



Nutritional epidemiology

Dietary assessments: use, design concepts, biological markers, pitfalls and validationn

Gunnar Johansson

Nutritional epidemiology.

Dietary assessments: use, design concepts, biological markers, pitfalls and validation

© Gunnar Johansson

ISBN 978-91-87045-07-3

E-published 2014

Halmstad University Press, 2014 | www.hh.se/hup

Contents

- 1 Who needs information on individuals and groups food and nutrient intake?5
 - 1.1 Education, research and health care..... 5
 - 1.2 The food industry..... 5
 - 1.3 Authorities and politicians..... 5
- 2 Dietary data on a national level. Per capita consumption.....7
- 3 Dietary data on the household level. Household based surveys 8
- 4 Food consumption of individuals. Dietary assessment methods.....9
 - 4.1 Prospective methods 9
 - 4.1.1 Menu records 9
 - 4.1.2 Estimated food records 9
 - 4.1.3 Weighed food records..... 9
 - 4.1.4 The duplicate portion technique 10
 - 4.1.5 Advantages with prospective methods..... 10
 - 4.1.6 Disadvantages with prospective methods 10
 - 4.2 Retrospective methods 10
 - 4.2.1 24 hour and 48 hour recalls 10
 - 4.2.2 Diet histories 11
 - 4.2.3 Food frequency questionnaires..... 12
 - 4.2.4 Assessment of intake in the distant past..... 12
- 5 Calculation of nutrient intake from data on food intake 13
- 6 Level of output data..... 14
 - 6.1 Level 1. Mean consumption of a group..... 14
 - 6.2 Level 2. Mean and distribution of consumption in a group 14
 - 6.3 Level 3. The relative magnitude of the consumption of an individual (rank order) 14
 - 6.4 Level 4. The absolute magnitude of the consumption of an individual..... 15
- 7 Number of days required to classify individuals' dietary intake..... 16
- 8 The validation of dietary assessment..... 17
 - 8.1 Relative validation 17
 - 8.1.1 Comparison of the mean and median values..... 17
 - 8.1.2 Regression and correlation 17
 - 8.1.3 Classification into fractiles 18
 - 8.1.4 Bland-Altman analysis..... 19
 - 8.2 Observation..... 20
 - 8.3 Biological markers for food intake 20
 - 8.3.1 Energy 22
 - 8.3.2 Protein 24
 - 8.3.3 Sodium 26
 - 8.3.4 Iodine, fluoride and chloride 26
 - 8.3.5 Selenium 26

8.3.6	Potassium	27
8.3.7	Verification of the completeness of 24 hour urine collections	27
8.3.8	Fibre.....	29
8.3.9	Cadmium and lead.....	30
8.3.10	Vitamin C.....	30
8.3.11	Fat and carbohydrates	31
8.3.12	Fatty acids	31
8.4	New biological markers for food intake	32
8.4.1	Sucrose and fructose.....	32
8.4.2	Thiamine	32
8.4.3	Whole grain intake of wheat and rye.....	32
8.4.4	Fruits and vegetables	32
9	Misreporting.....	33
10	Energy adjustments.....	37
11	Statistical corrections for the effect of measurement error.....	38
12	Estimating a sample size	39
13	Future research directions.....	40
14	Conclusions and practical recommendations.....	41

1 Who needs information on individuals and groups food and nutrient intake?

A number of institutions need information on the dietary intake of individuals and groups (1, 2). In order to know what should be changed and to know how to alter the attitudes toward food choices various medicine and behavioural science research institutions need knowledge of what people eat. The health care system as well as the food industry needs knowledge on how they could contribute towards better public health.

1.1 Education, research and health care

In education as well as in food information, data from dietary surveys is often basic information. For instance, it is important that food and nutrition books are based on accurate information on what people eat. To be able to give dietary counselling to various patient groups, such as to patients with diabetes and coronary heart disease, you need to know what these patients actually eat. In large-scale nutritional epidemiological studies it is important that accurate dietary data are used to investigate the relationship between diet and health/disease.

Sociologists, anthropologists, psychologists and economics are interested in studying individuals or people's dietary habits and factors that affect these habits. For instance, they could be interested in the content of the meal, where it is consumed and in what social context it is consumed. It can also be interesting to study the food choice people make in the grocery store or what meals they choose at restaurants. Further, it could be interesting to study in what way price, taste, cookery skills, nutrition knowledge etcetera affect the choices.

1.2 The food industry

The food industry needs knowledge about people's attitudes toward various foods, their preference of taste etcetera. The needs for the food industry are often restricted, for instance to specific products or even to a specific brand name.

1.3 Authorities and politicians

Health authorities are interested in dietary surveys, especially nutritional epidemiological studies regarding relationships between diet and disease/health. It is important that they know about connections between, for instance coronary heart diseases or various cancer forms. Health authorities have the responsibility for the entire population's health and therefore should have continuous information on how the consumption of various foods changes over time. In hospitals, dieticians have the responsibility to give correct dietary counselling to various patient groups.

Authorities that have the responsibility for the food to the elderly and children need information on consumption and the nutritional status of individuals in order to be able to suggest improvements and to aid in decision making.

Agricultural authorities need information on the population's dietary habits and changes over time in order to make political decisions and plan for production. Decision makers in the health care system have a responsibility for political decisions and need continuous information on population's dietary habits. Politicians in decision making positions for a number of nutritional economical areas, such as subsidising and deciding taxes on foods, can affect the consumption, which in turn may cause nutritional consequences.

Briefly dietary information on the national level and the household level will be given, but the focus will be on food consumption of individuals, i.e. dietary surveys.

2 Dietary data on a national level. Per capita consumption

Per capita consumption is the available amount of food/energy/nutrients per person and day. This data is obtained by taking the sum of food production plus imports minus the sum of exports and foods for animals. Losses due to storage, transportation, distribution etcetera should be drawn of the total amount, but is difficult to measure. These data can be obtained from national statistic bureaus, OECD Food Consumption Statistics and FAO Food Balance Sheets. This type of data can be used in ecological studies, to follow trends in food production and give a broad view for comparison between countries. One should bear in mind that the quality between countries can vary, which of course make comparisons difficult. Also, it is important to realise that it is foods available for consumption and not the amount that is really consumed.

3 Dietary data on the household level. Household based surveys

Household based surveys give information of the household's expenses for foods. This can later be transformed to foods. Usually a household member keeps record on all expenses and type of foods during a specific time period, usually one to four weeks and preferably evenly distributed during the year. This type of data can be used to monitor differences due to socioeconomic status, geographical area, type of family etcetera. Household surveys do not give information on the distribution of the consumption between the family members, cooking methods or losses. These surveys are often performed of economical reasons rather than nutritional reasons.

4 Food consumption of individuals. Dietary assessment methods

Methods used for measuring food consumption of individuals can be classified in several ways. One way is to classify them into prospective and retrospective methods. Another way of classification is into current diet and food/dietary habits. With current diet you mean the diet that you actually eat or have eaten. With food habits, the usual diet you mean the diet that you usually eat, but not necessarily the diet that you eat right now, for instance today. The current diet methods are the various record methods, the duplicate portion technique and 24/48 hour recall methods. The dietary habits methods are the various diet history methods and the food frequency questionnaires.

4.1 Prospective methods

In the prospective methods the diet are recorded when you consume the foods. The various methods differ in how precise you are in estimating the portion sizes. In general one records food for 3 to 7 days.

4.1.1 Menu records

A menu recording is a record of the respondent's food intake at the time of consumption regarding type of food, but does not indicate any portion sizes.

4.1.2 Estimated food records

In an estimated food record, respondents record the foods eaten at the time of consumption and estimate the portion sizes, for instance with the aid of pictures of foods, rulers, standard household measuring cups and spoons. Detailed descriptions of all foods and beverages, including brand names, and their method of preparation and cooking are recorded. For composite dishes the amount of each raw ingredient used in the recipe and, if possible, the final weight of the composite dish are recorded.

4.1.3 Weighed food records

In a weighed record all foods and beverages consumed are weighed by the subject, parent or caretaker at the time of consumption. Details of methods of food preparation, description of foods, and brand names should also be recorded. For composite dishes, weights of all raw ingredients used in the recipe should be noted, as well as the weight of the portion consumed. It can be both an advantage and disadvantage to weigh foods and beverages since you need a weighing scale. The advantage is that you do not need to estimate portion sizes and the portion sizes will be more accurate compared to the other methods. The primary disadvantage is that it is necessary to bring the scale with you all the time during the recording period.

4.1.4 The duplicate portion technique

In this method duplicate portions of all foods and beverages consumed are collected in a container and later analyzed.

4.1.5 Advantages with prospective methods

One of the major advantages with the prospective methods is that they are not affected by the memory since the foods and beverages are recorded at the time of consumption. Another advantage is that portion sizes can be more accurately estimated compared to retrospective methods. Concerning the duplicate portion technique it is an advantage that nutrients can be analyzed. An advantage, that is valid for all current methods, is that information regarding irregularities in food intake can be obtained.

4.1.6 Disadvantages with prospective methods

Probably the greatest disadvantage of prospective methods is that they often affect the consumption of foods and beverages. Several studies have demonstrated that the total intake will be low and that there is a selective underreporting (3). Another disadvantage is that you can only perform this method during a short time period, normally not more than a week. Since records and the duplicate portion technique are demanding methods for the subjects the drop-out rate may be high and one will get a selection of people who are probably more health and food conscious. A great disadvantage with the duplicate portion technique is of course that one must bring the container everywhere. This restricts the number of days of collection, as well as limiting the selection of people to those willing to put up with this inconvenience. Besides these disadvantages for the volunteers, the duplicate portion technique is a very expensive method and more expensive the more analyses one perform.

Prospective methods are not affected by the memory, except for the duplicate portion technique since subjects already at the grocery store have to remember to buy the double amount of foods and beverages that they normally need. Then they have to prepare the meal for double portions so that you can throw one portion into the container. Validation studies on the duplicate portion technique indicate an underestimation of the total food intake (4).

4.2 Retrospective methods

In retrospective methods information is collected on foods and beverages already consumed. Among the retrospective methods are the 24 and 48 hour recalls current diet methods and diet histories and food frequency questionnaires dietary habits methods.

Quantities of foods consumed can be obtained by food models, pictures of foods, rulers, standard household measuring cups and spoons.

4.2.1 24 hour and 48 hour recalls

In the 24 hour or 48 hour recall methods subjects are asked to recall the food intake during the previous 24 hour or 48 hour period or the preceding day(s) (5). Detailed descriptions of all foods and beverages consumed, including, if possible, brand names and cooking methods, are recorded by the interviewer. Mineral and vitamin supplement use is also noted. To cover a longer time period you can

repeat the interviews, and thereby get a better estimate of the individual's intake, so called repeated 24/48 hour recalls. The recalls can also be performed as telephone interviews.

The advantages of these methods are that they are simple, cheap and a large sample size is possible to obtain. In addition, as with all retrospective methods, the food habits are not affected by the methods.

The disadvantages with these methods are that the interviewer can affect the result and if the recalls are not repeated, the time period will be too short to be able to obtain information on the individual level (see table 1 and 2 and chapter 7). People have good as well as bad memory and the ability to estimate quantities vary. These factors affect the quality of the methods.

4.2.2 Diet histories

With this method, or rather methods, one will obtain an estimate of the individual's dietary habits during a long time period, from weeks up to one year. However, it has been demonstrated that describing food habits during a year is almost impossible, due to seasonal variations. Therefore, it is important to define the time period covered.

The method exists in many forms, which has led to that researchers often talk about the diet history family. They have in common that they describe habitual dietary habits. The diet history is considered to be a difficult method to perform. The method is very demanding for both the volunteers and the interviewers. For instance, the method is very time consuming (often several hours) and you need to have a good memory of what you have eaten and quantities consumed.

The method, first developed by Bertha Burke in 1947 (6) is made up of three parts. The first part consists of a collection of general information on the overall eating pattern. The general information obtained includes detailed descriptions of foods, their frequency of consumption, and usual portion sizes in common household measures. This first part also consists of a 24 hour recall.

