Mal J Nutr 17(1): 1 - 18, 2011

Development, Validity and Reproducibility of a Food Frequency Questionnaire in Pregnancy for the Universiti Sains Malaysia Birth Cohort Study

Loy SL[1], Marhazlina M[1], Nor Azwany Y[2] & Hamid Jan JM[1 *]

[1] *Nutrition Programme, School of Health Sciences, Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia*
[2] *Community Medicine Department, School of Medical Sciences, Universiti Sains Malaysia, Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia*

ABSTRACT

Introduction: This study aimed to develop and examine the validity and reproducibility of a semi-quantitative food frequency questionnaire (FFQ) among Malay pregnant women in Kelantan, Malaysia. **Methods:** A total of 177 Malay pregnant women participated in the validation study while 85 of them participated in the reproducibility study which was carried out in the antenatal clinic of Universiti Sains Malaysia Hospital. The newly developed FFQ was validated against two 24-hour dietary recalls (DR). The FFQ was repeated 20 to 28 days apart. **Results:** Results showed that the FFQ moderately over estimated the nutrient and food intakes compared to the DR. Spearman correlation coefficients for nutrients ranged from 0.24 (fat) to 0.61 (calcium) and for foods, ranged from 0.13 (organ meats, onion and garlic) to 0.57 (malt drink). For nutrients, 72 to 85% of women were classified into the correct quartiles from the FFQ and the DR while for foods, 67 to 85% of women were classified correctly. Bland-Altman plot showed relatively good agreement between these two dietary methods. The intra-class correlation (ICC) was used to estimate reproducibility. It ranged from 0.75 (vitamin C) to 0.94 (phosphorus) for nutrients while it ranged from 0.73 (confectionary) to 0.96 (coffee) for foods. **Conclusion:** On average, at least 90% of pregnant women were correctly classified into the quartiles for nutrients and foods from the two sets of the FFQ. The FFQ presented acceptable reproducibility and appears to be a valid tool for categorising pregnant women according to dietary intake.

Keywords: Assessment of nutritional status, diet, food frequency questionnaire, maternal nutrition.

INTRODUCTION

The diet-disease relationship is important as it helps in identifying ways to change or supplement a diet for health improvement (Emmett, 2009). This is particularly crucial in pregnancy as it helps to improve maternal health and birth outcomes. Variations in maternal dietary intake have shown potential programming effects on adult diseases (Sayer & Cooper, 2005). Thus, renewed interest in pregnancy nutrition has

[*] Correspondence author: Hamid Jan Jan Mohamed; Email: *hamidjan@kb.usm.my*

been generated by the 'developmental origins' hypothesis (Barker, 1992).

As maternal nutrition is a determinant in subsequent maternal health and fetal growth, assessment of maternal diet is crucial to investigate the long term effects on pregnancy outcomes. The food frequency questionnaire (FFQ) has been shown to be a valuable tool to assess diet in a wide variety of epidemiologic settings, including prospective studies in pregnancy (Knudsen *et al.*, 2008).

Dietary habits vary greatly depending on the demographic and cultural background of target participants. The FFQ must be tailored specifically to encompass the diet of the respective population. Although FFQs have been used to assess maternal nutrition by other investigators (Knudsen *et al.*, 2008), it has not been able to reflect the habitual dietary intake in Malaysian pregnant women. Any newly developed FFQ must be assessed for validity and reproducibility in nutritional epidemiology (Cade *et al.*, 2002).

FFQ validation is essential to minimise measurement errors of associations between diet variables and disease markers (Cade *et al.*, 2004). FFQ reproducibility is important to determine repeatability results in the same situation (Emmett, 2009). Validity and reproducibility of the FFQ for pregnant women have been examined in several studies (Erkkola *et al.*, 2001; Baer *et al.*, 2005). The 24-hour dietary recall (DR) has been reported as a suitable reference method to assess the validity of the FFQ (Cade *et al.*, 2002). However, a sufficient number of independent replicate 24-hour DRs is needed to represent average dietary intakes (Cade *et al.*, 2002). Since one day is unlikely to represent an individual's habitual intake, thus two days of 24-hour DR could be used to reduce the chance of variation between the dietary methods (Mouratidou, Ford & Fraser, 2006).

As an integral part of the Universiti Sains Malaysia (USM) birth cohort study, developing valid and reliable FFQ to elicit information about the usual dietary intake in pregnancy is essential. Hence, the aim of the present study was to validate the FFQ against the 24-hour DR among pregnant women. The FFQ was also designed to rank pregnant women according to their nutrient and food intake.

METHODS

Study subjects

The study subjects consisted of pregnant women attending antenatal clinic at USM Hospital, Kelantan, which is one of the states in North-east Peninsular Malaysia with a 95% majority Malay ethnicity. All the subjects were selected based on a purposive sampling technique. The inclusion criteria of the subjects were (1) Malaysian and Malay ethnicity; (2) aged 19 to 40 year old; (3) within gestational age of 12 to 22 weeks and 28 to 38 weeks based on last menstrual period or ultrasound scan; (4) able to give informed consent. Pregnant women with multiple gestations and pre-existing chronic diseases were excluded. A total of 178 pregnant women were recruited into the pilot study, 177 participated in the validation study while 85 were involved in the reproducibility study. The study was approved by the USM Human Research Ethics Committee.

Phase 1: Development of the semi-quantitative FFQ

Development of the FFQ followed three major steps. First, was the construction of a food list in the pilot 24-hour DR study. Second, was the prioritisation of the food list and categorisation of the food items while the third was the assignment of food frequency intake and portion size.

Step 1: Pilot 24-hour Dietary Recall Study

A pilot study on 24-hour DR was conducted among 178 pregnant women within the gestational weeks of 12 to 22 and 28 to 38 respectively, from August to September 2009.

This group of pregnant mothers who participated in the pilot study was different from those who participated in the validation study. Weekdays and weekend food intakes were obtained from the mothers through interview. A list of seasonal foods was asked at the end of the DR. To our knowledge, there has been no published data on the dietary habits of pregnant women in Malaysia; thus this data from the DR were used to develop a new structured semi-quantitative FFQ designed for use on Kelantan pregnant women.

