The possibility that diet may be important in the cause and prevention of cancer in humans has received major attention only in the past 25 years, despite long-standing knowledge that tumor incidence in animals can be influenced by nutritional manipulation (Tannenbaum, 1942). In an extensive review of causes of cancer, Doll and Peto (Doll and Peto, 1981) suggested that 35% of cancer deaths in the United States might be due to dietary factors. However, this estimate was highly uncertain, as they believed that the effect of diet could actually be as low as 10% or as high as 70%. The specific dietary factors that may cause or prevent cancer were also uncertain and have been the focus of much subsequent investigation.

Possible relationships of diet with specific cancer sites are discussed within the appropriate chapters of this book. For this reason and because information on diet and cancer is still rapidly evolving, this chapter will be devoted primarily to a review of approaches and methods that are used to study relationships of diet and cancer, including considerations of their strengths and limitations. Issues involved in epidemiological studies of diets are discussed in more detail elsewhere (Willett, 1998). This chapter will conclude by briefly noting some dietary factors that have received particular attention as being possibly related to risk of certain cancers.

GENERAL APPROACHES TO THE STUDY OF DIET AND CANCER

Hypotheses and supporting evidence relating dietary factors to cancer can be obtained from a variety of sources, including in vitro studies, animal experiments, small intervention studies using biochemical or molecular outcomes, epidemiological observations, and randomized trials using cancer or premalignant changes as the end points.

In Vitro Studies and Animal Experiments

Many substances that cause mutations among microorganisms also cause cancer in animals and humans. This observation underlies the usefulness of microbial mutagenicity tests (such as the Ames test), which have been widely used to study components of human diets. These tests are attractive because results are available in only days and at a relatively low cost. Although these tests are helpful in directing human research and elucidating mechanisms of action, they cannot by themselves provide information that is directly relevant to humans (Ames et al., 1987). For example, many substances that influence the risk of cancer are not mutagenic. They may act, for example, by affecting the permeability of host tissues to carcinogens, by altering hormonal balances that inhibit or promote tumor growth, by changing the immune response of the host, or by affecting the rate of cell division, which in turn influences the likelihood that a mutation is reproduced. Because these functions are not all replicated in bacterial testing systems, false-negative and false-positive results will appear.

Experimental exposure of laboratory animals to substances that may influence cancer incidence is more likely to simulate the effect of a chemical or food on the incidence of cancer in humans. High doses of potential carcinogens that do not reflect human experience have typically been used to induce tumors, and the response to different diets was then assessed. Animal models are now available with genetic alterations, such as a mutated tumor suppressor gene, that are more likely to simulate human carcinogenesis. However, species differ in many ways, such as the function of their enzymatic systems to activate or deactivate potentially carcinogenic substances. Such factors preclude direct extrapolation of findings from animal experiments to humans, even though they may provide critical direction for research and aid in the interpretation of epidemiological studies.

Metabolic and Biochemical Studies

Another approach to the study of diet and cancer involves metabolic or biochemical studies in humans. For example, Goldin and co-workers (Goldin et al., 1981) have studied the effect of diet on estrogen profiles, which in turn are thought to be related to the risk of breast cancer. Markers of DNA damage provide another salient outcome. These studies do not address the relations between dietary intake and the occurrence of cancer directly, but they can also be invaluable in the interpretation of other forms of evidence.

Epidemiological Studies

Epidemiological studies of diet and cancer constitute a relatively new area of research. Until the 1980s, many nutritionists and epidemiologists thought that the difficulties of assessing the diets of free-living human beings over extended periods of time made large-scale studies impossible. However, approaches for assessing dietary intake have been developed and have been shown to be informative. These include standardized questionnaires to assess intakes of foods from which nutrient intakes can be calculated, biochemical determinations of body tissues, and anthropometric measurements. Because the measurement of dietary intake is a central issue, these methods will be discussed in more detail later.

Correlation Studies

Early epidemiological investigations of diet and cancer consisted largely of "ecological" or "correlational" studies—comparisons of disease rates in populations with the population per capita consumption of specific dietary factors. The dietary information in such studies is usually based on "disappearance" data, meaning the national figures for food produced and imported minus the food that is exported, fed to animals, or otherwise not available for humans. Many of the correlations based on such information are remarkably strong; for example, the correlation between meat intake and incidence of colon cancer is 0.85 for men and 0.89 for women (Armstrong and Doll, 1975).

The use of international correlational studies to evaluate the relationships between diet and cancer has several strengths. Most importantly, the contrasts in dietary intake are typically large. For example, within the United States, most individuals consume between 25% and 45% of their calories from fat (Willett et al., 1987), whereas the mean fat intake for populations in different countries varies from approximately 15% to 42% of calories (Goodwin and Boyd, 1987). Second, the average of diets for persons residing in a country are likely to be more stable over time than are the diets of individual persons within the country; for most countries, the changes in per capita dietary intakes over a decade or two are relatively small. Finally, the cancer rates on which international studies are based are usually derived from relatively large populations and are therefore subject to only small random errors.
The primary problem of such correlational studies is that many potential determinants of cancer other than the dietary factor under consideration may vary between areas with a high and low incidence of disease. Such confounding factors can include genetic predisposition; other dietary factors, including the availability of total energy intake; and other environmental or lifestyle practices. For example, with few exceptions, countries with a low incidence of colon cancer tend to be economically undeveloped. Therefore, any variable related to industrialization will be similarly correlated with incidence of colon cancer. Indeed, the correlation between gross national product and colon cancer mortality rate is 0.77 for men and 0.69 for women (Armstrong and Doll, 1975). More complex analyses can be conducted of such ecological data that control for some of the potentially confounding factors. For example, McKeean-Eysen and Bright-See (McKeown-Eyssen and Bright-See, 1985) found that an inverse association of per capita dietary fiber intake and national colon cancer mortality rates persisted after adjustment for fat intake.

Most correlational studies are also limited by the use of disappearance data that are only indirectly related to intake and are likely to be of variable quality. For example, the higher “disappearance” of calories per capita for the United States compared with most countries is probably related in part to wasted food in addition to higher actual intake. In addition, aggregate data for a geographical unit as a whole may be only weakly related to the diets of those individuals at risk of disease. As an extreme example, the interpretation of correlational data regarding alcohol intake and breast cancer is complicated because, in some cultures, most of the alcohol is consumed by men, but it is the women who develop breast cancer. These issues of data quality can potentially be addressed by collecting information on actual dietary intake in a uniform manner from the population subgroups of interest. This has been done in a study conducted in 65 geographical areas within China that are characterized by an unusually large variation in rates of many cancers (Chen et al., 1990). Another serious limitation of the international correlational studies is that they cannot be independent of the procedure, which is an important part of the scientific process. Although the dietary information can be improved and the analyses can be refined, the data will not really be independent even as more information becomes available over time; the populations, their diets, and the confounding variables will be the same. Thus, it is not likely that many new insights will be obtained from further ecological studies among countries.

The role of correlational studies in nutritional epidemiology is controversial. Clearly, these analyses have stimulated much of the current research on diet and cancer, and in particular they have emphasized the major differences in cancer rates among countries. Traditionally, such studies have been considered the weakest form of evidence, primarily due to the potential for confounding by factors that are difficult to measure and control (Kinlen, 1983). Some have argued that such studies provide strong evidence for evaluating hypotheses relating diet to cancer (Hebert and Miller, 1988; Prentice et al., 1988). On balance, ecological studies have unquestionably been useful but are far from conclusive regarding the relationships between dietary factors and disease and may sometimes be highly misleading.

Special Exposure Groups
Subgroups within a population that consume unusual diets provide an additional opportunity to learn about the relation of dietary factors and disease. These groups are often defined by religious or ethnic characteristics and provide many of the same strengths as ecological studies. In addition, the special populations often live in the same general environment as the comparison group, which may somewhat reduce the number of alternative explanations for any differences that might be observed. For example, the observation that colon cancer mortality in the largely vegetarian Seventh-Day Adventists is only about half that expected (Phillips et al., 1980) has been used to support the hypothesis that meat consumption is a cause of colon cancer.

Findings based on special exposure groups are subject to many of the same limitations as ecological studies. Many factors, both dietary and nondietary, are likely to distinguish these special groups from the comparison population. Thus, another possible explanation for the lower colon cancer incidence and mortality among the Seventh-Day Adventist population is that differences in rates are attributable to a lower use of alcohol and tobacco or higher vegetable consumption. Given the many alternative explanations, such studies may be particularly useful when a hypothesized association is not observed. For example, the finding that the breast cancer mortality rate among the Seventh-Day Adventists is not appreciably different from the rate among the general United States population provides fairly strong evidence that eating meat is not a major cause of breast cancer.

