Validation of dietary data in pregnancy

Validation of the food frequency questionnaire developed for the Norwegian Mother and Child Cohort Study (MoBa)

Anne Lise Brantsæter

Institute of General Practice and Community Medicine,
University of Oslo

Division of Environmental Medicine, Department of Food Safety and Nutrition, Norwegian Institute of Public Health

2007

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 524

ISBN 978-82-8072-437-3

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS, Oslo, 2007.

Produced in co-operation with Unipub AS.
The thesis is produced by Unipub AS merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Unipub AS is owned by
The University Foundation for Student Life (SiO)

Contents

Abbreviations

MoBa:	The Norwegian Mother and Child Cohort Study.
FFQ:	Food Frequency Questionnaire. The Food Frequency Questionnaire is a retrospective method asking respondents to report their usual frequency of consumption of each food from a list of foods for a specific period. Food lists vary by the purpose of the study and study population. Frequency of consumption categories also vary by questionnaire but usually include per day, week or month.
FD:	Food Diary or Food Records (FR) are used to record food intake at the time of consumption over a number of days that are not necessarily sequential. Most studies ask respondents to enter descriptive and quantitative information in a hard copy format.
24-hour recall:	The 24-hour recall is a retrospective assessment method in which an interviewer prompts a respondent to recall and describe all foods and beverages consumed in the preceding 24 hours or the preceding day. Portion size estimating aides assist the respondent to recall amounts consumed.
Biomarker	As pertaining to diet: a biological parameter measurable in body tissue that relates to/reflects substances in the diet. Biomarkers may be categorised as concentrations or recovery biomarkers.
Validity:	The ability of an instrument to measure what it is intended to measure.
Test method:	Dietary assessment method being validated; in this study the MoBa FFQ.
Reference method:	Method against which the test method is being compared and validated. In this study several reference measures were used: a weighed food diary, motion sensor registration and several biomarkers in urine and blood.
EI	Energy Intake
TEE	Total Energy Expenditure
REE	Resting Energy Expenditure
PAL and PAL_{EI}	Physical Activity Level = TEE/REE and PAL_{EI} = EI/REE assuming energy balance.

Acknowledgments

The work presented in this thesis was carried out at the Department of Food Safety and Nutrition, Division of Environmental Medicine at the Norwegian Institute of Public Health during the years 2002 – 2007. My work was supported by a grant from the Research Council of Norway from 2002 – 2006 and the Norwegian Institute of Public Health in 2007. I want to thank these institutions for giving me the opportunity to carry out this work.

The Norwegian Mother and Child Cohort Study (MoBa) is the largest and most costly cohort study ever conducted in Norway. I was one of the first research fellows involved with MoBa data, and it has been stimulating to be part of a project that will be of interest for researchers in nutrition epidemiology for decades to come.

It is my pleasure to thank those who contributed to making this work possible. First of all, I want to thank my supervisors Helle Margrete Meltzer and Margaretha Haugen at the Norwegian Institute of Public Health. These two have been my mentors and have guided me every step of the way. They have given me their time, advice, knowledge, enthusiasm, insight and support, and shared my tears and joy. Both have extensive experience and knowledge of nutritional science, and they make a great team. Helle is a visionary and is able to handle hundreds of tasks simultaneously and still remain focused, sensitive and attentive to anyone in her proximity. Margaretha is exceptional when it comes to electronic handling of food and nutrient calculations and the scientific writing. I also want to thank Jan Alexander for his involvement, feedback and encouragement, and for providing excellent working facilities. I want to thank Tone Rasmussen, Hege Hjertholm and Ewa Andruchow for practical assistance and goodwill. I have made so many good friends among my colleagues at the Department of Food Safety and Nutrition, and I thank all of you for scientific and personal conversations along the way. Special thanks to Ingrid Fange, Helen Engelstad Kvalem, Elisabeth Elind, Hanne Torjusen, Hege Ølstørn and last, but not least, Anna-Pia Häggkvist.

I want to thank Jorid Eide, Piroozeh Nazem Barandeh and Lene Joa, midwifes at Bærum hospital, for helping to recruit participants to the validation study, and Edel Lilleås for drawing all the blood specimens. I particularly want to thank the 119 women who participated in the validation study who voluntarily invested their time, blood, urine and effort. Furthermore, I thank my collaborators and co-authors: Salka E Rasmussen at

the Danish Institute for Food and Veterinary Research, Sven Ove Samuelsen at the University of Oslo, Kåre Julshamn at the National Institute of Nutrition and Seafood Research, Tor-Arne Hagve at the Rikshospitalet University Hospital, Lage Aksnes at the University of Bergen, and Wenche Nystad, Hein Stigum and Per Magnus at the Norwegian Institute of Public Health. A special thanks to Trond Arild Ydersbond for his involvement and valuable contribution.

Many friends and relatives have encouraged me during the years I have been working with this thesis, and I would like to thank all of you. Special thanks to my mother, who has always believed in me and been there for me. My father taught me the value of systematic work, but sadly did not live to see the completion of this thesis. Finally, I want to thank my dear husband Arne and our three children Margrethe, Henrik and Thomas for their enduring love and support, and for contributing to everyday adventures and magic moments beyond validation studies.

Oslo, March 2007

Anne Lise Brantsæter

Summary

Due to recent insight into the relation between fetal nutrition and health of the child both early and later in life, assessment of maternal diet has become an integral part of pregnancy and birth cohort studies. A new food frequency questionnaire (FFQ) was developed for monitoring maternal diet in the Norwegian Mother and Child Cohort Study (MoBa). The basis for this thesis in nutritional epidemiology was to assess the relative validity of the new FFQ. A validation study was carried out with 119 women in MoBa. Reference measures were a 4-day weighed food diary (FD), a motion sensor for measuring total energy expenditure, and diet-related biological markers in urine and blood.

The dissertation contains five scientific papers. One presents the main methodological challenges when developing a new instrument for monitoring dietary intake in pregnant women. The other papers report on the validity of different foods and nutrients measured by the new FFQ relative to the reference measures.

The average correlation coefficient between the FFQ and FD was 0.48 for foods and 0.36 for nutrients which is accepted as an overall good agreement. Urinary recovery biomarkers and serum/plasma concentration biomarkers confirmed that the FFQ was able to distinguish between high and low intake of nutrients and foods. Furthermore, the biomarkers examined in the validation study confirmed differences in self-reported micronutrient intake between supplement and non-supplement users for vitamin D, beta-carotene, folate, n-3 fatty acids, flavonoids and iodine, showing that self-reporting is valid for classification of pregnant women according to dietary supplement.

Results from the validation study confirmed the applicability of biomarkers in the validation of dietary data in pregnant women. In conclusion, the results demonstrate that the MoBa FFQ provides valid intake estimates and is able to rank pregnant women according to dietary intake of foods and nutrients.

List of papers

Paper 1. Meltzer HM, Brantsæter AL, Alexander J, Ydersbond TA, Haugen M and the MoBa Dietary Support Group. *Methodological challenges when monitoring the diet of pregnant women in a large cohort study; experiences from the Norwegian Mother and Child Cohort Study.* Submitted.

Paper 2. Brantsæter AL, Haugen M, Alexander J, Meltzer HM. *Validity of a new Food Frequency Questionnaire for pregnant women in the Norwegian Mother and Child Cohort Study (MoBa).* Submitted.

Paper 3. Brantsæter AL, Haugen M, Rasmussen SE, Alexander J, Samuelsen SO and Meltzer HM. *Urine flavonoids and plasma carotenoids in the validation of fruit, vegetable and tea intake during pregnancy in the Norwegian Mother and Child Cohort Study (MoBa).* Public Health Nutrition, 2007 (in press).

Paper 4. Brantsæter AL, Haugen M, Julshamn K, Alexander J and Meltzer HM. *Evaluation of urinary iodine excretion as a biomarker for intake of milk and dairy products in pregnant women in the Norwegian Mother and Child Cohort Study (MoBa).* Submitted.

Paper 5. Brantsæter AL, Haugen M, Hagve TA, Aksnes L, Rasmussen SE, Julshamn K, Alexander J and Meltzer HM. *Self-reported dietary supplement use is confirmed by biological markers the Norwegian Mother and Child Cohort Study (MoBa).* Annals of Nutrition and Metabolism, 2007; 51 (in press).

1 Introduction

"The complexity of human diet represents a daunting challenge to anyone contemplating a study of its relation to disease. The foods we consume each day contain literally thousands of specific chemicals, some known and well quantified, some characterized only poorly, and others undescribed and presently unmeasurable" Walter C. Willett in Nutritional Epidemiology (1).

1.1 General background

The importance of nutrition during pregnancy has long been recognised, and yet our understanding is still limited. Maternal diet was regarded as an important environmental exposure in the Norwegian Mother and Child Cohort Study (MoBa), and a new Food Frequency Questionnaire (FFQ) was developed. The Norwegian Mother and Child Cohort Study, initiated by the Norwegian Institute of Public Health, is an ongoing, long-term prospective cohort study that aims to include 100 000 pregnancies by the end of 2007 (2). The objective of MoBa is to test specific aetiological hypotheses by estimating the association between exposures and diseases, aiming at prevention. Self-reported questionnaires were chosen as the main method to measure environmental exposures, including diet. The MoBa FFQ (Appendix 1) is a semi-quantitative questionnaire, designed to capture habitual diet and dietary supplement use during the first four months of pregnancy, when the fetus is most vulnerable (3;4). The development and rationale of the MoBa FFQ is explained in Paper 1 in this thesis.

Every new FFQ has to be validated to get an expression of the degree to which it is an accurate measure in the target population. Validity refers to the degree to which a questionnaire actually measures the aspect of diet that it was designed to measure. This implies that a comparison is made with a superior, although always imperfect standard. Validation is required before dietary intake can be used as exposure variables in epidemiological studies (1;5;6). A validation study of the MoBa FFQ was carried out and the results constitute the main basis of the present thesis. Papers 2 – 5 present the relative validity of different aspects of the FFQ.

1.2 Nutrition and pregnancy

1.2.1 Maternal diet and neonatal health

Nutrition during pregnancy plays an important role in the well-being of the mother and fetus, and may further influence the health of the children later in life (4). Some of the first scientific evidence of the influence of maternal nutrition on reproductive outcome came from studies of Dutch women who experienced food-restrictions during the Second World War (7). Subsequent evidence came from cross-sectional observational studies and controlled randomised trials of nutrient supplementation in pregnancy. The majority of studies used birth weight as a marker of the relative success of pregnancy. The interest in the diet of pregnant women in the developed world lessened when observational studies failed to identify any marked effects of diet on size at birth. Absence of an effect, together with the fact that the vast majority of babies born to women in the developed world fell within the "normal range" of birth weight, led to the belief that the fetus was a "perfect parasite" that was protected from wide variations in nutrient intake.

Recent epidemiological studies have, however, shown an inverse association between weight at birth and adult risk of development of diseases, and these associations are seen within the range of birth weights which is considered normal (8). Both maternal under- and over-nutrition reduce placental-fetal blood flow and stunt fetal growth (9). The fetal origins hypothesis, first advocated by Anders Forsdahl and later David Barker, states that impaired intrauterine growth and development may increase the risk of adult cardiovascular disease, type 2 diabetes, obesity and cancer through fetal programming at a critical time point (4;10). Recent reviews from the Dutch famine birth cohort have confirmed the theory of critical periods in development where timing of the nutrition insult determines which organ or metabolic system is affected (11;12). An alternative explanation to the fetal origins hypothesis is that there is a common underlying genetic basis to both reduced fetal growth and the risk of adult diseases (13). Genome-nutrient interactions and epigenetic mechanisms are still poorly understood, and is a complex and growing field of interest (9).

Weight at birth is a useful marker of conditions in the womb and of underlying events that result in both birthweight and in programming of postnatal physiology. However, maternal nutrition may affect the fetus even if birthweight is not affected. The importance of sufficient maternal folate in protection against neural tube defects in the developing fetus is an example that is well documented (14;15).

1.2.2 Recommended dietary intake in pregnancy

The percentage increase in estimated energy requirements during pregnancy is small relative to the estimated increased need for most other nutrients (Table 1). For women in well nourished populations, the increased energy requirement is often counterbalanced by decreased physical activity (16). Consequently, the nutrient density of the maternal diet becomes decisive.