Step 2: Food list identification and food group categorisation for FFQ

A total of 295 food items which covered food intakes on weekdays and weekends was derived from DR. The mean intake of each food item was calculated and the percentage of energy and nutrients were derived using the Block equation (Block *et al.*, 1985). The nutrients contributed by each food item were determined followed by calculating the cumulative percent distribution of total energy and each macronutrient and micronutrient until at least 90% of them had been included. The number of foods which contributed to 90% of total energy, protein, carbohydrate, fat, sodium, potassium, calcium, iron, phosphate, vitamin A, vitamin C, thiamin, riboflavin, niacin and dietary fibre were 77, 63, 62, 48, 43, 90, 61, 70, 70, 44, 33, 64, 68, 54 and 68 respectively. A total of 145 food items was determined from the analysis. These food items and mixed dishes were organised into 10 main food groups, which were: (1) Cereals and cereal products; (2) Meat and eggs, (3) Milk and milk products; (4). Nuts; (5) Vegetables; (6) Fruits and fruit juices; (7) Beverages; (8) Fats; (9) Sweet and baked goods; (10) Condiments. The arrangement of the food list was not in specific order as there is no major impact of nutrient estimates from the questionnaire (Wheeler *et al.*, 1994).

Food items within each food group were further grouped into smaller food categories based on similar food types for simpler administration. Those individual food items which were inappropriate to incorporate them into a certain food group were preserved. Finally, it yielded a FFQ comprising 82 food categories. Two summary questions on overall fruit and vegetable intakes per week were asked to reduce over- or under-estimation of fruit and vegetable intake. Questions on the overall cooking oil and salt used for family per month, number of cooking days per month and number of family members were also asked. The mean intake represented the individual daily oil and salt consumption.

.

Step 3: Assignment of food frequency and portion size

The frequency of intake was based on habitual intake over the past six months. There were four options in the category for frequency of intake, which were 'per day', 'per week', 'per month' and 'never'. Subjects were asked on the frequency of intake for each food category by responding to the number of times in either one of the four options. For seasonal food intake, subjects would respond to the frequency of intake at that time.

In order to obtain food portion size, we reviewed all reported portions based on DR and referred to the Malaysia food composition books and atlas (Tee *et al.*, 1997; Suzana *et al.*, 2009). One or few commonly consumed portion sizes were defined as the units of measurement for each food category. The portion size was based on the medium size.

Phase 2: Validation of the FFQ

Four local dieticians reviewed the FFQ to confirm content validity in October 2009. Comprehensibility and applicability of the newly developed FFQ were evaluated on 177 pregnant women to assess face validity. Overall, the FFQ at this initial stage was judged to have a good comprehensibility by 174 (98%) mothers and with adequate chosen portion sizes by 171 (97%) mothers. The mothers agreed that their usual food intakes were covered in the FFQ and the indicated portion sizes were compatible to their intakes. Only 4 (2%) mothers responded that the aided tools did not help in choosing portion size and 11 (6%) mothers stated that the FFQ was difficult to fill in and time consuming. On average, 20 to 30 min was spent to complete the FFQ administered by an interview. Based on the feedback from the dieticians and the pregnant women, modifications were made on food nomenclature, food category and portion size.

The questionnaires were administered on an interview basis to enhance the response rate and data accuracy. Both FFQ and DR interviews for each mother were conducted by the same interviewer to reduce inter-rater bias. Dietary data collection was carried out from November 2009 to April 2010.

Semi-quantitative FFQ administration

The FFQ was interview-administered to 177 pregnant women at 28 to 38 weeks of gestation by the dieticians. The subjects were first asked if they had consumed that particular food in the FFQ during the preceding six months of pregnancy. If they took the food, they were asked on how often the food was consumed and the serving size.

24-hour Dietary Recalls (DR)

The first 24-hour DR (DR1) was conducted immediately after the first FFQ (FFQ1) administration. There was no time gap between the FFQ1 and DR1 administration due to consideration of the follow-up rate. After 10 to 14 days of DR1 administration, the same subjects were re-contacted to carry out the second 24-hour DR (DR2). Estimated mean daily intake of nutrients and foods from the two DR were used as reference values to compare with the results from the FFQ.

Phase 3: Reproducibility of the FFQ

Repeated FFQ was assessed in 85 pregnant women as 68 of them refused to be followed-up for the reproducibility study, eight could not be traced after follow up and 16 were admitted to the ward for delivery after DR2 administration. The same FFQ (FFQ1 and FFQ2) was administered 20 to 28 days apart in the clinic.

Nutrient intake

Amount of daily food intake was calculated according to the formula: frequency of intake per day x serving size x total number of servings x weight of food in one serving (Wessek Institute of Public Health, 1995) and entered into the Nutritionist Pro™ software (Axxya Systems LLC., USA) to obtain the energy and nutrient values. For food items which are not available in the Malaysian Nutrient Composition of Foods such as goat milk, honeydew and dates, USDA nutrient database (U.S. Department of Agriculture, 2009) was referred. For mixed dishes that were not available in the database, local recipes were entered into the software. Assumptions about the relative frequencies of intake and portion sizes of the foods were made to compute the gram weights of each of the 82 food categories in FFQ (Cade *et al.*, 2002). No subject was removed due to implausibly low energy intake (<800 kcal).

Statistical analysis

Validity

The median and 5th and 95th percentiles were computed for the FFQ1 and the average DR. Relative differences were calculated between the two methods. Non-parametric statistical method was used as the nutrient and food intakes distributions were skewed. Wilcoxon's sign rank test was used to test the differences between the two dietary methods. Spearman correlation coefficients were computed to determine the strength of relation between the assessments for absolute values and energy-adjusted values. Total energy intake was adjusted using residual method described by Willett (1998). The ability of the FFQ to categorise individuals into quartiles of the nutrient and food distributions when compared with DR was evaluated using cross-classification analysis. Individuals were classified correctly if they were assigned into the same or within one quartile by both methods. Bland-Altman plot was performed to further assess the agreement between the dietary data obtained from the FFQ and DR, which was defined as limit of agreement (LOA; $\pm$ 2SD).

Reproducibility

The median and 5th and 95th percentiles were computed for the FFQ1 and the FFQ2. Relative differences between the two series of assessments were calculated. Wilcoxon's sign rank test was used to test their differences. Intra-class correlation (ICC) was used to measure the reproducibility between the two sets of FFQ. The dietary variables were log-transformed to meet the assumptions of normal distribution for the calculation of ICC. The reproducibility of absolute nutrient and food intakes from the FFQ1 and FFQ2 were further assessed by cross-classification analysis.

All statistical evaluations were computed using the Statistical Package for the Social Sciences (SPSS) version 12.0 (SPSS, Inc., Chicago, IL, USA). Two sided $p < 0.05$ was considered statistically significant at 95% confidence interval.

RESULTS

The selected characteristics of pregnant women in validation and reproducibility studies were similar, except for the educational level and incidence of hyperemis gravidarum during pregnancy. Higher educated pregnant women (47.0%) were more likely to participate in a reproducibility study. A lesser proportion of pregnant women who experienced hyperemis gravidarum during pregnancy participated in the reproducibility study (4.7%) compared to those in the validation study (7.3%) (Table 1). None of them reported smoking.