Migrant Studies and Secular Trends
Migrant studies have been particularly useful in addressing the possibility that the correlations observed in the ecological studies are due to genetic factors. For most cancers, populations migrating from an area with its own pattern of cancer incidence rates acquire rates characteristic of their new location (Adelstein et al., 1979; McMichael and Giles, 1988; Staszewski and Haenszel, 1965; Ziegler et al., 1993), although, for a few tumor sites, this change occurs only in later generations (Buell, 1973; Haenszel et al., 1972). Therefore, genetic factors cannot be primarily responsible for the large differences in cancer rates among these countries. Migrant studies may also be useful for examining the latency or relevant time of exposure.

Major changes in the rates of a disease within a population over time provide evidence that nongenetic factors play an important role in the etiology of that disease. In Iceland, for example, rates of breast cancer rose dramatically over the first half of this century (Bjarnason et al., 1974). These secular changes clearly demonstrate that environmental factors, possibly including diet, are primary causes of this disease, even though genetic factors may still influence who becomes affected given an adverse environment.

Case-Control and Cohort Studies
Many of the weaknesses of correlational studies are potentially avoidable in case-control studies (in which information about previous diet is obtained from diseased patients and compared to that of subjects without the disease) or cohort investigations (in which information on diet is obtained from disease-free subjects who are then followed to determine disease rates according to levels of dietary factors). In such studies, the confounding effects of other factors can be controlled either in the design (by matching subjects to be compared on the basis of known risk factors, or by restriction) or in the analysis (by any of a variety of multivariate methods) if information has been collected on the confounding variables. Furthermore, dietary information can be obtained for the individuals actually affected by disease rather than using the average intake of the population as a whole.

Case-control studies generally provide information more efficiently and rapidly than cohort studies because the number of subjects is typically far smaller and no follow-up is necessary. However, concerns have existed whether consistently valid results can be obtained from case-control studies of dietary factors and disease because of the inherent potential for methodological bias. This potential for bias is not unique for diet but is likely to be unusually serious for several reasons. Due to the limited range of variation in diet within most populations and some inevitable error in measuring intake, realistic relative risks in most studies of diet and disease are likely to be modest, say of the order 0.5 to 2.0. These relative risks may seem small but would be quite important because the prevalence of exposure is high. Given typical distributions of dietary intake, these relative risks are usually based on differences in means for cases and controls (or those who become cases and those who remain noncases in prospective studies) of only about 5%. Thus, a systematic error of even 3% or 4% can seriously distort such a relationship. In case-control studies, it is plausible that biases (due to selection or recall) of this magnitude could often occur, and it is extremely difficult to exclude the possibility that this degree of bias has occurred in any specific study. Hence, it is not surprising that case-control studies of dietary factors provide inconsistent findings.

The selection of an appropriate control group for a study of diet and cancer is also usually problematic. One common practice is to use...
provide the opportunity to obtain repeated assessments of diet over time and to examine the effects of diet on a wide variety of diseases, including total mortality, simultaneously. In addition to being less susceptible to bias, prospective cohort studies have the advantage that they may also be followed passively by means of disease registries and vital record listings (Stampfer et al., 1984). Because diet is particularly associated with the level of general health consciousness, the diets of those who participate are likely to differ from those who do not. For example, controls who participate are likely to be more health-conscious and thus to consume more fruits and vegetables and less animal fat, which would tend to create artifactual inverse and positive associations, respectively, with cancer risk. Direct evidence regarding the magnitude of biases in case-control studies of diet is limited. In two large prospective studies of diet and cancer, diets of patients with breast cancer and a sample of control participants were also assessed retrospectively. In one study, no evidence of recall bias was observed (Friedenreich et al., 1991), but in the other the combination of recall and selection bias did seriously distort associations with fat intake (Giovannucci et al., 1993b). Even if many studies arrive at correct conclusions, distortion of true associations in a substantial percentage would produce an inconsistent body of published data, making a coherent synthesis difficult or impossible for a specific diet and cancer relationship. Methodological sources of inconsistency may be particularly troublesome in nutritional epidemiology because of the inherent biological complexity resulting from nutrient–nutrient interactions. Because the effect of one nutrient may depend on the level of another (which can differ between studies and may not have been measured), such interactions may result in apparently inconsistent findings in epidemiological studies. Thus, compounding biological complexity with methodological inconsistency may result in an uninterpretable literature.

Prospective cohort studies avoid most of the potential sources of methodological bias associated with case-control investigations. Because the dietary information is collected before the diagnosis of disease, illness cannot affect the recall of diet. Although losses to follow-up that vary by level of dietary factors can result in distorted associations in a cohort study, follow-up rates tend to be rather high because participants have already provided evidence of willingness to participate, and they may also be followed passively by means of disease registries and vital record listings (Stampfer et al., 1984). In addition to being less susceptible to bias, prospective cohort studies provide the opportunity to obtain repeated assessments of diet over time and to examine the effects of diet on a wide variety of diseases, including total mortality, simultaneously.

The primary constraints on prospective studies of diet are practical. Even for common cancers such as those of the lung, breast, or colon, it is necessary to enroll tens of thousands of subjects. The use of structured, self-administered questionnaires has made studies of this size possible, although still expensive. Enough cohort studies have now become available to evaluate whether findings from case-control studies are replicable prospectively. Case-control studies of fat and breast cancer have been summarized in a pooled analysis by Howe and colleagues (Howe et al., 1990); a small but statistically significant positive overall association was seen, although highly significant heterogeneity was also observed among studies. In contrast, when large cohort studies were pooled, no association was seen, and there was no heterogeneity among studies (Hunter et al., 1996) (see Figure 21–1). Even more strikingly, positive associations have been seen consistently in case-control studies of dietary fat and lung cancer, whereas these associations are consistently absent in prospective studies (Smith-Warner et al., 2002). Similarly, suggestions of inverse associations between intake of foods high in beta-carotene and risk of lung cancer were not supported in a pooled analysis of prospective studies (Smith-Warner et al., 2002). These inconsistencies between case-control and cohort studies strongly suggest that concerns regarding the potential for biases in the former are justified, and that conclusions about diet and cancer should not rely on such studies. However, for diseases of somewhat low frequency, even very large cohorts will not accumulate a sufficient number of cases within a reasonable amount of time, and case-control studies will continue to play a role in nutritional epidemiology. Care in design and caution in the interpretation will be important. Due to current uncertainty about measuring diets in early life, whether either study design will be able to address the influence of childhood diet on disease occurring decades later is currently unclear.

**Controlled Trials**

The most rigorous evaluation of a dietary hypothesis is the randomized trial, optimally conducted as a double-blind experiment. The principal strength of a randomized trial is that potentially distorting variables should be distributed at random between the treatment and control groups, thus minimizing the possibility of confounding by these extraneous factors. In addition, it is sometimes possible to create a larger contrast between the groups being compared by use of an active intervention. Such experiments among humans, however, are best justified after considerable nonexperimental data have been collected to ensure that benefit is reasonably probable and that an adverse outcome is unlikely. Experimental studies are particularly practical for evaluating hypotheses that minor components of the diet, such as trace elements or vitamins, can prevent cancer, as these nutrients can be formulated into pills or capsules.

Even if feasible, randomized trials of dietary factors and disease are likely to encounter several limitations. The time between change in

![Figure 21-1. Case-control and cohort studies of dietary fat and breast cancer. Case-control data abstracted from Howe and colleagues (Howe et al., 1990) and prospective data from Hunter and colleagues. (Source: Hunter et al., 1996.)(Image)]
the level of a dietary factor and any expected change in the incidence of disease is typically uncertain. Therefore, trials must be of long duration, and the possibility that any lack of difference between treatment groups may be due to insufficient duration is difficult to exclude. Compliance with the treatment diet is likely to decrease during an extended trial, particularly if treatment involves a real change in food intake, and the control group may well adopt the dietary behavior of the treatment group if the treatment diet is thought to be beneficial. Such trends, which were found in the Multiple Risk Factor Intervention Trial of coronary disease prevention (Multiple Risk Factor Intervention Trial Research Group, 1982), may obscure a real benefit of the treatment.

A related potential limitation of trials is that participants who enroll in such studies tend to be highly selected on the basis of health consciousness and motivation. Therefore, it is possible that the subjects at highest potential risk on the basis of their dietary intake, and thus susceptible to intervention, are seriously underrepresented. For example, if low beta-carotene intake is thought to be a risk factor for lung cancer, and a trial of beta-carotene supplementation is conducted among a health-conscious population that includes few individuals with low beta-carotene intake, no effect might be observed simply because most members of the study population were already receiving the maximal benefit of this nutrient through their usual diet. In such an instance, it would be useful to measure dietary intake of beta-carotene before starting the trial. Because the effect of supplementation is likely to be greatest among those with low dietary intakes, it would be possible to exclude those with high intakes (the potentially nonsusceptibles) either before randomization or in subanalyses at the conclusion of the study. This requires, of course, a reasonable measurement of dietary intake.