Table 1 Energy and nutrient requirements during pregnancy (17)

Per day	Non-pregnant women	Pregnant women	Extra requirement for pregnancy
Energy (kJ)	9.2[a]	10.7	1.5 MJ and 2 MJ increase in 2. and 3. trimester
Protein (percent of energy)	10-20	10-20	1.1 g/kg body weight increase
Fat (percent of energy)	30	30	0.5 increase in n-3 FA
Added suger (percent of energy)	less than 10	less than 10	
Vitamin A (µg)	700	800	+ 100 µg
Vitamin D (µg)	7.5	10	+ 2.5 µg
Vitamin E (α-TE)	8	10	+ 2 mg
Thiamine (mg)	1.1	1.5	+ 0.4 mg
Riboflavin (mg)	1.3	1.6	+ 0.3 mg
Niacin (mg)	15	17	+ 2 mg
Vitamin B_6 (mg)	1.3	1.5	+ 0.2 mg
Folate (µg)	400[b]	500[b]	+ 100 µg,
Vitamin C (mg)	75	85	+ 15 mg
Calcium (mg)	800	900	+ 100 mg
Phosphorous (mg)	600	700	+ 100 mg
Potassium (g)	3.1	3.1	
Magnesium (mg)	280	280	
Iron (mg)	15	15[c]	+[c]
Zink (mg)	7	9	+ 2mg
Copper (mg)	0.9	1.0	+ 0.1 mg
Iodine (µg)	150	175	+ 25 µg
Selenium (µg)	40	55	+ 15 µg

[a]Body weight 63 kg, sedentary lifestyle

[b]Women are advised to use a dietary supplement to ensure sufficient intake during first weeks of pregnancy

[c]Supplementary iron needed for women with low iron stores

The main dietary recommendations for pregnant women are to consume a healthy, well-balanced diet with plenty of iron- and folate-rich foods, avoid alcohol and reduce the intake of sweet drinks and caffeine. Most women do not get the recommended amount of folate through food. With the convincing documentation of folic acid supplementation for prevention of neural tube defects, health authorities in many countries, therefore, recommend use of folic acid (synthetic folate) to all fertile women planning to or likely to become pregnant (18). Likewise, the physiological need for iron in some women can not be satisfied with food only in the last two thirds of pregnancy and supplemental iron is needed. The amount of iron recommended via supplements depends on an individual assessment of iron status (19;20).

Vitamin D is available only in a limited amount of foods, and for many pregnant women vitamin D status is maintained more by exposure to sun than through diet. Low vitamin D status is detrimental to both the mother and the fetus (21;22). Groups of women most vulnerable to vitamin D deficiency, for example those with scarce sunlight exposure, are recommended to take a vitamin D containing supplement. The use of cod liver oil, traditionally taken in Norway, provides vitamin D, vitamin A, vitamin E and n-3 fatty acids.

Suboptimal iodine intake in pregnant women has been described in many European countries (23;24). In many countries iodine fortification of salt is implemented to increase iodine intake, and pregnant women and women planning a pregnancy are encouraged to use an iodine-containing supplement (approximately 150 μg/day) (23;25). There are only limited data from well-controlled intervention studies with dietary supplements in well nourished populations, and with the exception of iron and folate, the evidence that nutrient supplements confer measurable benefit is not strong. Antioxidant and other nutrient supplementation have been investigated to study the impact of supplements on hypertensive disorders and other unfavourable pregnancy conditions, but the effects are equivocal (26;27). The MoBa FFQ includes detailed questions regarding frequency of use, dose and product names of the dietary supplements. For calculation of micronutrients such as vitamins, minerals, fatty acids and other bioactive substances supplied by dietary supplements in MoBa, a database containing details of the declared content of more than one thousand supplements was created. Thus, including dietary supplement use in the investigation of the relative validity of the MoBa FFQ is relevant (Paper 5).

1.2.3 Maternal weight gain and energy expenditure

A sufficient and balanced supply of energy and protein is the major determinant of weight gain during pregnancy (28). The maternal diet must provide sufficient energy and nutrients to meet the mother's usual requirements and the needs of the growing fetus. It must also enable the mother to lay down stores of nutrients required for fetal development and for lactation (19). An average weight gain of 12 kg (range 12 – 14 kg) for women with a healthy pre-pregnant weight has been shown to be associated with the lowest risk of complications during pregnancy and the lowest risk of a low birth-weight infant (29;30).

Excessive maternal weight gain and maternal obesity have serious adverse effects on the fetus (13;31). Staying physically active during the pregnancy may prevent excess weight gain. A number of trials have concluded that regular, moderate intensity exercise has no adverse effects on the health of the mother or the infant (32). On the contrary, it seems that regular physical activity in the time prior to and during pregnancy is associated with reduced risk of gestational diabetes, preeclampsia, hyper-lipidaemia and excessive fetal growth (33-37). Physical activity is the main determinant of total energy expenditure.

1.3 Dietary assessment and validation

1.3.1 Challenges related to dietary assessment in pregnancy

The complex relationship between maternal diet and birth outcomes emphasises the need for a consistent and thorough assessment of diet in pregnancy. The mixture of foods called "the diet" is a series of interrelated factors with large within- and between person variation making dietary assessment particularly challenging (38). Pregnancy makes dietary assessment even more complicated as the metabolic and physiological changes affect energy and nutrient needs, appetite and meal patterns (19). Pregnant women may develop food preferences and aversions due to changes in the sense of taste and smell. Nausea, which is estimated to occur in 60-80 percent of pregnancies, may begin as early as 4-6 weeks after conception, peaking around 8-12 weeks, and then declining (39). Heartburn and constipation are other common ailments that may trigger changes in usual food habits. Methodological challenges related to the assessment of diet in pregnancy are outlined in Paper 1.

1.3.2 Dietary assessment methods

Food frequency questionnaires are regarded as the primary method for dietary assessment in epidemiological studies and for surveillance of the intake of foods and nutrients in different populations and age groups (40). FFQs have been shown to be an appropriate method for assessing dietary information in a wide variety of epidemiological settings, including studies among pregnant women (41-44). In comparison with short-term records, the FFQ provides a better approximation of the habitual diet over a longer period (5). The usual frequency of consumption of different foods is reported. In addition questions on quantity may be added, as well as further aspects on diet composition. The advantages of FFQs are that they are easy to use and administer in large and

geographically widespread samples (5;45;46). Furthermore, they assess intake over an extended period of time and are less time-consuming for subjects to complete than many other dietary assessment methods. Data collection and processing can be administered in a standardised way and with relatively low cost (47). Disadvantages of FFQs include the amount of work required for questionnaire development and validation and the level of imprecision in the dietary estimates (40;46). Although FFQs provide reasonable estimates of food intake, the level of nutrient intake estimated by FFQ should be regarded as approximations. FFQs are better suited for ranking subjects according to food or nutrient intake than for estimating absolute intakes. However, for estimating relative risks in epidemiological studies, the correct classification of subjects according to intake is more important than the scale on which the ranking is made (45).

Every FFQ has to be validated to get an expression of the degree to which it is an accurate measure in the target population (5). Possible dietary reference methods in validation studies are food records or 24-hour recalls. Both are open-ended and accommodate diverse dietary patterns (5). Although 24-hour recalls are less demanding for the participant than dietary recording and less likely to influence the actual diet of the subjects, their sources of error seem to be more correlated with the error in dietary questionnaires (such as reliance upon memory, conceptualization of portion sizes and distortion of reported diet) (5;48). The weighed record has been considered the "gold standard" in dietary assessment and is the preferred reference method when validating food frequency questionnaires (5;40). The food diary is a very precise measurement of all food items over a given number of days and does not rely on memory or on the ability to estimate portion sizes. However, recording all food and drink is very demanding and includes the possibility to underreport intake or to eat differently during the recording period. To obtain good quality data from this method, participants must be trained to provide complete detail on mixed dishes, snacks and recipes, and follow-up is required to review records with the participants to complete the missing details. The accurate and complete documentation of intake requires motivated and literate participants (49).

Most FFQs estimate intake of vitamins and minerals from food only, leaving out information pertaining to the use of dietary supplements; whereas, in fact, the total intake of minerals and vitamins is of interest. Assessments of dietary supplement use in pregnant women have focused mainly on prevalence and patterns of supplement use (50;51).

Dietary assessment methods are associated with both random and systematic error. These errors arise from the use of food tables, assessment of the frequency of consumption, portion size, daily variation and failure to report usual diet, due to either changes in habits while taking part in the investigation or misreporting of food choice or amount (52-54).

1.3.3 Biomarkers of dietary intake

Biochemical measurements (biomarkers) of nutrient or dietary factors may provide useful, objective estimates of the dietary intake independent of the errors associated with self-reports. However, they are often expensive and nutrient specific, so may only be used to validate one nutrient at a time (5;55). The use of biomarkers in nutritional epidemiology is not new, although they have traditionally been referred to as indicators of nutritional status (53). Biomarkers may be measured in biological specimens such as blood, urine, saliva, human milk, nails, hair, fat tissue, skin, sweat, bone and teeth. In general there is a need to establish how tissue levels equate to consumption.

There are two types of biomarkers; First, there are quantitative (recovery) biomarkers, as in 24-hour urine collections which provide absolute values for comparison: for example, urinary nitrogen excretion for protein intake and doubly-labelled water for energy intake. Second, there are qualitative (concentration) biomarkers: for example, measurements in plasma and tissue which provide values for ranking individuals (52-54). The recovery biomarkers are accurately related to dietary intake, while most concentration biomarkers are only weakly associated to intake due to homeostatic control, regulated bioavailability, distribution, and metabolism. Analytical variation in biomarker assessment may also occur (55).

In validation studies, there is a need to be very clear about what the biomarker measures, to consider possible errors, the relevant time frame, and the relationship between biological variation in the biomarker and variation in dietary intake. However, there is a need for a greater variety of dietary biomarkers to be developed to reflect wider aspects of diet (54). Biomarkers for some food groups have been identified; for example, plasma carotenoids and urinary flavonoids for the intake of fruits and vegetables (56-58), and the fatty acid pentadecanoic acid (15:0), synthesised by the bacteria in the rumen, for the intake of dairy fat (59-61). Furthermore, plasma and tissue n-3 fatty acids has been used as a biomarker of marine fatty acids, contributed mainly by the intake of fatty fish (55;62-64).

1.3.4 The objectives of diet validation studies

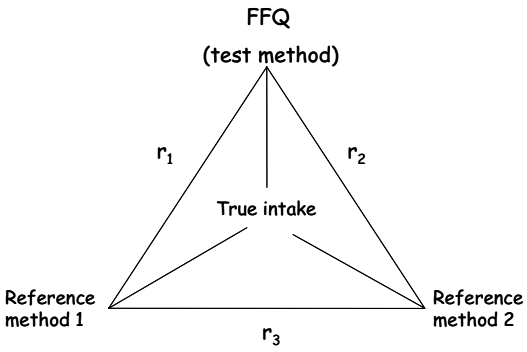
The purpose of diet validation studies is to estimate how well we can assess true dietary exposures in individuals using a particular diet assessment tool. Validation includes measurement of the true between-subject variation in the dietary factors of interest, and qualitative documentation that the dietary assessment method can detect the difference in diet that exists among subjects (1). Furthermore, a validation study can be used to identify subgroups among whom the questionnaire performs poorly and, hence might be excluded from the analysis in the main study. Finally, the validation process may be used to give quantitative assessment of exposure measurement error, so that measures of association, such as relative risks, can be corrected for measurement error (6;48). Validation studies are also referred to as calibration studies. The validation process should take into account the intended purpose of the study and assess the validity of food and food groups as well as nutrients.

1.3.5 Expressing the relative validity

Validity is an expression of the degree to which a method gives a true and accurate measure of what it is supposed to measure. Establishing validity requires a true, external reference measure against which the measurement can be compared. In nutrition, no such reference measure exists. Only the 'relative' validity of measurements can be assessed. Relative validity compares a new measurement method with one or more established methods believed to have a greater degree of face validity (6). Given the imperfect nature of measurement, validity is a matter of degree. The dietary assessment method being validated is called the test method, and the methods against which the test method is being compared and validated are called reference methods. If the new measure of the exposure gives the same results as a reference measure, then the new method is said to provide a valid measure. Reference methods can be other dietary methods as well as biomarkers. Errors and variations in dietary assessments will affect the validity of the measurements. The associations between the test method and reference methods can be examined and presented by different statistical approaches (1;5;40;65):

- Comparison of calculated absolute and energy adjusted intakes of foods and nutrients by the test method (FFQ) and a dietary reference method.
- Pair-wise correlations between intake estimates calculated by the FFQ and the reference measures. A correlation coefficient is an expression of the linear relationship between two sets of data.

- Ranking and cross classification.
- Bland-Altman plots. This is a plot of the differences between two measurements (y-axis) against the mean of the two methods (x-axis). This analysis assesses the agreement of the methods at the group level (defined as the difference between the two methods or relative bias), and the agreement in individuals (defined as the limit of agreement = plus or minus two standard deviations (± 2SD) of the bias) (66).
- The triangular method. This method makes use of three pair-wise correlations, for example, the correlations between a FFQ measure, a FD measure and biomarker measure, or the correlations between a FFQ measure and two independent biomarker measures. The triangular method, advocated by Ocke and Kaaks (67), calculates the correlation between the FFQ measure and the "true", but not known, intake, called a validity coefficient. The validity coefficients are calculated as: $VC_{FFQ}=\sqrt{r_1 \times r_2 / r_3}$, where the three pair-wise correlation coefficients are r_1, r_2 and r_3.