Validity

Tables 2 and 3 show median daily intake of nutrients and food, and the correlations between the two methods. The FFQ produced a significantly higher intake of most nutrients (average relative difference 34%) and foods (average relative difference 22%) than the DR. Both methods gave rather similar intakes for 3 nutrients and 12 foods. Compared with the DR, the most marked over estimation of nutrients by the FFQ was observed for vitamin C (97%). Reporting bias between the FFQ and the DR was larger for food intake compared to nutrient intake. Intake of fresh fruits showed the greatest overestimation by the FFQ (162%).

On average, Spearman correlation of absolute nutrient intake between the two dietary assessments was 0.46, ranging from 0.24 ($p<0.01$) (fat) to 0.61 ($p<0.01$) (calcium), whereas for food intake, it was 0.37, ranging from 0.13 ($p>0.05$) (organ meats, onion and garlic) to 0.57 ($p<0.01$) (malt drinks). The correlations for nutrients and foods were similar when pregnant women who were co-morbidities and hyperemesis gravidarum were excluded from the analysis (data not shown). The average correlations decreased after energy adjustment for intakes of nutrients (0.31) and food (0.32). The greatest changes in nutrients and food were observed in carbohydrate and coffee respectively.

For nutrients (Table 2),an average 79% of pregnant women, ranging from 72% to 85% were classified into the correct quartile while less than 10% of them were grossly misclassified when ranked by the FFQ and the DR. For foods (Table 3), an average 77% of pregnant women, ranging from 67 to 85% were correctly classified while less than 15% of them were grossly misclassified.

Table 1. Characteristics of pregnant women in validation study vs reproducibility study

	*Validation study (n=177)**	*Reproducibility study (n=85)**
	Mean (SD)	**Mean (SD)**
Age at recruitment (years)	29.9 (5.1)	30.1 (4.6)
Gestational age at recruitment (weeks)	32.9 (4.3)	32.1 (4.2)
BMI prior to pregnancy (kgm^{-2})	24.5 (5.2)	25.0 (5.2)
BMI at recruitment (kgm^{-2})	29.1 (5.1)	29.4 (5.1)
Weight gain rate (kg/week)	0.3 (0.2)	0.3 (0.2)
	n (%)	**n (%)**
Gravidity		
Primigravida	45 (25.4)	21 (24.7)
Multigravida	113 (63.8)	55 (64.7)
Grand multigravida	19 (10.7)	9 (10.6)
Occupation		
Employed	94 (53.1)	49 (57.6)
Unemployed	83 (46.9)	36 (42.4)
Education		
Primary school	3 (1.7)	2 (2.4)
Secondary school	97 (54.8)	43 (50.6)
University/ college	77 (43.5)	40 (47.0)
Monthly household income (Ringgit Malaysia)		
<1500	56 (31.8)	23 (27.1)
1500-3500	77 (43.8)	39 (45.9)
>3500	43 (24.4)	23 (27.1)
Health status		
Co-morbidities during pregnancy‡	36 (20.3)	16 (18.8)
Hyperemis gravidarum during pregnancy	13 (7.3)	4 (4.7)

SD, standard deviation; BMI, body mass index
* Number may not total 177 or 85 for each characteristic due to missing values
‡ Included gestational diabetes mellitus, pregnancy induced hypertension, anemia, renal infection

Spearman correlations were less than 0.3 when at least 10% of pregnant women were grossly misclassified, with the exception of fat and organ meats.

Bland-Altman plots were created to present the differences between the FFQ and the DR for absolute intakes of nutrient and food. The majority of the nutrients and foods showed a similar pattern to the plots as shown in Figure 1, for which the differences in absolute intake increased with increasing intake or the discrepancies between the two methods were equally distributed in either direction. The scattered plots were predominantly distributed within the 95% limit of agreement for all nutrients and foods. The positive and negative values shown by the differences between the assessments imply over- and under-reporting of the FFQ compared to the DR.

Reproducibility

Tables 4 and 5 present the median daily intakes of nutrient and food and the correlations between the FFQ1 and the FFQ2. The intake of most nutrients and foods were larger when estimated by the FFQ2 compared to the FFQ1, but the majority did

Table 2. Validation study: median daily intake of nutrient, relative difference, Spearman correlation coefficient and cross-classification for the comparison between FFQ1 and 24-hour DR in pregnant women (n=177).

Nutrient	*FFQ1*	*24-hour DR*			*Spearman correlation*		*Cross-classification into quartiles (%)*	
	Median (P_5, P_{95})	*Median (P_5, P_{95})*	*P value*†	*Relative difference*‡ *(%)*	*Unadjusted*	*Energy-adjusted*	*Correctly classified*§	*Grossly misclassified*¶
Energy (kcal/d)	2021 (1117, 3520)	1811 (1087, 2834)	<0.001	17	0.47**	-	77	5
Protein (g/d)	78.9 (39.4, 158.1)	69.9 (44.8, 108.8)	<0.001	23	0.42**	0.21**	74	3
Carbohydrate (g/d)	306.4 (159.8, 541.1)	245.6 (152.5, 416.0)	<0.001	30	0.48**	0.19*	75	2
Fat (g/d)	53.9 (27.0, 99.5)	56.5 (32.8, 103.9)	0.28	5	0.24**	0.01	72	8
Sodium (mg/d)	2586.2 (1065.1, 6349.8)	2122.9 (1040.6, 4528.7)	<0.001	40	0.39**	0.29**	73	4
Potassium (mg/d)	1647.1 (818.6, 3088.2)	1552.4 (844.6, 2794.2)	0.09	13	0.50**	0.29**	81	2
Calcium (mg/d)	830.6 (302.0, 1949.1)	626.4 (271.4, 1354.1)	<0.001	46	0.61**	0.56*	85	2
Iron (mg/d)	20.3 (8.7, 40.5)	18.8 (9.1, 36.6)	0.02	19	0.45**	0.40**	80	3
Phosphorus (mg/d)	967.7 (467.9, 1905.9)	1103.2 (661.6, 1935.7)	<0.001	-6	0.37**	0.22**	75	4
Vitamin A, RE (μg/d)	987.7 (406.8, 4906.8)	971.8 (452.3, 3704.2)	0.17	49	0.38**	0.30**	76	5
Vitamin C (mg/d)	127.2 (44.0, 334.1)	96.4 (23.7, 280.6)	<0.001	97	0.47**	0.28**	79	3
Thiamin (mg/d)	1.6 (0.7, 3.0)	1.1 (0.5, 2.1)	<0.001	55	0.58**	0.48**	84	3
Riboflavin (mg/d)	2.0 (0.9, 4.1)	1.6 (0.7, 2.9)	<0.001	47	0.55**	0.43**	83	3
Niacin (mg/d)	15.9 (7.0, 32.4)	12.8 (0.9, 22.6)	<0.001	34	0.47**	0.24**	81	2
Dietary fibre (g/d)	7.0 (2.0, 14.9)	5.5 (1.8, 12.0)	<0.001	43	0.54**	0.50**	83	3