Trials are sometimes said to provide a better quantitative measurement of the effect of an exposure or treatment because the difference in exposure between groups is better measured than in an observational study. Although this contrast may at times be better defined in a trial (it is usually clouded by some degree of noncompliance), trials still usually produce an imprecise measure of the effect of exposure because of marginally adequate sample sizes and ethical considerations that require stopping soon after a statistically significant effect is seen. For example, with a P-value close to 0.05, the 95% confidence interval will extend from no effect to a strong effect that is usually implausible. In an observational study, an ethical imperative to stop does not exist when statistical significance occurs; continued accumulation of data can provide increasing precision regarding the relation between exposure and disease. A trial can provide unique information on the latent period between change in an exposure and change in diet; since spontaneous changes in diet are typically not clearly demarcated in time, the estimation of latent periods for dietary effects will usually be difficult in observational studies.

Although all hypotheses would ideally be evaluated in randomized trials, this will sometimes be impossible for practical or ethical reasons. For example, our knowledge of the effects of cigarette smoking on risk of lung cancer is based on observational studies, and it is similarly unlikely that randomized trials could be conducted to examine the effect of alcohol use on human breast cancer risk. It remains unclear whether trials of sufficient size, duration, and degree of compliance can be conducted to evaluate many hypotheses that involve major behavioral changes in eating patterns, such as a reduction in fat intake (Michels and Willett, 1992).

### MEASUREMENT OF DIET IN EPIDEMIOLOGICAL STUDIES

The complexity of the human diet represents a daunting challenge to anyone contemplating a study of its relation to cancer. The foods we consume each day contain literally thousands of specific chemicals, some known and well quantified, some characterized only poorly, and others completely undescribed and currently unmeasurable. In human diets, intakes of various components tend to be intercorrelated. With few exceptions, all individuals are exposed; for example, everyone eats fat, fiber, and vitamin A. Thus, dietary exposures can rarely be characterized as present or absent; rather, they are continuous variables, sometimes with rather limited range of variation between persons. Furthermore, individuals are generally not aware of the content of the foods that they eat; hence, the consumption of nutrients is usually determined indirectly. The chemicals that constitute our food can be described by the non–mutually exclusive categories given in Table 21-1.

### Nutrients versus Foods

Throughout nutrition in general and in much of the existing cancer literature, diet has usually been described in terms of its nutrient content. Alternatively, diet can be described in terms of foods or food groups. The primary advantage of representing diets as specific compounds, such as nutrients, is that such information can be directly related to our fundamental knowledge of biology. From a practical perspective, the exact structure of a compound must usually be known if it is to be synthesized and used for supplementation. In epidemiological studies, measurement of total intake of a nutrient (as opposed to using the contribution of only one food at a time) provides the most powerful test of a hypothesis, particularly if many foods contribute to intake of that nutrient. For example, in a particular study it is quite possible that total fat intake could be clearly associated with risk of disease, whereas none of the contributions to fat intake by individual foods would be significantly related to disease on its own.

The use of foods to represent diet has several practical advantages when examining relationships with disease. Particularly when suspicion exists that some aspect of diet is associated with risk but a specific hypothesis has not been formulated, an examination of the relations of foods and food groups with risk of disease will provide a means to explore the data. Associations observed with specific foods may lead to a hypothesis relating to a defined chemical substance. For example, observations that higher intakes of green and yellow vegetables were associated with reduced rates of lung cancer led to the hypothesis that beta-carotene might protect DNA from damage caused by free radicals and singlet oxygen (Peto et al., 1981). The finding by Graham and co-workers (Graham et al., 1978) that intake of cruciferous vegetables was inversely related to risk of colon cancer supported the suggestion that indole compounds contained in these vegetables may be protective (Wattenberg and Loub, 1978).

Even more seriously than the lack of a well-formulated hypothesis, the premature focus on a specific nutrient that turns out to have no relation with disease may lead to the erroneous conclusion that diet has no effect. Mertz (Mertz, 1984) has pointed out that foods are not necessarily related to disease on its own. The use of foods to represent diet has several practical advantages when examining relationships with disease. Particularly when suspicion exists that some aspect of diet is associated with risk but a specific hypothesis has not been formulated, an examination of the relations of foods and food groups with risk of disease will provide a means to explore the data. Associations observed with specific foods may lead to a hypothesis relating to a defined chemical substance. For example, observations that higher intakes of green and yellow vegetables were associated with reduced rates of lung cancer led to the hypothesis that beta-carotene might protect DNA from damage caused by free radicals and singlet oxygen (Peto et al., 1981). The finding by Graham and co-workers (Graham et al., 1978) that intake of cruciferous vegetables was inversely related to risk of colon cancer supported the suggestion that indole compounds contained in these vegetables may be protective (Wattenberg and Loub, 1978).

More seriously than the lack of a well-formulated hypothesis, the premature focus on a specific nutrient that turns out to have no relation with disease may lead to the erroneous conclusion that diet has no effect. Mertz (Mertz, 1984) has pointed out that foods are not necessarily related to disease on its own. The use of foods to represent diet has several practical advantages when examining relationships with disease. Particularly when suspicion exists that some aspect of diet is associated with risk but a specific hypothesis has not been formulated, an examination of the relations of foods and food groups with risk of disease will provide a means to explore the data. Associations observed with specific foods may lead to a hypothesis relating to a defined chemical substance. For example, observations that higher intakes of green and yellow vegetables were associated with reduced rates of lung cancer led to the hypothesis that beta-carotene might protect DNA from damage caused by free radicals and singlet oxygen (Peto et al., 1981). The finding by Graham and co-workers (Graham et al., 1978) that intake of cruciferous vegetables was inversely related to risk of colon cancer supported the suggestion that indole compounds contained in these vegetables may be protective (Wattenberg and Loub, 1978).

Even more seriously than the lack of a well-formulated hypothesis, the premature focus on a specific nutrient that turns out to have no relation with disease may lead to the erroneous conclusion that diet has no effect. Mertz (Mertz, 1984) has pointed out that foods are not necessarily related to disease on its own. The use of foods to represent diet has several practical advantages when examining relationships with disease. Particularly when suspicion exists that some aspect of diet is associated with risk but a specific hypothesis has not been formulated, an examination of the relations of foods and food groups with risk of disease will provide a means to explore the data. Associations observed with specific foods may lead to a hypothesis relating to a defined chemical substance. For example, observations that higher intakes of green and yellow vegetables were associated with reduced rates of lung cancer led to the hypothesis that beta-carotene might protect DNA from damage caused by free radicals and singlet oxygen (Peto et al., 1981). The finding by Graham and co-workers (Graham et al., 1978) that intake of cruciferous vegetables was inversely related to risk of colon cancer supported the suggestion that indole compounds contained in these vegetables may be protective (Wattenberg and Loub, 1978).

### Table 21-1. Aspects of Diet Related to Cancer

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essential nutrients</td>
<td>Vitamins, specific fatty acids, amino acids, and minerals</td>
</tr>
<tr>
<td>Major energy sources</td>
<td>Proteins, carbohydrates, fats, and alcohol</td>
</tr>
<tr>
<td>Additives</td>
<td>Preservatives (nitrates, BHT, salt) and coloring and flavoring agents</td>
</tr>
<tr>
<td>Agricultural chemical contaminants</td>
<td>Pesticides, herbicides, fungicides, growth hormones</td>
</tr>
<tr>
<td>Microbial toxin contaminants</td>
<td>Aflatoxin</td>
</tr>
<tr>
<td>Inorganic contaminants</td>
<td>Cadmium, lead, polychlorinated biphenyls</td>
</tr>
<tr>
<td>Chemicals formed in cooking or processing of food</td>
<td>Heterocyclic amines from cooked meat</td>
</tr>
<tr>
<td>Natural toxins</td>
<td>Saffrole (in natural root beer), pyrrolizidine alkaloid (in comfrey tea), hydrzones (in mushrooms)</td>
</tr>
<tr>
<td>Other natural compounds</td>
<td>Protease inhibitors, indoles, cholesterol</td>
</tr>
</tbody>
</table>
Epidemiological analyses based on foods, as opposed to nutrients, are generally most directly related to dietary recommendations because individuals and institutions ultimately manipulate nutrient intake largely by their choice of foods. Even if the intake of a specific nutrient is convincingly shown to be related to risk of cancer, this is not sufficient information on which to make dietary recommendations. Because foods are an extremely complex mixture of different chemicals that may compete with, antagonize, or alter the bioavailability of any single nutrient contained in food, it is not possible to predict with certainty the health effects of any food solely on the basis of its content of one specific factor. For example, there is concern that high intake of nitrates may be deleterious, particularly with respect to gastrointestinal cancer. However, the primary sources of nitrates in our diets are green, leafy vegetables, which, if anything, appear to be associated with reduced risk of cancer at several sites. Similarly, because of the high cholesterol content of eggs, their avoidance has received particular attention in diets aimed at reducing the risk of coronary heart disease. Per capita consumption of eggs declined by 25% in the United States between 1948 and 1980 (Welsh and Marston, 1982). However, eggs are more than cholesterol capsules; they provide a rich source of essential amino acids and micronutrients and are relatively low in saturated fat. It is thus difficult to predict the net effect of egg consumption on risk of coronary heart disease, much less the effect on overall health, without empirical evidence.

