Chapter 2

Measuring intake of fruit and vegetables

This chapter describes methods for estimating fruit and vegetable intake: household measures, questionnaire measures of usual or habitual intake and recording of actual or current intake (Table 7). These methods are used for various purposes, including nutrition surveillance, epidemiological research (case–control and cohort studies) and methodological research for validation of other dietary methods. They can also be used in clinical trials and intervention studies as well as for clinical evaluation.

Household measures of food availability

Household dietary surveys, household budget surveys and food balance sheets are used at the national or population level to estimate intake for nutrition surveillance and monitoring. They provide a broad view of the availability and consumption of fruit and vegetables. These survey methods provide what are technically considered crude measures of dietary intake, expressed at the household or per capita level.

Household dietary surveys

This method involves the compilation of an inventory of all foods present in the household at the beginning and at the end of the survey, complemented by the report of the amounts of foods purchased or otherwise obtained or consumed elsewhere and of the amount of edible food wasted or otherwise disposed of in the intervening survey period (Cresta *et al.*, 1969; Burke & Pao, 1976). The data may be recorded by weight and/or estimated on the basis of household measures and units, or as a combination. This method, fairly common in the past, is now more rarely used. The information obtained refers to the household and is expressed as per capita consumption. Expressing total consumption on the basis of consumption units, determined according to the estimated energy requirements of the individual members of the household, can provide some approximation of individual consumption. This procedure however ignores perforce the possible non-proportional distribution of various foods among the members of a household, and no statistical method can fully correct for this.

Household budget surveys

Another source of information on nationally representative dietary patterns is household budget surveys (HBS) (Trichopoulou *et al.*, 1999). These surveys are regularly conducted in most of the developed countries and in several developing ones. The sampling unit is the household, and the surveys are conducted principally for the purpose of monitoring the expenditure of families. Purchases of food are recorded as part of the overall purchases of the family and translated at a second stage into amounts. In some countries, foods are reported also as quantities. Socio-demographic information is also obtained, such as the educational level and employment of the head and other members of the family, the composition of the household, the urban, rural or semi-urban location of the household. Since each country has its own procedures and protocols for these surveys, the disparity of the collected data precludes comparison between countries. The diversity concerns not only the sampling methods but also the duration of the survey period, the number and details of the foods recorded, the inclusion or omission of foods consumed outside the home and the level of aggregation of individual food items into larger groups. In 1993, the European Commission funded a project (DAFNE, DAta Food NEtworking) that undertook to create an European data bank on food availability for human consumption, exploiting HBS data. In 1998, DAFNE harmonized the HBS of 10 European countries (Belgium, Germany, Greece, Hungary, Ireland, Luxembourg, Norway, Poland, Spain, United Kingdom), thus making available a set of data that provides an insight into national food habits and their distribution on the basis of socio-economic, educational and demographic parameters.

Food balance sheets

The food balance sheets (FBS) of the United Nations Food and Agriculture Organization (FAO) provide a unique set of data on food intake, collected

Table 7. Methods for estimation of dietary fruit and vegetable intake in different settings

Method	Measurement of consumption	National surveillance	Observational epidemiology[a]	Validation for FFQ
Household measures of food availability				
Household dietary surveys	Food inventory (disappearance)	V		
Household budget surveys	Expenditure	V		
Food balance sheets	Food disappearance	V		
Questionnaires of usual intake for individuals				
Diet history	Usual intake (past, time varies)		v	
FFQ—long	Usual intake (past, time varies)	v	V	
FFQ—brief	Usual intake (past, time varies)	v	v	
Recording of actual intake				
24-hours recall	Actual intake (specific time-point)	V	v	V
Food record	Actual intake (specific time-point)	V	v	V

[a] Case–control and cohort studies
v, occasionally used; V, frequently used

year after year with a unified and consistent method. Details of this method are available on the internet (http://www.fao.org/waicent/faostat/-agricult/fbs-e.htm). The information provided by FBS is in fact an estimate of the quantity of the various food commodities available for human consumption, after accounting for post-harvest losses. Post-harvest losses are particularly important for perishable foods, including fruit and vegetables, especially in developing countries. However, they do not account for wastage of edible foods at the household level. Thus, FBS data are more correctly referred to as disappearance or availability figures. The information is at a country level and provides no insight into intra-country differences in food consumption, between either socioeconomic groups or diverse ecological or geographical zones, nor into seasonal variations of the total food supply. A serious limitation of the FBS is the level of aggregation. The category "Vegetables", for example, includes a great variety of specific vegetable commodities, but it is not possible to retrieve any information on these.

The accuracy of FBS depends on the reliability of the underlying basic statistics on the supply and utilization of foods transmitted by each country, and varies therefore between countries. The developing regions of the world tend to have poorer statistics, and their FBS therefore have a larger margin of uncertainty.

Despite these limitations, FBS have the advantage that – having been regularly tabulated every year with a unified and unchanging technique since 1961 – they are the only source of information on worldwide time trends and country differences.

Methods to measure dietary intake at the individual level

Two main approaches are used to estimate dietary intake at the individual level. Questionnaires can be used to obtain information on usual intake during the preceding months or years either as quantities and frequencies of specific foods consumed (quantitative food frequency questionnaires) or the frequencies only (food frequency questionnaires). Alternatively, subjects are asked to report from memory the precise amounts of different foods actually eaten over the last 24 hours (24-h diet recall method) or to record all that they eat at the time of consumption (food consumption diaries or weighed food consumption records).