1.3.6 Validation studies in pregnant women

Food and nutrient estimates based on FFQ methods have been validated against other dietary methods in many different population groups and found to correlate with these, generally in the range of 0.4 to 0.7 (46;48). FFQ estimates also correlate to a varying degree with biological markers of dietary intake. Validation studies of dietary assessment methods in pregnant women are limited. A literature review of such studies published in 2003 identified ten studies where FFQs were validated in groups of pregnant women (68). Since then, at least three more studies have been conducted (69-71). These studies

are difficult to compare because of differences in the FFQ instruments, reference methods, and because the studies cover various periods of pregnancy. Three validation studies in European pregnant women found that the FFQ overestimated energy intake (41;42;72), whereas, a study in highly educated white pregnant women in the US found that the FFQ underestimated energy intake compared with food records (43). Correlation coefficients between the FFQ and the food record ranged from 0.62 to 0.68 for fats in a Belgian study (n = 26) (72), and from 0.27 to 0.37 for macronutrients in a UK study (n = 569) (42). In a Finish study, the average correlation was 0.47 for foods and 0.37 for nutrients (n = 113) (41), while in the US study correlation coefficients were larger than 0.5 for energy and 7 out of 15 nutrients (n = 56) (43). All studies concluded that the FFQ generally classified the women into the same or adjacent nutrient intake category as the food record.

FFQ validation studies that have included biological markers have shown differences in biomarker levels across varying levels of dietary intake of relevant nutrients and foods for alpha carotene, lycopene, lutein, gamma-tocopherol, and n-3 fatty acids (69;73). Correlation coefficients of 0.23 ($p<0.001$) between serum vitamin C and vitamin C intake (n = 569) (42) and of 0.41 between serum folate and folate intake (n = 2026) (74) have been reported.

Although most of the validation studies in pregnant women collected information on supplement intake, validation of reported supplement intake was not discussed. To our knowledge, the present validation of dietary supplement use in pregnant women is the first to be published (Paper 5).

Total energy expenditure has been examined in two studies in pregnant women which also evaluated total energy intake (75;76). In both studies the self-reported energy intake from food records under-reported total energy expenditure compared with the doubly labelled water method, indicating that dietary intake records can be misleading with regard to total energy intake. The inclusion an objective measure of energy expenditure as a reference for evaluating the reported energy intake by the FFQ and the FD was therefore given priority when planning the MoBa validation study.

2 Aims and research questions

A new food frequency questionnaire was developed for use among pregnant women in the Norwegian Mother and Child Cohort Study (MoBa). The aim of the MoBa FFQ evaluated in this thesis was to be able to rank individuals according to the level of nutrient and food intake. Paper 1 elaborates on the methodological challenges of dietary assessment and validation in pregnant women, and explains the design decisions involved in the development of the MoBa FFQ.

The overall aim of the present work was to validate the MoBa FFQ by comparing the self-reported FFQ data with relevant reference measures (weighed food diary and biomarkers) in pregnant women. The independent scientific contribution of this thesis is the planning, implementation and analysis of the validation study among participants in MoBa. The main research questions addressed in this thesis are:

- Is the MoBa FFQ a valid tool for ranking pregnant women according to the level of nutrient and food intakes relative to a 4-day weighed food record? (Paper 2)

- Is the MoBa FFQ a valid tool for measuring the intake of fruit, juice and vegetables relative to biomarkers in urine and plasma? (Paper 3)

- Is the urinary iodine excretion a useful biomarker in the validation of milk/dairy product intake in pregnant women, and is the MoBa FFQ a valid tool for measuring the intake of milk and dairy products? (Paper 4)

- Is self-reported dietary supplement use in pregnancy confirmed by biological markers in blood and urine, and is the MoBa FFQ a valid tool for measuring the intake of micronutrients provided by dietary supplements? (Paper 5)

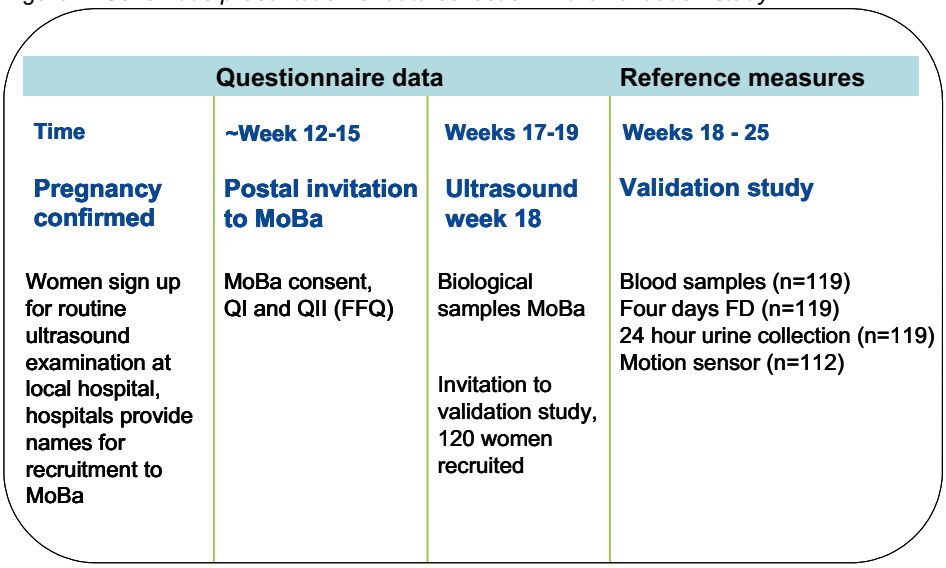
3 Subjects and methods

3.1 Study design

This thesis is based on results from the validation study carried out in a subgroup of pregnant women within the Norwegian Mother and Child Cohort Study (MoBa). The validation study was carried out as soon as participants had been recruited to the main study (MoBa). Strict control of the time between the test- (FFQ) and reference measures (validation study) was not possible due to postal distribution of the MoBa questionnaires (Figure 1). The completed questionnaires were not available for analysis until 3 months after the final inclusion of all participants in the validation study. Women fill in the date when completing the MoBa FFQ. The average time interval between completion of the FFQ and entry into the validation study was 24 days (standard deviation 12 days, range 1-59 days).

Participants in the validation study were asked to keep a 4-day weighed food diary (FD) and to provide a 24-hour urine collection and a blood sample. They were given detailed information and materials for data and urine collection. Self-reported weight, height and age were recorded in the food diaries. Data pertaining to smoking and education were collected from MoBa questionnaire 1.

Figure 1. Schematic presentation of data collection in the validation study



3.2 Study population

3.2.1 The Norwegian Mother and Child Cohort study (MoBa)

The Norwegian Mother and Child Cohort Study (MoBa) is a pregnancy cohort that in the period 1999 - 2006 has included > 75 000 pregnancies, and that aims to include 100 000 by the end of 2007 (2). Pregnant women are recruited to the study by postal invitation after they have signed up for the routine ultrasound examination in their local hospital. Participants are asked to provide biological samples and to answer questionnaires covering a wide range of information up to age of 7 years for the child. The study has been approved by the regional committee for ethics in medical research and the Data Inspectorate. The cohort database is linked to the Medical Birth Registry of Norway (77). The study is the largest and most costly of its kind in Norway ever, and has imposed multiple challenges economically, logistically and scientifically.

Dietary assessment in MoBa is comprised of the maternal diet in pregnancy and the diet of the child during the first years of life. In this thesis only the maternal diet is considered. The development and choice of dietary assessment method in MoBa is described in Paper 1.

3.2.2 Validation study participants

Healthy pregnant women in MoBa, assigned to Bærum Hospital (Norway) were invited to participate in a validation study when they came for routine ultrasound examination at around 18 weeks of gestation. Written invitations were given by midwives to eligible women (Appendix 2). Exclusion criteria were hyperemesis and anorexia. Subjects had to have completed the FFQ before inclusion. The four midwives who assisted in distributing invitations were instructed to invite all healthy women. However, the exact number of women that actually received the invitation was not recorded. Approximately 800 MoBa participants attended the clinic during the twelve month inclusion period (15 January 2003 – 1 February 2004). Participation in the validation study imposed a considerable burden due to the time and effort involved. Obtaining the planned sample size of 150 women was not accomplished. According to the midwives, the main reason for not choosing to participate in the study was the time and effort demanded. As a reward for participation in the validation study, each woman received a personal letter of dietary advice based on their food diary, a book about pregnancy, and the kitchen scale used for the weighed food diary.

3.2.3 Study approval

In the written declaration in which they agreed to participate in MoBa, women are made aware of the possibility that they may be contacted to participate in sub-studies. The validation study required a separate approval by the Regional Ethics Committee of Southern Norway. A new informed approval had to be signed by each woman who agreed to participate in the validation study. This approval stated the right of each person to withdraw from the validation study at any given time and that all information would be handled by non-identifiable identity-numbers only (Appendix 3).

3.3 Self-reported measures

3.3.1 Calculation of food and nutrient intakes

Details of the FFQ developed for use in MoBa is described in Paper 1. In brief, the MoBa FFQ is a semi-quantitative questionnaire that comprises 340 questions and asks what the mother has eaten since she became pregnant, and covering the habitual diet and the use of dietary supplements. The FFQ also includes questions relating to dietary habits and dietary changes due to the current pregnancy. Respondents are asked to fill in the mean intake of the food items eaten since becoming pregnant (78) (Appendix 1). The questionnaires were optically read. In the FFQ, portion size was only given for units of fruit, bread (slices) and liquids (cups/glasses). For dinners, vegetables, cakes and snacks the standard Norwegian portion size is used (79), although adjusted for some fruit and vegetable portion sizes reported in the validation study (food diary), and also adjusted for potatoes, rice and cereals according to more recent portion size estimations (80).

In the 4-day weighed food diary (FD) participants recorded the weight and detailed description of all foods and beverages consumed during three weekdays and one weekend day. Participants also reported intake of dietary supplements by brand name, frequency and dose. FoodCalc (81) and the Norwegian food composition table (82) were used for calculating the amounts of nutrients and foods reported by the FFQ and FD. For the calculation of nutrients supplied by the dietary supplements, a database containing details of the declared content of the supplements was developed.

3.3.2 Lifestyle and demographic measures

In the first MoBa questionnaire answered at the same time in pregnancy as the MoBa FFQ, participants reported marital status, parity, education (highest completed education), health status, and smoking habits.

Self-reported anthropometric measures (height and weight) were used for calculation of body mass index, resting metabolic rate, and for evaluation of weight change in the time period between filling in the questionnaire and entering the validation study. Self-reported weight at three different points in time was recorded; weight prior to this pregnancy, weight at the time of filling in the FFQ and weight at the time of inclusion in the validation study.

3.4 Objective measures

3.4.1 Motion sensor assessment of total energy expenditure

The motion sensor ActiReg® was used for four days of activity registration. ActiReg® is an instrument for measuring physical activity level and total energy expenditure by combined recording of body position and motion (83). Daily total energy expenditure (TEE) was computed with ActiCalc software. Resting energy expenditure (REE) was calculated with the WHO expert group standard equation (84), using weight and height at the time of inclusion. Physical activity level (PAL) is the ratio of TEE to REE.

3.4.2 Biomarker sampling and analysis

For the analysis of urinary excretion of nitrogen, iodine, and flavonoids each participant provided one 24-hour urine collection taken on a weekday. On the first morning of the urine collection, participants were asked to discard their first urine specimen and collect all specimens for the next 24 hours, including the first urine specimen the following day. More details pertaining to the urine collection are reported in Papers 2 and 3. For the analysis of biomarkers in blood, each participant provided three vials of blood (EDTA-blood, heparin blood and whole blood) at the time of inclusion. EDTA-blood was analysed within 24 hours, while heparin- and whole blood were separated into aliquots of serum and plasma within two hours of venipuncture and stored at -70°C until analysis. Erythrocytes for lipid analysis were washed with 9% NaCl, resuspended in NaCl and stored at -70°C until analysis.

Total urinary nitrogen was determined by the Kjeldahl technique (The Norwegian Institute for Food and Environmental Analysis, Oslo, Norway). Details of the other biomarker analyses are described in Papers 3, 4 and 5.

3.5 Statistical analysis

In all the papers, the p-values were two-sided, and a 5 % level of significance was used. Statistical analyses were performed with SPSS (version 12.0-14.0, SPSS Inc.). The statistical program R (85) was used in Paper 1 for the modelling of gamma curves, and in papers 3 and 4 for maximum likelihood estimation[1] and the bootstrap procedure[2] needed to estimate confidence intervals for the triangular validity coefficients.

Assumptions of normality were checked by the visual inspection of Q-Q plots and by the evaluation of the skewness and kurtosis of the variable. Most nutrient and food intakes were not normally distributed and non-parametric or parametric statistical tests were used as appropriate. The agreement between the two dietary methods and between the intake estimates and biomarker concentrations in the validation study is presented by Spearman rank correlation coefficients. Differences between the intakes estimated with the FFQ and the FD were tested with Wilcoxon's signed rank test (paired data), while differences in the intakes between groups were tested with the Mann-Whitney U-test (unpaired data) (Paper 2–5). Variance component analysis was used to calculate the within-person and between- person variation for the selected nutrients calculated by the FD (Paper 2). Linear regression analysis was used to identify the correlates of urinary iodine excretion (Paper 4) and to examine the impact of supplement use on biomarker measures (Paper 5). All linear regression models were checked for possible violations from the model assumptions (constant linearity and homoscedacity). Further details of the statistical analysis are described in each paper and in section 5.1.4.

---

[1] Maximum likelihood estimation is a statistical method used to make inferences about parameters of the underlying probability distribution from a given data set.

[2] Bootstrapping is used for estimating the sampling distribution of an estimator by sampling with replacement from the original sample.

