg GR, Ashford J, Sawyer M and Whitehead RG (1985). Unexpected low levels of energy expenditure in healthy women. *Lancet* **i:**1419-1422.

Prentice AM, Lucas A, Vasquez-Velasquez L, Davis PSW and Whitehead RG (1988). Are current dietary guidelines for young children a prescription for overfeeding? *Lancet* **2**:1066-1069.

Raheja BS, Pathak RB, Vispute KV and Koppikar GD (1989). Dietary habits and prevalence of Type II NIDDM in Indian subjects. *J. Diab. Assoc. India* **29**:3-8.

Ramachandran A, Jali MU, Mohan V, Snehalatha C and Viswanathan M (1988). High prevalence of diabetes in an urban population in South India. *Br. Med. J.* **297**:587-590.

Ramachandran A, Dharmaraj D, Snehalatha C and Viswanathan M (1992). Prevalence of glucose tolerance in Asian Indians. *Diabetes Care* **15**:1348-1355.

Ramachandran A (1993). Epidemiology of diabetes in Indians. *Int. J. Diab. Dev. Countries* **13**:65-67.

Ray TK, Mansell KM and Knight LC (1983). Long term effects of dietary fibre on glucose tolerance and gastric emptying in non insulin dependent diabetes patients. *Am. J. Clin. Nutr* **37**: 376-381.

Randrianjohany A, Balkau B, Cubeau J, Ducimetieve P, warnet JM and Eschwege E (1993). The relationship between behavioral pattern, overall and central adiposity in a population of healthy French men. *Int. J. Obes* **17**:651-655.

Rothenberg E (1994). Validation of FFQ with the 4 day record method and analysis of 24 hour urinary nitrogen . *Eur. J. Clin. Nutr.* **48**:725-735.

Samanta A, Campbell JE, Spalding DL, Panja KK and Neogi SK (1987). Dietary habits of Asian diabetics in a general practice clinic. *Hum. Nutr. Appl. Nutr.* **41A**:160-163.

Schoenberg RJ and Arminger G (1989). LINCS: Linear covariance structure analysis user's guide. RJS Software, Kensington.

Schofield WN, Schofield C and James WPT (1985). Basal metabolic rate. Review and prediction together with an annotated bibliography of source material. *Hum. Nutr. Clin. Nutr.* **39C (Suppl 1)**:5-96.

Sevak L, McKeigue PM and Marmot MG (1994). Relationship of hyperinsulinaemia to dietary intake in South Asians and European men. *Am. J. Clin. Nutr.* **59**:1069-1074.

Silman A, Loysen E, DeGraff W and Sramek M (1985). High dietary fat intake and cigarette smoking as risk factors for ischaemic heart disease in Bangladeshi male immigrants in East London. *J. Epid. Comm. Health.* **39**:301-303.

Simmons D, Williams DRR and Powell MJ (1989). Prevalence of diabetes in a predominantly Asian community: preliminary findings of the Coventry diabetes study. *Br. Med. J.* **298**:18-21.

Simmons D (1992). Parity, ethnic groups and prevalence of type 2 diabetes: The Coventry Diabetes Study. *Diab. Med.* **9**:706-7-9.

Simmons D, Williams DRR and Powell MJ (1992). Prevalence of diabetes in different regional, religious South Asian communities in Coventry. *Diab. Med.* **9**:428-431.

Simmons D and Powell MJ (1993). Metabolic and clinical characteristics of South Asian and Europeans in Coventry. *Diab. Med.* **10**:751-758.

Smith Z, Knight T, Sahota P, Kernohan E and Baker M (1993). Dietary patterns in Asian and Caucasian men in Bradford:

differences and implications for nutrition education. *J. Hum. Nutr. Diet.* **6:**323-333.

Soares MJ, Francis DG and Shetty PS (1993). Predictive equation for basal metabolic rate of Indian males. *Eur. J. Clin. Nutr.* **47:**389-394.

Sorensen AW, Calkins BM, Conolly MA and Diamond E (1985). Comparison of nutritional intake determined by four dietary intake instrument. *J. Nutr. Educ.* **3:**92-99.

Stephens WP, Klimiuk PS, Warrington S and Taylor JL (1982). Observations on the dietary practice of Asians in the United Kingdom. *Hum. Nutr. Appl. Nutr.* **36:**438-444.