Given the strengths and weaknesses of using nutrients or foods to represent diet, it appears that an optimal approach to epidemiological analyses will employ both. In this way, a potentially important finding is least likely to be missed. Moreover, the case for causality is strengthened when an association is observed with overall intake of a nutrient and also with more than one food source of that nutrient, particularly when the food sources are otherwise different. This provides, in some sense, multiple assessments of the potential for confounding by other nutrients; if an association was observed for only one food source of the nutrient, other factors contained in that food would tend to be similarly associated with disease. As an example, the hypothesis that alcohol intake causes breast cancer was strengthened by observing not only an overall association between alcohol intake and breast cancer risk, but also by independent associations with both beer and liquor intake, thus making it less likely that some factor other than alcohol in these beverages was responsible for the increased risk.

One practical drawback of using foods to represent diet is that their large number and complex, often reciprocal, interrelationships that are largely due to individual behavioral patterns. Many reciprocal relationships emerge upon perusal of typical data sets; for example, eaters of dark bread tend not to eat white bread, margarine users tend not to eat butter, and skim milk users tend not to use whole milk. This complexity is, of course, one of the reasons to compute nutrient intakes that summarize the contributions of all foods.

An intermediate solution to the problem posed by the complex interrelationships among foods is to use food groups or to compute the contribution of nutrient intake from various food groups. For example, Manousos and co-workers (Manousos et al., 1983) combined the intakes of foods from several predefined groups to study the relation of diet to risk of colon cancer; they observed increased risk among subjects with high meat intake and with low consumption of vegetables. The computation of nutrient intakes from different food groups is illustrated by the partitioning of dietary fiber, an extremely heterogeneous collection of substances, into fiber from grains, fiber from fruits, and fiber from vegetables. This circumvents the inadequacy of food composition databases and also provides information in a form that is directly useful to individuals faced with decisions regarding choices of foods.

In general, maximal information will be obtained when analyses are conducted at the levels of nutrients, foods, and food groups.
subject, thus limiting its application to those who are literate and highly motivated. In addition, the effort involved in keeping diet records may increase awareness of food intake and induce an alteration in diet. However, diet recording has the distinct advantages of not depending on memory and allowing direct measurements of portion sizes.

The validity of 24-hour recalls has been assessed by observing the actual intake of subjects in a controlled environment and interviewing them the next day. In such a study, Karvetti and Knuts (Karvetti and Knuts, 1985) observed that subjects both erroneously recalled foods that were not actually eaten and omitted foods that were eaten; correlations between nutrients calculated from observed intakes with calculations from the recalled information ranged from 0.58 to 0.74. In a similar study among elderly persons, Madden and co-workers (Madden et al., 1976) found correlations ranging from 0.28 to 0.87. Relatively few validation studies have been conducted of diet recordings. In a comparison of nitrogen intake calculated from diet records with intake based on analyses of replicate meals, Bingham and Cummings (Bingham and Cummings, 1985) found a correlation of 0.97.

The most serious limitation of the 24-hour recall method is that dietary intake is highly variable from day to day. Diet records reduce the problem of day-to-day variation because the average of a number of days is used. For nutrients that vary substantially, however, even a week of recording will still not provide an accurate estimate of an individual’s intake (Beaton et al., 1979). The variability in intake of specific foods is even greater than for nutrients (Salvini et al., 1989), so that only very commonly eaten foods can be studied by this method. The problem of day-to-day variation is not an issue if the objective of a study is to estimate a mean intake for a population, as might be the goal in an ecological study. However, in case-control or cohort investigations, accurate estimation of individual intakes is necessary.

Practical considerations and issues of study design further limit the application of short-term recall and diet record methods in epidemiological studies. Because they provide information on current diet, their use will typically be inappropriate in case-control studies because the relevant exposure will have occurred earlier and diet may have changed as a result of the cancer or its treatment. A few exceptions may occur, such as in the case of very early tumors or premalignant lesions. Although the average of multiple days of 24-hour recalls or diet recording could theoretically be used in prospective studies of diet and cancer, the costs may be prohibitive because of the large numbers of subjects required and substantial expense involved in collecting this information and processing it. These methods, however, can play an important role in the validation or calibration of other methods of dietary assessment that are more practical for epidemiological studies.

**Food Frequency Questionnaire.** Because short-term recall and diet record methods are generally expensive, inappropriate for assessment of past diet, and may be unrepresentative of usual intake, investigators have sought alternative methods for measuring long-term dietary intake. Burke (Burke, 1947) developed a detailed dietary history interview that attempted to assess an individual’s usual diet; this included a 24-hour recall, a menu recorded for 3 days, and a checklist of foods consumed over the preceding month. This method was time-consuming and expensive because a highly skilled professional was needed for both the interview and processing of information. The checklist, however, was the forerunner of the more structured dietary questionnaires in use today. During the 1950s, Stephanik and Trulson (Stephanik and Trulson, 1962), Heady (Heady, 1961), Wiehl and Reed (Wiehl and Reed, 1960), and Marr (Marr, 1971) developed food frequency questionnaires and evaluated their role in dietary assessment. Heady, using diet records collected by British bank clerks, demonstrated that the frequencies with which foods were used correlated highly with the total weights of the same foods consumed over a several-day period, thus providing the theoretical basis for the food frequency method. Multiple investigators have converged toward the use of food frequency questionnaires as the method of dietary assessment best suited for most epidemiological studies of diet and cancer. During recent years, substantial refinement, modification, and evaluation of food frequency questionnaires have occurred, so that data derived from their use has become considerably more interpretable.

The basic food frequency questionnaire consists of two components: A food list and a frequency response section for subjects to report how often each food was eaten. Questions related to further details of quantity and composition may be appended. A basic decision in designing a questionnaire is whether the objective is to measure intake of a few specific foods or nutrients or whether a comprehensive assessment of dietary intake is desired. A comprehensive assessment is generally desirable whenever possible. It is often impossible to anticipate at the beginning all the questions regarding diet that will appear important at the end of a study; a highly restricted food list may not have included an item that is, in retrospect, important. Furthermore, total food intake, represented by energy consumption, may be related to disease outcome and thus confound the effects of specific nutrients or foods. Even if total energy intake is not related to a disease outcome, adjustment for total intake may increase the accuracy of specific nutrient measurements (Willett, 2001, Willett and Stampfer, 1986). Nevertheless, epidemiological practice is usually a compromise between the ideal and reality, and it may simply not be possible to include a comprehensive diet assessment in a particular interview or questionnaire, especially if diet is not the primary focus of the study.

Because diets tend to be reasonably correlated from year to year, most investigators have asked subjects to describe their frequency of using foods in reference to the preceding year. This provides a full cycle of seasons so that, in theory, the responses should be independent of the time of year. In case-control studies, the time frame could be in reference to a period or a specified number of years previously.

Typically, investigators have provided a multiple-choice response format, with the number of options usually ranging from five to ten. Another approach is to use an open-ended format and provide subjects the option of answering in terms of frequency per day, week, or month (Block et al., 1986). In theory, an open-ended frequency response format might provide for some enhanced precision in reporting because the frequency of use is truly a continuous rather than a categorical variable. However, the overall increment in precision is unlikely to be large, because the estimation of the frequency that a food is used is inherently an approximation.

Several options exist for collecting additional data on serving sizes. The first is to collect no additional information on portion sizes at all; that is, to use a simple frequency questionnaire. A second possibility is to specify a portion size as part of the question on frequency; for example, to ask how often a "glass" of milk is consumed rather than only how often milk is consumed, which has been termed a semi-quantitative food frequency questionnaire. A third alternative is to include an additional question for each food to describe the usual portion size, in words (Block et al., 1986), using food models (Morgan et al., 1978), or pictures of different portion sizes (Hankin et al., 1983). Because most of the variation in intake of a food is explained by frequency of use rather than differences in serving sizes, several investigators have found that portion size data are relatively unimportant (Block et al., 1990, Pickle and Hartman, 1985; Samet et al., 1984), Cummings and co-workers (Cummings et al., 1987) found that adding questions on portion sizes to a simple frequency questionnaire only slightly improved estimation of calcium intake; others have found that the use of food models in an in-person interview did not increase the validity of a self-administered, semiquantitative food frequency questionnaire (Hernandez-Avila et al., 1988). These findings have practical implications because the cost of data collection by mail or telephone is far less than the cost of personal interviews, which are necessary if food models are to be used for assessing portion sizes. Cohen and co-workers (Cohen et al., 1990) also found that the portion size information included in the Block questionnaire added only slightly to validity assessed by comparison with diet records (average correlation 0.41 without portion sizes and 0.43 with portion sizes).