4 Summary of results

Paper 1: Methodological challenges when monitoring the diet of pregnant women in a large cohort study; experiences from the Norwegian Mother and Child Cohort Study. Meltzer HM, Brantsæter AL, Alexander J, Ydersbond TA, Haugen M and the MoBa Dietary Support Group.

Many challenges and decisions were encountered when a new dietary assessment instrument was developed for the monitoring of diet in the women participating in the Norwegian Mother and Child Cohort Study. The foundation and rationale behind the development of the new food frequency questionnaire is described and discussed. Results from the first 40 000 women who completed the new FFQ, and the results from a validation study, indicate that the MoBa FFQ strikes a reasonable balance between the potentially conflicting methodological and scientific interests.

Paper 2: Validity of a new Food Frequency Questionnaire for pregnant women in the Norwegian Mother and Child Cohort Study (MoBa). Brantsæter AL, Haugen M, Alexander J, Meltzer HM.

The relative validity of nutrients and food groups calculated with the MoBa FFQ was evaluated. The reference measures were a 4-day weighed food diary (FD), motion sensor assessment of total energy expenditure, one 24-hour urine collection and a venous blood specimen. One hundred and nineteen women participated in the validation study and 112 women completed the motion sensor registration. The average correlation coefficient between the FFQ and FD for daily intake was 0.36 for nutrients and 0.48 for foods. Density adjusted correlations for macronutrients were $r = 0.44$ for protein, $r = 0.39$ for fat and $r = 0.36$ for carbohydrates ($p<0.001$ for all). The correlations for energy and protein intake between the FFQ and reference measures were influenced by pregnancy related nausea. On the average 68 % of the participants were classified into the same or adjacent quintiles when classified by the FFQ and the FD. The results of the validation study indicate that the MoBa FFQ gives reasonable valid intake estimates and is a valid tool for ranking pregnant women according to low and high intakes of energy, nutrients and foods.

Paper 3. Urine flavonoids and plasma carotenoids in the validation of fruit, vegetable and tea intake during pregnancy in the Norwegian Mother and Child Cohort Study (MoBa). Brantsæter AL, Haugen M, Rasmussen SE, Alexander J, Samuelsen SO and Meltzer HM.

The intake of fruit, vegetables and tea estimated by the MoBa FFQ was compared with urinary flavonoid excretion, plasma carotenoid concentration and intake measured by the 4-day weighed food diary (FD) in a validation study. The triangular method was applied to calculate FFQ validity coefficients using two independent biomarkers. The FFQ estimate of fruit intake was significantly correlated with urine phloretin (r = 0.33), citrus fruit/juice with urine hesperetin (r = 0.44), cooked vegetables with plasma α-carotene (r = 0.37), and tea with urine kaempferol (r = 0.41) (P<0.01 for all). Significant correlations between the FFQ and the FD were found for fruit (r = 0.39), vegetables (r = 0.34), juices (r = 0.50) and tea (r = 0.53). The FFQ validity coefficient was 0.65 for citrus fruit/juice and 0.59 for cooked vegetables as calculated by the triangular method. The results indicate that the MoBa FFQ provides valid estimates of the fruit, juice, vegetable and tea intake of pregnant Norwegian women, and that it can be used to rank individuals within the distribution.

Paper 4: Evaluation of urinary iodine excretion as a biomarker for intake of milk and dairy products in pregnant women in the Norwegian Mother and Child Cohort Study (MoBa). Brantsæter AL, Haugen M, Julshamn K, Alexander J and Meltzer HM.

The aim of this study was to explore the use of a 24-hour hour urinary iodine excretion as a biomarker for dairy product intake in pregnant women. Iodine was analysed in 24-hour urine samples. Dietary intake of milk and other food groups were estimated by a food frequency questionnaire (FFQ) and by a 4-day weighed food diary (FD). Using linear regression, predictors of urinary iodine excretion were identified. The triangular method was applied to calculate validity coefficients.

Significant predictors of 24-hour urinary iodine excretion were: reported intake of milk/dairy products, iodine containing supplements and intake of fruit/vegetables. Fish/seafood intake and time of the year influenced the 24-hour urinary iodine excretion, although not significantly. The validity coefficients observed for intake of milk and dairy products was 0.65, 0.94 and 0.52 by the FFQ, FD and 24-hour urinary iodine excretion respectively. The present study showed that 24-hour urinary iodine excretion is a useful biomarker in the validation of milk and dairy product intake in Norway.

Paper 5. Self-reported dietary supplement use is confirmed by biological markers in the Norwegian Mother and Child Cohort Study (MoBa). Brantsæter AL, Haugen M, Hagve TA, Aksnes L, Rasmussen SE, Julshamn K, Alexander J and Meltzer HM.

In a validation study including 119 pregnant women in MoBa, the relation between self-reported dietary supplement use and relevant biomarkers in biomarker-supplement and non-supplement users was examined. Biomarker concentrations and dietary intake differed significantly between the supplement and non-supplement users for vitamin D, carotenoids, folate, the n-6/n-3 fatty acid ratio and iodine ($p<0.05$ for all variables). Flavonoid excretion was higher in flavonoid-supplement users ($p<0.05$). Significant correlations between the total dietary intake (food and supplements) and biomarker concentrations were found for vitamin D ($r = 0.45$, $p<0.001$), folate ($r = 0.26$, $p = 0.005$), the n-6/n-3 fatty acid ratio ($r = 0.36$, $p<0.001$) and iodine ($r = 0.42$, $p<0.001$). The biochemical indicators examined in the validation study confirmed differences in the self-reported micronutrient intake between supplement and non-supplement users, and showed that the MoBa FFQ is a valid tool for estimating the intake of micronutrients supplied by dietary supplements.

5 General discussion

The work included in this thesis draws attention to the validity of the MoBa FFQ relative to several reference measures. In developing the FFQ, the goals have been to achieve a good classification of dietary intake, rather than a precise numerical estimation, and to record dietary patterns for future testing of a broad range of hypotheses. The relative validity varies for the different nutrients and food items. A more comprehensive discussion of the results is found in each paper.

5.1 Methodological considerations

5.1.1 Sample size and selection

The sample size of 119 is reasonable for a validation study (48;86), and the subjects came from the population for which the questionnaire was designed. Our initial aim was to include 150 women. However, recruitment was slower than anticipated, and the overall participation rate was low (15- 20%). The main reason given by those who chose not to participate was the extra burden of time and work involved in the validation study, as they already were involved in the MoBa study.

The study sample was not random, as all subjects were recruited from Bærum Hospital. Validation study participants were slightly older, better educated and included fewer smokers than a larger sample of MoBa participants (Table 2). In planning the validation study we did not aim for a representative sample, as the purpose of the validation study was to investigate the agreement between the FFQ, FD and biomarkers within the same individuals. However, it is possible that the associations between these measures may have been influenced by a selection or information bias. Subjects misreporting certain foods in the FFQ would be likely to misreport the same foods in the food diary. Furthermore, when a study-population is self-selected this may reduce the range of intakes and thus the correlation coefficients, since the magnitude of the correlation coefficient depends upon the range of intakes (between-subject variation) (66;87).

Table 2 Characteristics of the women included in the validation study (n =119) and of a nationwide sample of women in the Norwegian Mother and Child Cohort Study (MoBa) (n =39375)[a]

		Validation study Mean ± SD (min, max)	MoBa Mean ± SD (min, max)
N		119	40792
Age		31.2 ± 4.1 (23, 44)	29.6 ± 4.6 (14, 47)
BMI prior to pregnancy	Kg/m^2	23.2 ± 3.6 (17, 43)	24.3 ± 4.3 (13, 56)
		%	%
Age group	<20	0	2.4
	20 – 24	5	15.9
	25 – 29	41.2	39.6
	30 – 34	31.1	27.8
	35+	22.7	14
BMI prior to pregnancy	<20	16	11.6
	20 - 24	63	53.4
	25 – 29	15.1	22
	30+	5.9	10
Parity	0	55.5	44.2
	1	21	35.6
	2 +	23.5	19.3
Marital status	Married	60.5	49.7
	Living together	37.8	46.7
	Single	1.7	2.3
Smoking habits prior to pregnancy	Never	79.8	69.2
	Occasional	10.9	9.6
	Daily	9.2	20.3
Smoking habits during pregnancy	Never	97.5	89.4
	Occasional	0	3.2
	Daily	2.5	6.7
Education	≤12 years	16.8	39.3
	13 – 15 years	48.7	40.6
	>16 years	34.5	17.7
	Other or missing	0	2.4
Nausea during pregnancy	Yes	76.5	72.3
Nausea at time of FFQ	Yes	15.1	14.6
Work prior to pregnancy	< 30 hours	16.0	23.1
	> 30 hours	75.6	60.9
	Student	6.7	8.7
	Housewife	1.7	7.3
Work in pregnancy	< 30 hours	23.5	27.4
	> 30 hours	66.4	54.6
	Student	8.4	10.7
	Housewife	1.7	7.3

[a]*The MoBa file included FFQ data for 40786 subjects, while background variables were available for n=39375 (96.5%) of these*

The 119 participants in the validation study varied with regard to age, pre-pregnant body mass index (BMI), parity, marital status, education and smoking. The participants were sufficiently heterogeneous to reflect known differences in food intake related to smoking

and education (88;89), as the intake of fruit, juice and vegetables were significantly greater in non-smoking women (p=0.015) than in smokers, and in women with education > 12 years compared to women with =< 12 years of education.

Due to the logistics, there was a time lapse between completion of the FFQ and participation in the validation that allowed for weight changes to occur in the individuals and influenced the correlation of total energy intake between the FFQ and FD (Paper 2-4). Likewise, nausea, which was reported by a majority of the participants (77%), influenced the correlations (Papers 2 and 4).

5.1.2 The dietary reference method

Four days of weighed food diary was chosen as the dietary reference method in this validation study. Recording, that is, weighing and measuring of food, drink and dietary supplements does not rely on memory or an ability to estimate portion sizes and satisfies the criteria of independence of errors associated with the two methods compared (FFQ and FD) (5). However, common errors in both test and reference methods are the use of national food composition tables and a tendency to misreport food intake. The alternative to a weighed food record would be repeated 24-hour recalls. Weighed records are usually recommended over 24-hour recalls although the recalls are less demanding for the participants (5). Keeping a food diary is demanding and includes the possibility of under-reporting intake or eating differently during the recording period. Comparison of the FFQ and FD energy intakes with total energy expenditure (motion sensor) in our study showed larger underreporting by the FD than by FFQ (Section 5.3). Many questions relating to the food diary came up, indicating that conscientious food recording was difficult. Participants were asked not to alter their food habits during the recording days, but many admitted that snacks were difficult to record.

Because of day to day variation in the consumption of many foods, four days is a rather small number of days to use as a reference measure for the habitual food intake reported by an FFQ. Four days of food recording can be sufficient if the sample size is large, however, foods eaten rarely will not be accurately assessed (86). Upon collecting dietary data at the individual level there is a trade-off between the burden one can impose on the included subjects and the accuracy, and hence the usefulness of the data. As several biomarkers and activity registration were included as reference methods in addition to the food diary, four days was considered the acceptable burden that could be imposed on the study participants.

The calculated daily intake of nutrients and foods was larger by the FFQ than by the FD, as is the case in most validation studies (90-92). Differences between intakes calculated with the two methods were statistically significant for nutrients more often than for food groups (Paper 2 and 3). The FFQ covers a longer time span and may thus be a better reflection of habitual intake than the FD, since the use of the questionnaire is based on the concept that the average diet over time is a better measure of exposure than intake over a few specific days (45). Hence, the two methods are not really comparable, and this is why biomarkers were included as additional and important reference methods. However, comparison of intakes calculated by the two methods provides useful insight into the strengths and weaknesses of the FFQ relative to the FD.

Misreporting of energy intake has been described in free-living individuals in many population groups including pregnant women (75;93;94). Physiological and behavioural characteristics of under-reporters have been identified (95;96). In dietary assessment it is not possible to distinguish whether under-recording or under-eating is underlying the under-reporting. In the present study weight change > l kg occurred in 77 out of 119 in the time between filling in the FFQ and the FD. Two of these lost weight, while the remaining 75 gained weight. In light of this, energy intakes by the food diary would be expected to exceed the FFQ intakes and be similar or higher than the estimated energy expenditure. However, energy intake by the FD was 91% of the estimated energy expenditure and energy intake by the FFQ was 96.5% of the estimated energy expenditure, indicating that at the group level participants under-reported their total energy intake more in the FD than in the FFQ.

Four days is not sufficient to reflect true long-term intake, especially for food groups and nutrients that are not eaten on a daily basis. For selected nutrients an attenuation factor (AF) was calculated based on the day-to day variation in the FD.

The correction of the observed correlations for the attenuating effect of random within-person error can be written as $r_{true} = r_{observed} \times \sqrt{1 + \frac{S_w^2 / S_b^2}{n}}$, where S^2_w is the within-person variation and S^2_b is the between-person variation and n is the number of days (1;40). Correlations were strengthened when attenuation factors were applied (Paper 2).