Questionnaire methods

Comprehensive descriptions and discussions of these methods as well as summaries of strengths and weaknesses of each method have been published (Margetts & Nelson, 1991; National Cancer Institute, 1994; Thompson & Byers, 1994; Willett, 1998a, b).

The most commonly used methods to assess dietary intake in cohort and case–control studies of cancer are food frequency questionnaires (FFQ) and the diet history. In cohort studies, the aim is to assess habitual current diet. In case–control studies, the aim is to assess habitual diet during a reference period before the onset of disease. In order to ascertain individual exposure to fruit and vegetables and other dietary components, information on intake needs to be obtained. However, accurately quantifying and classifying an individual's exposure is complex; measures that provide an estimate of usual intake are designed to minimize the effect of intra-individual variation.

The questionnaires used have differed widely between studies. They vary in the length of the food list, the number of questions, the fruits and vegetables included, how the instrument is structured, what other dietary information is obtained, the method used to address portion sizes and quantification of the data. There is no universally accepted questionnaire, standard interview, database or calculation system for use in epidemiological studies. Most FFQs or diet history questionnaires and interview methods are study-specific, being tailored to specific research questions and to the population being studied. Dietary methods are continually being refined based on methodological research. The many resulting variations in methods can affect estimates of dietary intake of fruits and vegetables in epidemiological studies and their relation to disease outcome.

During surveys with the FFQ and diet history, individuals provide information about intake of specific foods, food groups, dietary practices and/or food preparation methods. The information may be obtained by interview, by self-administered questionnaire or through a combination of these methods. The respondent may be the designated participant or a surrogate respondent. The data obtained are then reduced to summary measures using defined algorithms and food and nutrient databases.

Diet history

A diet history is information about usual intake of the individual's whole diet, usually obtained by interview (Burke, 1947). Detailed information is collected for a specified time period on the type, amount and frequency of foods eaten as well as food preparation practices. Typically a food list is used. Recipe information may be obtained, as well as meal-by-meal information about the time, place and content of meals. There is often a crosscheck feature to ensure complete determination of intake and to check for potential overreporting or double counting by the participant. Data may be collected in written form or directly on a computer using a special program (McDonald *et al.*, 1991).

The strength of this method is that detailed quantified information is collected about usual dietary intake for an extended period of time. Compared with data from the recording and recall methods described later, a diet history covers a longer period of time and provides estimates of usual intake. It provides information on specific fruits and vegetables and about seasonal intake, as well as their consumption in mixed dishes. The method is time-consuming for the respondent and the investigators, but may be less conceptually demanding for respondents than food records or FFQs.

Food frequency questionnaire

Food frequency questionnaires (FFQs) have been the most commonly used method to assess dietary exposure in cohort and case–control studies. Respondents are asked to report their usual daily, weekly or monthly frequency of consumption of each item on a list of specific foods over a recent period of about a year. FFQs were developed during the 1950s and 1960s as the most cost-effective method for large epidemiological studies. Initial versions of the FFQ were designed only to rank individuals according to their relative level of dietary consumption expressed in quantiles, and only the frequency of food consumption was requested of the study subjects. Such questionnaires are reported as non-quantitative FFQs. During the 1980s and 1990s, variants of the FFQ were developed to allow its use in different study contexts and populations and to improve the estimation of individual absolute intake. Different questionnaire designs including standard or individual portion size estimates for all or selected items of the food list can lead to inconsistent reporting. These questionnaires may be described as "semi-quantitative" or "quantitative" FFQs (or dietary questionnaires). Over the last 20 years, there has been a clear methodological shift in epidemiological research from basic FFQs to more quantitative questionnaires, including the so-called dietary history questionnaires (see above).

The FFQ is usually self-administered in cohort studies. Respondents may receive the questionnaire along with any associated instructions and visual aids by mail and are asked to complete it at home and return it by mail. They may also complete the questionnaire at a research study centre; in this case verbal instructions can be provided and the questionnaire may be reviewed and clarified before the participant leaves the centre. In case–control studies, an FFQ may be administered by interviewers.

A core feature of the FFQ is usually a closed list of foods. The length of the list varies considerably between studies. The items included on the list depend on the nature of the investigation (particular foods and nutrients may be of interest); it must be borne in mind that a very detailed questionnaire places a heavy demand on the respondents. In cancer epidemiology, there are hypotheses about the effects of overall intake of fruit and vegetables as well as regarding the effects of individual fruits, vegetables or subcategories. Inaccurate estimates of intake can result from an incomplete listing of fruits and vegetables, while if key fruits or vegetables are neglected or if fruits or vegetables are grouped inappropriately (see Chapter 1), important information regarding intake may be lost.

Krebs-Smith *et al.* (1995) compared data from three surveys in the USA, in which FFQs had different numbers of questions relating to intake of fruit and vegetables. The values for median frequency of total fruit and vegetable intake differed between the surveys, and were associated with the number of questions asked. This pattern was also apparent for total fruit and total vegetables. The pattern did not appear to be accounted for by survey year, differences in the seasons covered or differences in the distribution of subjects by age and sex. In one of the surveys, the responses to a summary question "About how many servings of fruits and vegetables do you eat per day or per week?" indicated a median frequency of consumption substantially lower than that obtained by summing the responses to all individual questions about fruit and vegetable intake. In a pooled analysis of cohort studies of breast cancer and intake of fruit and vegetables, there was a more than four-fold variation between studies in the number of questions about fruit and vegetable intake (Smith-Warner *et al.,* 2001a). The median intake increased with the number of items on the questionnaire. Thus, the number of questions asked is a potential source of heterogeneity between studies, and has implications for the categorization of reported intakes if data from different studies are combined.