Stuff JE, Garza G, O'Brian Smith E, Nichols BL and Montandon CM (1983). A comparison of dietary methods in nutritional studies. *Am. J. Clin. Nutr.* **37:**300-306.

Tan SP, Wenlock RW and Buss DH (1985). Immigrant foods. Second supplement to Mc Cance and Widdowson's The Composition of Food. London: H.M.S.O.

Tanhanand S and Vannasaeng S (1988). *Diabetes in Asia.* World Book of Diabetes in Practice, Vol 3,327-334. Elsevier Science Publisher BV (Biomedical Division).

Thai AC, Yeo PBB and Lun KC (1987). Changing prevalence of diabetes mellitus in Singapore over a ten year period. *J. Med. Assoc. Thailand.* **70 (2):**54-58.

Thompson RL and Margetts BM (1993). Comparison of a FFQ with a 10 day weighed record in cigarette smokers. *Int. J. Epid.* **22:**824-833.

Tjonneland A, Overad K, Haraldsdottir J, Bang S, Ewertz M and Jensen OM (1991). Validation of a semiquantitative FFQ developed in Denmark. *Int. J. Epid.* **20:**906-912.

Tylavsky FA and Sharp GB (1995). Misclassification of nutrients and energy intake from use of close ended questions in epidemiological research. *Am. J. Epid.* **142:**342-352.

van Staveren WA, de Boer JO and Burema J (1985). Validity and reproducibility of dietary history method estimating the usual food intake during one month. *Am. J. Clin. Nutr.* **42:**554-559.

van't Veer P, Kardinaal AFM, Bausch-Goldbohm RA and Kok FJ (1993). Biomarkers for validation. *Eur. J. Clin. Nutr.* **47:**S58-S63.

Verma NPS, Mehta SP, Madhu S, Mather HM and Keen H (1986). Prevalence of known diabetes in an urban Indian environment:The Darya Ganj Diabetes Survey. *Br. Med. J.* **293:**423-424.

Waldhausl W, Rathauser K and Konjati M (1989). Increase of insulin sensitivity and improvement of intravenous glucose tolerance by fish oil in healthy man. In: *Health effects of fish and fish oils.* ed RK Chandra. St John's Art Biomedical. pp 171-187.

Wharton PA and Eaton PM (1983). Sorrento Asian food tables: Food tables, recipes and customs of mothers attending Sorrento Maternity Hospital, Birmingham, England. *Hum. Nutr. Appl. Nutr.* **37A:**378-402.

Wheeler C, Rutihauser I, Conn J and O'Dea K (1994). Reproducibilty of a meal based food frequency questionnaire. The influence of format and time interval between questionnaires. *Eur. J .Clin. Nutr.* **48:**795-809.

WHO study group report on diabetes mellitus. (1985) WHO Technical Report Series 727, WHO, Geneva.

Willett WC (1994). Future directions in the development of food frequency questionnaire. *Am. J. Clin. Nutr. 59 (Suppl)*:171S-174S.

Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH and Speizer FE (1985). Reproducibility and validity of a semiquantattive FFQ. *Am. J. Epid.* **122**:51-65.

Willett WC (1990). Reproducibility and validity of food frequency questionnaire. In *Nutritional epidemiology*. ed WC Wilett. Oxford University Press. New York. pp 92-126.

Willett WC (1990). Correction for the effects of measurement error. In *Nutritional epidemiology*. ed WC Willett. Oxford University Press. New York. pp 272-291.

Willett WC (1990). Implications of total energy intake for epidemiological analyses. In *Nutritional epidemiology*. ed WC Willett. Oxford University Press. New York. pp 245-271.

Wolk A, Bergstrom R, Adami HO, Ohlander EM, Bruce A, Holmberg L and Bergkvist L (1994). Self administered questionnaire: The effect of different designs on food and nutrient estimates. *Int. J. Epid.* **23**:570-576.

Yarnell JWG, Ichily AM, Milbank JE, Sweetnam PM and Walker CL (1983). A short dietary questionnaire for use in an epidemiology survey: comparison with weighed dietary records. *Hum. Nutr. Appl. Nutr.* **37A**:103-112.