Statistically significant correlations between the FFQ and FD calculated intakes were found for all nutrients and food groups (Paper 2). The correlations were generally stronger for food groups than for nutrients. Energy adjustment improved some of the correlations, but not all.

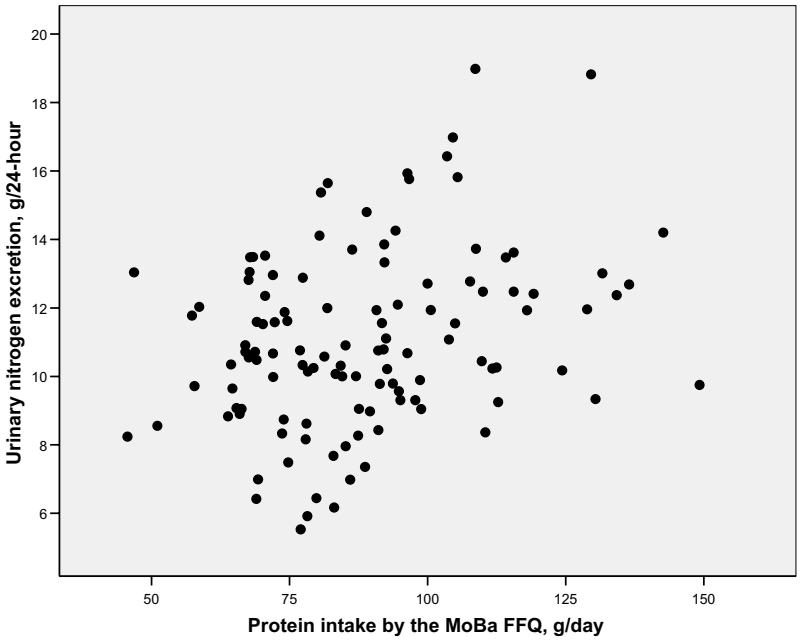
Classification into quintiles showed that on average two thirds of subjects were classified by the FFQ to within ± one fifth of their recorded intake (FD), whether absolute or density adjusted intakes were examined, and that few subjects (<10%) were grossly misclassified into opposite quintiles by the two methods (Paper 2).

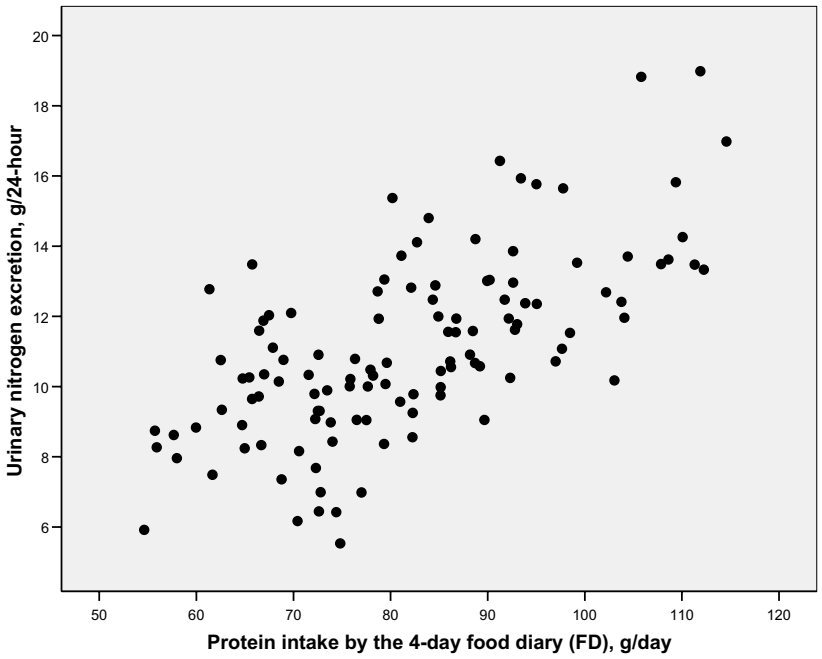
5.1.3 Biological markers

Biological markers are useful to overcome the problem of correlated errors between dietary assessment methods. The usefulness of biological markers is primarily for use as independent reference measures in the validation of dietary measures rather than as replacements of dietary measures (97). Urinary nitrogen excretion is the most well-known biological marker, with results from published metabolic studies showing high correlation between daily nitrogen (protein) intake and daily urinary nitrogen excretion (55). However, nitrogen balance is a condition for the use of this biomarker that is not present during pregnancy. Furthermore, the use of an external marker such as *Para*-aminobenzoic acid to verify completeness of urine collection (98) was considered unacceptable in this population.

One 24-hour urine collection cannot be expected to fully reflect habitual protein intake, due to the day to day variation in protein intake and urinary nitrogen excretion. In spite of this, we found statistically significant correlations between a single 24-hour urinary nitrogen excretion and protein intake calculated by the FFQ and the FD (Figure 2). The correlation between the 24-hour urinary nitrogen excretion and protein intake was $r=0.27$ ($p=0.004$) for the FFQ and $r=0.65$ ($p<0.001$) for the FD ($n=117$ singleton pregnancies). The 24-hour urine collection was done close to the FD registration and a stronger association between the FD and urinary nitrogen was expected. When participants with a weight change larger than 1 kg (between the FFQ and FD assessment) were excluded from the analysis, the correlation between the urinary nitrogen excretion and protein intake by the FFQ was $r=0.34$ ($p=0.030$). Similarly, when participants reporting nausea were excluded, the correlation was $r=0.58$ ($p=0.001$), demonstrating how factors relating to pregnancy influenced the validation.

Figure 2. Urinary nitrogen excretion (y-axis) plotted against protein intake by the 4-day weighed food diary (x-axis)(upper figure) and plotted against protein intake by the MoBa FFQ (lower figure) in 117 singleton pregnancies





When all participants were included, the correlation between the 24-hour urinary iodine excretion and the iodine intake was r=0.42 for the FFQ and r=0.53 for the FD (Paper 4), The correlations between the 24-hour urine flavonoid excretion and the intake of fruit, vegetables and tea were in the range 0.20 to 0.40 for the FFQ and 0.20 to 0.60 for the FD (Paper 3). Lower correlation coefficients for the FFQ than for the FD were expected due to the time lapse between the assessment methods and the superior detail of a FD.

Significant correlations between the urinary excretion of flavonoids and the calculated intakes of fruit, vegetables and tea were found for both the FFQ and the FD despite only a single 24-hour urine collection for each participant (Paper 3). The amounts of excreted flavonoids were in good accordance with previous reports of urinary flavonoid excretion in non-pregnant women in Denmark and Finland (57;99).

The correlations between serum, plasma and erythrocyte biomarkers and dietary intake of nutrients were strongly influenced by micronutrients and fatty acids supplied by dietary supplements and were used for assessing the relative validity of self-reported dietary supplement use (Paper 5).

In the present validation study, several factors may have attenuated the associations between biological markers and the dietary assessment. Blood samples were not drawn while the subjects were fasting. The blood samples were collected in the evening when groups of 5-10 participants met with the project worker. This was arranged so that the participants did not have to take time off from work in the mornings. Fasting blood samples obtained in the morning would have been optimal, because the circadian ('about 24-hour') variation in the hormones that effect food intake and metabolism may also effect biomarker concentrations. The impact of using non-fasting blood specimens is uncertain. Ahn *et al*. found no evidence of circadian variation in folate pharmacokinetics (100), while it has been reported for lipids and cholesterol (101).

Non-fasting blood samples have been used in a validation study of carotenoid intake by FFQ (102). The analysis of serum carotenoids was based on a single blood sample for each participant. However, contrary to urine measurements, it has been shown that a single sample may accurately rank individuals (56;103). Dixon *et al.* reported no advantages of 2 blood samples over 1, thus suggesting a reasonable stable ranking of individuals for carotenoids and tocopherol in women with only one blood sample (103). The plasma levels of carotenoids in the validation study corresponded well with levels found in female participants in other European counties (104).

There is a great need for a larger variety of dietary biomarkers to be developed to reflect wider aspects of diet. To our knowledge, Paper 4 is the first publication to use urinary iodine excretion as a biomarker in the validation of milk and dairy product intake by FFQ. In spite of having only a singe 24-hour urine collection, and in spite of iodine not being exclusively found in dairy products, fair correlations were shown between dairy product intake estimated by the FFQ, FD and this biomarker.

The correlations between estimated intake and biomarkers are generally weaker than correlations between two dietary methods. This is because biomarker concentrations will be influenced by factors unrelated to intake, such as individual differences in absorption, metabolism and distribution, which in again is influenced by age, weight, smoking status and other factors (56).

The time-period for which a biomarker reflects dietary exposure differs for the different biomarkers and for the different body tissues. The fat soluble carotenoids are fruit and vegetable constituents with an estimated half-life of 1-3 months (105), while flavonoids are fruit and vegetable constituents with short half-lives, often no more than a few hours (99;106). As a result plasma carotenoids would reflect the intake of fruit and vegetables over several weeks, while the 24-hour urinary excretion of flavonoids would only reflect the intake of fruits and vegetables from the present and previous day or meal. Multiple urine collections for each person would be needed to provide a reasonable average representing their habitual intake. The 24-hour urinary nitrogen and iodine excretion are markers of a short-time intake, while plasma, serum and erythrocyte substances represent intakes over longer periods of time. Both short-term and long-term biomarkers are included in this thesis. In our study, the fact that urinary flavonoids and plasma carotenoids were mutually correlated strengthened our results and showed that even a single urine collection was of value (Paper 3).

5.1.4 Statistical issues

Correlation coefficients have been used extensively in the presentation of the results from this validation study. Correlation coefficients are convenient because they only need a single number for each food or nutrient to reveal something about the relationship between two measurements, and they facilitate comparison with other FFQ validation studies. Although the correlation coefficients seem to be a good way to present validity, there are several important issues that must be considered when evaluating them. *First*, a correlation coefficient is only an indicator of the relative agreement between two

variables and reveals nothing about the absolute agreement between them (66). Food intake measured with two dietary methods may have large differences in the absolute intake of food, but still have a relatively high agreement if the individuals were ranked similarly according to food intake estimated from the two methods. *Second*, the correlation coefficients are influenced by correlated errors in the dietary methods: for example, correlations between an FFQ and the average of 24-hour recalls tend to be higher than the correlations between an FFQ and a food record because both the FFQ and the 24-hour recalls rely on memory (5;107). *Third*, the range of intakes (between-subject variation) will influence the magnitude of the correlations, and give lower correlation coefficients if the between-person variation is small (66;87). This is more likely to occur in a self-selected study sample (as in this study) than in a random study sample. *Fourth*, the two most commonly used methods, Pearson correlation of log-transformed data or Spearman correlation do not necessarily give the same result. When both methods were applied to the same data Pearson correlation coefficients were generally higher than the Spearman correlation coefficients (65). This results suggest that log transformation did not remove the influence of outlying data points on the Pearson coefficients, and that the Spearman coefficients may be more reliable since they use rank order and are, therefore, not as sensitive to extreme values as the Pearson coefficient (65).

Finally, there are no set rules as to what constitutes a satisfactory level of correlation. Most authors draw the conclusion that the higher the correlation coefficients, the better the validity of their questionnaire. However, the following question has been raised: "How good is good enough, and how bad is too bad for dietary measures?" (38). Recent reviews of validation studies have shown that, regardless of how detailed the FFQ is and regardless of how many days of food records or recalls there are, there may be a "ceiling of validity" between questionnaires and reference methods of approximately 0.7 (108). Correlation coefficients in the range of 0.5 – 0.7 may seem low to scientists in other fields, but appear to be the best attainable in dietary validation. Correlations in the range of 0.30 – 0.49 are considered fair, while correlations < 0.3 are considered poor because they are too low to detect a potential risk between the measure and an epidemiological outcome (5;65;107;109). For correlations between dietary intakes and biochemical concentration biomarkers, the magnitude of the correlations will tend to be modest, even when the dietary measurements are highly accurate and precise (56). This is because biomarker concentrations are influenced by factors other than dietary intake, as described in section 5.1.3.

Both the correlation coefficient and the agreement of classification into categories indicate the relative agreement between two measurements, or the questionnaire's ability to rank subjects according to exposure. The MoBa FFQ was able to distinguish between high and low consumers according to the FD, even when the magnitude of the correlation was as modest as 0.29 (Table 3). The results regarding total intake of fish/seafood in Table 3 have not yet been published.