Some FFQs provide only a list of foods, without portions specified. Others provide a portion size with each item and the respondent reports the frequency of intake of such a portion. Estimates of servings of food/food groups and nutrient intake are obtained by summing the reported frequencies (and the nutrient levels for each) over all foods. Intake is usually expressed as a mean number of servings per day or as a mean nutrient amount per day (Thompson & Byers, 1994; Willett, 1998a).

The portion sizes typically reflect some standardized approach with common household units (such as cups, ounces or grams) as reported in nutrient databases, although in some cases they may reflect typical local portion sizes. Some FFQs allow the respondent, for each item, to choose a portion size from a list or to record his or her own portion size. FFQs may incorporate questions regarding usual portion sizes for some food items. For studies of fruit and vegetables, this may be particularly important in populations that are relatively well fed and that have access to a wide variety of foods. In general, the ranking of individuals according to intake of specific nutrients seems to be determined largely by reported frequency of intake, with little contribution of inter-individual variation in portion size (Samet *et al.*, 1984; Humble *et al.*, 1987; Hunter *et al.*, 1988; Flegal & Larkin, 1990; Tjonneland *et al.*, 1992; Noethlings *et al.*, 2003), although there are exceptions (Clapp *et al.*, 1991; Block, 1992).

In a study in which cognitive interviewing was used, respondents tended to skip portion-size questions after completing frequency questions (Subar *et al.*, 1995). In a study of women in Sweden who were randomly allocated to receive different questionnaires, mean frequency of consumption was significantly lower for vegetables (and other foods) when portion size questions were included (Kuskowska-Wolk *et al.*, 1992). Moreover, there was an adverse effect on response rate. In an investigation of the validity of using pictures to estimate portion size, 103 volunteers were offered standard dishes, and the weight of the food eaten was compared with weight estimated by recall the next day with the aid of pictures (Faggiano *et al.*, 1992). There was a tendency to overestimate portion size among those who ate smaller portions and to underestimate portion size among those who ate larger portions. However, Blake *et al.* (1989) reported that there were no differences in the ability to estimate portion size between normal-weight and overweight subjects.

It is difficult with FFQs to capture information about fruit and vegetable intake consumed in the form of mixed dishes. Such dishes may be listed as "mixed dishes", pasta dishes, soups, vegetable soups, stews, casseroles, Chinese dishes, ethnic foods, salads etc. The actual fruit and vegetable content of these items varies greatly and no estimate of specific fruits and vegetables will be available. In many cuisines, mixed dishes contribute a large proportion of fruit and vegetable intake. The FFQ method requires respondents to integrate the fruit and vegetable intake from these foods into their report of the separate fruit and vegetable items.

Respondents are asked to report their usual intake for a specified time period. The time frame used is often one year, but varies between studies from as little as one month up to 3–10 years. It is assumed that intake over a recent one-year period reflects longer-term intake.

FFQs can be structured in several ways, most commonly by food group, but sometimes by meal. Cognitive testing indicates that many individuals, when asked to report their usual intake of fruit and vegetables, do so by recalling a typical day (Thompson *et al.*, 2000). When 874 subjects in the USA were randomly assigned to receive one of two brief questionnaires designed for surveillance of fruit and vegetable intake, the questionnaire subdivided to assess intake in different parts of the day gave the best agreement with true habitual intake estimated on the basis of two 24-hour recalls (Thompson *et al.*, 2000). In a study in France of two groups of 20 volunteers, Boutron *et al.* (1989) compared data on intake of foods and nutrients obtained with two interviewer-administered questionnaires, one structured by meals and the other by broad food groups, and a 14-day dietary record. The questionnaire structured by meals gave better correlation with the dietary record than the questionnaire structured by food groups, when the data were analysed either in terms of the relative ranking of subjects or in terms of correlation with absolute intake.

Two other aspects of FFQ structure are whether food items are grouped or listed separately and whether closed or open-ended questions are used (Kuskowska-Wolk *et al.*, 1992; Tylavsky & Sharp, 1995; Subar *et al.*, 2000; Thompson *et al.*, 2002).

Brief food frequency questionnaires

Brief food frequency questionnaires are sometimes used, containing a very abbreviated list of foods. The questionnaire may focus on a specific food group or a limited number of food groups or food items. The food list may comprise groupings of foods and be aimed at characterizing some major dietary components such as fat. Respondents are asked to report their usual frequency of intake for the specified time period, as described above. Such instruments have been used to estimate fat and calcium intake, as well as intake of servings of fruit and vegetables (Block *et al.*, 1990, 2000; Willett, 1998a). As a part of the Behavioral Risk Factor Surveillance System (BRFSS), the US Centers for Disease Control and Prevention (CDC) use a brief telephone-administered questionnaire to assess fat intake with 13 questions and fruit and vegetable intake with six questions (Serdula *et al.*, 1993). In efforts to assess changes in fruit and vegetable intake in response to intervention programmes, a variety of brief questionnaires addressing fruit and vegetable intake have been developed and validated in conjunction with 5-A-Day research programmes (see Chapter 3) and community campaigns (Domel *et al.*, 1993a). Brief methods have been developed in efforts to apply a common measure across studies, reduce cost and participant burden, and to enhance the number and type of individuals who can be reached. Kristal *et al.* (2000) compared the validation data from these studies in which 24-hour dietary recalls, food records or serum carotenoid concentrations were used as criterion measures. The validation studies differed in distributions of participants' age, race/ethnicity, sex and socioeconomic status. Mean intakes of total fruit and vegetables based on the 5-A-Day brief method were consistently lower than those from either a much longer FFQ (3.11 versus 4.06), 24-h recalls (3.32 versus 4.07) or food records (3.11 versus 3.46; all $p < 0.01$), and this was due primarily to underestimation of vegetable intake with the brief FFQ method.