The second part consists of a questionnaire on the frequency of consumption of specific food items, used to verify and clarify the information on the kinds and amounts of foods given as the usual intake in the first component. This second part serves as a "cross-check" for the information on usual intake obtained from the first part.

The third part consists of a three-day food record using household measures. Since this last part showed to be the least helpful it is often abandoned.

An advantage with these methods is that relatively long time periods can be studied and thereby the intake on the individual level can be obtained (see table 1 and chapter 7). Another advantage is that the interviewer can contribute to a good communication; thus, a low drop-out rate can be obtained. It is also likely that these methods give a more valid intake compared to other methods (7).

Disadvantages with the methods are that they are very labour intensive, which affects both the interviewer and the volunteer, and unsuitable for large surveys. Also, the results obtained depend on the skill of the interviewer, which impacts the quality of the interview. Researchers will not obtain information on the day-to-day variation in food intake and, as with all retrospective methods, good memory is important for estimating frequencies and portion sizes.

4.2.3 Food frequency questionnaires

These questionnaires are designed to obtain qualitative, descriptive information about usual food consumption pattern (8). The first component of the questionnaire is a list of foods, beverages, supplements, etcetera; and the second part is a set of frequency-of-use response categories. Sometimes there is a third part with portion sizes, especially for foods that can easily be quantified, such as number of eggs, apples, bananas, slices of bread, cups of coffee and tea etcetera.

These questionnaires are common when the number of participants is large, such as in epidemiological investigations studying associations between dietary habits and disease. The time period for the consumption should be defined and should be at least one month. If the survey is repeated it should be during the same time of the year.

The semi-quantitative food frequency questionnaire is a special type of food frequency questionnaire. It has portion sizes of the food items of interest, with or without the use of food models or photographs. With this procedure nutrient scores of each subject can be computed by multiplying the relative frequency that each food item is consumed by the nutrient content of the average portion size specified. A large number of nutritional epidemiological studies, for instance investigating the relationship between diet and cancer, are based on semi-quantitative food frequency questionnaires.

Advantages with food frequency questionnaires are that, compared with the other methods, they are not a heavy burden for the volunteers. You do not need skilled personnel, you will get a relatively low drop-out rate and avoid interviewer bias. In general, they are relatively cheap, can be filled out quickly and are not difficult to analyze.

The above mentioned advantages are valid when you already have a questionnaire. However, it is very time consuming to construct and validate this type of questionnaire (9). Other disadvantages are that it is almost impossible to obtain the total food intake, which sometimes is forgotten. In addition, skilled personnel are crucial when constructing a questionnaire, as well as producing nutrient scores and performing energy adjustments.

4.2.4 Assessment of intake in the distant past

These data are derived from some form of food frequency questionnaire or diet history interview. The methods are primarily used in case-control studies when dietary information is needed from the distant past. Results from studies in which current as well as past diet were assessed reveal that the recalled diet agrees more closely with the current diet than with the past diet (10-12). This demonstrates that the current diet is having a strong influence on the recalled diet.

5 Calculation of nutrient intake from data on food intake

When dietary data is collected, often it should be converted into energy and nutrients. For this the researcher needs a food composition table/nutrient database. National food composition tables/nutrient databases are available in many countries. The obvious choice is of course to use the one from your own country. However, for international studies, or for those in which it is planned to compare the data with that obtained in other studies, the best alternative may be to use international tables/databases.

If the dietary survey has portion sizes in volumes, these must be converted into weights, which can be a time consuming labour. Depending on the quality of the programme various calculations may be performed, such as energy and nutrients per day, nutrients per 10 MJ, meal pattern, and food frequencies. The dietary data may also be transferred to a statistical programme for various statistical calculations and for instance comparison with other data, such as clinical findings.

6 Level of output data

The level of the output data decides what conclusions one can draw from a dietary survey concerning how well the description is on the individuals or groups dietary data. There are four levels, which are described below (13).

6.1 Level 1. Mean consumption of a group

If you only want to know the mean consumption of a group without the distribution of consumption in the population you can stop at level 1. There are few surveys that only require data on level 1 since you do not get much information. For instance, the information can be used when you want to investigate the effect of an intervention.

When you talk about a group it could be interesting to know how large this group should be for valid information on the mean value. This depends on which foods or nutrients you wish to obtain information on and how homogenous the group is. Usually a group of 50 individuals is enough to give a reasonably good mean value for the energy intake (2).

6.2 Level 2. Mean and distribution of consumption in a group

On level 2 you will obtain information on the mean value and the distribution within a population. If you need to determine the size of a so called “risk group” you should be on level 2. However, it is not possible to establish whether a particular individual belongs to the group or not.

In many surveys when you compare the intakes of two or more groups and in intervention studies you should preferably be on level 2. In an intervention study you would like to obtain information which makes it possible to interpret the results, for instance if there are the same mean values there can be differences in the distribution. This level is also used when you would like evaluate the consumption of a group with various recommendations. The distribution within the group can give you an idea on how many individuals are at risk with a critical low or high intake. You will get information on the size of the group but without being able to pick out the correct individuals.

6.3 Level 3. The relative magnitude of the consumption of an individual (rank order)

It is on level 3 that a person can be ranked in comparison with others. This usually means that the person is classified as belonging to a certain fractile, such as a quartile of the distribution (see table 1). There is often an interest in extreme groups, such as those with highest and lowest intake of a certain nutrient or food. Rank order data are often used in epidemiological studies where you study the relationship between diet and disease.

Table 1. Number of days necessary to correctly rank individuals into thirds and fifths, respectively, and not more than 1 % in the opposite third or fifth, respectively (14).

Variable	Number of days	
	Thirds	Fifths
Energy	7	14
Protein	7	17
Fat	7	15
Fat (E %)	17	27
Carbohydrate	3	6
Alcohol	14	30
Vitamin C	14	30
Calcium	5	12

6.4 Level 4. The absolute magnitude of the consumption of an individual

Finally we are on level 4, an estimate of the absolute level of consumption. However, it is debatable whether it is possible to obtain the true absolute value, in other words a value that is both representative of the individual and measured without a systematic error. Nevertheless, it is a dream that you should reach this level even if there are many pitfalls in dietary assessments. People not involved in dietary survey research often take for granted that you are on level 4, which can lead to serious mistakes. Even if it is debatable whether we can reach this level, it is clear that this level can be reached for certain days (see table 2).

Data on this level can be used for several purposes. It can be used for evaluating individual's intake in relation to a certain recommendation. It can also be used for relating dietary intake to health/disease. Thirdly, this level is used to evaluate the effects of an intervention.

Table 2. Number of days required for 95 % probability that the sample average is within ± 20 % of the true individual mean for 80 % of the individuals (14).

Variable	Number of days
Energy	7
Protein	8
Carbohydrate	8
Fat	11
Fat (E %)	5
Alcohol	305
Thiamine	27
Vitamin C	53
Iron	13
Calcium	29

7 Number of days required to classify individuals' dietary intake

There is a normal biological variation in food intake due to the fact that individuals do not eat the same foods every day, for instance during work and leisure time, seasonal variations, etcetera. There is also an artificial variation due to that we cannot measure exactly what we want to measure. It can be due to that the food tables show mean values instead of the real value for the food eaten, lack of motivation from the volunteer, poor memory, etcetera. The interviewer and the food models can also produce errors.

Thus, a variation in food intake exists that must be considered when characterising an individual's intake. When you would like to determine how many days are needed to describe an individual's dietary intake, there are certain questions one needs to answer first.

- What is the purpose of the study?
- What nutrients are interesting to study?
- What precision is desirable?
- How homogenous is the population?

See table 1 and 2 for information on how many days you need to be on level 3 and level 4, respectively. To be able to calculate the number of days to describe an individual's dietary intake, which is performed in table 1 and 2 there are equations. To rank an individual with a certain precision the equation below is used (15).

$$n = r^2 / (1 - r^2) * (SD_w^2 / SD_b^2)$$

n = number of days needed per person

r = rank correlation coefficient between the observed and true intakes

SD_w = The standard deviation for within-subject variation (day-to-day or intra-individual variation)

SD_b = The standard deviation for the between-subject variation (inter-individual)

To be able to calculate the number of days that represent a dietary intake on an individual level the following equation is used (16).

$$n = [z^2 \times (CV_w / 100)^2] / D^2$$

n = number of days needed per person

z = the normal deviate for the percentage of times the measured value should be within a specified limit (1.96 for 95 %)

CV_w = the within-person coefficient of variation

D = the specified limit (as a percentage of long-term true intake, often 20 %, i.e. 0.20)

8 The validation of dietary assessment

The validity of a dietary survey method may be defined as the degree to which the survey method can estimate the true dietary intake. We have valid dietary data when the person 1) has eaten as usual and 2) reports this. The validity of measurements of dietary intake in free living individuals is difficult to assess because all methods rely on information given by the subjects themselves, which may not be correct. In an attempt to determine objective measures of validating dietary assessments, the search has begun for objective measures using biological specimens that closely reflect food intake, but which do not rely on reports of food consumption. Negligence in performing a validation study may lead to false associations between diet and diseases/disease markers. Poor dietary data may be one reason for the current confusion in many areas of the diet and health field.

One major problem that may arise when a validation study is completed, is if the validated method is good enough to be used in the main study. In other words, when is the method valid (or good enough)? Unfortunately there are no clear answers to this question. There are four alternatives. 1. Use the method. 2. Abandon the method and find an alternative method to measure the dietary exposure. 3. Modify the method according to the results from the validation study. 4. Choose another method and validate the new method. In reality there are not enough economical or time resources for the fourth alternative.

8.1 Relative validation

In a relative validation you make a comparison with a reference measure, often a more extensive method. Then you measure the difference between the test and reference method according to the methods described below. For a more detailed explanation of the statistical methods used in relative validation techniques the reader is referred to statistical books and reference 17.

8.1.1 Comparison of the mean and median values

The simplest comparison is a non-paired comparison, i.e. a comparison of the group mean value and the standard deviation or the median value and percentiles. You can also add a measurement within individuals, a so called paired comparison. Comparisons between test and reference methods of energy and nutrient intake can be examined by a Student's t-test (undertaking log-transformations where appropriate). For comparisons of foods, the distribution is less likely to be parametric, non-parametric tests are more likely to be appropriate.

8.1.2 Regression and correlation

The correlation coefficient r is a measure of the linear relationship between two variables. Correlation describes a relationship and regression describes both a relationship and predicts an outcome. You can compare two methods by performing a scatter plot and insert a regression line. A correlation coefficient can be obtained, which gives a value between -1 and +1 as a measure of the relationship between the methods, where 0 demonstrates no relationship and +1 a perfect relationship between the methods. The r value is positive when the slope of the regression line is directed upwards and negative

if it is directed downwards. The **Pearson** product-moment correlation coefficient is the “usual” correlation coefficient and is used for normally distributed material. The **Spearman** correlation coefficient is a non-parametric version.

A drawback with the correlation coefficient is that it describes only one aspect of agreement that relates to ranking. Poor agreement can exist between test and reference methods even when correlation coefficients are high. The slope of the regression line may be different from unity, which means that test and reference measurements are different. Thus, a high r-value does not necessarily mean that the methods measure the same thing. In figure 1, the coefficient of determination is high (0.99), but in general method 1 yields higher values. In an extreme case, there can be a horizontal or vertical line, which yields high r-values, even though the methods do not measure the same intake. Another drawback with the correlation coefficient is that if two methods measure roughly the same intake, but there is no variation in intake data, the r-value will be low.

Due to the above problems with the correlation coefficient some authors have suggested that a better measure of association is the **intra-class correlation** (17). This is because the intra-class correlation coefficient takes into account both the degree of correlation and the size of disagreement within pairs.