FFQ1, food frequency questionnaire during 1st administration; DR, dietary recall; P_5, 5th percentile; P_{95}, 95th percentile; RE, retinol equivalents

† Wilcoxon signed rank test to test for differences between FFQ1 and DR

‡ Relative difference = [(FFQ1 – DR)/ DR]*100

§ Correctly classified if classified into same or adjacent (+/ - 1) quartiles

¶ Grossly misclassified if classified into opposing quartiles

* P<0.05; ** P<0.01

Table 3. Validation study: median daily intake of food, relative difference, Spearman correlation coefficient and cross classification for the comparison between FFQ1 and 24-hour DR in pregnant women (n=177).

Food group	*FFQ1*	*24-hour DR*			*Spearman correlation*		*Cross-classification into quartiles (%)*	
	Median (P_5, P_{95})	*Median (P_5, P_{95})*	*P value*† (%)	*Relative difference*‡	*Unadjusted adjusted*	*Energy-*	*Correctly classified*§	*Grossly misclassified*¶
Cereal products	431.0 (195.1, 786.7)	391.0 (204.0, 633.8)	<0.001	23	0.45**	0.45**	79	5
Poultry	17.1 (1.0, 120.0)	59.0 (0, 181.5)	<0.001	-43	0.49**	0.48**	79	4
Red meats	3.0 (0, 42.9)	0 (0, 61.5)	0.62	-45	0.33**	0.31**	80	1
Organ meats	1.7 (0, 57.9)	0 (0, 33.0)	<0.001	-47	0.13	0.10	74	3
Fish and other seafood	96.6 (18.9, 325.2)	72.4 (0, 224.0)	<0.001	76	0.45**	0.38**	77	5
Eggs	13.9 (1.6, 48.3)	11.5 (0, 58.1)	0.15	7	0.33**	0.26**	73	9
Dairy products	49.2 (0, 272.5)	22.0 (0, 256.0)	<0.001	84	0.55**	0.46**	82	7
Nuts and legume	2.0 (0, 32.1)	0 (0, 47.7)	0.10	-30	0.30**	0.25**	77	5
Leafy vegetables	18.4 (1.4, 98.3)	19.3 (0, 97.3)	0.46	53	0.32**	0.25**	76	7
Cruciferous vegetables	7.1 (0, 50.0)	8.3 (0, 49.0)	0.25	18	0.38**	0.38**	72	8
Legume vegetables	2.0 (0, 17.5)	0 (0, 24.0)	0.99	-46	0.36**	0.29**	80	2
Tuber vegetables	4.0 (0, 51.8)	0 (0, 64.6)	0.08	-2	0.37**	0.26**	81	3
Raw vegetables ‘Ulam’	1.3 (0, 17.0)	0 (0, 25.7)	0.08	-37	0.38**	0.25**	77	3
Other vegetables	14.7 (0, 103.0)	16.0 (0, 90.1)	0.95	18	0.25**	0.23**	72	10
Onion and garlic	8.0 (0, 36.0)	12.5 (0, 38.2)	0.02	24	0.13	0.15*	69	12
Fresh fruits	144.0 (14.1, 495.4)	128.0 (0, 425.7)	0.01	162	0.39**	0.23**	77	5
Dried fruits	1.8 (0, 37.3)	0 (0, 21.8)	<0.001	52	0.42**	0.41**	79	3
Pickled fruits	0 (0, 16.0)	0 (0, 8.7)	0.01	-38	0.36**	0.34**	80	0
Juices	63.8 (0, 296.3)	0 (0, 291.0)	<0.001	2	0.38**	0.37**	82	1
Tea	51.4 (0, 540.0)	90.0 (0, 273.0)	0.04	5	0.53**	0.57**	75	7
Coffee	0 (0, 180.0)	0 (0, 180.0)	0.27	-38	0.48**	0.05	83	0
Malt drinks	5.1 (0, 36.0)	4.5 (0, 24.0)	0.001	20	0.57**	0.50**	84	4
Cordials	0 (0, 174.3)	0 (0, 300.0)	0.10	-47	0.50**	0.41**	85	0

Table 3: Continued

Food group	*FFQ1*	*24-hour DR*	*P value*†	*Relative difference*‡ *(%)*	*Spearman correlation*		*Cross-classification into quartiles (%)*	
	Median (P_5, P_{95})	*Median (P_5, P_{95})*			*Unadjusted*	*Energy-adjusted*	*Correctly classified*§	*Grossly misclassified*¶
Soybean milk	33.3 (0, 250.0)	0 (0, 227.5)	<0.001	10	0.46**	0.45**	84	0
Fats and oils	25.2 (7.2, 63.1)	22.5 (2.5, 51.7)	0.02	66	0.15*	0.13	70	11
Confectionary	52.1 (5.6, 170.1)	38.0 (0, 180.7)	<0.001	123	0.37**	0.36**	79	7
Biscuits	8.6 (0, 74.4)	0 (0, 48.6)	<0.001	25	0.41**	0.34**	84	2
Sugar	15.0 (0, 90.0)	17.5 (0, 49.5)	0.001	70	0.40**	0.41**	75	7
Salt	1.9 (0.7, 6.7)	2.0 (0.5, 3.8)	0.04	27	0.15*	0.16*	67	10
Condiments	15.3 (1.0, 78.6)	4.8 (0, 53.2)	0.55	153	0.34**	0.30**	72	6

FFQ1, food frequency questionnaire during 1st administration; DR, dietary recall; P_5, 5th percentile; P_{95}, 95th percentile; RE, retinol equivalents