These methods have many limitations in the context of epidemiological investigations aimed at understanding associations between fruit and vegetables and cancer risk. They yield very limited information about intake of specific food items crucial to hypotheses about diet and cancer. If they are limited to a single nutrient or food group, information on the total diet and other potential dietary confounders is not available. Because they focus on a few items (particularly fat, fruit and foods that receive a great deal of media attention regarding health consequences), brief questionnaires may suffer from biased reporting based on the subjects' perceptions of what they ought to eat (social desirability bias) and general overreporting of fruit and vegetable intake.

Recording-based measures of actual intake

The 24-hour dietary recall

The aim of the 24-hour dietary recall is to estimate actual dietary intake. An interview is conducted either in person or by telephone, often by a dietitian. The respondent is asked to recall and then report all foods and beverages consumed in the previous 24 hours (sometimes in the preceding day). Respondents are asked to report the amount they consumed typically in household units or weights if known and to provide information about food preparation, brand names and recipes. Photographs of portion sizes may be used. Respondents are asked to report any items added to foods such as condiments, salt, sugar or fats. The interview is usually structured with probes to help the individual remember foods consumed and to provide detailed descriptions of these foods. Data may be recorded using paper forms and subsequently coded and entered on a computer or may be directly entered on the computer with the help of specialized software.

Because the recall covers a recent time period, issues related to memory are reduced. Respondents are not required to be literate and the burden on them may be much lower than with self-administered dietary methods. This generally improves the participation rate. A major strength of the 24-hour recall is that detailed information about all fruit and vegetables and other foods consumed and their specific form (cooked, raw) can be obtained (assuming that a comprehensive food data-base is used). The major limitation regarding fruit and vegetable intake is the short time period covered, since there is considerable day-to-day and season-to-season variation in both the types and the amounts of fruit and vegetables consumed. When only one day of intake is sampled, this approach does not provide a reliable estimate of an individual's intake over longer periods (Beaton *et al.*, 1979, 1983; Todd *et al.*, 1983). Obtaining repeat 24-hour recalls reduces this problem greatly.

The 24-hour recall method was used in some early case–control and cohort studies, before development and widespread use of FFQs, and in clinical trials where the primary purpose of the dietary data was to characterize group intakes. Dietary recall data from clinical trials have been used to evaluate cohorts for subsequent investigations related to diet and cancer.

Food records

Food records are detailed meal-by-meal recordings of the types and quantities of food and drink consumed during a specified period, typically 3–7 days. For a weighed record, or weighed inventory record, the subject weighs all foods consumed during the specified period. A variant of this method does not require the subject to weigh the foods, but to report quantities in terms of household measures or using food models or photographs. This provides detailed information on actual food intake. By having respondents record their intake at the time of consumption, recall problems are minimized and more details about each food item may be available. Such methods may place a considerable burden on the subjects, limiting their application to literate respondents who are highly motivated, and may therefore introduce selection bias, while compliance may produce alterations in diet (Bathalon *et al.*, 2000).

Quantification of fruit and vegetable portions

In general, recalling and reporting sizes of portion sizes of foods consumed is a difficult cognitive task; respondents often have difficulty in estimating weights, volumes and dimensions (Thompson *et al.*, 1987; Smith *et al.*, 1991). Methods to help respondents with reporting and quantification have been developed and good questionnaire design can also improve estimation of fruit and vegetbale intake. Respondents may, depending on the method used, report consumption in units they are most comfortable with or they may have to convert their concept of portion size to those used on a questionnaire. They also may have to adjust their frequency reporting to those specified. Fruits and vegetables vary greatly in size, shape and seasonal availability, how they are prepared and the form consumed. Quantities for fruits and vegetables can be obtained as servings as defined by the respondent, in household units such as cups, or in pieces such as one apple, with dimension descriptions or by weight.

There are many differences between how fruit and vegetables are eaten that affect portion size specification and quantification. Because of the ways fruits are prepared and eaten, they may be easier than vegetables to remember and to quantify. Fruits are often eaten as the single item or combined with other fruits as in a fruit salad or fruit cup. Although pieces of fruit vary in size, there is some uniformity due to modern horticultural and retailing practices for grading and selling fruit based on size. Furthermore, fruit is often consumed as fruit juice, again a discrete item that may be easy to recall and quantify. When juice is sold in individual portions, there is also some standardization of the amount sold. Because fruits are often eaten in specific contexts such as a snack or as a dessert, they may be easier to recall and quantify than vegetables.