Food frequency questionnaires are extremely practical in epidemiological applications because they are easy for subjects to complete,
often as a self-administered form. Processing is readily computerized and inexpensive, so that even prospective studies involving tens of thousands of subjects are feasible.

Validity of Dietary Assessment Methods

The interpretation of epidemiological data on diet and cancer depends directly on the validity of the methods used to measure dietary intake. This is particularly true when no association is found because one possible explanation could be that the method used to measure diet was not able to discriminate among persons. A substantial body of evidence has accumulated regarding the validity of food frequency questionnaires, which has been the method most commonly used in epidemiological studies.

In evaluating the validity of a dietary assessment method, the choice of a standard for comparison is a critical issue. As for all variables, no perfect standard exists. Thus, a desirable feature for the comparison method is that its errors be independent from the method being evaluated so that an artificial correlation will not be observed (Willett, 1998). For this reason, biochemical indicators of diet are probably the optimal standard. Their greatest limitation is that markers of diet do not exist for most of the nutrients of current interest, such as total fat, fiber, and sucrose intake. Moreover, the available biochemical indicators of diet are likely to be quite imprecise measures of diet because they are influenced by many factors such as differences in absorption and metabolism, short-term biological variation, and laboratory measurement error. However, the capacity to demonstrate a correlation between a questionnaire estimate of nutrient intake and a biochemical indicator provides useful qualitative evidence of validity. Such correlations have been reported for questionnaire estimates of dietary carotenoids, folate, vitamin E, vitamin B6, and specific fatty acids (Willett, 1998).

If a biomarker is sensitive to change in a dietary factor, even if not specific to that dietary factor, it can still provide a quantitative estimate of the measurement error in a dietary method. For example, on the basis of many metabolic studies, increasing intake of fat reduces fasting plasma triglyceride levels. Demonstration that fat intake measured by a dietary questionnaire is inversely related to fasting triglyceride levels provides qualitative information on the validity, and comparing the slope from a regression of plasma triglycerides on fat intake with the slope anticipated from controlled feeding studies provides a quantitative assessment of measurement error (Willett et al., 2001). In an application of this approach, the slope using food frequency questionnaires was not substantially lower than that from controlled feeding studies—it was actually somewhat larger. Thus, associations between fat intake estimated by the questionnaire and cancer risks are not likely to be seriously underestimated due to measurement error.

Kipnis and colleagues (Kipnis et al., 2003) have used doubly labeled water and urinary nitrogen to assess protein intake that was also measured by 24-hour recalls and a food frequency questionnaire. This confirmed the reduction in error achieved by adjustment for total energy intake. Although these authors found evidence of correlation in errors between the 24-hour recalls and the food frequency questionnaire, the estimates of validity for the questionnaire after accounting for correlated errors were similar to those reported in previous validation studies (Willett, 2003).

Validation studies of dietary questionnaires have also been conducted by comparing computed intakes with those based on other dietary assessment methods. Among the possible comparison methods, diet records are particularly attractive because their errors are likely to be minimally correlated with errors in the food frequency questionnaires, as they do not depend on memory, and scales can be used to assess portion sizes. Several of the most comprehensive validation studies involving comparisons of questionnaires completed at about a 1-year interval, before and after multiple diet records collected during the intervening months, are summarized in Table 21–2 (Block et al., 1990; Goldbohm et al., 1994; Pietinen et al., 1988a; Pietinen et al., 1988b; Rimm et al., 1992; Willett et al., 1985b). Similar degrees of misclassification were seen in these studies; for questionnaires completed at the end of the 1-year recording of diet (which corresponds to the time frame of the questionnaires), correlations adjusted for total energy intake tended to be mainly between 0.5 and 0.7.

Although the degree of measurement error associated with nutrient estimates calculated from food frequency questionnaires appears to be similar to that for many epidemiological measures, these errors will lead to important underestimates of relative risks. Less commonly appreciated, the errors will also result in observed confidence intervals that are inappropriately narrow; this is of particular concern when no association is seen because the entire interest is then in the range of possible relative risks that are reasonably excluded by the data. In part generated by the interest in diet and cancer and the recognized issue of measurement error in assessing dietary intake, considerable effort has been directed to the development of methods that provide corrected estimates of relative risks and confidence intervals based on quantitative assessments of measurement error (Byar and Gail, 1989; Rosner et al., 1992; Willett, 1998). Thus, validation studies of dietary questionnaires can provide important estimates of error that can be used to quantitatively interpret the influence of error on observed associations. Based on such analyses, it can be shown that important associations will generally not be missed by typical dietary questionnaires.

Table 21–2. Comparison of Food Frequency Questionnaires with Other Dietary Assessment Methods

<table>
<thead>
<tr>
<th>Source</th>
<th>Population</th>
<th>Comparison Methods</th>
<th>Interval Between Methods</th>
<th>Reference Period</th>
<th>Range of Correlations</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willett et al. (1985b)</td>
<td>Registered nurses (N = 194)</td>
<td>Diet record</td>
<td>1 month to 1 year</td>
<td>Previous year</td>
<td>0.36 Vitamin A without supplements to 0.75 vitamin C</td>
<td>Adjustment for energy had little effect on correlations</td>
</tr>
<tr>
<td>Pietinen et al. (1988a, 1988b)</td>
<td>Finnish men (N = 189)</td>
<td>Twelve 2-day diet records (vs. 273-item questionnaire)</td>
<td>1–6 months</td>
<td>1 year</td>
<td>0.51 vitamin A to 0.73 polyunsaturated fat</td>
<td>Correlations were similar in low-fat and usual diet group. Variable portion sizes added little to correlations.</td>
</tr>
<tr>
<td>Block et al. (1990)</td>
<td>260</td>
<td>Three 4-day diet records</td>
<td>1–12 months</td>
<td>6 months</td>
<td>0.37 vitamin A to 0.74 vitamin C with supplements, average = 0.55</td>
<td>Mean correlation increased to 0.65 with adjustment for variation in diet records</td>
</tr>
<tr>
<td>Rimm et al. (1992)</td>
<td>127 U.S. health professionals</td>
<td>Two 1-week diet records</td>
<td>1–12 months</td>
<td>1 year</td>
<td>0.28 for iron to 0.86 for vitamin C with supplements, average = 0.59</td>
<td>Adjustment for sex and energy intake had little effect except for fat intake, changing from 0.72 to 0.52</td>
</tr>
<tr>
<td>Goldbohm et al. (1994)</td>
<td>59 men 48 women</td>
<td>Three 3-day diet records</td>
<td>3–15 months</td>
<td>1 year</td>
<td>0.33 for B1 to 0.75 for polyunsaturated fat, average = 0.64</td>
<td></td>
</tr>
</tbody>
</table>
PART III: THE CAUSES OF CANCER

(Rosner et al., 1989), although sample sizes for studies will need to be several times larger than those estimated, assuming that measurement error did not exist (Walker and Blettner, 1985).

Caution is necessary when interpreting validating studies when the standard method is based on calculated nutrient intakes, such as diet records. For nutrients that vary substantially from one specimen of a food to another, the calculated values may be highly correlated because intakes of each of the foods is measured well by both methods, but neither method is actually measuring the nutrient well. This can occur when the nutrient contents of the food vary greatly with the soil on which it was grown or raised, as is true for selenium, or when the nutrient is sensitive to degradation during processing or storage.

Biochemical Indicators of Diet

The use of biochemical measurements made on blood or other tissues to indicate intake of a nutrient is attractive because this does not depend on the memory or knowledge of the subject. Furthermore, such measurements can be made in retrospect; for example, using blood specimens that have been collected and stored for other purposes.

Choice of Tissues for Analysis. Most commonly, serum or plasma has been used in epidemiological studies to measure biochemical indicators of diet. However, consideration should also be given to red blood cells, subcutaneous fat, hair, and nails. The choices should be governed by the ability of the tissue to reflect dietary intake of the factor of interest; the time-integrating characteristics of the tissue; practical considerations in collecting, transporting, and storing the specimen; and cost. These considerations are examined in detail elsewhere for a number of dietary factors (Hunter, 1998); some general comments are provided here.