† Wilcoxon signed rank test to test for differences between FFQ1 and DR

‡ Relative difference = [(FFQ1 – DR)/ DR]*100

§ Correctly classified if classified into same or adjacent (+/ - 1) quartiles

¶ Grossly misclassified if classified into opposing quartiles

* *P*<0.05; ** *P*<0.01

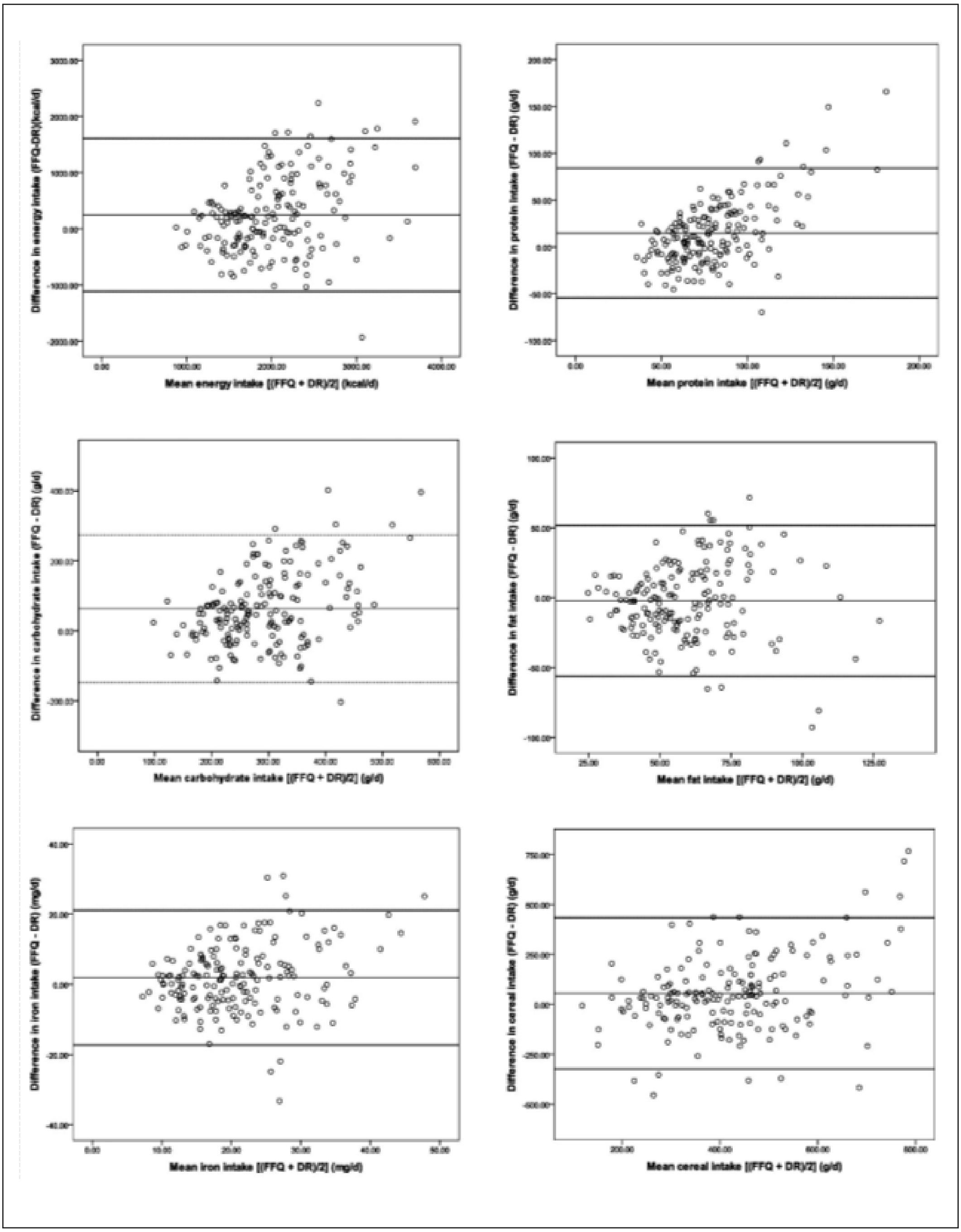

Figure 1. Bland-Altman plot showing agreement between the FFQ1 and the average of 24-h DR for energy, protein, carbohydrate, fat, iron and cereal intakes. The solid line represents the mean difference in absolute intake between the two dietary assessment methods, while the dashed lines represent the limits of agreement (±2SD).

Table 4. Reproducibility study: median daily intake of nutrient, relative difference, Spearman correlation coefficient and cross classification for the comparison between FFQ1 and FFQ2 in pregnant women (n=85).

Nutrient	*FFQ1*	*FFQ2*		*Relative difference*† *(%)*	*Intraclass correlation coefficient*‡	*Cross-classification into quartiles (%)*	
	Median (P_5, P_{95})	*Median (P_5, P_{95})*	*P value**			*Correctly classified*§	*Grossly misclassified*¶
Energy (kcal/d)	2069 (1252, 3599)	2036 (1210, 3411)	0.78	0.03	0.93	96	1
Protein (g/d)	76.7 (43.1, 177.2)	85.7 (44.3, 153.4)	0.15	7	0.91	91	2
Carbohydrate (g/d)	306.4 (174.1, 548.7)	306.4 (181.1, 514.2)	0.13	-2	0.92	94	0
Fat (g/d)	53.8 (26.8, 97.4)	52.1 (28.6, 102.7)	0.30	5	0.89	94	1
Sodium (mg/d)	2771.7 (1076.3, 6458.0)	2560.6 (1375.3, 5895.6)	0.93	6	0.86	92	2
Potassium (mg/d)	1671.5 (861.0, 3035.9)	1788.8 (828.6, 2997.1)	0.60	3	0.85	87	2
Calcium (mg/d)	890.6 (312.2, 1838.3)	891.3 (319.3, 1932.0)	0.98	6	0.92	97	0
Iron (mg/d)	20.1 (10.3, 42.4)	21.0 (10.3, 45.1)	0.31	15	0.84	95	0
Phosphorus (mg/d)	1018.4 (504.7, 2159.7)	1002.8 (521.9, 1869.2)	0.56	3	0.94	94	1
Vitamin A, RE (µg/d)	986.4 (424.2, 5286.4)	959.0 (432.8, 3666.1)	0.24	3	0.91	98	0
Vitamin C (mg/d)	132.6 (46.8, 335.6)	132.6 (46.8, 305.5)	0.32	6	0.75	92	4
Thiamin (mg/d)	1.7 (0.7, 3.0)	1.6 (0.7, 2.8)	0.83	3	0.93	99	0
Riboflavin (mg/d)	2.2 (0.9, 4.0)	2.2 (1.0, 4.2)	0.72	12	0.87	98	1
Niacin (mg/d)	16.4 (7.1, 35.2)	17.1 (8.0, 31.0)	0.03	11	0.91	94	0
Dietary fiber (g/d)	7.4 (2.4, 16.3)	7.6 (2.0, 13.8)	0.61	3	0.87	94	1

FFQ1, food frequency questionnaire during 1st administration; FFQ2, food frequency questionnaire during 2nd administration; P_5, 5th percentile; P_{95}, 95th percentile; RE, retinol equivalents

* Wilcoxon signed rank test to test for differences between FFQ1 and FFQ2

† Relative difference = [(FFQ2 – FFQ1)/ FFQ1]*100

‡ Based on log-transformed values

§ Correctly classified if classified into same or adjacent (+/ - 1) quartiles

¶ Grossly misclassified if classified into opposing quartiles

Table 5. Reproducibility study: median daily intake of food, relative difference, Spearman correlation coefficient and cross classification for the comparison between FFQ1 and FFQ2 in pregnant women (n=85).