Vegetable consumption varies much more. As noted in Chapter 1, there is a wide variety of vegetables consumed by humans and even what is defined as a vegetable varies according to the cultural and research settings. Food preparation and culinary practices vary greatly for vegetables and this affects how they can be quantified. Vegetables may be consumed as a single item (a carrot, corn, artichoke, potato), but are commonly served after some preparation (chopping, slicing, cooking etc.) and as mixtures (soup, stew, pasta dishes, stir fry); many are included in recipes in forms that may not be easily identified by respondents (tomato sauce, chopped onion or garlic). They may also be served as accompaniments to foods in sauces, relishes or sandwiches. These varied ways of serving and eating vegetables make recall and quantification more difficult for respondents and complicate conversion of data from diet assessments to food consumption amounts (either servings or weights) to be used in statistical analyses. There are many nutrients and phytochemicals of interest in cancer epidemiology and even small

amounts of specific fruits or vegetables may contribute importantly to total intake of these. If important sources in the diet are not identified, it may be impossible to adequately classify individual exposure.

Visual aids have been developed to help respondents estimate the amount of foods consumed or the portions typically eaten, including for fruit and vegetables (Margetts & Nelson, 1991; Riboli & Kaaks, 1997). Such visual aids can be used in conjunction with any dietary assessment method. Three-dimensional aids such as food models, actual plates, cups, glasses, spoons or portions of real food displayed in service ware may be shown to respondents during an interview. Two-dimensional printed aids are used in many settings and frequently with FFQs and dietary recalls conducted by telephone. These may be diagrams of food portions (such as portions of meat) or household utensils (measuring cups or spoons) and dishes with portion size indications noted, or be pictures of actual foods on or in appropriate service ware; these may include pictures of several different portion sizes. One study found no great difference between mean intakes reported with use of three-dimensional aids compared with those obtained using two-dimensional diagrams (Posner *et al.*, 1992). A benefit of photographs is they can show regional foods and can display foods in a familiar context, both of which may improve recall and quantification, and this approach has been used with good results in several studies (Pietinen *et al.*, 1988). Visual aids may be used in the interview setting or be provided to participants (by mail or other means) to refer to when they are completing a questionnaire or record.

Some research protocols use more extensive procedures and ask respondents to either measure typical amounts of foods they consume, measure the volume of their usual service ware or weigh their foods before consumption. Training of respondents on how to estimate and report their intake has been shown to improve reporting and portion estimation (Bolland *et al.*, 1988).

Measurement error and validity

Sources of error

Many factors affect the accuracy with which the intake of fruit and vegetables can be measured and contribute to measurement error. Respondent factors and factors associated with the measurement techniques are the two main sources. Respondent factors that may contribute to error include: memory, socio-demographic factors such as age, gender, education, literacy, ethnicity, occupation, cultural background, disease or health status, knowledge and attitudes. Even individuals able to accurately recall their food intake may be influenced by factors such as social desirability that affect how and what intake they report. For fruit and vegetables, respondents may overreport consumption because high intake of these foods is perceived as healthy (Margetts & Nelson, 1991; Hankin & Wilkens, 1994; Willet & Lenart, 1998).

Dietary changes during prospective studies need to be considered in the design and analysis of longitudinal studies, particularly very long ones. In the Potsdam cohort of the EPIC study, 47% of the participants reported making some type of dietary change during the first two years of the study (Bergmann & Boeing, 2002). The reported changes tended to be consistent with dietary guidelines; increased fruit and vegetable consumption and lower fat intake were the most common changes noted.

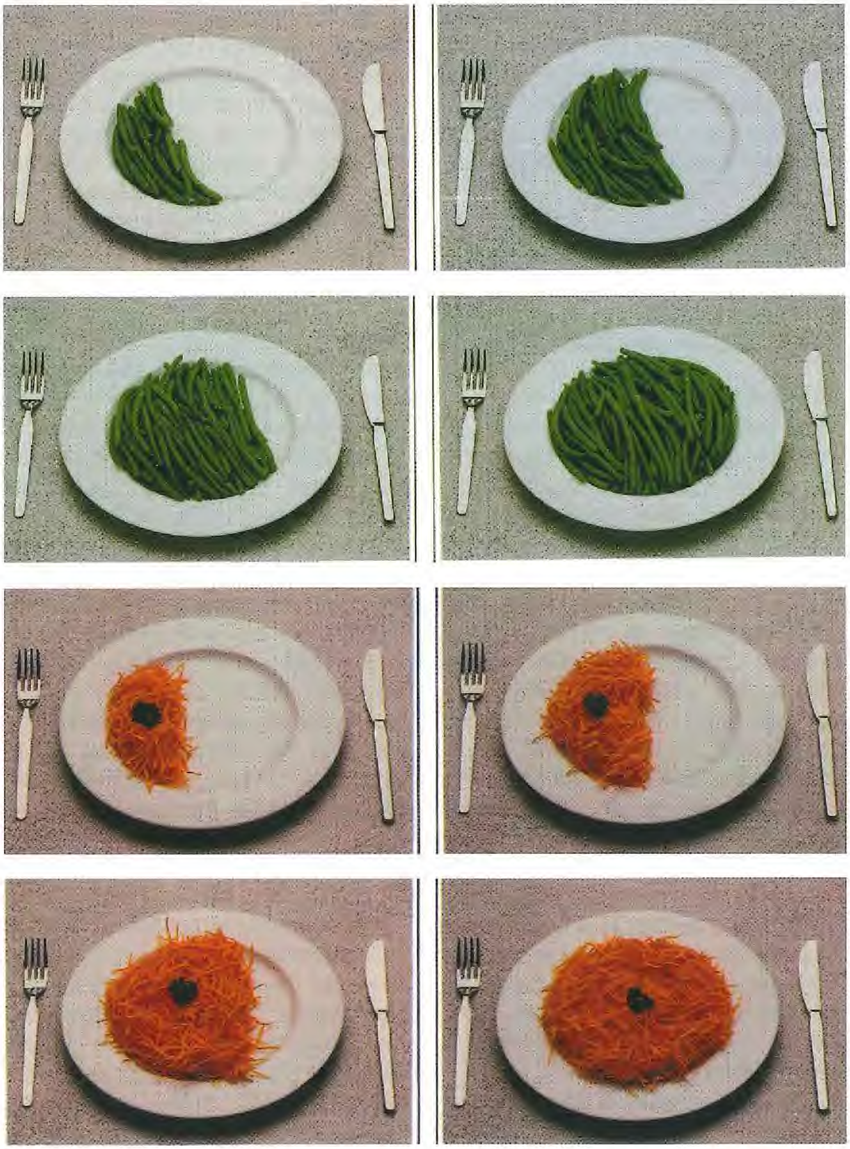
Method-related errors can arise through aspects of questionnaire construction (composition of the food list, specification of portion sizes, grouping of foods into a single item, the order of questions) and the database used to calculate nutrients, food group coding and fruit and vegetable classifications (Margetts & Nelson, 1991).