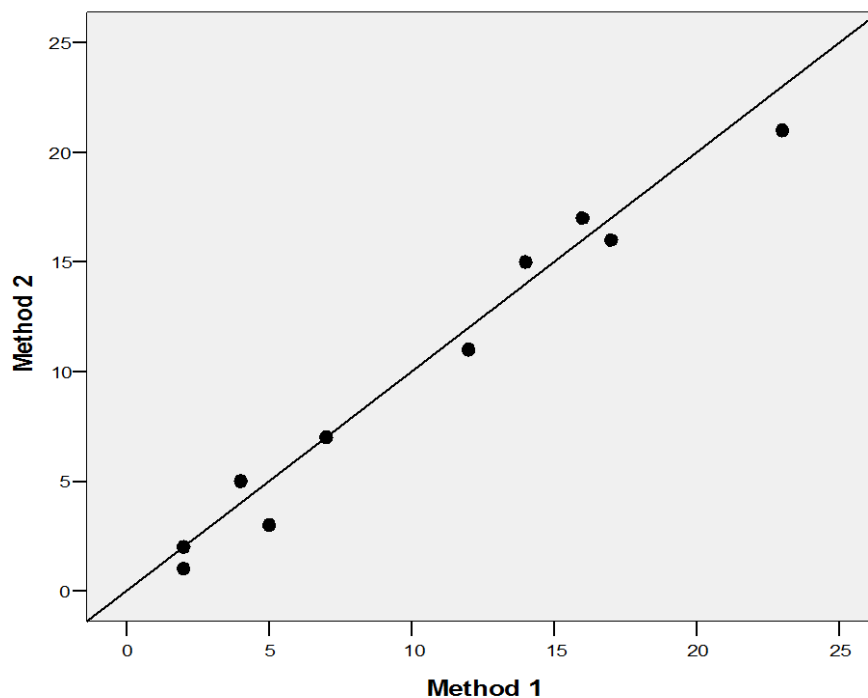


Figure 1. A theoretical comparison between two dietary assessment methods, method 1 and method 2, where the intakes are plotted and a regression line fitted.

8.1.3 Classification into fractiles

Classification into fractiles, usually tertiles, quartiles or quintiles is a relative validation method. You rank the intake from lowest to highest values and then group the data into fractiles and compare the percentage that will be in the same or opposite fractile. Tables 3 and 4 demonstrate the extent to which

each method is able to classify individuals into the same quartile of intake and to misclassify into opposite quartiles according to Bingham et al. (18).

Table 3. Comparison of the percentage classification of results from four different dietary assessment methods into the same quartile of the distribution of energy and nutrient intake as that obtained from the mean of 16 days weighed records obtained from 150 women aged 50-65 years.

	Diet history	24-hour recalls	FFQ1	FFQ2
Energy	82	61	47	58
Protein	52	58	64	49
Calcium	47	41	49	46
Fibre	39	58	51	41

FFQ is Food Frequency Questionnaire

Table 4. Comparison of the percentage misclassification of results from four different dietary assessment methods into extreme quartiles of the distribution of energy and nutrient intake, compared with that obtained from the mean of 16 days weighed records obtained from 150 women aged 50-65 years.

	Diet history	24-hour recalls	FFQ1	FFQ2
Energy	2	6	4	8
Protein	0	1	6	9
Calcium	7	4	9	6
Fibre	3	8	1	4

FFQ is Food Frequency Questionnaire

8.1.4 Bland-Altman analysis

A method for assessing the extent of differences that has been rather popular lately is the approach suggested by Bland and Altman (19). You plot the difference against the sum of each pair of observations. This makes no assumption about which of the methods yields the better measure and assesses only the level of agreement. Figure 2 demonstrates that the variation around the mean difference is wide and when there are low mean values the reported energy expenditure (EE) is high and with high mean values the reported EE is low. This is in comparison with measured EE. The figure demonstrates that the method to measure EE at low physical activity is overestimated and the physical activity at high EE is underestimated (flat slope syndrome). You can in the figure see both the difference between the methods and a trend from low to high mean values. In this type of figure it is left open to interpretation by the investigator on which is the best method. In this case we know that the doubly labelled water technique is regarded as the “golden standard” for measuring EE (see chapter 8.3.1).

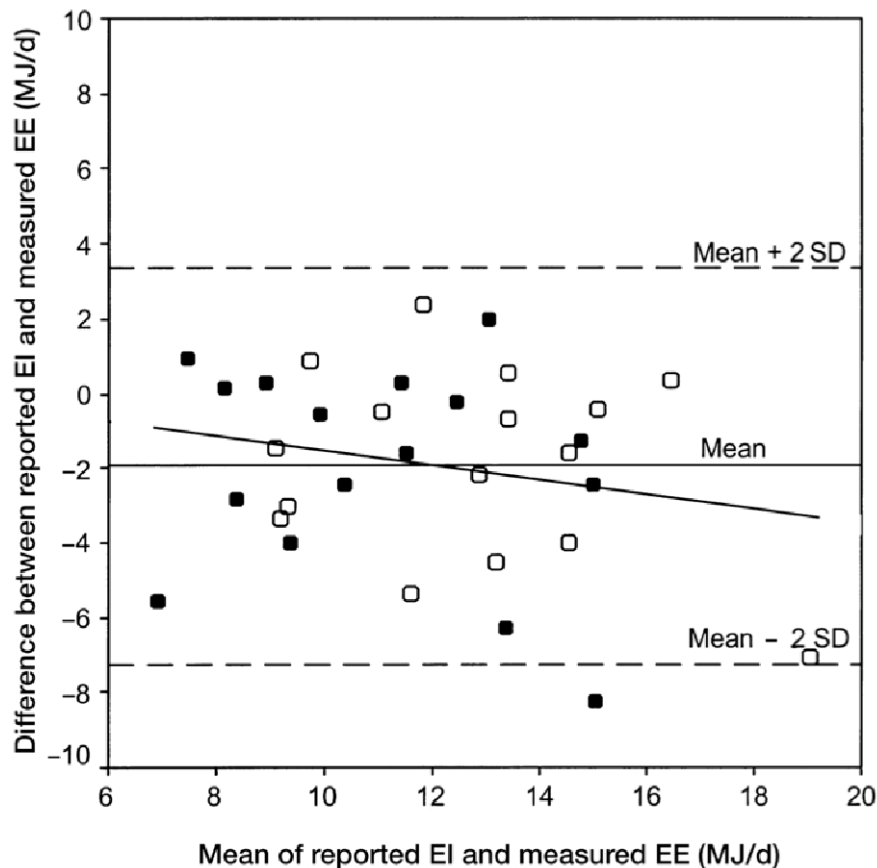


Figure 2. Difference between reported (during diet-history interview) energy intake (EI) and measured (by the doubly labeled water method) energy expenditure (EE) plotted against the mean of reported EI and measured EE in 16 omnivores (□) and 16 vegans (▪). The solid horizontal line represents all subjects ($r^2 = 0.0450$). Negative values for the difference indicate that reported EI was lower than measured EE. $r = -0.212$ (NS). Linear regression equation: $y = 0.396 - 0.196x$.

(Larsson CL, Westerterp KR, Johansson GK. Validity of reported energy expenditure and energy and protein intakes in Swedish adolescent vegans and omnivores. *Am J Clin Nutr.* 2002 1;75(2):268–74.)

8.2 Observation

Observation means that you observe what people eat. However, it is not often that you are able to see what people eat. It is at certain occasions, such as patients in a hospital, children at a day care centre, people in prison, etcetera. Even at these occasions it is difficult to see what people eat all day. It is primarily at meals that this is possible. To conclude, this type of validation is in most cases not very feasible.

8.3 Biological markers for food intake

A biological marker for food intake is a marker in any “specimen” that gives a predictive response to a given dietary component. “Specimen” in this context can be blood, urine, faeces, hair, nails etcetera.

Biological markers are used...

- As a replacement for a dietary assessment.
- For validating a dietary assessment.

Requirements for an ideal biological marker are:

- Should be possible to measure in urine, blood etc.
- The relation intake to marker should be the same over the entire intake range (not measure nutritional status).
- The relation should not change due to diseases, age etc.
- Reflect the relevant time period.
- Demonstrate no inter-individual difference.
- Demonstrate the same answer regardless of food source.
- Demonstrate the same answer regardless of cooking method.

However, all markers are not perfect. There are limitations and some markers are only valid in certain conditions. Pros (+) and cons (-) with biological markers are:

- + **Reflects a "true" dietary intake.** In best case they indicate a true dietary intake, during the condition that the sampling and analyses are correctly performed. You can at least say that the markers show an intake that is independent of information from the volunteer.
- +/- **Reflects a relevant time period.** This depends on what time period you are interested in. The energy validation with the doubly-labelled water (DLW) method reflects the latest two to three weeks. The protein validation reflects roughly the last week. The fibre validation reflects the last few days. Vitamin C in plasma reflects approximately the last month, in whole blood the last three months and in leucocytes the last four months.
- +/- **Accuracy.** It depends on the relevant time period.
- **Markers are not always available.** There are not markers for all nutrients and food components. Sometimes there may be a risk that you measure the nutritional status instead of the dietary intake.
- **There are no markers for foods.** However, in a near future there may be markers for foods, such as fruits and vegetables (see chapter 8.4).
- +/- **Cost.** Some markers are very expensive. For instance the energy validation with the DLW method cost approximately SEK 7000 per person. However, there are some cheap markers such as sodium and potassium.
- **Bad compliance.** There is a risk for bad compliance since it causes some problems to collect several days of complete 24 hour urine samples, several days of complete faecal collections, etcetera. This makes it difficult to obtain a random sample of the population.
- + **Good compliance.** Some volunteers feel selected and important when they should give samples, such as blood samples. Therefore, compliance can be high when you use biological markers.

Biological markers that either reflect the intake in a good way, are frequently used or simple to use are presented in the next chapters. More biomarkers are presented in reference 21. However, they are often not as simple, frequently used or good as those presented below.

8.3.1 Energy

The energy intake is a surrogate measure of the total quantity of food intake and can therefore be seen as the most important validation, especially since an underestimation of the energy intake results most likely in an underestimation of many other nutrients that are related to the total food intake. For instance, this can lead to an overestimation of the number of people with a specific nutrient deficiency. Since the validation of the energy intake (EI) is based on the validation of the energy expenditure (EE), estimation of the EE should be included in dietary surveys (22). See also chapter 8.3.1.3. The Goldberg cut-off.

8.3.1.1 The doubly labelled water method

Validation of the energy intake with the doubly labelled water (DLW) method has been performed on people since 1986 (23-25). The validation is based on the assumption that people are in energy balance, i.e. the EI equals the EE. The volunteers are given a carefully weighed oral dose of $^2\text{H}_2^{18}\text{O}$ and are then required to donate timed urine samples over the next two to three weeks. Measurements of the disappearance of the two isotopes are calculated. The method is easy to perform for the subjects. However it is an expensive method. The method is described in detail in references 24 and 25.

In order to identify records of poor validity a confidence limit for valid reports has to be defined. The ratio EI:EE is expected to be 1.00 (or 100 %), but variations in EI and EE are inevitable. Therefore an equation for the 95 % confidence limit (95% CL) of agreement between EI and EE is set up according to the equation below:

$$95\% \text{ CL} = 2 \times \sqrt{[(\text{CV}_{\text{wEI}}^2/d) + \text{CV}_{\text{wEE}}^2 - 2r \times (\text{CV}_{\text{wEI}}/d) \times \text{CV}_{\text{wEE}}]} \quad (23)$$

There exists also another version of the equation:

$$95\% \text{ CL} = 2 \times \sqrt{[(\text{CV}_{\text{wEI}}^2/d) + \text{CV}_{\text{wEE}}^2]} \quad (22)$$

CV_{wEI} is the within-subject coefficient of variation for daily EI (23 %), d is the number of days of diet records, CV_{wEE} is the within-subject coefficient of variation for repeat DLW-EE (8,2 %) and r is the correlation between EI and EE (0,425). The figures in the equation and motivation for the selected figures are discussed in reference 23. As an example of the application of the equation, for a 7-day record $\pm 18\%$ is accepted as a valid record.