Food group	*FFQ1*	*FFQ2*	*P value**	*Relative difference[†] (%)*	*Intraclass correlation coefficient[‡]*	*Cross-classification into quartiles (%)*	
	Median (P_5, P_{95})	*Median (P_5, P_{95})*				*Correctly classified[§]*	*Grossly misclassified[¶]*
Cereal products	472.7 (244.5, 854.4)	437.5 (226.1, 776.0)	0.24	1	0.80	89	4
Poultry	17.1 (0.3, 111.0)	17.1 (2.0, 111.0)	0.43	57	0.90	92	0
Red meats	4.0 (0, 50.1)	4.0 (0, 37.7)	0.68	62	0.84	91	1
Organ meats	1.7 (0, 62.9)	0 (0, 42.9)	0.33	21	0.93	95	0
Fish and other seafood	94.4 (19.4, 332.4)	94.4 (21.6, 273.2)	0.44	17	0.92	92	2
Eggs	14.7 (1.8, 51.7)	14.7 (0.5, 53.7)	0.81	42	0.79	85	6
Dairy products	51.0 (0, 368.0)	49.2 (0, 345.9)	0.88	25	0.91	97	0
Nuts and legume	1.5 (0, 32.6)	2.0 (0, 41.2)	0.57	68	0.86	88	5
Leafy vegetables	18.4 (2.4, 109.6)	18.4 (1.6, 114.6)	0.31	7	0.90	95	0
Cruciferous vegetables	7.1 (0, 60.0)	7.1 (0, 67.5)	0.97	31	0.90	95	0
Legume vegetables	3.3 (0, 19.7)	3.3 (0, 14.8)	0.13	106	0.82	93	4
Tuber vegetables	3.1 (0, 45.5)	2.9 (0, 25.8)	0.46	51	0.87	91	1
Raw vegetables 'Ulam'	1.3 (0, 21.3)	1.3 (0, 18.0)	0.11	-3	0.90	94	1
Other vegetables	14.7 (0, 103.0)	14.7 (0, 94.7)	0.49	11	0.88	93	1
Onion and garlic	10.8 (0, 41.0)	9.0 (0, 41.0)	0.04	-8	0.89	95	4
Fresh fruits	161.2 (32.4, 471.8)	166.5 (34.2, 487.0)	0.74	24	0.88	92	2
Dried fruits	2.0 (0, 38.8)	2.1 (0, 36.4)	0.41	62	0.83	86	4
Pickled fruits	0 (0, 12.0)	0 (0, 17.6)	0.89	41	0.75	87	0
Juices	57.1 (0, 324.3)	57.1 (0, 400.0)	0.89	50	0.83	91	4

Table 5: Continued

Food group	FFQ1	FFQ2	P value*	Relative difference† (%)	Intraclass correlation coefficient‡	Cross-classification into quartiles (%)	
	Median (P_5, P_{95})	Median (P_5, P_{95})				Correctly classified§	Grossly misclassified¶
Tea	51.4 (0, 540.0)	51.4 (0, 498.0)	0.50	20	0.86	91	2
Coffee	0 (0, 156.9)	0 (0, 180.0)	0.35	-13	0.96	95	0
Malt drinks	5.1 (0, 36.0)	5.1 (0, 27.2)	0.46	97	0.85	91	1
Cordials	0 (0, 134.3)	0.3 (0, 114.3)	0.45	83	0.90	95	0
Soybean milk	35.7 (0, 243.6)	28.6 (0, 250.0)	0.002	-12	0.94	97	0
Fats and oils	24.8 (5.6, 53.5)	22.6 (6.6, 64.1)	0.50	17	0.90	92	1
Confectionary	52.6 (5.8, 168.9)	54.6 (0.1, 186.5)	0.29	29	0.73	90	2
Biscuits	10.3 (0, 72.0)	10.3 (0, 88.8)	0.88	23	0.86	92	2
Sugar	15.0 (0.6, 81.0)	15.0 (0, 70.5)	0.85	28	0.87	92	2
Salt	2.0 (0.8, 7.0)	2.1 (0.9, 8.0)	0.67	11	0.75	94	1
Condiments	16.0 (1.2, 117.4)	14.7 (1.1, 103.4)	0.72	125	0.86	88	5

FFQ1, food frequency questionnaire during 1st administration; FFQ2, food frequency questionnaire during 2nd administration; P_5, 5th percentile; P_{95}, 95th percentile; RE, retinol equivalents

* Wilcoxon signed rank test to test for differences between FFQ1 and FFQ2

† Relative difference = [(FFQ2 – FFQ1)/ FFQ1]*100

‡ Based on log-transformed values

§ Correctly classified if classified into same or adjacent (+/ - 1) quartiles

¶ Grossly misclassified if classified into opposing quartiles not report significant differences between the two series of assessment, with the exception of niacin for nutrient and soybean milk, onion and garlic for food. The ICC between nutrients measured approximately one month apart by the FFQ ranged from 0.75 (vitamin C) to 0.94 (phosphorus) (Table 4). The ICC for foods were generally similar to nutrients and ranged from 0.73 (confectionary) to 0.96 (coffee) (Table 5). Overall, the mean ICC for both nutrients and foods was 0.87. On average, classifications into quartiles for nutrients and foods showed that at least 90% of the subjects were correctly classified, while less than 2% of them were grossly misclassified.

DISCUSSION

This is perhaps the first validation study of FFQ conducted among Malaysian pregnant women. Food list in the FFQ was developed using data-based approach which better reflects representative dietary data among Kelantan pregnant women instead of using an adapted questionnaire. Pregnant women in early and late trimesters were included in the FFQ development process in order to compile a comprehensive food list which covered habitual intake throughout the pregnancy. It is complicated to assess dietary intake in pregnancy due to sequential physiological change throughout the gestational period (Coad & Dunstall, 2005). The present study showed relatively low to moderate agreement between dietary data derived from the FFQ and DR. The sample size of 177 subjects was sufficient to assess validity of a questionnaire (Willett, 1998). No respondent was excluded from the study on the basis of misreporting as indicated in other studies (Wei *et al.,* 1999; Baer *et al.,* 2005). Therefore, any bias among the included subjects which cannot be corrected by subject exclusion with improbable energy intake was avoided (Brantsaeter, 2007). Moreover, the high or low levels of energy intake reported by some of the women should not be viewed solely as misreporting but may due to appetite and meal pattern changes in pregnancy. Poslusna *et al.* suggested that energy adjustment approach using a residual model was more appropriate to handle misreporting and to examine a diet-disease relationship instead of subject exclusion (Poslusna *et al.,* 2009).