Validity

A major issue is whether a study aimed to rank individuals according to relative dietary intake or to provide a measure of absolute intake of fruit and vegetables. There continues to be debate about whether FFQs (and to some extent diet history interviews) can accurately assess absolute intake of foods, or can only classify individuals in terms of their relative intake (Block, 2001; Byers, 2001; Willett, 2001d). Validation studies have been used to investigate the extent of misclassification, and this information has sometimes been used to adjust for misclassification within studies (Rosner & Gore, 2001).

Because there is no gold standard for validation of diet methods (Mertz, 1992), a variety of reference methods have been used in validation studies, including 24-hour diet recalls, food consumption records and biological markers (Bingham *et al.*, 1997; Ocké *et al.*, 1997b; Pisani *et al.*, 1997; Riboli & Kaaks, 1997; Smith-Warner *et al.*, 1997; Field *et al.*, 1998; Thompson *et al.*, 2000).

In relation to relative intake, there have been many studies of the reproducibility of dietary instruments, that is, the extent to which different methods applied to the same individuals result in the same ranking of individuals. In order to check that a method is not "consistently wrong", there is a need to compare it with a reference method, usually repeated food records or 24-hour diet recalls.



Examples of food portion sizes from a dietary questionnaire (EPIC), Lyon

The majority of validation studies have compared questionnaire estimates of food consumption or nutrient intake with assessments of current or very recent diet by the reference method. Such an approach is satisfactory for a cohort study investigating the relationship between diet current at the time of assessment and future disease. However, in case–control studies, the relationship between disease and past diet is under investigation, so that, in theory, the reference method should have been applied in the past. In practice, this has rarely been done, and the investigator is forced to make assumptions about the relationship between current and past diet. In studies of chronic disease such as cancer, an additional difficulty is that the disease process itself may have an effect on diet.

Besides various interview and questionnaire methods to assess individuals' intake of fruit and vegetables, biochemical markers of dietary intake have also been proposed (Kaaks *et al.*, 1997a; Hunter, 1998; Crews *et al.*, 2001). Examples of such markers are vitamin C and different types of carotenoid, which can be measured in blood or in adipose tissue (carotenoids only). Measurements of biomarkers can provide complementary information to help assess the performance of different dietary methods, as they should be more objective, depending less on subjects' memory or overall response or cooperation in a study. However, despite initial hopes that markers could be identified that would correlate highly with subjects' true intake of specific dietary compounds or of specific foods, many studies have shown rather weak correlations (Kaaks *et al.*, 1997a; Polsinelli *et al.*, 1998; McEligot *et al.*, 1999; Crews *et al.*, 2001; El Sohemy *et al.*, 2002). Although there may be exceptions (e.g., blood lycopene level as a specific marker of tomatoes and tomato products), these low correlations make markers less attractive than traditional dietary assessment methods as the main exposure assessment method for epidemiological studies (see Chapter 4). Nevertheless, markers can be of some use as an additional reference measurement in validation studies (see also below).