Table 3 Spearman correlation coefficients (r) between FFQ intakes and reference measures (FD and biomarkers) and median level of food intake by the FD and median biomarker level according to quintiles of FFQ intake. N differs because women using a dietary supplement containing the biomarker were excluded before analysis

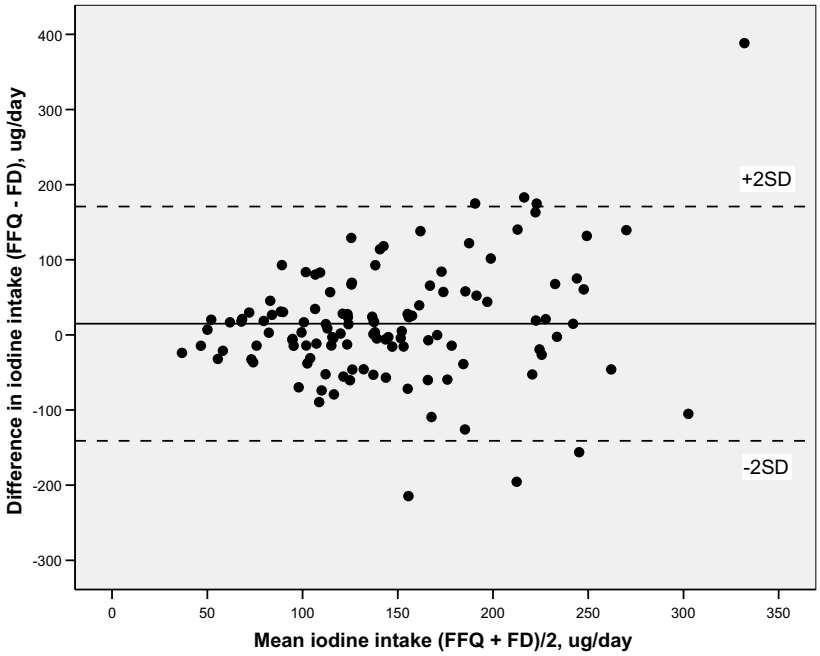
Food group	n	r	Q1 (FFQ)	Q2	Q3	Q4	Q5	Q1 vs Q5[a]	*p* for trend[b]
Dairy products									
Intake by FD (g/day)	119	0.58**	197	289	443	356	600	<0.001	<0.001
Urinary iodine (µg/24h)	84	0.34**	82	128	113	107	152	<0.001	0.002
Citrus fruit and juice									
Intake by FD (g/day)	119	0.39**	63	106	116	202	304	<0.001	<0.001
Urinary hesperetin (µg/24h)	99	0.44**	58	205	416	985	1521	<0.001	0.001
Cooked vegetables including roots									
Intake by FD (g/day)	119	0.29**	44	51	54	66	67	0.026	0.089
Alpha-carotene (µmol/L)	106	0.37**	0.083	0.096	0.100	0.152	0.131	0.006	0.004
Fish, fish products and seafood									
Intake by FD (g/day)	119	0.29**	20	38	53	38	67	0.009	0.045
Whole blod Arsenic (µg/L)	119	0.38**	1.19	1.54	1.46	2.04	3.59	0.001	0.001

* p< 0.05
**p< 0.001
[a]Mann-Whitney U-test
[b]Kruskall-Wallis H-test

With a p-value less than 0.05, correlations were considered significantly different from zero. However, due to multiple comparisons significances were interpreted with care, and relevant p-values specified. With multiple comparisons a Bonferroni correction can be used. This would reduce the overall probability of finding significant associations by chance. This correction is rarely used when several foods/nutrient intakes estimated by one method is compared to the same intakes estimated by a second method. Instead of using a fixed correction for multiple comparisons in the papers included in this thesis we have specified the p-value by indicating whether the significance was at the 0.05 level or at the 0.01 level, or by stating the exact p-values when relevant.

The Bland-Altman plots for investigation of absolute agreement between methods is especially important for evaluating agreement of absolute intakes. From this plot it is possible to observe the magnitude of disagreement, outliers and trends. Two Bland-Altman plots are presented in Paper 2. The Bland-Altman plot of iodine intake is shown below (Figure 3).

Figure 3. Bland-Altman plot showing the difference in total iodine intake (including iodine from supplements) calculated by the food frequency questionnaire (FFQ) and the food diary (FD) plotted against the mean intake by the two methods. The solid line is the mean difference between the two methods and the dashed lines denote ± 2 standard deviations (SD). N=119



The mean difference showed that the estimated intake of iodine was larger with the FFQ than with the FD, but the mean difference was small (15 µg). Furthermore, the plot reveals that the observed differences between the FFQ and the FD were negative as well as positive, implying that participants both under- and over-reported intakes with the FFQ in compared with the food diary. However, nearly all the observations fell within the limit of agreement, defined as the difference ±2 SD of the difference.

Using the method of triads resulted in validity coefficients for the FFQ that were higher than the correlation coefficients between the FFQ and each of the biomarkers. Despite the fact that the triangular method states the need for independent measurements, two dietary methods and one biomarker have been used in the triangular equation in other published studies (102;110-112). To our knowledge Paper 3 is the first study that has applied three totally independent measurements with this equation. In spite of the fact that the three pair-wise correlation coefficients were only modest, they served to improve the robustness of the validation process. We also calculated triangular validity coefficients using the FFQ, FD and one biomarker, as used in other studies, for fruit and vegetables, and for evaluation of urine iodine as a biomarker for intake of dairy products (Paper 4). For urinary iodine excretion, the validity coefficient for the total iodine intake was higher than the validity coefficients for the intake of milk/dairy products, indicating this biomarker has higher validity for calculating total iodine intake than for predicting the intake of dairy products. The practical use of urinary iodine excretion as a biomarker is not for the prediction of milk/dairy product intake as a substitute for dietary assessment, but as an independent reference measure in addition to the dietary method.

Further considerations and statistical issues related to the triangular method are discussed in Paper 3.

5.2 FFQ controversy

The applicability and accuracy of the FFQ method is controversial (38;108;113-115). Most researchers agree that an FFQ is appropriate for obtaining usual, long-term dietary information from literate persons when the objective is to rank subjects and evaluate associations between dietary habits and health outcomes. However, the strongest criticism has come from cancer researchers, where imprecision in the dietary estimates attenuate risk estimates and thereby reduce the power of the study (114;116;117). Scientists in basic and experimental nutrition can measure dietary exposure precisely, but the outcomes are less precise. In observational epidemiology the situation is turned around. The measurements of dietary exposure are rather crude, but the outcomes are specific, like disease or mortality. Thus, observational studies are important for issues that cannot be studied through experimental studies (38;40). In spite of the low precision in the FFQ estimates, FFQs have provided insight into population health and disease that later have been confirmed by other types of studies (38;108). Rather than examining

single or few nutrients in relation to disease, food intake and food pattern analyses have come forward as a better approach (118). Analyses of food groups can reveal diet-disease relationships that do not relate to any known chemical substance, while food patterns allow for studies of the cumulative effects of several nutrients, substances and foods (119-122).

Some of the criticism against FFQs related to low precision is probably less relevant to the MoBa questionnaire. We are only focusing on a relatively short time period, resulting in less recalling, estimation and abstraction for the participants. Furthermore, the MoBa questionnaire aims for good classification rather than precise numerical estimation of dietary intake.

Much of the disagreement regarding FFQs is related to the use of energy adjustment; that is whether the food frequency reports should be regarded as representation of dietary composition or as estimates of absolute intake (38;108). The strategy of adjusting for total energy intake depends on the assumption that all nutrients, and consequently foods, are misreported in approximately similar proportions. Studies of misreporting have shown that fat and/or carbohydrate intake tend to be underestimated more than protein intake (123;124). In the present validation study, density adjustment improved the correlation between FFQ and FD estimates for macronutrients, but not for micronutrients, fruit, vegetables or other major food groups. The main argument for using energy (or density) adjusted food and nutrient intakes is to study dietary composition rather than absolute intake (125).

The FFQ method is relevant for many purposes, and is more or less suitable depending on the research questions and the characteristics of the target population. The approach is the most practical in large observational studies and for dietary surveillance in different population groups. Food-frequency questionnaires critically depend on the participants' long-term knowledge of their own dietary patterns, and their willingness to report this, and are intended to measure average intakes. Another way to look at food-frequency questionnaires is to acknowledge that they have less to do with memory for what was consumed than with subjective assumptions about the nature of the habitual diet. Food frequencies, much like food preferences or body image, appear to be a measure of attitude. As such, they may not be subject to absolute validation procedures (47).

5.3 Quality of reported intakes

Misreporting of dietary intake is seen with all dietary assessment methods (1;96;126) and has been described in various populations, including pregnant women (75;76). Evaluating the validity of reported EI provides a valuable check on the general quality of the dietary data in any study (94). The ratio of reported energy intake to estimated or measured resting energy expenditure (EI/REE=PAL_{EI}) can be applied to determine if the EI reported by an individual or a group is sufficient to sustain life. The derivation of cut-offs based on EI/REE was published by Goldberg *et al.* and later modified by Goldberg and Black (127;128). The Goldberg cut-off values have mainly been used to identify low energy reporters and to identify characteristics of low energy reporters within population groups, but not to exclude poor data (95;96;124;129;130). For exclusion of biological improbable intakes, a lower limit of 2500 kJ/day and an upper limit of 15000 kJ/day (alternatively 500 and 3500 kcal) were applied in the Nurses' Health study (131), the European Prospective Investigation into Cancer and Nutrition study (EPIC) (132) and in the Norwegian Women and Cancer Study (133) without further justification than biological plausibility. In his book, Nutritional Epidemiology (1), Walter Willett denotes this as "the arbitrary allowable range for women".

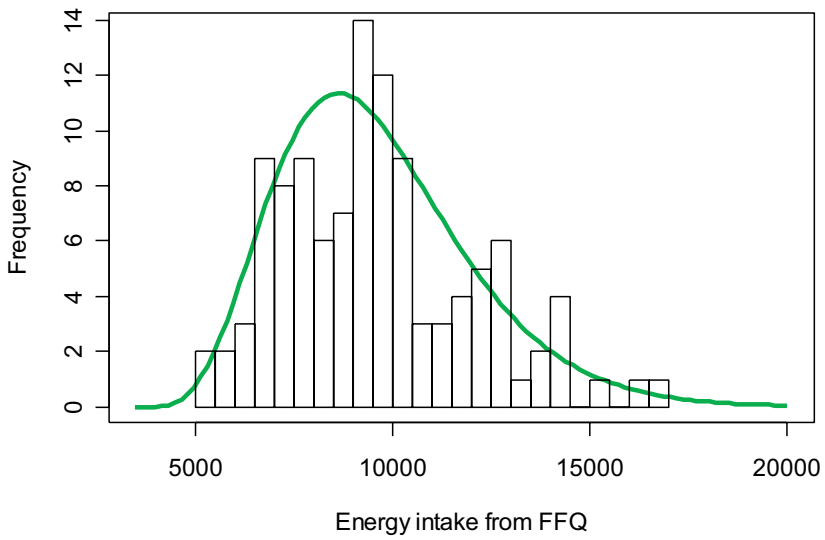
Assessment of total energy expenditure by motion sensor was included as an independent reference measure in the validation study in order to explore the possible misreporting of energy intake. In the study sample, the average motion sensor PAL (TEE/REE) was 1.55 (range 1.29 – 1.96), the average FFQ PAL_{EI} was 1.70 (range 0.70 - 2.70) and the average FD PAL_{EI} was 1.50 (range 0.90 - 2.10). The FFQ energy intake constituted 96% of the total energy expenditure (motion sensor) and the FD energy intake constituted 91% of the total energy expenditure. These results indicate that under-reporting was more extensive with the FD than with the FFQ. No participants were excluded from the validation study on the basis of misreporting, as the aim was to examine the relative validity of the FFQ independent of the quality of the reports.

In an attempt to establish study-specific energy cut-off values, rather than the arbitrary cut-offs used by Willett and others, we examined energy intake and expenditure in the validation group (n = 112) and compared their energy intake with the larger pregnancy population (n = 40 786). The averages of energy consumption as estimated from motion sensor data, 10.02 ± 1.02 MJ/day, corresponded well with intake means calculated from food diaries (FD, 9.09 ± 1.35 MJ/day) and food frequency questionnaires

(FFQ, 9.61 ± 2.43 MJ/day) for the validation group. The range of the FFQ calculated intakes was 5.00 to 16.67 MJ/day, while the range of energy expenditures was 7.87 to 12.88 MJ/day. The mean FFQ intake also corresponded well with the MoBa population mean (9.93 MJ/day) (n = 40786), while the standard deviation in the entire study population was somewhat higher (3.30 MJ/day). This indicates that the validation group did not fully mirror the variation in the whole population, but the validation group data may still be used for establishing cut-off points. Figure 2 shows the histogram of the FFQ intakes with a gamma curve fitted by maximum likelihood estimation. A 99.8% confidence interval (CI) from this curve is (4.52, 19.70) MJ/day, and when 4.5 and 20 MJ/day are used as cut-off points for the population data, we exclude approximately 1 percent at each end. The 99.8% CI from the validation group thus becomes a 98% CI for the population. In our study, a limit of 15 MJ would exclude 5.4% of all subjects, and our energy expenditure data indicate that an intake on this level is not improbable.

It should be kept in mind that exclusion of subjects with improbable energy intakes will not correct for any misreporting or bias among the included subjects. As in other epidemiological studies, the exclusion of dietary reports based on relative misreporting is not going to be feasible in MoBa, and the energy adjusted intakes have to be considered when food composition is more appropriate than absolute intakes.

Figure 4. The distribution of energy intakes (kJ/day) calculated from the FFQ, with gamma distribution fitted by maximum likelihood estimation.



Measurement error in dietary assessment and adjusting for measurement error is a difficult, complex and important issue in nutritional epidemiology (134;135). Different approaches for estimating the effect of measurement error have been investigated, and needs to be further investigated within the MoBa dietary data. However, further investigation of error structures and methods to adjust for these are beyond the scope of this thesis.