The FFQ often over estimated nutrient and food intakes as shown in our study and by others (Erkkola *et al.,* 2001). The highest over-estimation which was seen in vitamin C and fruit intake indicates that cross-check questions may not contribute much to better estimation of the consumption compared to individual intake. It reveals difficulty in translating the reported dietary values into the actual dietary intakes. However, over-estimation may not be problematic in epidemiologic studies aimed at investigating diet-disease associations provided individuals' ranking according to dietary intake is valid (Beaton, 1991). Thus, precise numerical estimation of dietary intake has not been the focus. The over-estimation of intake can be explained by the long list of food items and predefined portion sizes in the FFQ compared to the actual consumption in DR.

Spearman coefficients were used to measure the strength of relation between the FFQ and the DR since it may be more reliable as it uses rank order and is not as sensitive to extreme values as the Pearson coefficients (Masson *et al.,* 2003). The poor correlations for fat in nutrient, onion and garlic, fats and oils and salt in food can be explained by the difficulty in estimating the amount consumed since they are more likely to act as seasonings in cooking and the measurements are not straightforward. Irregular intake of organ meats may raise the same situation. In order to improve the validity in future studies, total serving size of those food items used in each cooking will be specified in the FFQ for better estimation instead of individual amounts. The average correlation coefficients between the FFQ and the DR for the absolute nutrient intakes were stronger than in the food intakes in our study

(0.46 vs. 0.37). The weaker correlations in food may reflect higher variations in food intake than in the nutrient intake. However, contradictory findings were reported in validation studies for pregnant Finnish women (0.37 vs. 0.47) (Erkkola *et al.,* 2001) and Norwegian women (0.36 vs. 0.48) (Brantsaeter *et al.,* 2008). A similar mean of correlations for nutrients were found in previous validation studies among pregnant Portuguese women (0.41) (Pinto *et al.,* 2010) and Massachusetts women (0.47) (Wei *et al.,* 1999), while a lower mean correlation for nutrients was found in a validation study among Sheffield pregnant women (0.20) (Mouratidou *et al.,* 2005). In general, it was difficult to compare the studies since limited validation studies have been conducted among pregnant women especially among the Asian population. Differences in FFQ designations, reference methods, periods of administration and target populations may lead to discrepancies in study findings.

After energy adjustment, the correlations between the two methods decreased compared with the unadjusted correlations. Consistent findings were reported by other studies in China (Cheng *et al.,* 2008). Cheng *et al.* suggested that energy adjustment might over-adjust the nutrient intake for a population with staple food as the main contributor to the total energy intake and this is reflected in fluctuations in measurement errors by the methods (Cheng *et al.,* 2008). The same situation occurred in our study whereby the main contribution to total energy was shown by the staple food, cereals (data not shown). This explains why energy-adjusted values were not used for further agreement analysis due to inflated correlation coefficients.

We found similar correlations between the two methods for absolute nutrient and food intakes when pregnant women who reported hyperemesis gravidarum were excluded (data not shown). The present findings were different from other validation studies whereby the correlations became stronger after excluding those pregnant women with nausea or vomiting (Brantsaeter *et al.,* 2008). This might be due to the small proportion of women who were excluded and which did not result in significant changes in correlations. We only considered women who experienced severe vomiting until ward admission instead of all incidences of pregnancy-related nausea or vomiting.

Hankin *et al.* suggested that correlation coefficients between the dietary methods were considered poor if <0.30, fair if 0.30 – 0.49 and good if >0.50 in most dietary validations (Hankin *et al.,* 1991). In our study, the overall correlations were considered fair for both nutrients and foods. However, correlations alone are not sufficient to represent the performance of the FFQ due to high day-to-day variability caused by appetite changes in pregnancy. Moreover, it has been shown to be flawed as only the degree of correlations between the two methods are measured but not the agreements (Cade *et al.,* 2002). Cross classification and Bland-Altman plot are the more appropriate methods to assess the relative validity of FFQ with the DR (Pinto *et al.,* 2010). Both methods showed a satisfactory level of agreement between the dietary assessments.

Cross-classification gives a much clearer and undistorted picture of how well the FFQ performs than the correlation coefficients (Cade *et al.,* 2002). The degrees of misclassification were small for both nutrient and food intakes. More than two-thirds of the pregnant women were correctly classified into the quartiles by the two methods. Similar findings were shown by other validation studies among pregnant women but these were based on cross-classification over quintiles for nutrient intake (Brantsaeter *et al.,* 2008; Pinto *et al.,* 2010). It was only a study showing cross-classification into quintiles for food intake in pregnancy, with similar findings (Erkkola *et al.,* 2001). These results indicate that the FFQ is capable of ranking pregnant women in relation to their intake and this is important for investigating diet-disease relationship in future maternal cohort studies.

The Bland-Altman plot was used to obtain further information regarding the agreement between the FFQ and the DR based on graphical interpretation. A systematic increase or decrease in the differences between the methods with increasing intake revealed that individuals who consumed higher amounts of these nutrients or foods reported more errors during diet assessments. Most micronutrients and micronutrient-rich foods do not demonstrate a similar trend with increasing intake compared to macronutrients and energy-yielding foods which show more precise dietary assessments. These findings are consistent with the study conducted among pregnant women by Brantsaeter *et al.*(2008).

The validation study was designed by administrating the FFQ1 prior to the DR1 and DR2, followed by the FFQ2 at the end. The use of FFQ before and after the DR may provide minimal and maximal estimates of true validity (Erkkola *et al.*, 2001). The DR administration in between two sets of FFQ in this study helped to improve the accuracy of reporting in the latter FFQ2. However, the high degree of correlations and agreements as well as low degree of within person variations between the FFQ1 and FFQ2 for nutrients and foods, show that the true validity of FFQ1 is not much different from that of the FFQ2. In other words, both FFQ1 and FFQ2 showed similar validity against the DR. The considerablly high correlations and percentages in correct quartiles compared to other studies (Erkkola *et al.*, 2001) may be partly due to the short duration of administration between FFQ1 and FFQ2. As reviewed by Cade, correlations were higher for repeat administration 1 month or less apart compared to 6 months to 1 year apart (Cade *et al.*, 2002). In general, the FFQ is found to be reproducible.