Red Blood Cells. For a number of dietary factors, red cells are less sensitive to short-term fluctuations in diet than plasma or serum and may thus provide a better index of long-term exposure. Nutrients that can be usefully measured in red cells include some fatty acids, folic acid, and selenium.

Subcutaneous Fat. Composed primarily of fatty acids, the adipose tissue turns over slowly among individuals with relatively stable weight. For at least some fatty acids, the half-life is of the order 600 days, which makes this an ideal indicator of long-term diet in epidemiological studies. Fat-soluble vitamins such as retinol, vitamin E, and carotenoids are also measurable in subcutaneous fat, although their relations to diet are yet to be clearly established.

Hair and Nails. Hair and nails incorporate many elements into their matrix during formation, and for many heavy metals these may be the tissues of choice because these elements tend to be rapidly cleared from the blood. Nails appear to be an excellent tissue for the assessment of long-term selenium intake because of their capacity to integrate exposure over time (Morris et al., 1983). Because the hair and nails can be cut at various times after formation (a few weeks for hair close to the scalp and approximately 1 year for the great toe), an index of exposure can be obtained that may be little affected by recent experiences. This can be a particular advantage in the context of a case-control study of diet and cancer. Contamination poses the greatest problem for measurements in hair due to its extensive exposure to the environment and great surface area; these problems are generally much less for nails but still need to be considered.

Limitations of Biochemical Indicators. Although the use of biochemical indicators for assessing diet is attractive, no practical indicators exist for many of the dietary factors implicated in the etiology of cancer. Even when tissue levels of a nutrient can be measured, these levels are often highly regulated and thus reflect dietary intake poorly; blood retinol, cholesterol, and sodium are good examples. Just as with dietary intake, the blood levels of some nutrients fluctuate substantially over time, so that one measurement may not provide a good reflection of long-term intake. Furthermore, experience has provided sobering evidence that the tissue levels of many nutrients can be affected by the presence of cancer, even several years prior to diagnosis (Wald et al., 1986), rendering the use of many biochemical indicators treacherous in most case-control studies. Despite these limitations, careful application of biochemical indicators can provide unique information about dietary intake, particularly for nutrients that cannot be accurately calculated from data on food intake.

Anthropometry and Measures of Body Composition

The influence of energy balance at various times in life is likely to have important effects on the incidence of some cancers. Energy balance is better reflected by measurements of body dimensions and composition than by assessments based on the difference between energy intake and expenditure (largely physical activity), because both of these variables are measured with considerable error (Willett, 1998).

The most common use of anthropometric measurements is in the calculation of obesity using either indices such as Quetelet index or body mass index (weight divided by the second power of height), or relative weight (weight standardized for height). Remarkably valid estimates of weight and height can be obtained even by questioning individuals (Stunkard and Bauman, 1981), including their recall for several decades earlier (Rhoads and Kagan, 1983). Thus, estimates of obesity can be obtained easily for large prospective investigations or retrospectively in the context of case-control studies. The major limitation of obesity estimates based on height and weight is that an assessment of weight cannot differentiate between fat and lean body mass. For this reason, these are imperfect measures of obesity. Until recently, studies of the validity of BMI as a measure of obesity have used as a “gold standard” body fat expressed as a percent of total weight, usually determined by underwater weighing. However, BMI is actually a measure of fat mass adjusted for height rather than a measure of percent body fat. When fat mass, determined from densitometry, is adjusted for height and used as the standard, the correlation with BMI is approximately 0.90, indicating a substantially higher degree of validity than has generally been appreciated (Spiegelman et al., 1992). Moreover, in the same study, fat mass adjusted for height was correlated more strongly with biologically relevant variables such as blood pressure and fasting blood glucose than was percent body fat. Using the simple measure of BMI, much has been learned about the relation of positive energy balance with cancer, as well as with coronary heart disease. The electrical impedance method (Hodgdon and Fitzgerald, 1987) is a simple method of measuring adiposity that could potentially be used in epidemiological studies, but this needs further evaluation.

The use of one or a small number of skin-fold thicknesses does not appear to be appreciably more accurate than weight and height in the estimation of overall adiposity but can provide additional information on the distribution of body fat. The ratio of waist-to-hip circumferences has received considerable attention in relation to cardiovascular disease, diabetes, and blood pressure (Bjorntorp, 1987; Hartz et al., 1984). This ratio has also been of interest with respect to hormonally sensitive cancers, as it has been suggested that central fat functions differently than peripheral fat with respect to estrogen metabolism (Vague, 1956).

Height has often been ignored as a variable of potential interest in epidemiological studies. However, it can provide unique information on energy balance during the years before adulthood, a time period that may be important in the development of some tumors that occur many years later. For example, in many studies height has been positively associated with risk of breast cancer (Swanson et al., 1988). Furthermore, this information can be valid even in the context of case-control studies because height will usually be unaffected even if illness has caused recent weight loss. Caution is indicated, however, if associations are not found with height because it is possible that, in some populations, few individuals will have been sufficiently deprived of energy intake during development to reduce their longitudinal growth. In such populations, height will primarily reflect genetic factors. Further information on the measurement and interpretation of body dimensions and composition is provided elsewhere (Lohman et al., 1988; Willett, 1998).
METHODOLOGICAL ISSUES IN NUTRITIONAL EPIDEMIOLOGY

Between-Person Variation in Dietary Intake

In addition to the availability of a sufficiently precise method for measuring dietary intake, an adequate degree of variation in diet is necessary to conduct observational studies within populations; if no variation in diet exists among persons, no associations can be observed. Some have argued that the diets within populations such as the United States are too homogeneous to study relationships with diet (Goodwin and Boyd, 1987; Hebert and Miller, 1988; Prentice et al., 1988). The true between-person variation in diet is difficult to measure directly, but it cannot be measured by the questionnaires used by epidemiologists because the observed variation will combine true differences with those due to measurement error; more quantitative methods must be used for this purpose. The fat content of the diet varies less among persons than does any other specific nutrient (Beaton et al., 1979; for men in our prospective study (Rimm et al., 1992), the mean fat intake assessed by the mean of two 1-week diet records for those in the top quintile was 40% of calories, whereas for those in the bottom quintile 24% of calories were derived from fat. Although this is not a large range of fat intake and is certainly smaller than the variation among countries, it is of considerable interest because it corresponds closely to the changes recommended by many organizations. Other nutrients vary much more among persons than does total fat intake (Beaton et al., 1979; Willett, 1998).

Evidence that measurable and informative variation in diet exists within the United States population is provided by several sources. First, the correlations between food frequency questionnaires and independent assessments of diet found in the validation studies noted above could not have been observed if variation in diet did not exist. For the same reason, the correlations between questionnaire estimates of nutrient intakes and biochemical indicators of intake provide solid evidence of variation. In addition, the ability to find associations between dietary factors and incidence of disease (particularly when based on prospective data) indicates that measurable and biologically relevant variation exists. For example, reproducible relationships have been demonstrated between specific types of dietary fat and cardiovascular disease (Hu et al., 1997; Kushi et al., 1985; Shekelle et al., 1981) and between dietary fiber and risk of coronary heart disease (Prasad et al., 1992) and diabetes (Hu et al., 2001).

Although accumulated evidence has indicated that informative variation in diets exists within the United States population and that these differences can be measured, it is important that findings be interpreted in the context of that variation. For example, a lack of association with fat intake within the range of 25% to 40% of energy should not be interpreted to mean that fat intake has no relation with risk of disease under any circumstances. It is possible that the relation is nonlinear and that risk changes at lower levels of fat intake; for example, at 20% of total energy.

Implications of Total Energy Intake

Energy balance is likely to have important associations with some cancers; however, this cannot be studied directly by examining the relation of energy intake with risk of cancer because energy intake largely reflects factors other than over- or undereating in relation to requirements (Willett and Stampfer, 1986). The implications of total energy intake can be appreciated by realizing that variation among persons is, to a large degree, secondary to differences in body size and in physical activity. Persons also appear to differ somewhat in metabolic efficiency; inefficient persons require higher energy intake for the same level of function. However, these differences in metabolic efficiency are not practically measurable in epidemiological studies. Because virtually all nutrient intakes tend to be correlated with total energy intake, much of the variation in intake of specific nutrients is secondary to factors that may be unrelated to risk of disease. The effect of extraneous variation is, of course, to increase misclassification of diet and attenuate associations. Although the interrelations of diet and factors that determine variation in total energy intake are complex and beyond the scope of this discussion, failure to adjust for total energy intake may result in lack of significant associations.