5.4 Usefulness of the motion sensor assessment

Wearing the motion sensor turned out to be the reference measure most difficult to comply with. Of the 119 participants in the validation study, 112 completed four days of motion sensor assessment. Three women discarded the motion sensor due to discomfort while wearing it, three women did not wear the motion sensor according to the instructions and one woman did not wear it because of fear of radiation on the fetus. The data obtained from the 112 who completed the assessment did, however, add valuable insight into the quality of energy intake reported by the FFQ and FD.

6 Conclusions and future perspectives

Validation is required before dietary data derived from a new instrument can be used in studies of dietary exposure and disease. Few studies have validated FFQs for assessment of dietary intake in pregnancy. In spite of limitations related to pregnancy and imperfect reference methods, the present study shows that the MoBa FFQ is a valid and useful tool for estimating habitual intake in order to rank individuals according to the level of nutrient and food intake. Intakes reported by the FFQ correlated significantly with intakes reported by a 4-day weighed food diary and with relevant biomarkers of intake. Furthermore, cross-classification of the FFQ intakes and reference measures showed that the FFQ is able to rank individuals according to high or low intake and, even more important, the degree of misclassification is small (Paper 2-5).

The present study has demonstrated the utility of biological markers in the validation of dietary intake. For the first time, two independent biomarkers (urine flavonoids and plasma carotenoids) were combined in the validation of fruit, vegetable and tea intake. Since only a few previous studies have used urinary flavonoid excretion in the validation of a food frequency questionnaire, our results confirm the value of this biomarker. Furthermore, our results demonstrate the potential applicability of urinary iodine excretion as a biomarker in the validation of milk/dairy product intake in pregnant Norwegian women. Further studies aiming at establishing urinary iodine excretion as a biomarker of milk/dairy products are warranted. There is a widespread use of dietary supplements during pregnancy in many countries. We have showed that self-reported dietary supplement use in pregnant Norwegian women was confirmed by relevant biomarkers in urine and plasma. To our knowledge, this is the first time dietary supplement use has been validated in pregnant women.

The process of validating the new MoBa FFQ has been challenging and time-consuming, but is an absolute condition for future use of the dietary data. The work presented in this thesis is the foundation for further dietary research in MoBa. At present, the nutritional data in the Norwegian Mother and Child Cohort study is among the largest epidemiological databases worldwide containing extensive information on maternal dietary exposure. In the future, these data will be used in numerous studies examining the impact of dietary factors on maternal and child health.

7 Reference list

1. Willett WC. Nutrition Epidemiology. New York, Oxford: Oxford University Press, 1998.

2. Magnus P, Irgens LM, Haug K, Nystad W, Skjaerven R, Stoltenberg C. Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). Int J Epidemiol 2006;35:1146-50.

3. Moore VM, Davies MJ. Diet during pregnancy, neonatal outcomes and later health. Reprod Fertil Dev 2005;17:341-8.

4. Godfrey KM, Barker DJ. Fetal programming and adult health. Public Health Nutr 2001;4:611-24.

5. Cade J, Thompson R, Burley V, Warm D. Development, validation and utilisation of food-frequency questionnaires - a review. Public Health Nutr 2002;5:567-87.

6. Nelson M. The validation of dietary assessment. In: Margetts BM, Nelson M, eds. Design Concepts in Nutritional Epidemiology. Oxford: Oxford University Press 1997:241-72.

7. Smith CA. Effects of wartime starvation in Holland on pregnancy and its products. Am J Obstet Gynecol 1947;53:599-608.

8. Barker DJ. Deprivation in infancy and risk of ischaemic heart disease. Lancet 1991;337:981.

9. Wu G, Bazer FW, Cudd TA, Meininger CJ, Spencer TE. Maternal nutrition and fetal development. J Nutr 2004;134:2169-72.

10. Vangen S, Nordhagen R, Lie KK. [Revisiting the Forsdahl-Barker hypothesis]. Tidsskr Nor Laegeforen 2005;125:451-3.

11. Painter RC, Roseboom TJ, Bleker OP. Prenatal exposure to the Dutch famine and disease in later life: an overview. Reprod Toxicol 2005;20:345-52.

12. Susser M, Stein Z. Timing in prenatal nutrition: a reprise of the Dutch Famine Study. Nutr Rev 1994;52:84-94.

13. Henriksen T. Nutrition and pregnancy outcome. Nutr Rev 2006;64:S19-S23.

14. Medical Research Council. Prevention of neural tube defects. Lancet 1991;338:131-7.

15. Wald NJ. Folic acid and the prevention of neural-tube defects. N Engl J Med 2004;350:101-3.

16. Durnin JV. Energy requirements of pregnancy. Diabetes 1991;40:Suppl-6.

17. Becker W, Alexander J, Andersen S et al. Nordic nutrition recommendations 2004; Integrating nutrition and physical activity. Copenhagen: Nordic Council of Ministers, 2004.

18. Eichholzer M, Tonz O, Zimmermann R. Folic acid: a public-health challenge. Lancet 2006;367:1352-61.

19. Picciano MF. Pregnancy and lactation: physiological adjustments, nutritional requirements and the role of dietary supplements. J Nutr 2003;133:1997S-2002S.

20. Allen LH. Multiple micronutrients in pregnancy and lactation: an overview. Am J Clin Nutr 2005;81:1206S-12S.

21. Holick MF. Vitamin D: importance in the prevention of cancers, type 1 diabetes, heart disease, and osteoporosis. Am J Clin Nutr 2004;79:362-71.

22. Bodnar LM, Simhan HN, Powers RW, Frank MP, Cooperstein E, Roberts JM. High prevalence of vitamin D insufficiency in black and white pregnant women residing in the northern United States and their neonates. J Nutr 2007;137:447-52.

23. Zimmermann M, Delange F. Iodine supplementation of pregnant women in Europe: a review and recommendations. Eur J Clin Nutr 2004;58:979-84.

24. Delange F. Iodine deficiency in Europe and its consequences: an update. Eur J Nucl Med Mol Imaging 2002;29 Suppl 2:S404-S416.

25. Hess SY, Zimmermann MB, Torresani T, Burgi H, Hurrell RF. Monitoring the adequacy of salt iodization in Switzerland: a national study of school children and pregnant women. Eur J Clin Nutr 2001;55:162-6.

26. Morris CD, Jacobson SL, Anand R et al. Nutrient intake and hypertensive disorders of pregnancy: Evidence from a large prospective cohort. Am J Obstet Gynecol 2001;184:643-51.

27. Roberts JM, Balk JL, Bodnar LM, Belizan JM, Bergel E, Martinez A. Nutrient involvement in preeclampsia. J Nutr 2003;133:1684S-92S.

28. Prentice AM. Can maternal dietary supplements help in preventing infant malnutrition? Acta Paediatr Scand Suppl 1991;374:67-77.

29. Hytten FE, Leitch I. The Physiology of Human Pregnancy, 2nd edition. Oxford: Blackwell Publishing, 1971.

30. World Health Organization. Maternal anthropometry and pregnancy outcomes. A WHO Collaborative Study. Bull World Health Organ 1995;73 Suppl:1-98.

31. Deruelle P, Houfflin-Debarge V, Vaast P, Delville N, Helou N, Subtil D. [Maternal and fetal consequences of increased gestational weight gain in women of normal prepregnant weight]. Gynecol Obstet Fertil 2004;32:398-403.

32. Leiferman JA, Evenson KR. The effect of regular leisure physical activity on birth outcomes. Matern Child Health J 2003;7:59-64.

33. Dempsey JC, Sorensen TK, Williams MA et al. Prospective study of gestational diabetes mellitus risk in relation to maternal recreational physical activity before and during pregnancy. Am J Epidemiol 2004;159:663-70.

34. Sorensen TK, Williams MA, Lee IM, Dashow EE, Thompson ML, Luthy DA. Recreational physical activity during pregnancy and risk of preeclampsia. Hypertension 2003;41:1273-80.

35. Butler CL, Williams MA, Sorensen TK, Frederick IO, Leisenring WM. Relation between maternal recreational physical activity and plasma lipids in early pregnancy. Am J Epidemiol 2004;160:350-9.

36. Alderman BW, Zhao H, Holt VL, Watts DH, Beresford SA. Maternal physical activity in pregnancy and infant size for gestational age. Ann Epidemiol 1998;8:513-9.

37. Campbell MK, Mottola MF. Recreational exercise and occupational activity during pregnancy and birth weight: a case-control study. Am J Obstet Gynecol 2001;184:403-8.

38. Byers T. Food frequency dietary assessment: how bad is good enough? Am J Epidemiol 2001;154:1087-8.

39. Lacroix R, Eason E, Melzack R. Nausea and vomiting during pregnancy: A prospective study of its frequency, intensity, and patterns of change. Am J Obstet Gynecol 2000;182:931-7.

40. Margetts BM, Nelson M. Design Concepts in Nutritional Epidemiology. Oxford: Oxford University Press, 1997.

41. Erkkola M, Karppinen M, Javanainen J, Rasanen L, Knip M, Virtanen SM. Validity and reproducibility of a food frequency questionnaire for pregnant Finnish women. Am J Epidemiol 2001;154:466-76.

42. Robinson S, Godfrey K, Osmond C, Cox V, Barker D. Evaluation of a food frequency questionnaire used to assess nutrient intakes in pregnant women. Eur J Clin Nutr 1996;50:302-8.

43. Brown JE, Buzzard IM, Jacobs DR, Jr. et al. A food frequency questionnaire can detect pregnancy-related changes in diet. J Am Diet Assoc 1996;96:262-6.

44. Greeley S, Storbakken L, Magel R. Use of a modified food frequency questionnaire during pregnancy. J Am Coll Nutr 1992;11:728-34.

45. Willett WC. Food Frequency Methods. In: Willett W, ed. Nutrition Epidemiology. New York, Oxford: Oxford University Press 1998:74-100.

46. Subar AF. Developing dietary assessment tools. J Am Diet Assoc 2004;104:769-70.

47. Drewnowski A. Diet image: a new perspective on the food-frequency questionnaire. Nutr Rev 2001;59:370-2.

48. Willett WC, Lenart E. Reproducibility and validity of Food-Frequency Questionnaires. In: Willett W, ed. Nutritional Epidemiology. New York, Oxford: Oxford University Press 1998:101-47.

49. Tucker KL. Assessment of usual dietary intake in population studies of gene-Diet interaction. Nutr Metab Cardiovasc Dis 2006.

50. Vollset SE, Lande B. Knowledge and attitudes of folate, and use of dietary supplements among women of reproductive age in Norway 1998. Acta Obstet Gynecol Scand 2000;79:513-9.

51. Maats FH, Crowther CA. Patterns of vitamin, mineral and herbal supplement use prior to and during pregnancy. Aust N Z J Obstet Gynaecol 2002;42:494-6.

52. Black AE, Bingham SA, Johansson G, Coward WA. Validation of dietary intakes of protein and energy against 24 hour urinary N and DLW energy expenditure in middle-aged women, retired men and post-obese subjects: comparisons with validation against presumed energy requirements. Eur J Clin Nutr 1997;51:405-13.

53. Bates JC, Thurnham DI, Bingham SA, Margetts BM, Nelson M. Biochemical markers of nutrient intake. In: Margetts BM, Nelson M, eds. Design Concepts in Nutritional Epidemiology. Oxford: Oxford University Press 1997:170-240.

54. Bingham SA. Biomarkers in nutritional epidemiology. Public Health Nutr 2002;5:821-7.

55. Hunter D. Biochemical indicators of dietary intake. In: Willett WC, ed. Nutritional Epidemiology. New York, Oxford: Oxford University Press 1998:174-243.

56. van Kappel AL, Steghens JP, Zeleniuch-Jacquotte A, Chajes V, Toniolo P, Riboli E. Serum carotenoids as biomarkers of fruit and vegetable consumption in the New York Women's Health Study. Public Health Nutr 2001;4:829-35.

57. Nielsen SE, Freese R, Kleemola P, Mutanen M. Flavonoids in human urine as biomarkers for intake of fruits and vegetables. Cancer Epidemiol Biomarkers Prev 2002;11:459-66.

58. Krogholm KS, Haraldsdottir J, Knuthsen P, Rasmussen SE. Urinary total flavonoid excretion but not 4-pyridoxic Acid or potassium can be used as a biomarker for the intake of fruits and vegetables. J Nutr 2004;134:445-51.

59. Wolk A, Vessby B, Ljung H, Barrefors P. Evaluation of a biological marker of dairy fat intake. Am J Clin Nutr 1998;68:291-5.

60. Wolk A, Furuheim M, Vessby B. Fatty acid composition of adipose tissue and serum lipids are valid biological markers of dairy fat intake in men. J Nutr 2001;131:828-33.

61. Biong AS, Veierød MB, Ringstad J, Thelle DS, Pedersen JI. Intake of milk fat, reflected in adipose tissue fatty acids and risk of myocardial infarction: a case-control study. Eur J Clin Nutr 2006;60:236-44.

62. Andersen LF, Solvoll K, Drevon CA. Very-long-chain n-3 fatty acids as biomarkers for intake of fish and n-3 fatty acid concentrates. Am J Clin Nutr 1996;64:305-11.