An example of validation with DLW (PAL_{mea} and EE_{mea} , respectively) of the EI measured with a diet history (FIL and EI_{rep} , respectively) and reported EE with a physical activity interview (PAL_{rep}) is given in tables 5 and 6. FIL and PAL can be replaced by EI and EE, respectively by multiplying FIL and PAL , respectively with the basal metabolic rate (BMR).

Table 5 Validation of reported energy expenditure and energy and protein intakes in Swedish vegans and omnivores by using the doubly labeled water method and 24-h urine collections in 1997–1998¹

	Females		Males	
	Vegans (n = 7)	Omnivores (n = 7)	Vegans (n = 9)	Omnivores (n = 9)
BMR (MJ) ^{2,3}	6.78 ± 0.63	6.11 ± 0.50	7.43 ± 0.44	7.68 ± 0.48
FIL ²	1.16 ± 0.344	1.69 ± 0.48	1.58 ± 0.24	1.71 ± 0.29
PAL _{reported} ²	1.48 ± 0.114	1.66 ± 0.09	1.61 ± 0.17	1.68 ± 0.16
PAL _{measured} ^{2,5}	1.41 ± 0.224	1.84 ± 0.44	1.87 ± 0.39	2.05 ± 0.33
FIL/PAL _{measured}	0.84 ± 0.25	0.92 ± 0.19	0.87 ± 0.19	0.85 ± 0.16
PAL _{reported} /PAL _{measured} ²	1.08 ± 0.21	0.93 ± 0.15	0.89 ± 0.17	0.83 ± 0.10
(N _{reported} × 0.81)/N _{measured} ^{2,6}	0.78 ± 0.154	1.02 ± 0.11	1.01 ± 0.22	0.99 ± 0.147

¹ $\bar{x} \pm$ SD. Analyses were performed with a two-factor ANOVA with diet and sex as the 2 factors. BMR, estimated basal metabolic rate; FIL, food intake level (reported energy intake/BMR); PAL, physical activity level (energy expenditure/BMR); N, nitrogen.

² Significant interaction effect of diet and sex, $P < 0.05$.

³ Significant main effect of diet, $P < 0.05$.

⁴ Significantly different from omnivores of the same sex, $P < 0.05$.

⁵ Significant main effect of sex, $P < 0.05$.

⁶ 81% of reported nitrogen intake divided by measured nitrogen excretion in urine.

⁷ $n = 8$.

(Larsson CL, Westerterp KR, Johansson GK. Validity of reported energy expenditure and energy and protein intakes in Swedish adolescent vegans and omnivores. *Am J Clin Nutr*. 2002 1;75(2):268–74.)

8.3.1.2 Other indirect methods

The DLW method is the most accurate method for measuring EE. Unfortunately, this method is very expensive and skilled personnel are required. Therefore other methods are used to estimate EE, such as minute-by-minute heart rate monitoring, accelerometers, pedometers, activity diaries, physical activity questionnaires (26-30). These methods have advantages and disadvantages, but are not described here. Reviews of these methods are given in references 26 and 27.

8.3.1.3 The Goldberg cut-off technique

The Goldberg cut-off for energy intake:basal metabolic rate was introduced 1991 (31). The basic principle for the equation is that you compare FIL (EI:BMR) with PAL (Physical Activity Level = EE:BMR) and add a confidence limit and taking into account variation in BMR, dietary intake and physical activity. See the equation below (32).

BMR can be measured with direct or indirect calorimetry, but also from equations based on age, sex, body weight (and sometimes height) (33). The Goldberg cut-off is of course not a biological marker. A PAL value, which is a surrogate measure for EE, is necessary for the equation.

The Goldberg cut-off for $FIL = PAL \times \exp [SD_{\min} \times (S/100)/\sqrt{n}]$

$$S = \sqrt{(CV_{wEI}^2)/d + CV_{wB}^2 + CV_{tP}^2}$$

PAL should be measured as accurately as possible. SD_{\min} is -2 for the lower 95 % confidence limit and +2 for the upper 95 % confidence limit. S is the factor that takes account of the variation in intake, BMR and energy requirements, where ...

- CV_{wEI} = The within-subject daily variation in energy intake (23 %).
- d = the number of days of diet assessment.
- CV_{wB} = The variation in basal metabolic rate. The CV of repeated BMR measurements or the precision of estimated compared with measured BMR (4 % for measured BMR and 8.5 % for estimated BMR).
- CV_{tP} = The between-subject variation in physical activity. The CV derived from the mean and sd of a study and includes true between-subject variation, an element of within-subject variation and methodological errors (15 %).

The figures in the equations are discussed in reference 32.

8.3.2 Protein

Nitrogen in urine has been suggested as a biological marker for protein intake already 1924 (34). "Laws of nutrition, relevant to protein" was set up for more than 100 years ago by Langworthy (1898) and are still mainly valid:

- All nitrogen is supplied by food, that is, none from atmosphere.
- All nitrogen is excreted in urine and faeces, none as gaseous nitrogen.
- The animal adjusts itself to its nitrogen intake and comes into N-balance, in which state the intake and output are equal.
- The body comes into nitrogen equilibrium at different levels of protein intake.
- As furnishers of energy the different nutrients may replace each other and it is theoretically and within certain limits, unimportant which nutrient supplies the necessary energy.

The basic assumption for protein validation is that the individual is in nitrogen balance and that there are complete 24 hour urine collections (35). Isaksson suggested 1980 an equation for protein validation (36):

$$\text{Protein}_U = 6,25 \times (\text{N}_u + 2)$$

Protein_U is the protein intake according to the validation method, 6.25 is the factor for converting protein from nitrogen and is based on a mixed diet. The factor refers to the nitrogen content of proteins. The factor 6.25 may sometimes be changed when, for instance an “extreme” diet is being validated, such as a vegan diet, etcetera. The factor 2 stands for extra renal nitrogen losses.

Five years later, 1985, Bingham and Cummings published a study where they carefully tested the protein validation method (35) and presented the equation below.

$$\text{N}_d/\text{N}_u = 0.81 \text{ or } \text{Protein}_u = 7,72 \times \text{N}_d$$

N_d stands for nitrogen in the diet. The equation demonstrates that 81 % of the nitrogen intake is excreted in the urine. The difference between the methods of Isaksson and Bingham is that the extra renal loss is a constant factor in Isaksson’s equation, 2, but a proportion in Bingham’s equation. Examples of protein/nitrogen validations are presented in tables 5 (20) and 6 (38).

There is a daily variation in protein intake as well as in nitrogen excretion, even though the variation in nitrogen excretion is not as large as in protein intake (35). Bingham and Cummings demonstrated that on average 18 days of dietary assessment are needed to assess dietary intake to $\pm 5\%$ of the habitual mean. They also demonstrated that eight 24 hour urine collections have been shown to be sufficient to estimate dietary nitrogen (protein) intake to within $81 \pm 5\%$ (1 SD) of the habitual dietary intake. Therefore they suggest the collection of eight 24 hour urine samples for protein validation on an individual level. However, eight days of urine collection is not feasible in all studies and less precision is sometimes necessary for practical reasons (20).

Due to the difficulty of obtaining eight complete 24 hour urine collections, replacement with partial collections (overnight collections) has been investigated, yielding poor results (37). Unfortunately, complete 24 hour urine collections are necessary for validation of protein, sodium and potassium (21), and probably for all markers in 24 hour urine collections, even though this has not been studied for all biomarkers.

Table 6. Validation of reported intakes of energy, nitrogen, sodium, and potassium in 30 vegans and 29 omnivores with de use of the doubly labeled water method and 24-h urine collections¹

	Females		Males	
	Vegans (n=15)	Omnivores (n=15)	Vegans (n=15)	Omnivores (n=14)
EI _{rep} /EE _{mea} ²	0.84 ± 0.25	0.92 ± 0.19	0.87 ± 0.19	0.85 ± 0.16
[N _{rep} x 0.81]/N _{mea} ³	1.00 ± 0.32	0.98 ± 0.13	1.14 ± 0.27	1.04 ± 0.14
Na _{rep} /Na _{mea} ⁴	0.78* ± 0.42	1.02 ± 0.21	1.02 ± 0.32	1.15 ± 0.26
[K _{rep} x 0.73 or 0.77] /K _{mea} ⁶	0.90 ± 0.30	0.88 ± 0.19	1.02 ± 0.27	0.92 ± 0.18

¹ $\bar{x} \pm SD$. EI, energy intake; rep, reported; EE, energy expenditure; mea, measured; N, nitrogen. The Mann-Whitney *U* test was used for statistical comparisons between groups with different diets.

² Reported energy intake of 32 subjects (16 vegans and 16 omnivores; 44% female) divided by energy expenditure, as measured by the doubly labeled water method (13).

³ Reported nitrogen intake times the urinary excretion factor of nitrogen, 0.81 (11), divided by measured nitrogen in urine.

⁴ Reported sodium intake divided by measured sodium in urine.

⁵ Significantly different from female omnivores ($P < 0.05$).

⁶ Reported potassium intake times the urinary excretion factor of potassium, 0.73 for vegans and 0.77 for omnivores (15), divided by measured potassium in urine.

(Larsson CL, Johansson GK. Dietary intake and nutritional status of young vegans and omnivores in Sweden. *Am J Clin Nutr.* 2002 1;76(1):100–6.)

8.3.3 Sodium

The excretion of sodium in urine is a good indicator of dietary intake (39). Faecal and other excretion of sodium is minimal. On average, 95 – 98 % of the dietary intake is excreted via the urine. Therefore, in validation studies the approximation that intake equals excretion is used. Unfortunately, the within-person variability in sodium excretion is large and eight complete urine collections are needed for a validation on an individual level (40). See table 6 for an example of sodium validation. Since sodium intake is difficult to assess with dietary assessment methods it may sometimes be better to use urinary excretion as a measurement method.

8.3.4 Iodine, fluoride and chloride

The urinary content virtually equals intakes for iodine, fluoride and chloride, which make them good biomarkers over the short-term (hours to days) (21). Since a very high proportion of the intake appears rapidly in the urine, 24 hour urine collections is the first choice of measurement of intake, instead of a traditional dietary assessment method.

8.3.5 Selenium

Urine provides a reasonable marker of short-term intake (hours to days), since urine is the major route of excretion (21). Plasma levels reflect intake to a degree if the range of variation is large. Hair and toe-nails are possible markers for long-term intake but are regarded as crude markers. Contamination of hair samples by shampoos must be controlled for.

8.3.6 Potassium

Urine is the major route of excretion for potassium (21). The faecal excretion of potassium may vary between approximately 10-30 %, and increases with increasing fibre intake (39, 41). For instance, a study on vegetarians and omnivores demonstrated a faecal excretion of 27 % and 23 %, respectively (42). Therefore, different correction factors for faecal losses for vegetarians (0.73) and omnivores (0.77) were used (see table 6).

8.3.7 Verification of the completeness of 24 hour urine collections

Validation of the intakes of protein, sodium, potassium, iodine, fluoride, chloride and other components is based on the assumption that the 24 hour urine collections are complete. An incomplete sample makes the biological marker invalid. Table 7 demonstrates that when there are low PABA values (indicates incomplete collections) it looks like the dietary intake is overreported. This is because the biological marker indicates a too low value, and thus, the ratio will be too high since the biological marker is the denominator.

Unfortunately, incomplete urine collections are a common problem. One study demonstrated that on average four out of ten people had incomplete urine collections (PABA less than 85 %) and in one study 71 % had incomplete samples (43). This clearly demonstrates the importance of having an objective marker for verification of the completeness of 24 hour urine collections. Without a complete urine collection there will not be a valid biological marker.

Table 7. The ratio of dietary intake according to the food record method or duplicate diet technique to dietary intake according to the biological marker is presented in relation to the PABA recovery. The first (55%) and second (90%) groups are the total group of women in the Swedish study divided into two PABA recovery groups (less than 70% and over 85%). The third group is the total number of women adjusted for PABA recoveries under 85% up to 93% as described in this study.