Limitations in the reference measures and time frame of assessments may attenuate our results. No objective measurement was used to validate the FFQ as indicated in another study (Brantsaeter *et al.*, 2008). The 24-hour DR is based on subjective measurement and not considered as a gold standard in dietary assessment as weighed records. However, it was chosen due to its high response rate. Its less demanding effect may reduce the possibility of under-reporting or usual meal changes and give higher face validity. The use of two 24-hour DR during the third trimester is inadequate to capture the complete picture of usual dietary intake throughout pregnancy due to high day-to-day variability, especially when dealing with pregnancy symptom variations. Multiple recalls which cover longer periods of gestation should have been conducted to give a better daily estimation and represent a more relevant time span between the two methods. However, these are beyond our capability as the time factor was the main restriction in this project. In general, this questionnaire provides new information on the validity of nutrient and food intakes among Malay pregnant women in Kelantan state. However, we believe that this FFQ can serve as the basic tool for further dietary validation in a multi-ethnic pregnant population in Malaysia.

In conclusion, a wide range of nutrient and food intakes showed relative good validity and reproducibility in the study. Thus, it suggested that this FFQ is a valid tool to collect dietary data and rank individuals by relative level of intake for Malay pregnant women in a prospective study.

ACKNOWLEDGEMENTS

We are grateful to the research team, hospital staff and pregnant women for their cooperation and support. We would like to acknowledge Universiti Sains Malaysia (USM) for funding the study via the USM Research University Grant (1001/PPSK/

811087). We would also like to thank USM for fellowship funding provided to Loy SL and Universiti Sultan Zainal Abidin for funding Marhazlina M. A special appreciation goes to Juliana S, Mohd Hairil Nizam A H, Monaliza A and Nor Elyana A from USM Dietetics Programme for their review in FFQ development.

REFERENCES

Beaton GH (1991). Interpretation of Results from Diet History Studies. In: T*he Dietary History Method. Proceedings of the 2nd Berlin Meeting on Nutritional Epidemiology*, October 1989. Kohlmeier L (ed). Smith-Gordon, London.

Baer HJ, Blum RE, Rockett HR, Leppert J, Gardner JD, Suitor CW & Colditz GA (2005). Use of a food frequency questionnaire in American Indian and Caucasian pregnant women: a validation study. *BMC Public Health* 5: 135.

Barker DJP (1992). Fetal and Infant Origins of Adult Disease. British Medical Journal Books, London.

Block G, Dresser CM, Hartman AM & Carroll MD (1985). Nutrient sources in the American diet: quantitative data from the NHANES II survey. I. Vitamins and minerals. *Am J Epidemiol* 122: 13-26.

Brantsaeter AL (2007). Validation of dietary data in pregnancy: validation of the food frequency questionnaire developed for the Norwegian mother and child cohort study (MoBa). Norwegian Institute of Public Health. *http://www.duo.uio.no/publ/allmen/2007/62649/DUO_Brantsaeter_17x24.pdf* [Retrieved June 2010].

Brantsaeter AL, Haugen M, Alexander J & Meltzer HM (2008). Validity of a new food frequency questionnaire for pregnant women in the Norwegian Mother and Child Cohort Study (MoBa). *Matern Child Nutr* 4: 28-43.

Cade J, Thompson R, Burley V & Warm D (2002). Development, validation and utilisation of food-frequency questionnaires - a review. *Public Health Nutr* 5: 567-587.

Cade JE, Burley VJ, Warm DL, Thompson RL & Margetts BM (2004). Food-frequency questionnaires: a review of their design, validation and utilisation. *Nutr Res Rev* 17: 5-22.

Cheng Y, Yan H, Dibley MJ, Shen Y, Li Q & Zeng L (2008). Validity and reproducibility of a semi-quantitative food frequency questionnaire for use among pregnant women in rural China. *Asia Pac J Clin Nutr* 17: 166-177.

Coad J & Dunstall M (2005). Anatomy and Physiology for Midwives (2nd ed.). Elsevier Health Sciences, London.

Emmett P (2009). Assessing diet in longitudinal birth cohort studies. *Paediatr Perinat Epidemiol* 23 Suppl 1: 154-173.

Erkkola M, Karppinen M, Javanainen J, Rasanen L, Knip M & Virtanen SM (2001). Validity and reproducibility of a food frequency questionnaire for pregnant Finnish women. *Am J Epidemiol* 154: 466-476.

Hankin JH, Wilkens LR, Kolonel LN & Yoshizawa CN (1991). Validation of a quantitative diet history method in Hawaii. *Am J Epidemiol* 133: 616-628.

Knudsen VK, Orozova-Bekkevold IM, Mikkelsen TB, Wolff S & Olsen SF (2008). Major dietary patterns in pregnancy and fetal growth. *Eur J Clin Nutr* 62: 463-470.

Masson L, McNeill G, Tomany J, Simpson J, Peace H, Wei L, Grubb DA & Bolton-Smith C (2003). Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic *Public Health Nutr* 6: 313-321.

Mouratidou T, Ford F & Fraser RB (2006). Validation of a food-frequency questionnaire for use in pregnancy. *Public Health Nutr* 9: 515-522.

Pinto E, Severo M, Correia S, dos Santos Silva I, Lopes C & Barros H (2010). Validity and reproducibility of a semi-quantitative food frequency questionnaire for use among Portuguese pregnant women. *Matern Child Nutr* 6: 105-119.

Poslusna K, Ruprich J, de Vries JH, Jakubikova M & van't Veer P (2009). Misreporting of energy and micro-nutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice. *Br J Nutr* 101 Suppl 2: S73-85.

Sayer AA & Cooper C (2005). Fetal programming of body composition and musculo-skeletal development. *Early Hum Dev* 81: 735-744.

Suzana S, Noor Aini MY, Nik Shanita S, Rafidah G & Roslina A (2009). Atlas of Food Exchanges & Portion Sizes (2nd ed.). MDC Publishers Sdn Bhd (91168-A), Kuala Lumpur.

Tee ES, Ismail MN, Mohd Nasir A & Khatijah I (1997). Nutrient Composition of Malaysian Foods (4th ed.). Malaysian Food Composition Database Programme, Institute for Medical Research Kuala Lumpur.

U. S. Department of Agricultue (USDA) (2009). SR 22, Release 22, U. S. Department of Agriculture Nutrient Database for Standard Release. Washington DC.

Wei EK, Gardner J, Field AE, Rosner BA, Colditz GA & Suitor CW (1999). Validity of a food frequency questionnaire in assessing nutrient intakes of low-income pregnant women. *Matern Child Health J* 3: 241-246.

Wessek Institute of Public Health, University of Southampton, 1995.

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

Willett W (1998). Nutritional Epidemiology (2nd ed.). Oxford University Press, New York.