When total energy intake is related to risk of disease—for example, when physical activity is protective—failure to consider total energy intake in the analysis can be particularly serious because it will confound associations with specific nutrients. The example of coronary heart disease is instructive: Because of the inverse relation with total energy intake, specific nutrients such as saturated fat will also tend to be inversely related to risk. Adjustment for total energy intake is necessary to avoid misleading conclusions; several statistical methods can be used (Willett, 1989a; Willett, 1998). However, the most commonly employed method, division by total energy, which is also called a nutrient density, is not an adequate solution because the inverse of energy intake can then be confounding; this may or may not be a serious problem in any particular study. Appropriate adjustment for energy intake can be a nontrivial issue in some studies; the direction of association can be reversed, such as in the relation between saturated fat intake and myocardial infarction (Gordon et al., 1981) and between fiber intake and risk of colon cancer (Lyon et al., 1987). Unfortunately, total energy intake has been either not measured or, when associated with disease, not appropriately accounted for in many studies on diet and cancer, thus rendering the interpretation unclear.

Nutrient intakes adjusted for total energy can be viewed conceptually as measures of nutrient composition rather than as measures of absolute intake. Measures of nutrient composition are most relevant to personal decisions and public health policy because individuals must alter nutrient intakes primarily by manipulating the composition of their diets rather than their total energy intake. This reasoning underlies, for example, the use of fat as a percent of calories in expressing a dietary objective (Committee on Diet Nutrition and Cancer et al., 1982; U.S. Department of Agriculture & U.S. Department of Health and Human Services, 2000).

ASSOCIATIONS OF DIET WITH SPECIFIC CANCERS

In this section, several aspects of diet that have been considered in relation to cancer etiology will be noted. This is intended to provide a sense of the scope of the issues and direction to other sources of information, including the chapters in this book relating to specific cancers. A comprehensive, critical review would require a volume in itself.

Energy Balance

In animal studies, restriction of energy intake sufficient to reduce growth consistently and substantially reduces the occurrence of many tumors (Ross and Bras, 1971; Weindruch and Walford, 1982). In humans, excessive energy intake, as assessed by indices of obesity, is strongly related to risks of cancers of the endometrium, colon, breast (after menopause), kidney, esophagus (adenocarcinoma), and biliary system (International Agency for Research on Cancer, 2002), and recent evidence suggests that greater adiposity is also associated with death due to leukemia, lymphoma, myeloma, and cancer of the pancreas (Calle et al., 2003). Among premenopausal women, obesity appears to be protective for breast cancer (Huang et al., 1997; Le Marchand et al., 1988), perhaps because heavier women have more anovulatory menstrual cycles, and a high BMI at age 18–20 continues to predict lower risk of breast cancer through the postmenopausal years. Because weight in young women is inversely related to risk of breast cancer, and weight after menopause is positively related to risk, weight gain during midlife has been consistently and strongly associated with postmenopausal breast cancer (Marchand et al., 1988), perhaps because heavier women have more overweight adipose tissue (International Agency for Research on Cancer, 2002).

Height, which in part reflects energy balance during development, is positively related to risk of breast cancer (Swanson et al., 1988; Tretti, 1989; Valoraas et al., 1969; van den Brandt et al., 2000). Height has also been associated with colon cancer (Chute et al., 1991), but the data are limited.
Dietary Fat

Fat intake has been singled out (Committee on Diet Nutrition and Cancer et al., 1982; National Research Council (U.S.) & Committee on Diet and Health, 1989) as the aspect of diet most importantly related to risk of cancer.

The most important sites for which associations have been suggested are cancers of breast, colon, and prostate. Much of the support for these relationships is based on the striking international correlations between per capita fat intake and rates of these malignancies (Armstrong and Doll, 1975; Prentice et al., 1988). Positive associations with breast cancer have been seen in some case-control studies (World Cancer Research Fund & American Institute for Cancer Research, 1997), and in a pooled analysis a highly significant positive association was seen (Howe et al., 1990). However, prospective studies have consistently shown no positive association (Hunter et al., 1996). In an updated pooled analysis of large prospective studies, the RR for a 25 g/day increment in energy-adjusted total fat intake was 1.03 (Smith-Warner et al., 2001a).

This discordance between the case-control and cohort studies suggests that concerns about recall and selection biases in case-control studies of diet are justified. In a recent analysis limited to young women, intake of animal fat, but not vegetable fat, while premenopausal was associated with greater risk of breast cancer (Cho et al., 2003), suggesting that meat and dairy products may contain constituents other than fat that influence risk of breast cancer. For colon cancer, associations with red meat intake, and processed meat in particular, have been seen frequently (Norat et al., 2002), but fat intake itself does not appear to be related to risk (Howe et al., 1997; World Cancer Research Fund & American Institute for Cancer Research, 1997). Associations have also been reported between animal fat intake and risk of prostate cancer (Giovannucci et al., 1993a; Graham et al., 1983; Heshmat et al., 1985; Kolonel et al., 1988; Le Marchand et al., 1994; World Cancer Research Fund & American Institute for Cancer Research, 1997), but there is little evidence that vegetable fats increase risk. This suggests that fat per se is not an etiologic factor for prostate cancer, although the prospective studies are few. Overall, the more powerful and less biased studies have not supported an association between fat intake during midlife and cancer incidence. The effects of diet during earlier periods of life remain to be explored.

Fiber

Interest in the relation between fiber intake and colon cancer is largely due to Burkitt's observation of low rates of colon cancer in areas of Africa where fiber consumption and stool bulk were high (Burkitt, 1971). Fiber has been hypothesized to dilute potential carcinogens and speed their transit through the colon. Inverse associations with total fiber intake have been seen in most case-control studies (Howe et al., 1992; Trock et al., 1990). When specific sources of fiber have been examined, fiber intake from fruits or vegetables have been most consistently associated with lower incidence, whereas fiber from cereals has either been related to an increased risk or not associated with colon cancer (Willett, 1989b). However, dietary fiber has not been associated with risk of colon cancer in most prospective studies (Fuchs et al., 1999; Terry et al., 2001). A modest inverse association was observed in a multicenter cohort study in Europe (Bingham et al., 2003), but the confounding by intake of folic acid (see below) was not considered, which is plausible because whole grains, fruits, and vegetables are the primary dietary sources of both nutrients.

Folic Acid

Folic acid deficiency has long been known to cause tumors in animals, possibly by influencing gene expression through DNA methylation or by increasing the incorporation of uracil in DNA (Blount and Ames, 1994). Considerable evidence from both case-control and cohort studies supports an inverse association between folate intake and risk of colon cancer (Giovannucci et al., 2002a), and this association appears stronger among regular alcohol consumers (Giovannucci et al., 1995).

An association between a functional polymorphism in the folate acid metabolizing gene, methylene tetrahydrofolate reductase, and incidence of colon cancer adds support for a causal relationship. An inverse association between folate intake or blood level and risk of breast cancer among regular alcohol consumers has also been reported in multiple studies (Rohan et al., 2000; Zhang et al., 1999; Zhang et al., 2003).

Preformed Vitamin A and Carotenoids

Because vitamin A plays a central role in regulating cell differentiation, reason exists to suspect a relation to cancer incidence. In some animal models, preformed vitamin A or chemical analogues can inhibit tumor development, even when administered well after the carcinoma (Lippman and Meyskens, 1989). Vegetable precursors of vitamin A, the carotenoids, have been less studied in animals but have reduced tumor incidence in some models, particularly skin cancers (Mathews-Roth, 1989). Whether carotenoids act by virtue of conversion to retinol, the main circulating form of vitamin A with physiologic activity, or by other mechanisms, such as being an antioxidant or quencher of singlet oxygen (Peto et al., 1981), remains uncertain.

Epidemiological studies of preformed vitamin A or carotenoids in relation to cancer incidence have used either questionnaire assessments of diet or blood measurements. The distinction between these two forms of vitamin A has practical implications because preformed vitamin A is obtained only from foods derived from animal sources or vitamin supplements, whereas carotenoid precursors of vitamin A are obtained almost entirely from plant sources. The initial reports of an inverse relation between total vitamin A intake and risk of lung cancer (Bjelke, 1975; Mettlin et al., 1979) did not clearly distinguish between these sources. In a subsequent prospective study by Shekelle and coworkers (Shekelle et al., 1981), an apparent protective effect of total vitamin A for lung cancer was found to be entirely attributable to carotenoid sources; preformed vitamin A was unrelated to risk of this disease. Subsequent case-control studies based on dietary intake (Gregor et al., 1980; Hinds et al., 1984; Samet et al., 1985; Wu et al., 1985; Ziegler et al., 1984) and beta-carotene measurements in prospectively collected blood (Menkes et al., 1986; Nomura et al., 1985; Staehlin et al., 1984) provided evidence for a protective relationship between carotenoid intake and risk of lung cancer after controlling for cigarette smoking but little support for any relationship with preformed vitamin A (Hunter and Willett, 1994; Willett, 1990).