Variable	Mean PABA recovery value		
	55% (n=28)	90% (n=21)	Adjusted (n=73)
Protein*	1.78	0.91	0.86
Protein†	1.14	0.90	0.85
Sodium‡	1.19	0.88	0.84
Potassiumx§	1.15	0.87	0.82

* The protein intake according to the food record/protein intake according to urinary nitrogen, 6.25 * (nitrogen in urine + 2).

† The protein intake according to the duplicate diet/protein intake according to urinary nitrogen, 6.25 * (nitrogen in urine + 2).

‡ The sodium intake according to the food record/sodium in urine.

§ The potassium intake according to the food record/potassium in urine and faeces.

(Johansson G, Bingham S, Vahter M. A method to compensate for incomplete 24-hour urine collections in nutritional epidemiology studies. Public Health Nutr. 1999;2(4):587-91.)

8.3.7.1 Creatinine

Before the PABA method (see chapter 8.3.7.2) was introduced creatinine was used to check on the completeness of urinary collections (41). Creatinine excretion is dependent upon the creatinine intake, mainly from meat in the diet and endogenous creatinine production. The main conclusion from reference 41 was that the urinary creatinine excretion is not a valid marker for verification of the

completeness of 24 hour urine collections. This is because the variation in creatinine excretion is large, which makes it impossible to draw any conclusion on the individual level. In reference 44 they found that out of 21 incomplete urine collections, only 3 would have been detected by the creatinine method. One study demonstrated the extreme result that the group with the highest urine volume had the lowest creatinine excretion (45).

8.3.7.2 PABA

Para-amino benzoic acid (PABA) have been used since 1983 for verification of the completeness of 24 hour urine collections (46). The method consists of three tablets of 80 mg PABA taken during the day of urine collection, at breakfast, lunch and dinner, which is quantitatively excreted within 24 hours. Statistically, PABA recoveries below 85 % with the colorimetric method (43) and 78 % with the HPLC method (47) are considered to be incomplete. PABA has no convincing side-effects or pharmacological effect or known ill-effects from overdosage; nor is there any known drug interaction. The only known major problem is that drugs containing paracetamol are detected as PABA with the colorimetric method. Therefore, the researcher has to ask which drugs the volunteer is taking. Without doing this the PABA recoveries may be very high, far above 100 %, which make them easy to detect. A solution to this problem is to use the more recently developed HPLC method, where paracetamol is not detected as PABA (47).

8.3.7.2.1 *The PABA compensation method*

Before the method to compensate for incomplete 24 hour urine collections was introduced, incomplete collections could not be used (46), i.e. PABA recovery below 78 % with the HPLC method (47) and below 85 % with the colorimetric method (46). Thus, one lost a part of the participants, probably those with the lowest motivation. Since incomplete urine collections are such a common problem, up to 71 % in published studies (43), a method that could compensate for incomplete collections would be of great practical and economical value. Therefore, a method has been developed to make use of incomplete urine collections (43). The method is based on linear regression equations for urinary excretion of nitrogen, sodium and potassium in relation to PABA recovery, where values between 50 % and 84 % are compensated to a theoretical complete collection. The equation below is used:

$$y = a + \beta * x$$

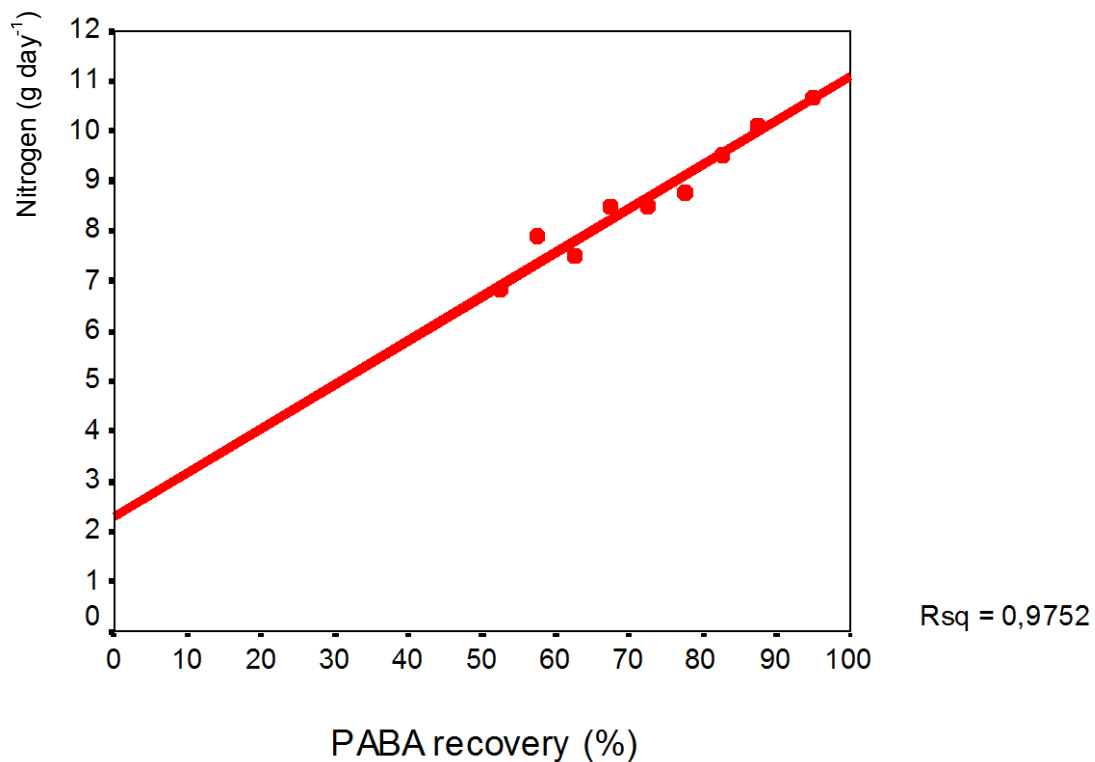
where y is the analyte in urine, in this case nitrogen, sodium or potassium, a is the analyte in urine when the PABA recovery is zero, β is the slope of the curve and x is the PABA recovery value (93 % minus the measured value). The value 93 % is from the study of Bingham and Cummings, where complete urine collections had an average value of 93 % (46). The linear regression equations for urinary nitrogen, sodium and potassium in relation to PABA recovery are:

$$\text{nitrogen: } y = 2.3 + 0.088 * x \text{ (r= 0.99) (see figure 3)}$$

$$\text{sodium: } y = 45 + 0.82 * x \text{ (r= 0.87)}$$

$$\text{potassium: } y = 19 + 0.60 * x \text{ (r= 0.93)}$$

As an example, if a urine sample had a PABA recovery of 63 % and a nitrogen content of 14.5 g, the linear regression equation ($\beta * x$) will give $0.088 * (93-63) = 2.6$ g. Thus, the adjusted nitrogen output would be $14.5 + 2.6 = 17.1$ g.



Cases weighted by N4

Figure 3. The relationship between PABA recovery (%) and nitrogen output in urine (g day⁻¹). The PABA recovery values have been divided into 5% intervals from 50% to 90% and one interval between 90% and 110%. The number of individuals for each PABA interval is 10, 12, 15, 13, 12, 19, 37, 63 and 131, respectively (total, n = 312). The correlation coefficient r is 0.99 when a weighted calculation is performed, i.e. all 312 values are taken into account in the regression analysis

(Johansson G, Bingham S, Vahter M. A method to compensate for incomplete 24-hour urine collections in nutritional epidemiology studies. *Public Health Nutr.* 1999;2(4):587–91.)

8.3.8 Fibre

Several studies have shown that the faecal weight increases with the fibre intake. Cummings et al. found a linear relationship between the faecal weight and the fibre intake and on average the faecal weight increased with 5 g for each g NSP (non-starch polysaccharides) (48). The linear regression equation is: $y = 5.3 x + 38$, where y = the faecal weight and x = the fibre weight.

Bingham found that fibre from mixed sources increased the faecal weight with 3.5 g (49). The linear regression equation is: $y = 3.5 x + 40$, where y = the faecal weight and x = the fibre weight. This equation is based on the English food tables. Johansson et al. modified this equation to be suitable for Swedish conditions, the fibre intake = $[(\text{faecal weight} - 40)/3.5] \times 0.86$. The figure 0.86 stands for the adjustment from the English food tables to the Swedish food tables. See table 8 for an example of fibre validation.

A major problem with fibre validation is that it is necessary to collect faecal samples for several days, since the day-to-day variation is large both in frequency and weight, which makes fibre validation unacceptable for many people (see also chapter 8.3.9 for validation of cadmium and lead). Preferably, orally taken radio-opaque pellets should be used as markers for complete faecal collections (51).

Table 8. Validation of the fibre intake before (0 month) and 3, 6 and 12 months after the dietary shift from a mixed diet to a lacto-vegetarian diet. The results are presented as mean values \pm 95 % confidence limits (50).

	0 month	3 months	6 months	12 months
Diet¹ (g/day)	19 \pm 1,5	31 \pm 6,0	29 \pm 5,2	28 \pm 4,5
Biol^{2,3}	20 \pm 7,4	37 \pm 10	34 \pm 9,7	27 \pm 10
Diet¹/ Biol^{2,3}	0,96	0,83	0,86	1,06

¹ Fibre intake according to the dietary survey

² Fibre intake according to the biological marker

³ Biol = [(faecal weight – 40)/3,5] x 0,86

(Johansson G, Callmer E, Gustafsson JA. Validity of repeated dietary measurements in a dietary intervention study. *Eur J Clin Nutr.* 1992;46(10):717–28.)

8.3.9 Cadmium and lead

Cadmium and lead, as well as fibre, are examples of biomarkers where faecal collections are necessary to obtain. Approximately 85 % of the intake of lead and 95 % of the intake of cadmium are excreted via faeces, which make them suitable as markers for dietary intake (52-54). It is rather a measurement of intake than a biological marker for intake, since cadmium and lead could not be found in food tables. Even if lead and cadmium could be found in the food tables it would not be a good way of measuring the intake since there is a large variation in foods. For these compounds there are only the duplicate portion technique and faecal samples that could be used for measuring dietary intake. Faecal analyses are less expensive than the duplicate portion technique and faecal collections do not affect the dietary intake to the same extent as the duplicate portion technique. The intake of lead can be estimated as $Pb_{Faeces}/0.85$ and the intake of cadmium as $Cd_{Faeces}/0.95$ (54).

8.3.10 Vitamin C

Of all the vitamins, vitamin C is considered to have the best biomarkers. Plasma is a relatively good marker for intakes between approximately 30 and 90 mg per day (21). At low intakes buffy coat is a better marker than plasma. Buffy coat is the upper, lighter portion of the blood occurring when coagulation is delayed or when blood has been centrifuged. One problem with vitamin C validation is that age, sex, smoking habits, infections, and etcetera is affecting the relationship between intake and marker. Therefore vitamin C validation should preferably be performed within a certain group of people. Another problem with vitamin C validation is that the vitamin is not stable and the sample needs to be taken care of in a special way. Preferably you use the vitamin C marker to categorize people into “risk groups” rather than provide a specific intake. Table 9 demonstrates that vitamin C is higher in women than in men, both measured by the dietary assessment method and in plasma (55).

Table 9. Daily intakes of vitamin C measured by a 7-day diary and plasma vitamin C in 1240 women and 877 men in Norfolk, England. The results are presented as means and standard deviations.

	Women	Men	P-value
7-day diary (mg)	87 (48)	81 (47)	<0,01
Plasma (µmol/l)	58 (21)	47 (19)	<0,001

(Bingham SA, Welch AA, McTaggart A, Mulligan AA, Runswick SA, Luben R, et al. *Nutritional methods in the European Prospective Investigation of Cancer in Norfolk. Public Health Nutr.* 2001 Jun;4(3):847–58.)

8.3.11 Fat and carbohydrates

There are no markers for total fat and carbohydrate intake.

8.3.12 Fatty acids

Fatty acids validations are examples of relative validations (21). The fatty acid intake is presented as a percentage of the total fatty acids. This means that if one fatty acid increases, others have to decrease. It is a qualitative marker, in contrast a quantitative marker, where the absolute intake can be estimated.