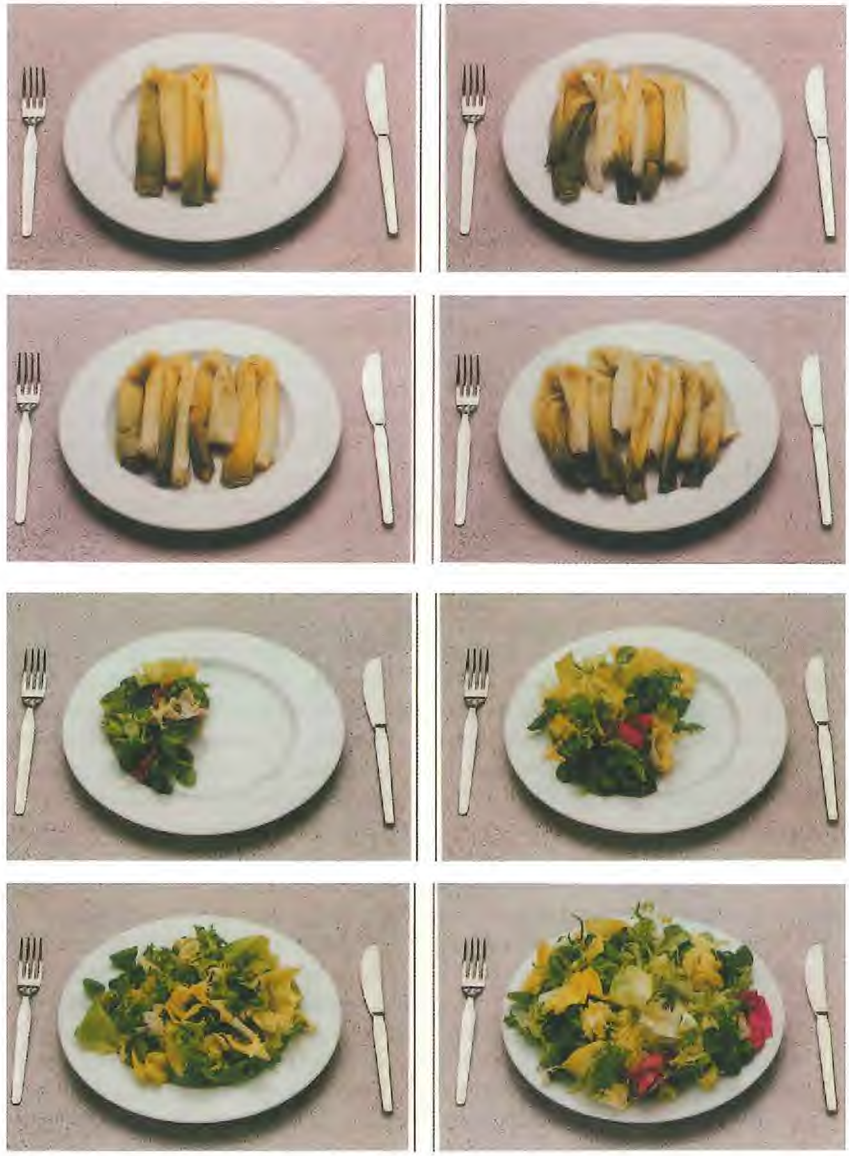
Effects of dietary measurement error

Dietary intake assessments are generally imperfect and generally contain errors. The overall measurement error

– simply defined as the difference between individuals' measured and true intake levels – can be decomposed into systematic and random components (Kaaks *et al.*, 1994a; Willett, 1998c; Kipnis *et al.*, 1999).

Constant or proportional scaling biases may occur when, on average, study subjects tend to over- or underestimate intake by, respectively, a constant amount or by an amount that is proportional to the subjects' true intake levels. This type of error may cause bias in relative risk estimates for quantitative differences in dietary intake levels expressed on an interval scale. In addition, between-population differences in scaling errors may complicate the pooling of data across different populations (Kaaks *et al.*, 1994b). Scaling errors, however, will not affect relative risk estimates for subjects classified into diferent quantile categories of the population distribution of intake levels.

In contrast to scaling errors, random (between-subject) error components are neither constant nor structurally related to subjects' true intake level (Kaaks *et al.*, 1994a) and generally tend to lead to underestimation ("attenuation") of measures of association between diet and disease, with a substantial loss of statistical power to detect these associations. The underestimation of relative risk and associated loss of statistical power depend on the correlation, ρ_{QT}, between the questionnaire measurements of intake and the true habitual intake levels. Assuming that both true and measured intake levels follow an approximately normal distribution, the relationship between the relative risk observed for quantile categories (e.g., quartiles or quintiles) of intake measurements and the relative risk for the same quantile categories of true intake levels can be written as (de Klerk *et al.*, 1989)



$$RR_O = (RR_T)^{\rho_{QT}} \qquad [1]$$

From this mathematical relationship [1] of estimated versus true relative risks for given proportions (e.g., quintiles) of the population ranked into low and high intake levels, it follows that estimates of population attributable risk, as well as relative risk, will also be biased by random measurement error.

Relative risks estimated for a quantitative difference in intake levels expressed on an absolute (interval) scale (e.g., relative risk for a 100 gram increase in total vegetable intake) will be also biased by random error. Here, the mathematical relationship between true and estimated relative risks can be written as

$$RR_O = (RR_T)^{\rho^2_{QT}} \qquad [2]$$

Approaches to evaluating impact of dietary assessment error

To correct for attenuation bias in measures of diet–disease associations, the correlation coefficient ρ_{QT} can be estimated in validation studies and, using either equation [1] or [2], corrected relative risk estimates can be obtained from initial, 'crude' estimates based on questionnaire assessments (de Klerk *et al.*, 1989; Rosner *et al.*, 1989; Kaaks *et al.*, 1995). Especially within prospective cohort studies, validation studies have been increasingly included as a standard part of the overall design (Willett *et al.*, 1985; Colditz *et al.*, 1986; Goldbohm *et al.*, 1994; Margetts & Pietinen, 1997; Stram *et al.*, 2000; Hankin *et al.*, 2001; Slimani *et al.*, 2002). Validity is estimated for measurements obtained by a given questionnaire within a specific study context, rather than for the method itself, which may not perform the same way in other contexts. It is crucial that validity studies be conducted in a representative subsample of the main study population.