63. Olsen SF, Hansen HS, Sandstrøm B, Jensen B. Erythrocyte levels compared with reported dietary intake of marine n-3 fatty acids in pregnant women. Br J Nutr 1995;73:387-95.

64. Fuhrman BJ, Barba M, Krogh V et al. Erythrocyte membrane phospholipid composition as a biomarker of dietary fat. Ann Nutr Metab 2006;50:95-102.

65. Masson LF, McNeill G, Tomany JO et al. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: use of correlation coefficients and the kappa statistic. Public Health Nutr 2003;6:313-21.

66. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.

67. Baer HJ, Blum RE, Rockett HR et al. Use of a food frequency questionnaire in American Indian and Caucasian pregnant women: a validation study. BMC Public Health 2005;5:135.

68. National Cancer Institute. Risk Factor Monitoring and Methods: Validation Studies in Pregnant Populations [online]. Available at: http://riskfactor.cancer.gov/tools/children/review/agegroups /pregnancy/ (Accessed August 2006).

69. Fawzi WW, Rifas-Shiman SL, Rich-Edwards JW, Willett WC, Gillman MW. Calibration of a semi-quantitative food frequency questionnaire in early pregnancy. Ann Epidemiol 2004;14:754-62.

70. Mouratidou T, Ford F, Fraser RB. Validation of a food-frequency questionnaire for use in pregnancy. Public Health Nutr 2006;9:515-22.

71. Mikkelsen TB, Osler M, Olsen SF. Validity of protein, retinol, folic acid and n-3 fatty acid intakes estimated from the food-frequency questionnaire used in the Danish National Birth Cohort. Public Health Nutr 2006;9:771-8.

72. De Vriese SR, De Henauw S, De Backer G, Dhont M, Christophe AB. Estimation of dietary fat intake of Belgian pregnant women. Comparison of two methods. Ann Nutr Metab 2001;45:273-8.

73. Parra MS, Schnaas L, Meydani M, Perroni E, Martinez S, Romieu I. Erythrocyte cell membrane phospholipid levels compared against reported dietary intakes of polyunsaturated fatty acids in pregnant Mexican women. Public Health Nutr 2002;5:931-7.

74. Siega-Riz AM, Savitz DA, Zeisel SH, Thorp JM, Herring A. Second trimester folate status and preterm birth. Am J Obstet Gynecol 2004;191:1851-7.

75. Forsum E, Kabir N, Sadurskis A, Westerterp K. Total energy expenditure of healthy Swedish women during pregnancy and lactation. Am J Clin Nutr 1992;56:334-42.

76. Goldberg GR, Prentice AM, Coward WA et al. Longitudinal assessment of energy expenditure in pregnancy by the doubly labeled water method. Am J Clin Nutr 1993;57:494-505.

77. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. Acta Obstet Gynecol Scand 2000;79:435-9.

78. Norwegian Institute of Public Health. MoBa Food Frequency Questionnaire [online]. Available online at: http://www.fhi.no/morogbarn

79. Blaker B, Aarsland M. Mål og vekt for matvarer [Household measures and weights of foods]. Oslo: Landsforeningen for kosthold og helse [National Association for Nutrition and Health], 1989.

80. National Association for Nutrition and Health. Mat på data. [A nutrient intake calculation program]. Oslo: National Association for Nutrition and Health [Landsforeningen for kosthold og helse], 2004.

81. Lauritsen J. FoodCalc [online]. Available at: http://www.ibt.ku.dk/jesper /foodcalc (Accessed February 2004).

82. Rimestad AH, Borgejordet Å, Vesterhus KN et al. Den store matvaretabellen [The Norwegian Food Table]. Oslo: Statens råd for ernæring og fysisk aktivitet, Statens næringsmiddeltilsyn, Institutt for ernæringsforskning, 2001.

83. Hustvedt BE, Christophersen A, Johnsen LR et al. Description and validation of the ActiReg(R): a novel instrument to measure physical activity and energy expenditure. Br J Nutr 2004;92:1001-8.

84. World Health Organisation. Energy and protein requirements. Report of a joint FAO/WHO/UNU expert consultation. Geneva: WHO Technical Report No. 724, 1985.

85. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna 2005. Available online at: http://www.r-project.org/.

86. Stram DO, Longnecker MP, Shames L et al. Cost-efficient design of a diet validation study. Am J Epidemiol 1995;142:353-62.

87. Delcourt C, Cubeau J, Balkau B, Papoz L. Limitations of the correlation coefficient in the validation of diet assessment methods. Epidemiology 1994;5:518-24.

88. Trygg K, Lund-Larsen K, Sandstad B, Hoffman HJ, Jacobsen G, Bakketeig LS. Do pregnant smokers eat differently from pregnant non-smokers? Paediatr Perinat Epidemiol 1995;9:307-19.

89. Pollard J, Greenwood D, Kirk S, Cade J. Lifestyle factors affecting fruit and vegetable consumption in the UK Women's Cohort Study. Appetite 2001;37:71-9.

90. Nelson M, Bingham SA. Assessment of food consumption and nutrient intake. In: Margetts BM, Nelson M, eds. Design Concepts in Nutritional Epidemiology. Oxford: Oxford University Press 1997:123-69.

91. Andersen LF, Solvoll K, Johansson LR, Salminen I, Aro A, Drevon CA. Evaluation of a food frequency questionnaire with weighed records, fatty acids, and alpha-tocopherol in adipose tissue and serum. Am J Epidemiol 1999;150:75-87.

92. Andersen LF, Lande B, Trygg K, Hay G. Validation of a semi-quantitative food-frequency questionnaire used among 2-year-old Norwegian children. Public Health Nutr 2004;7:757-64.

93. de Vries JH, Zock PL, Mensink RP, Katan MB. Underestimation of energy intake by 3-d records compared with energy intake to maintain body weight in 269 nonobese adults. Am J Clin Nutr 1994;60:855-60.

94. Livingstone MB, Black AE. Markers of the validity of reported energy intake. J Nutr 2003;133 Suppl 3:895S-920S.

95. Maurer J, Taren DL, Teixeira PJ et al. The psychosocial and behavioral characteristics related to energy misreporting. Nutr Rev 2006;64:53-66.

96. Westerterp KR, Goris AH. Validity of the assessment of dietary intake: problems of misreporting. Curr Opin Clin Nutr Metab Care 2002;5:489-93.

97. Pollard J, Wild CP, White KL, Greenwood DC, Cade JE, Kirk SF. Comparison of plasma biomarkers with dietary assessment methods for fruit and vegetable intake. Eur J Clin Nutr 2003;57:988-98.

98. Bingham SA. Urine nitrogen as a biomarker for the validation of dietary protein intake. J Nutr 2003;133 Suppl 3:921S-4S.

99. Nielsen SE, Freese R, Cornett C, Dragsted LO. Identification and quantification of flavonoids in human urine samples by column-switching liquid chromatography coupled to atmospheric pressure chemical ionization mass spectrometry. Anal Chem 2000;72:1503-9.

100. Ahn E, Kapur B, Koren G. Study on circadian variation in folate pharmacokinetics. Can J Clin Pharmacol 2005;12:e4-e9.

101. Romon M, Le Fur C, Lebel P, Edme JL, Fruchart JC, Dallongeville J. Circadian variation of postprandial lipemia. Am J Clin Nutr 1997;65:934-40.

102. McNaughton SA, Marks GC, Gaffney P, Williams G, Green A. Validation of a food-frequency questionnaire assessment of carotenoid and vitamin E intake using weighed food records and plasma biomarkers: The method of triads model. Eur J Clin Nutr 2005;59:211-8.

103. Dixon LB, Subar AF, Wideroff L, Thompson FE, Kahle LL, Potischman N. Carotenoid and tocopherol estimates from the NCI diet history questionnaire are valid compared with multiple recalls and serum biomarkers. J Nutr 2006;136:3054-61.

104. Al Delaimy WK, van Kappel AL, Ferrari P et al. Plasma levels of six carotenoids in nine European countries: report from the European Prospective Investigation into Cancer and Nutrition (EPIC). Public Health Nutr 2004;7:713-22.

105. Burri BJ, Neidlinger TR, Clifford AJ. Serum carotenoid depletion follows first-order kinetics in healthy adult women fed naturally low carotenoid diets. J Nutr 2001;131:2096-100.

106. Erlund I, Kosonen T, Alfthan G et al. Pharmacokinetics of quercetin from quercetin aglycone and rutin in healthy volunteers. Eur J Clin Pharmacol 2000;56:545-53.

107. Willett WC. Overview of Nutritional Epidemiology. In: Willett W, ed. Nutritional Epidemiology. New York, Oxford: Oxford University Press 1998:1-17.

108. Willett W. Invited commentary: a further look at dietary questionnaire validation. Am J Epidemiol 2001;154:1100-2.

109. Hankin JH, Wilkens LR, Kolonel LN, Yoshizawa CN. Validation of a quantitative diet history method in Hawaii. Am J Epidemiol 1991;133:616-28.

110. Ocke MC, Kaaks RJ. Biochemical markers as additional measurements in dietary validity studies: application of the method of triads with examples from the European Prospective Investigation into Cancer and Nutrition. Am J Clin Nutr 1997;65:Suppl-1245S.

111. Kabagambe EK, Baylin A, Allan DA, Siles X, Spiegelman D, Campos H. Application of the method of triads to evaluate the performance of food frequency questionnaires and biomarkers as indicators of long-term dietary intake. Am J Epidemiol 2001;154:1126-35.

112. Andersen LF, Veierød MB, Johansson L, Sakhi A, Solvoll K, Drevon CA. Evaluation of three dietary assessment methods and serum biomarkers as measures of fruit and vegetable intake, using the method of triads. Br J Nutr 2005;93:519-27.

113. Block G. Invited commentary: another perspective on food frequency questionnaires. Am J Epidemiol 2001;154:1103-4.

114. Kristal AR, Peters U, Potter JD. Is it time to abandon the food frequency questionnaire? Cancer Epidemiol Biomarkers Prev 2005;14:2826-8.

115. Fraser GE. A search for truth in dietary epidemiology. Am J Clin Nutr 2003;78:521S-5S.

116. Kristal AR, Potter JD. Not the time to abandon the food frequency questionnaire: counterpoint. Cancer Epidemiol Biomarkers Prev 2006;15:1759-60.

117. Willett WC, Hu FB. Not the time to abandon the food frequency questionnaire: point. Cancer Epidemiol Biomarkers Prev 2006;15:1757-8.

118. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002;13:3-9.

119. Greenwood DC, Cade JE, Draper A, Barrett JH, Calvert C, Greenhalgh A. Seven unique food consumption patterns identified among women in the UK Women's Cohort Study. Eur J Clin Nutr 2000;54:314-20.

120. Kant AK. Dietary patterns and health outcomes. J Am Diet Assoc 2004;104:615-35.

121. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: a review. Nutr Rev 2004;62:177-203.

122. Shi Z, Hu X, Yuan B, Pan X, Dai Y, Holmboe-Ottesen G. Association between dietary patterns and anaemia in adults from Jiangsu Province in Eastern China. Br J Nutr 2006;96:906-12.

123. Heitmann BL, Lissner L, Osler M. Do we eat less fat, or just report so? Int J Obes Relat Metab Disord 2000;24:435-42.

124. Olafsdottir AS, Thorsdottir I, Gunnarsdottir I, Thorgeirsdottir H, Steingrimsdottir L. Comparison of women's diet assessed by FFQs and 24-hour recalls with and without underreporters: associations with biomarkers. Ann Nutr Metab 2006;50:450-60.

125. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr 1997;65:Suppl-1228S.

126. Kubena KS. Accuracy in dietary assessment: on the road to good science. J Am Diet Assoc 2000;100:775-6.

127. Goldberg GR, Black AE, Jebb SA et al. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr 1991;45:569-81.

128. Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. Int J Obes Relat Metab Disord 2000;24:1119-30.

129. Black AE, Cole TJ. Biased over- or under-reporting is characteristic of individuals whether over time or by different assessment methods. J Am Diet Assoc 2001;101:70-80.

130. Johansson L, Solvoll K, Bjorneboe GE, Drevon CA. Under- and overreporting of energy intake related to weight status and lifestyle in a nationwide sample. Am J Clin Nutr 1998;68:266-74.

131. Schulze MB, Manson JE, Ludwig DS et al. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. JAMA 2004;292:927-34.

132. Davey GK, Spencer EA, Appleby PN, Allen NE, Knox KH, Key TJ. EPIC-Oxford: lifestyle characteristics and nutrient intakes in a cohort of 33 883 meat-eaters and 31 546 non meat-eaters in the UK. Public Health Nutr 2003;6:259-69.

133. Hjartaker A, Lund E. Relationship between dietary habits, age, lifestyle, and socio-economic status among adult Norwegian women. The Norwegian Women and Cancer Study. Eur J Clin Nutr 1998;52:565-72.

134. Day NE, Wong MY, Bingham S et al. Correlated measurement error--implications for nutritional epidemiology. Int J Epidemiol 2004;33:1373-81.

135. Wong MY, Day NE, Bashir SA, Duffy SW. Measurement error in epidemiology: the design of validation studies I: univariate situation. Stat Med 1999;18:2815-29.