It is important to take a baseline sample for each “tissue” to be able to measure the change. This is because the relative amounts of fatty acids vary between “tissues” and individuals. When validating fatty acids it is also important to have knowledge about the fatty acid metabolism and factors that can affect it. For instance, fatty acids can be elongated and synthesized endogenously. Unfortunately, there are to my knowledge, no studies that have investigated the relationship between a controlled dietary intake and biomarkers for fatty acids. There are only studies that have compared reported intake with biomarkers.

The best biomarkers are for fatty acids that to a small extent or not at all can be synthesized in the body (56). The relationships between the dietary fatty acids and the relative amount of fatty acids in adipose tissue is most pronounced for linolic acid (18:2, n-6) and the long chain fatty acids eicosapentaenoic acid (EPA, 20:5, n-3) and docosahexaenoic acid (DHA, 22:6, n-3). Saturated fatty acids usually demonstrate weak but significant relationships, for instance myristic acid (14:0). The transfatty acids are usually well correlated with dietary intake. In general, the correlations between fatty acids in dietary surveys and in adipose tissue are stronger compared to the relationships in plasma and erythrocytes, which have a faster turnover. See table 10 for the time period of dietary intake that the fatty acids in various “tissues” reflect.

Table 10. Time period of dietary intake that the fatty acids in various “tissues” reflect.

“Tissue”	Reflected time period
Plasma triglycerides	Hours
Cholesteryl esters Phospholipids in red cells	Days
Red cell membranes Platelets	Weeks
Adipose tissue	Months or years

8.4 New biological markers for food intake

Below some recently developed biological markers for food intake are presented, which can be potential valuable markers in the near future.

8.4.1 Sucrose and fructose

In two studies with 25 individuals carried out in a metabolic suite, total sugars intake was significantly correlated with the sum of sucrose and fructose in urine (57). In the regression equation, 100 mg of sucrose and fructose in urine predicted approximately 200g of sugars intake.

8.4.2 Thiamine

In the same studies as in chapter 8.4.1, urinary thiamine was also highly correlated with thiamine intake. On average, $25 \pm 8\%$ (range 11.9 to 41.5%) of thiamine intake was excreted in the urine when subjects consumed their habitual diet. Similar promising results were presented at the same conference (58) and they are more promising than earlier results presented in reference 21.

8.4.3 Whole grain intake of wheat and rye

Alkylresorcinols are found in high contents in the outer parts of wheat and rye kernels (59). Since they are only found in whole grain or bran of wheat and rye and not in any other commonly consumed foods, they can be regarded as a biomarker for intake of foods containing these compounds.

8.4.4 Fruits and vegetables

Polyphenolic compounds are found in fruits and vegetables and have therefore been tested as potential markers for the intake of fruits and vegetables (60). Preliminary data indicates that the citrus polyphenols correlate significantly with the intake of oranges and orange juice determined by 4-days weighed food record. The study also shows that phloretin is a specific marker of fruit intake and that isorhamnetin is a marker for vegetable consumption. Kaempferol was found to be a specific marker for tea intake.

9 Misreporting

The most prevalent misreporting is underreporting of the dietary intake, which is a serious problem in nutritional epidemiological research. Underreporting is probably the main reason for the current confusion in the area of the importance of nutritional factors for health and disease. Black et al. presented in a study that 64 % of the food records, 88 % of the 24 hour recalls and 25 % of the diet histories were underreported (7). In a study with the doubly labelled water technique nine out of ten studies demonstrated a bias towards underestimation (61). This is a remarkably high percentage since there are often highly motivated volunteers in doubly labelled water studies.

Table 11 indicates that it is common with unrealistically low energy intakes, especially among obese people (62). The FIL values in the table should be compared with PAL values, and should have the same values if there are correct estimates of energy intake and energy expenditure. For instance, a PAL value of 1.2 indicates that the individual is bed-bound or chair-bound. The PAL value 1.35 is the former Goldberg cut-off 1 which is derived from whole-body calorimetry studies where an overall mean PAL of 1.35 was obtained. A room for whole-body calorimetry studies is a very small room, and thus, there are almost non-existent opportunities for physical activity. PAL values between 1.5 and 1.8 are “normal” PAL values for a Western sedentary life-style. To conclude some studies, great predictors of underreporting are obesity, weight consciousness and dieting (61).

Table 11. Proportion (%) underreporters by 24-hour recalls after stratification for gender and BMI. Two cut-off levels were applied for FIL (food intake level equivalent to reported energy intake/estimated basal metabolic rate) (62).

Strata	Underreporters (%)	
	FIL<1.2 ^a	FIL<1.35 ^b
<i>Gender</i>		
Men (n=94)	44 ^{NS}	61 ^{NS}
Women (n=99)	47	72
<i>BMI (kg m⁻²)</i>		
<25 (n=105)	32 ^{p<0.001}	53 ^{p<0.001}
25-30 (n=32)	56	79
>30 (n=16)	88	94

^a FIL<1.2 correspond to a PAL for a chair-bound or bed-bound person (survival limit).

^b FIL<1.35 correspond to a PAL for lowest possible free-living sedentary lifestyle.

p-values obtained with Chi²-testing among the groups. NS for non-significant.

(Johansson G, Wikman A, Ahrén AM, Hallmans G, Johansson I. Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living. Public Health Nutr. 2001;4(4):919–27.)

The quality of the diet may be different in various groups with different energy intake, which is demonstrated in table 12 (62). The question is if this is due to that they eat differently or if it is a question of different abilities to report the food intake. Unfortunately, several studies have demonstrated that there are differences in the reporting ability between the groups. This is serious,

since opposite associations between diet and other factors may be found if incorrect dietary data are used (64, 65).

Table 12. Composition of the diet at different level of FIL (food intake level equivalent to reported energy intake/estimated basal metabolic rate). Data are expressed as mean (SD) (62).

	FIL			ANOVA P value
	<1.2 (n=88)	1.2-1.35 (n=40)	>1.35 (n=65)	
Energy (MJ)	6.2 (1.4) ^{a,b}	8.1 (1.4) ^{a, c}	10.3 (2.0) ^{b,c}	<0.001
Protein (g/10 MJ)	97 (15) ^a	92 (11)	86 (10) ^a	<0.001
Fat (g/10 MJ)	90 (11) ^a	93 (13)	97 (9) ^a	<0.01
Carbohydrates (g/10 MJ)	282 (27)	275 (35)	278 (29)	NS
Sucrose (g/10 MJ)	50 (20) ^a	45 (13) ^b	57 (16) ^{a,b}	<0.01

^{a-c} Numbers (within a line) sharing the same superscript differ significantly, by at least $P < 0.05$, when tested with a multiple mean test (Tukey's test), applied after the ANOVA had indicated a significant difference among the groups.

NS = not significant

(Johansson G, Wikman A, Ahrén AM, Hallmans G, Johansson I. Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living. *Public Health Nutr.* 2001;4(4):919–27.)

In a study individuals were asked if they have altered their dietary intake during the food recording period (66). Those admitting that they have changed their dietary intake and claimed that they became more conscious, were those with the lowest energy intake (FIL = 1.10). Those admitting that they have changed their dietary intake and claimed that it was too much hassle to perform the survey had the highest energy intake (FIL = 1.50). Those who claimed that they had not altered the dietary intake had a FIL value of 1.23. This indicates that it is not meaningful to ask whether people have altered their dietary intake during the recording period. Reasons for these reporting problems in dietary surveys may be of a psychological nature and have been studied, among others, by Taren et al. (67). They showed that social desirability and self image of body shape were associated with reporting accuracy. Social desirability predicts underreporting and especially underreporting of so-called “socially undesirable food items”. An example of this is a consistent overreporting of culturally approved, “healthful” behaviours (whole grains, fruits, vegetables, skim milk), and underreporting of culturally undervalued, “indulgent” behaviours (alcohol consumption, high-fat snacks, sugared cereals, cookies) (68). These results indicate that underreporting is selective rather than general, i.e. some foods are underreported, while other foods are not or even overreported. Table 13 gives an example of selective reporting, since fibre is overreported and the other components are underreported (69).

Table 13. Comparison of the ratio dietary intake to biological marker for energy, protein, sodium, potassium and fibre for 34 women eating a mixed diet. (69).

Variable	Ratio
Energy	0,84
Protein	0,80
Sodium	0,93
Potassium	0,83
Fibre	1,20

As a consequence of the selective underreporting of “socially undesirable food items” (62, 68, 70, 71), meals and snacks have been studied with regard to reporting accuracy. Popitt et al. demonstrated that meals were reported relatively well, while snacks were not (see table 14) (71). Also, the reporting accuracy seems to be independent of dietary assessment method, so that when an individual reports poorly with one method it is likely that the reporting will be poor even with other methods, a so-called “subject-specific bias” (22).

Table 14. Reported energy intake during meals, between meals (snacks) and total reported energy intake compared with measured energy intake (100 %) in 33 women (18 obese and 15 non-obese) (71).

	Meals	Snacks	Total
Energy intake	96 %	64 %	86 %

The selective reporting of foods distorts the information on nutrient intake (see table 15) (71). Further, the selective reporting has consequences on how you can deal with the problem.

Table 15. Reported intake of macronutrients compared to measured intake (100 %) in 33 women (18 obese and 15 non-obese) (71).

Macronutrient	Percent of measured intake
Protein	101 %
Fat	91 %
Carbohydrate	80 %
Alcohol	73 %
Total	86 %

So far, results have been presented that have demonstrated that underreporting is prevalent, is selective rather than general, has psychological reasons, and is common in certain types of people, such as obese subjects. These facts need an explanation for further understanding and how to overcome the problem of misreporting. One answer is that people are aware that they are judged on the basis of the foods they eat (72). Emotional and moral feelings influence the perceived value of making a true record. The more we tell people that they should reduce their fat intake, the more they tell us how little of it they eat, although actual intakes do not appear to change or even increase. The following quotation from an obese woman illustrates how feelings affect what one wants to report: “I did not want to perform a food record, but at the end I did it. I do not want to know that I eat 5000 kcal per day. How fun is that? I already know that I eat the wrong things” (73).

Other reasons for not stating a true food intake is presented by Vuckovic et al. as “simplifying food intake” (74). Below are quotes from focus group participants that illustrate “simplifying” strategies.

Use of simpler foods:

- “I still ate the same amount as what I have written in the journal, but it was simply just a sandwich instead of cooking something. I didn’t have to mess around.”
- “When I ate out, I tried to make sure it was something straightforward: it was rice, it was beets. Not anything really complicated, just so it would be easier for the record.”
- I didn’t want to cook. I don’t cook that much anyway, but I didn’t want to cook and write down the ingredients”.

Choice of foods with defined portion sizes:

- “On a Sunday I went to a conference, and I didn’t have to prepare anything. It just so happened I just had 2 (Burger King) Whoppers for the whole day, and I was so happy that I didn’t have to write anything down except the 2 Whoppers. So I was all done for the whole day.”

Less frequent eating and consumption of fewer snacks:

- “I think sometimes I ate less than usual. I went to eating 3 times a day instead of eating between meals, so I wouldn’t have to write that down.”
- “If I wanted a piece of candy, I did think, “Oh I’ve got to write it down”. Sometimes I would eat less for that reason.”
- “I thought, “I don’t want to bother with writing it down so I’ll just forego it.”

Not eating in restaurants:

- “I only ate out once during the time when I probably would have eaten out twice because it was kind of a hassle to record every item and get people to give you a recipe and things.”