Most validation studies have been based on a comparison with repeated daily intake methods for a number of days. The correlation ρ_{QT} can then be estimated by

1. calculation of a crude correlation coefficient ρ_{QR} between questionnaire measurements and individuals' average intake estimates from several days of food consumption records;
2. estimation of the residual error variance in the reference measurements (average food consumption records) themselves, and calculation of an attenuation coefficient by which the estimate ρ_{QR} would need to be corrected, to yield a more unbiased estimate of ρ_{QT} (Rosner & Willett, 1988).

The second step of this estimation procedure should correct for residual random error in the reference measurements. In the early 1990s, this approach of estimating ρ_{QT} was extended, using models in which subjects' true dietary intake levels are considered a 'latent variable' (Plummer & Clayton 1993a, b; Kaaks *et al.*, 1994a).

The most important assumption that underlies any type of validity study is that different types of measurement being compared – from questionnaires, recording methods or biomarkers – will be correlated exclusively because they all measure the same underlying latent variable (true intake). This means that random errors must be uncorrelated between the different types of measurement compared (Plummer & Clayton 1993a, b; Kaaks *et al.*, 1994a, 2002). Unfortunately, there is increasing evidence that generally errors may not be entirely independent between questionnaire assessments of habitual dietary intake and measurements obtained by a recording method, assessing actual food consumption on a number of days. In particular, it has been shown that individuals vary systematically in their tendency to over- or underreport dietary intakes, not only when using the same measurement method, but even when different questionnaire and/or recording methods are used (Livingstone *et al.*, 1990; Black & Cole, 2001; Livingstone & Black, 2003). Thus, errors that are random between individuals may be partially systematic within subjects ("subject-specific biases") and this will result in positive correlations between random errors in different intake measurements from the same individual. A positive correlation between random errors of questionnaire measurements and the reference measurements used tends to cause overestimation of ρ_{QT}. On the other hand, a positive correlation between random errors of replicate dietary intake records, as the reference, can lead to incomplete adjustment for attenuation bias in estimates of ρ_{QT}. In practice, it is difficult to predict the balance between the two possible and opposite biases in estimating ρ_{QT} (Kipnis *et al.*, 2001; Kaaks *et al.*, 2002). This problem of correlated measurement errors can only be partially, if at all, overcome by the use of available biomarkers, depending on the type of nutrient of food group considered (Plummer & Clayton, 1993a,b; Kaaks *et al.*, 1994a, 2002; Kipnis *et al.*, 2001).

Estimated validity of measured fruit and vegetable consumption

Table 8 shows the estimated correlation ρ_{QT} for total fruit and total vegetable intake, from a number of validity studies. Correlation coefficients were within a range of about 0.30 to 0.76, and were generally estimated with rather wide confidence intervals, due to the limited size of studies (generally 50–150 subjects). From equation [2], it can be estimated that, with a correlation of $\rho_{QT} = 0.30$ and a true relative risk of 3.0 between highest and lowest exposure categories (e.g., quintiles of the intake distribution), the observed relative risk would be as low as 1.10. For a correlation of $\rho_{QT} = 0.7$, the estimated relative risk would be less attenuated but still only 1.7. Thus, as illustrated by this numerical example, there will generally be considerable attenuation bias in relative risk estimates for quantile categories of intake and this may lead to substantial loss of statistical power to detect a real association.

Table 8. Estimated validity of fruit and vegetable intake for questionnaires (FFQ) for selected study populations

Population, reference	No. items on FFQ	Reference method	No. repeated measures	Type of correlation*	Men			Women		
					N	Fruit	Vegetables	*N*	Fruit	Vegetables
The Netherlands cohort, Goldbohm *et al.*, 1994	150 (21 veg., 8 fruits)	Diet records	3, 3 days	S, c	107 (59 M + 48 W)	0.60	0.38	–	–	–
Hawaii cohort, Hankin *et al.*, 1991	Diet history (47)	Food records	4, 1 week	ICC	128	0.60	0.39	134	0.34	0.19
EPIC – France, van Liere *et al.*, 1997	Diet history (101)	24-h DR	12	S, c	–	–	–	115	0.44	0.50
EPIC – Germany, Bohlscheid-Thomas *et al.*, 1997; Kaaks *et al.*, 1997	158	24-h DR	12	S, c	49	0.33	0.39	55	0.45	0.53
EPIC – Italy, Pisani *et al.*, 1997	47	24-h DR	12	S, c	47	0.56	0.30	150	0.39	0.45
EPIC – Netherlands, Ocké *et al.*, 1997b	79	24-h DR	12	S, c	63	0.68	0.38	58	0.56	0.31
EPIC – Spain, The EPIC group of Spain, 1997	Diet history (17)	24-h DR	12	P, c	46	0.76	0.73	45	0.66	0.65
EPIC – Sweden, Kaaks *et al.*, 1997b	130	24-h DR	12	S, c	44	0.72	0.42	559	0.62	0.49
Health professionals cohort, Feskanich *et al.*, 1993	122	Diet records	2, 1 week	P, d	127	0.75#	0.46#	–	–	–
Minnesota Cancer Prevention diet intervention trial, Smith-Warner *et al.*, 1997	153 (33 veg, 18 fruits)	Diet records	5, 3 days	P, d	101 (71 M + 30 W)	0.67	0.32	–	–	–
Finnish lung cancer intervention trial, Pietinen *et al.*, 1988	276	Food records	12, 2 days	P, d	158	0.69	0.58	–	–	–

*: S, Spearman; P, Pearson; ICC, intraclass correlation coefficient; c, crude; d, deattenuated
Median of reported values for individual fruits or vegetables DR, dietary recall