After this report on the size and nature of the underreporting problem, the next question to ask is: What shall we do about it? One alternative is to leave the underreporters out. Then one would lose valuable information and introduce a selection bias. The underreporters may have different dietary habits and higher prevalence of certain diseases. Another problem with leaving the underreporters out is to correctly identify them. For instance, if you discard those people with low FIL values you run the risk of keeping underreporters with relatively high FIL values, but even higher PAL values (see chapter 8.3.1). You need good data on PAL to be able to detect underreporters. You can perform analyses with and without the underreporters to see whether there will be differences in the conclusions. However, it is difficult to use this information in a constructive way, except to conclude that the underreporters are different in some aspects. Energy adjustments can be performed but they do not eliminate bias due to selective underreporting (75). They can even make the problem worse (22). Of course, do they not provide correct estimates of absolute nutrient intake, which may be useful information. If the data are going to be compared with a specific recommendation, the number of individuals will vary depending on the choice of cut-off value. In addition, the selective underreporting may cause misclassification within the distribution of values when ranking people in nutritional epidemiological studies, which will have effects on the conclusions drawn from the study. To conclude this section, underreporting introduces serious problems, which will be a challenge for researchers to understand, estimate, and make use of the error structure during analysis.

10 Energy adjustments

Usual reasons for performing energy adjustments are to exclude differences in nutrient intake due to body size, physical activity and metabolic efficiency (76). Many nutrients are related to the total energy intake, which has the implication that those who have a total high food intake will usually also have a total high nutrient intake. For example, it could be more relevant to relate an energy adjusted fat intake to a disease, rather than the total fat intake. For instance, this is because tall, small and physical active women have a higher total fat intake compared with sedentary women, without eating a fat diet. Energy adjusted nutrient intake is a way in nutritional epidemiological studies to investigate the effect of a nutrient independently from the energy intake.

The probably most common way of energy adjustment is the nutrient density method, where the nutrient value is divided by the total energy intake. A similar approach for macronutrients is to express intake as a percentage of total energy intake (E%). The nutrient density has two components: the nutrient intake and the inverse of total energy intake. This can lead to problems when the total energy intake is related to the investigated disease.

Another energy adjustment is the so-called residual method or the Willett residual method (76). These energy adjusted nutrient intakes are computed as the residuals from the regression model with total energy intake as the independent variable and absolute nutrient intake as the dependent variable. In this model, the energy adjusted nutrient intakes are not associated with energy intake. For instance, if the nutrient is fat and because fat and energy are highly correlated, adjustment for total energy intake reduces the variation in fat intake radically. For skewed nutrient variables Willett suggests a log transformation of the nutrient intake. Usually, the nutrient density model and the residual model yield the same associations in nutritional epidemiological studies. In the case of underreporting, especially selective underreporting, energy adjustment models may result in dubious relationships between diet and disease.

Willett describes three additional energy adjustment methods. They are “the standard multivariate method”, “the energy decomposition method” and “the multivariate nutrient density method”. Since these methods are more complicated than those described above, the reader is referred to reference 76 for detailed explanations.

11 Statistical corrections for the effect of measurement error

If the effect of measurement error is possible to assess, then it may be suitable to consider a statistical correction. The statistical methods are likely to become a frequent procedure in the future. However, eminent statistical knowledge is required to elaborate further on this issue. Therefore, statistically interested and knowledgeable readers are referred to references 77-79.

12 Estimating a sample size

Ethics boards, funding agencies and research review panels often request that a researcher perform a power analysis, for example to determine the minimum number of people needed for a study to be informative. Power analysis can be used to calculate the minimum sample size required so that one can be reasonably likely to detect an effect of a given size. Power analysis can also be used to calculate the minimum effect size that is likely to be detected in a study using a given sample size. Power is the probability of finding a difference that does exist, i.e. the power of a statistical test is the probability that the test will reject the null hypothesis when the null hypothesis is actually false (i.e. the probability of not committing a Type II error, or making a false negative decision). The probability of a Type II error occurring is referred to as the false negative rate (β). Power is equal to $1 - \beta$. The likelihood of declaring a difference that does not exist is known as a Type I error. Most study designers assess the power of their tests using 0.80 as a standard for adequacy. This convention implies a four-to-one tradeoff between β -risk and α -risk, i.e. $\beta = 0.2$ ($1 - 0.8$) and $\alpha = 0.05$. Usually you use a two-tailed test i.e. you are open for a change in both directions.

A statistical significance criterion (p-value) is a statement of how unlikely a result must be, if the null hypothesis is true, to be considered statistically significant. The most commonly used criteria are probabilities (p) of 0.05 (5%, 1 in 20), 0.01 (1%, 1 in 100), and 0.001 (0.1%, 1 in 1000). In nutritional epidemiology a p-value of 0.05 is the most common. If the criterion is 0.05, the probability of obtaining the observed effect when the null hypothesis is true must be less than 5 % (0.05). The magnitude of the effect of interest (the magnitude of the difference between the groups, tests, experiments etc.) in the population can be quantified in terms of an effect size. There is greater power to detect larger effects, i.e. when there is a large effect or a large difference you do not need as many people as if there was a small effect or a small difference.

The design of an experiment or observational study influences the power. For instance, in a two-sample testing situation with a given total sample size n , it is optimal to have equal numbers of observations from the two populations being compared (as long as the variances in the two populations are the same).

When you want to estimate your sample size you can use statistical programmes that you can buy from statistical software companies, such as SPSS, or are freely available on the internet. There are four components (three except the sample size) when you want to calculate your sample size. These are:

- sample size, the number of units, people, animals etc.
- effect size, the effect of the treatment/intervention.
- alpha level (α , or significance level), or the odds that the observed result is due to chance.
- power, or the odds that you will observe a treatment effect when it occurs.

13 Future research directions

If dietary surveys are to be improved, we must understand which foods, meals and snacks are misreported (22). Further, we must understand why people misreport. If the sources of bias in dietary assessment are to be identified, nutritionists must go beyond the narrowly mechanistic focus of dietary intake measurement and examine the cultural, social and psychological context of such reporting. For instance, the significance of social desirability suggests that it is vital to train volunteers in such a way that they become comfortable reporting foods that may be considered socially undesirable. As a researcher it is important to make the subjects comfortable with trainers and to emphasize during the interview that items such as large and extra-large portion sizes, snacks, high fat foods, and sweets are acceptable to report since they are part of the social norm (67). It is important to realize that the relation between interviewer and volunteer may play a critical role in how dietary data are reported. Many aspects of behaviour may contribute to the problem, such as general conduct towards subjects, unconscious body language, phraseology, voice intonation, unconscious attitudes, and etcetera (22). Also, it is important to identify underreporters at the individual level (22). Thus, a good measurement of PAL/EE is crucial to obtain. Even though, the above factors are considered, there will always be error in dietary surveys. By understanding more about factors that create bias in dietary assessments, investigators can design more effective instruments, provide appropriate instruction to volunteers, and better interpret results (74). If these aspects are considered, new dietary effects may be discovered in old as well as in new areas of nutritional research.

14 Conclusions and practical recommendations

1. **Create a good atmosphere for obtaining valid dietary information.** Why should people reveal their food habits if they will feel bad about it? Therefore, it is important to make an effort to create a good atmosphere between investigator and volunteer. People should feel comfortable to talk about their dietary intake. There is a video-tape that can be used, which specifically promoted honesty because it explicitly dealt with issues such as consumption of liquor and energy-dense “bad” foods (67). Another approach is to discuss the value of true dietary information, both for the volunteer and for the study.
2. **Include biological markers for food intake.**
3. **Critical evaluation of energy intake data using fundamental principals of energy physiology. Derivation of cut-off limits to identify under-recording (7, 32).** As an alternative to the DLW method the Goldberg cut-off can be used. Information on age, sex, body weight, height and PAL/EE should be included in the dietary survey. Thus, no samples, such as blood or urine, need to be obtained from the individuals.
4. **Include questions on weight consciousness and dieting (7, 32)** in order to evaluate FIL in relation to these known risk factors for misreporting. Ask about lowest and highest body weight during the last years.
5. **Energy adjustments (76)** in order to investigate the effects of nutrients independently from energy intake. However, in the case of underreporting, especially selective underreporting, energy adjustment models may result in dubious relationships between diet and disease.
6. **Statistical corrections for the effect of measurement error (77-79).** Excellent knowledge in statistics is necessary for this approach to be used. Otherwise, a co-operation with a statistician is recommended for these analyses.

References

1. Permanente nordiske utvalg for næringsmiddelspørsmål. Valg av metode ved kostholdsundersøkelser. [Oslo]: [Nordisk ministerråd]; 1986.
2. Nes M, editor. Kostholdsundersøkelser : hvorfor og hvordan? København: Nordisk Ministerråd; 1987.
3. Macdiarmid JI, Blundell JE. Dietary under-reporting: what people say about recording their food intake. *Eur J Clin Nutr.* 1997 18;51(3):199–200.
4. Stockley L. Changes in habitual food intake during weighed inventory surveys and duplication diet collections. A short review. *Ecol Food Nutr.* 1985;17(3):263–9.
5. Callmer E, Hagman U, Haraldsdóttir J, Løken EB, Seppänen R, Trygg K. Standardisering av 24-timmars intervju. *Vår föda.* 1986;38(Suppl. 4).
6. Burke BS. The dietary history as a tool in research. *J Am Diet Assoc.* 1947;23:1041–6.
7. Black AE, Goldberg GR, Jebb SA, Livingstone MB, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *Eur J Clin Nutr.* 1991;45(12):583–99.
8. Hagman U, Haraldsdóttir J, Rastas M, Sarlio-Lähteenkorva S, Seppänen R, Trygg K. Riktlinjer för användning av livsmedelsfrekvensmetoder. *Vår föda.* 1989;41(Suppl. 1):127–38.
9. Sarlio-Lähteenkorva S. Short methods for dietary assessment: A review of the methodological studies. *Vår föda.* 1989;41(Suppl. 1):91–110.
10. Bakkum A, Bloemberg B, van Staveren WA, Verschuren M, West CE. The relative validity of a retrospective estimate of food consumption based on a current dietary history and a food frequency list. *Nutr Cancer.* 1988;11(1):41–53.
11. Rohan TE, Potter JD. Retrospective assessment of dietary intake. *Am J Epidemiol.* 1984;120(6):876–87.
12. Willett WC, Sampson L, Browne ML, Stampfer MJ, Rosner B, Hennekens CH, et al. The use of a self-administered questionnaire to assess diet four years in the past. *Am J Epidemiol.* 1988;127(1):188–99.
13. Callmer E, Løken EB, Solvoll K. Selecting a method for a dietary survey. *Näringsforskning.* 1985;29:43–52.
14. Callmer E, Gustafsson L. Personal communication.

15. Nelson M, Black AE, Morris JA, Cole TJ. Between- and within-subject variation in nutrient intake from infancy to old age: estimating the number of days required to rank dietary intakes with desired precision. *Am J Clin Nutr.* 1989 Jul;50(1):155–67.
16. Beaton GH, Milner J, Corey P, McGuire V, Cousins M, Stewart E, et al. Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am J Clin Nutr.* 1979;32(12):2546–59.
17. Nelson M. The validation of dietary assessment. In: Margetts BM, Nelson M, editors. *Design concepts in nutritional epidemiology.* 2. ed. Oxford: Oxford University Press; 1997.
18. Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, et al. Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. *Br J Nutr.* 1994;72(4):619–43.
19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986 8;1(8476):307–10.
20. Larsson CL, Westerterp KR, Johansson GK. Validity of reported energy expenditure and energy and protein intakes in Swedish adolescent vegans and omnivores. *Am J Clin Nutr.* 2002 1;75(2):268–74.
21. Bates CJ, Thurnham DI, Bingham SA, Margetts BM, Nelson M. Biochemical markers of nutrient intake. In: Margetts BM, Nelson M, editors. *Design concepts in nutritional epidemiology.* 2. ed. Oxford: Oxford University Press; 1997.
22. Black AE. The sensitivity and specificity of the Goldberg cut-off for EI:BMR for identifying diet reports of poor validity. *Eur J Clin Nutr.* 2000;54(5):395–404.
23. Livingstone MBE, Black AE. Markers of the validity of reported energy intake. *J Nutr.* 2003;133 Suppl 3:895S–920S.
24. Speakman JR. *Doubly labelled water : theory and practice.* London: Chapman & Hall; 1997.
25. Forsum E. Bestämning av energiomsättning med dubbelmärkt vatten. *Näringsforskning.* 1989;33(3):95–8.
26. Sjöström M, Ekelund U, Yngve A. Assessment of physical activity. In: Gibney MJ, editor. *Public health nutrition.* Oxford, UK :: Blackwell Science; 2004. p. 83–105.
27. Nilsson A, Hurtig-Wennlöf A, Ekelund U. Objektiva metoder för bestämning av fysisk aktivitet. *Sven Idrottsforskning.* 2001 1;(2):8–12.
28. Andrén Aronsson C, Wirfält E, Gullberg B, Hedblad B. An evaluation of different methods to assess physical activity in a sub-sample of the Malmö Diet and Cancer (MDCS) cohort. *ICDAM.* Copenhagen, Denmark; 2006. p. 23.

29. Matthiessen J, Biloft-Jensen A, Banke Rasmussen L, Velsing Groth M, Fagt S. Validation of a new physical activity questionnaire for assessment of energy expenditure. ICDAM. Copenhagen, Denmark; 2006. p. 88.
30. Johansson G, Hagfors L, Westerterp K. Validation with doubly labelled water of a very short questionnaire on physical activity. The 8th Nordic Nutrition Conference. Tønsberg, Norway; 2004. p. 101.
31. Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, 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(12):569–81.
32. 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(9):1119–30.
33. Committee on Medical Aspects of Food Policy. Panel on Dietary Reference. Dietary reference values for food energy and nutrients for the United Kingdom. London: HMSO; 1991.
34. Denis W, Borgstrom P. A study of the effect of temperature on protein intake. *J Biol Chem.* 1924;61(1):109–16.
35. Bingham SA, Cummings JH. Urine nitrogen as an independent validity measure of dietary intake: a study of nitrogen balance in individuals consuming their normal diet. *Am J Clin Nutr.* 1985;42(6):1276–89.
36. Isaksson B. Urinary nitrogen output as a validity test in dietary surveys. *Am J Clin Nutr.* 1980;33(1):4–5.
37. Bingham SA. Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr.* 2003;133 Suppl 3:921S–924S.
38. Larsson CL, Johansson GK. Dietary intake and nutritional status of young vegans and omnivores in Sweden. *Am J Clin Nutr.* 2002 1;76(1):100–6.
39. Schachter J, Harper PH, Radin ME, Caggiula AW, McDonald RH, Diven WF. Comparison of sodium and potassium intake with excretion. *Hypertension.* 1980;2(5):695–9.
40. Liu K, Stamler J. Assessment of sodium intake in epidemiological studies on blood pressure. *Ann Clin Res.* 1984;16 Suppl 43:49–54.
41. Cummings JH, Hill MJ, Jenkins DJ, Pearson JR, Wiggins HS. Changes in fecal composition and colonic function due to cereal fiber. *Am J Clin Nutr.* 1976;29(12):1468–73.
42. Johansson G, Callmer E, Gustafsson JA. Validity of repeated dietary measurements in a dietary intervention study. *Eur J Clin Nutr.* 1992;46(10):717–28.

43. Johansson G, Bingham S, Vahter M. A method to compensate for incomplete 24-hour urine collections in nutritional epidemiology studies. *Public Health Nutr.* 1999;2(4):587–91.
44. Bingham SA, Cummings JH. The use of creatinine output as a check on the completeness of 24-hour urine collections. *Hum Nutr Clin Nutr.* 1985;39(5):343–53.
45. Knuiman JT, Hautvast JG, van der Heyden L, Geboers J, Joossens JV, Tornqvist H, et al. A multi-centre study on completeness of urine collection in 11 European centres. I. Some problems with the use of creatinine and 4-aminobenzoic acid as markers of the completeness of collection. *Hum Nutr Clin Nutr.* 1986;40(3):229–37.
46. Bingham S, Cummings JH. The use of 4-aminobenzoic acid as a marker to validate the completeness of 24 h urine collections in man. *Clin Sci Lond Engl* 1979. 1983;64(6):629–35.
47. Jakobsen J, Ovesen L, Fagt S, Pedersen AN. Para-aminobenzoic acid used as a marker for completeness of 24 hour urine: assessment of control limits for a specific HPLC method. *Eur J Clin Nutr.* 1997;51(8):514–9.
48. Cummings JH, Bingham SA, Heaton KW, Eastwood MA. Fecal weight, colon cancer risk, and dietary intake of nonstarch polysaccharides (dietary fiber). *Gastroenterology.* 1992;103(6):1783–9.
49. Bingham S. Low-residue diets: a reappraisal of their meaning and content. *J Hum Nutr.* 1979;33(1):5–16.
50. Johansson G, Callmer E, Gustafsson JA. Validity of repeated dietary measurements in a dietary intervention study. *Eur J Clin Nutr.* 1992;46(10):717–28.
51. Cummings JH, Jenkins DJ, Wiggins HS. Measurement of the mean transit time of dietary residue through the human gut. *Gut.* 1976;17(3):210–8.
52. Hammond B, Gartsid PS. Biological media – their advantages and pitfalls when used in biological monitoring – faeces. In: Clarkson TW, editor. *Biological monitoring of toxic metals.* New York: Plenum; 1988. p. 641–8.
53. Vahter M, Berglund M, Friberg L, Jorhem L, Lind B, Slorach S, et al. Dietary intake of lead and cadmium in Sweden. *Vår föda.* 1990;42(Suppl. 2):1–16.
54. Vahter M, Johansson G, Akesson A, Rahnster B. Faecal elimination of lead and cadmium in subjects on a mixed and a lactovegetarian diet. *Food Chem Toxicol.* 1992;30(4):281–7.
55. Bingham SA, Welch AA, McTaggart A, Mulligan AA, Runswick SA, Luben R, et al. Nutritional methods in the European Prospective Investigation of Cancer in Norfolk. *Public Health Nutr.* 2001 Jun;4(3):847–58.
56. Kohlmeier L, Kohlmeier M. Adipose tissue as a medium for epidemiologic exposure assessment. *Environ Health Perspect.* 1995;103 Suppl 3:99–106.

57. Tasevska N, Runswick SA, Bingham SA. The development of new biomarkers of nutritional intake. ICDAM. Copenhagen, Denmark; 2006. p. 7.
58. Runswick SA, Tasevska N, Bingham S. Urinary thiamine as a biomarker for estimates of thiamine intake. ICDAM. Copenhagen, Denmark; 2006. p. 34.
59. Landberg R, Lindo A-M, Adlercreutz H, Vessby B, Kamal-Eldin A, Åman B. Alkylresorcinols as biomarkers for whole grain intake of wheat and rye. ICDAM. Copenhagen, Denmark; 2006. p. 33.
60. Struntze Krogholm K, Poulsen L, Kristensen M, Brantsaeter AL, Rasmussen SE. New and validated biomarker for intake of fruits and vegetables. ICDAM. Copenhagen, Denmark; 2006. p. 60.
61. Livingstone MBE, Black AE. Markers of the validity of reported energy intake. *J Nutr.* 2003;133 Suppl 3:895S–920S.
62. Johansson G, Wikman A, Ahrén AM, Hallmans G, Johansson I. Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living. *Public Health Nutr.* 2001;4(4):919–27.
63. Hirvonen T, Männistö S, Roos E, Pietinen P. Increasing prevalence of underreporting does not necessarily distort dietary surveys. *Eur J Clin Nutr.* 1997;51(5):297–301.
64. Rosell MS, Hellénus M-LB, de Faire UH, Johansson GK. Associations between diet and the metabolic syndrome vary with the validity of dietary intake data. *Am J Clin Nutr.* 2003 Jul;78(1):84–90.
65. Macdiarmid JI, Vail A, Cade JE, Blundell JE. The sugar-fat relationship revisited: differences in consumption between men and women of varying BMI. *Int J Obes Relat Metab Disord.* 1998;22(11):1053–61.
66. Macdiarmid JI, Blundell JE. Dietary under-reporting: what people say about recording their food intake. *Eur J Clin Nutr.* 1997 18;51(3):199–200.
67. Taren DL, Tobar M, Hill A, Howell W, Shisslak C, Bell I, et al. The association of energy intake bias with psychological scores of women. *Eur J Clin Nutr.* 1999 Jul;53(7):570–8.
68. Ritenbaugh CK, Harrison GG. Reactivity of garbage analysis. *Am Behav Sci.* 1984 1;28(1):51–70.
69. Johansson G, Akesson A, Berglund M, Nermell B, Vahter M. Validation with biological markers for food intake of a dietary assessment method used by Swedish women with three different dietary preferences. *Public Health Nutr.* 1998;1(3):199–206.

70. Lafay L, Mennen L, Basdevant A, Charles MA, Borys JM, Eschwège E, et al. Does energy intake underreporting involve all kinds of food or only specific food items? Results from the Fleurbaix Laventie Ville Santé (FLVS) study. *Int J Obes Relat Metab Disord.* 2000;24(11):1500–6.
71. Poppitt SD, Swann D, Black AE, Prentice AM. Assessment of selective under-reporting of food intake by both obese and non-obese women in a metabolic facility. *Int J Obes Relat Metab Disord.* 1998;22(4):303–11.
72. Blundell JE. What foods do people habitually eat? A dilemma for nutrition, an enigma for psychology. *Am J Clin Nutr.* 2000;71(1):3–5.
73. Olsson H, Rosendahl E. Kvalitativ studie av överviktiga kvinnors tankar och känslor kring sig själva, sin övervikt, sitt beteende och sin överviktsbehandling. C-uppsats. [Umeå]: Umeå universitet. Institutionen för kostvetenskap; 2000.
74. Vuckovic N, Ritenbaugh C, Taren DL, Tobar M. A qualitative study of participants' experiences with dietary assessment. *J Am Diet Assoc.* 2000;100(9):1023–8.
75. Stallone DD, Brunner EJ, Bingham SA, Marmot MG. Dietary assessment in Whitehall II: the influence of reporting bias on apparent socioeconomic variation in nutrient intakes. *Eur J Clin Nutr.* 1997;51(12):815–25.
76. Willett W, Stampfer M. Implications of total energy intake for epidemiological analyses. In: Willett M, editor. *Nutritional epidemiology.* 2. ed. New York ;: Oxford University Press; 1998. p. 273–301.
77. Day N, McKeown N, Wong M, Welch A, Bingham S. Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium. *Int J Epidemiol.* 2001;30(2):309–17.
78. Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E, et al. Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutr.* 2002;5(6A):915–23.
79. Willett M, editor. Correction for the effects of measurement error. *Nutritional epidemiology.* 2. ed. New York : Oxford University Press; 1998. p. 302–20.

Appendix 1

The physical activity questionnaire used in this study.

Describe your physical activity at work (even work at home, sick leave at home and studying, for instance in a university).

1. **Very light**, e.g. sitting at the computer most of the day or sitting at a desk.
2. **Light**, e.g. light industrial work, sales, or office work that comprises light activities.
3. **Moderate**, e.g. cleaning, staffing at kitchen, or delivering mail on foot or by bicycle.
4. **Heavy**, e.g. heavy industrial work, construction work, or farming.

Describe your physical activity at leisure time. If the activities vary between summer and winter, try to give a mean estimate.

1. **Very light**. Almost no activity at all.
2. **Light**, e.g. walking, nonstrenuous cycling, or gardening approximately once a week.
3. **Moderate**. Regular activity at least once a week, e.g. walking, bicycling, or gardening or walking to work 10-30 min/day.
4. **Active**. Regular activities more than once a week, e.g. intense walking or bicycge or sports.
5. **Very active**. Strenuous activities several times a week.

Appendix 2

The scheme for estimating physical activity levels.

Physical activity in leisure time	Physical activity at work			
	Very light	Light	Moderate	Heavy
Very light	1.4	1.5	1.6	1.7
Light	1.5	1.6	1.7	1.8
Moderate	1.6	1.7	1.8	1.9
Active	1.7	1.8	1.9	2.1
Very active	1.9	2.0	2.2	2.3