Large cohort studies have been less supportive of a protective role for beta-carotene and risk of lung cancer; little overall relationship has been seen, although a small increase in risk at very low intakes could not be excluded (Michaud et al., 2000).

Neither preformed vitamin A nor carotenoid intake has been consistently associated with risk of colon cancer, although these relationships have been examined in several studies.

Randomized trials using supplements of beta-carotene have not supported a protective effect. In a randomized study among patients with previous skin cancer, beta-carotene did not prevent the development of new skin cancer (Greenberg et al., 1990). Similarly, beta-carotene— as well as vitamins C and E—did not influence the recurrences of adenomatous colon polyps (Greenberg et al., 1994). In the large Finnish trial among men at high risk of lung cancer (Group, 1994), a statistically significant 18% increase in incidence of this malignancy was seen in those randomized to beta-carotene, and in a trial among persons at high risk for lung cancer, a combined beta-carotene/preformed vitamin A supplement also increased risk by 28% (Omenn et al., 1996). In the Physicians' Health Study, no effect of beta-carotene supplements on lung or overall cancer incidence was seen (Hennekens et al., 1996). Many explanations have been offered for the lack of benefit (or even harm) in the randomized trials, including interference with the metabolism of other beneficial carotenoids, the relatively high dose of beta-carotene in the supplements, or induction of carcinogen-activating enzymes. Also, the lack of benefit from beta-carotene supplements does not exclude the possibility that other protective nutrients or biologically active constituents in fruits and vegetables could be protective. However, the premise that greater consumption...
of fruits and vegetables is protective has not been supported by the results of large cohort studies of diet and lung cancer: apart from a possible small increase in risk with very low intake, the dose-response relationship is flat (Smith-Warner et al., 2003). Thus, the basis for the enthusiasm for finding a powerful chemopreventive factor in fruits and vegetables appears to have been flawed. The findings from case-control studies could have been due to biases discussed earlier, and the studies based on beta-carotene levels in prospectively collected blood could have resulted from incomplete control of confounding by smoking because smoking reduces beta-carotene levels (Willett, 1998). Although the best epidemiologic evidence and randomized trials both indicate no benefit of high beta-carotene intake for lung cancer, it is sobering to consider the resources that have been invested in this hypothesis.

**Vitamin C**

Vitamin C has been hypothesized to reduce cancer risk by its antioxidant properties, by blocking the conversion of nitrates and nitrogen-containing compounds to carcinogens under conditions found in the stomach (Mirvish et al., 1972) and in food (Raineri and Weisburger, 1975), and by other mechanisms (Cameron et al., 1979). In case-control studies, evidence for a protective effect of vitamin C has been seen for laryngeal cancer (Graham et al., 1981), oral cancer (Winn et al., 1984), esophageal cancer (Mettlin et al., 1981; Ziegler et al., 1981), stomach cancer (Correa et al., 1985), and cervical dysplasia (Romney et al., 1985; Wassertheil-Smoller et al., 1981). As discussed elsewhere (Block and Menkes, 1989), the interpretation of these data, while supporting a protective effect of vitamin C, is also compatible with the possibility that other factors in fruits and vegetables are the primary cause of reduced cancer risk. However, given the experience with beta-carotene and lung cancer, supportive data from large cohort studies or randomized trials would be needed to be confident of these relationships.

**Vitamin E**

The concept that vitamin E might reduce risk of human cancer derives from its role as a potent intracellular antioxidant and from suggestive animal studies (Bieri et al., 1983; Mersens et al., 1989). Wald reported an extremely strong inverse relation between prediagnostic blood levels of vitamin E and risk of breast cancer (Wald et al., 1984); however, this was later found to be an artifact of differential handling of case and control specimens (Wald et al., 1988). In a separate study population, Wald reported a protective association between prediagnostic blood vitamin E levels and cancer at all sites combined but provided evidence that this was likely due to an effect of preclinical disease on vitamin E levels rather than a preventive effect of this micronutrient (Wald et al., 1987). Also using blood collected from a large cohort, Menkes and co-workers (Menkes et al., 1986) reported an inverse association with lung cancer; this was not confirmed in a similar study by Nomura and co-workers (Nomura et al., 1985).

Although associations have generally not been seen for other cancer sites (Nomura et al., 1985), an inverse association was reported in a Finnish cohort between prediagnostic serum vitamin E levels and cancers at all sites combined, which was particularly strong for non-smoking-related cancers (Knekt et al., 1988). In the recent Finnish trial (Group, 1994), vitamin E supplementation did not influence risk of lung cancer and was associated with a lower risk of prostate cancer, but no relation was seen in the United States cohort study (Chang et al., 1999). The overall data thus remain unclear at this time; this may be the result of only modest numbers of specific cancers in the studies that have been reported, but a large beneficial effect of vitamin E against cancer seems unlikely.

**Selenium**

That higher intake of selenium might reduce the risk of some human cancers has been suggested by numerous animal experiments (Combs and Combs, 1986) and ecological studies in which indices of selenium intake have been inversely related to risk of cancer, both internationally (Schrauzer et al., 1977) and within the United States (Shamberger et al., 1976). The correlational data have been particularly strong for cancers of the colon and breast. Inverse associations have been seen between serum selenium levels and subsequent risk of all cancers combined in several large prospective studies (Kok et al., 1987; Salonen et al., 1984; Salonen et al., 1985; Willett et al., 1983), but no evidence of a protective effect has been found in other similar investigations (Coates et al., 1988; Menkes et al., 1986; Nomura et al., 1987; Peleg et al., 1985; Virtamo et al., 1987). In two prospective studies that together included a large number of breast cancer cases, no relation was found between selenium in nails and subsequent risk of this malignancy (Hunter et al., 1990); van Noord et al., 1987). A strong inverse association between nail selenium levels and risk of lung cancer was seen in a prospective study from Holland (van den Brandt et al., 1993), and an inverse association between nail selenium levels and risk of prostate cancer was seen in U.S. men (Yoshizawa et al., 1998). In a randomized trial of selenium supplementation for prevention of skin cancer, no benefit was seen for this end point (Clark et al., 1996). Although based on a small number of cases, large reductions in the incidence of prostate and colon cancer were seen. A randomized trial is now ongoing, using a 2 x 2 factorial design that will evaluate the chemopreventive effects of selenium and vitamin E (Klein, 2003).

**Calcium**

Higher calcium intake has reduced bowel tumors in animals (Lamprecht and Lipkin, 2001), possibly by precipitation of bile acids that are thought to promote tumorigenesis. In a randomized trial, calcium supplementation reduced recurrence by about 20% (Baron et al., 1999). In large prospective studies, high calcium intake has consistently been associated with a modest and nonsignificantly lower risk of colon cancer. In a pooled analysis of cohort studies, those with the highest calcium intake had a 20% lower risk of colon cancer that was statistically significant (Cho et al., 2004). Thus, the convergence of many lines of evidence support a modest benefit of higher calcium intake in colon cancer prevention. However, the practical implications are currently unclear because high calcium intake or dairy product consumption has been associated with higher risk of prostate cancer in many studies (Chan and Giovannucci, 2001; Giovannucci, 2002b).

**Salt**

Ecological associations have led to the suggestion that high salt intake—salt being traditionally used in many societies for the preservation of food—might increase the risk of gastric cancer (Howson et al., 1986; Joossens and Geboers, 1987). Salt is hypothesized to act as a local irritant that may compromise the gastric mucosal barrier and thus facilitate the action of local carcinogens. Positive associations of salt intake with risk of gastric cancer have been seen in several case-control studies (Fontham et al., 1986; Haenszel et al., 1972; You et al., 1988). This hypothesis is compatible with the striking decline in gastric cancer in most industrialized countries over this century as refrigeration reduced the need of salt for preservation. However, economic improvements also enhanced the supply of fresh fruits and vegetables on a year-round basis and improved hygiene, which may have reduced the transmission of Helicobacter infection. A strong association between salted fish intake during childhood and risk of nasopharyngeal cancer has been noted among several Chinese populations (You et al., 1988).

**Fruits and Vegetables**

Higher intake of fruits and vegetables has been widely believed to reduce risk of many specific cancers (Block et al., 1992; World Cancer Research Fund & American Institute for Cancer Research, 1997), and an increase in consumption has been a key element of most dietary recommendations and strategies to prevent cancer. However, several factors may have led to an overstatement of the potential benefits. First, most of the studies supporting such conclusions have been case-
control investigations and are thus susceptible to bias. Reporting of evidence may have been selective, in part because several dozen fruit and vegetable items are typically included in dietary questionnaires, and publication of all associations is not feasible. Thus, “positive” findings supporting prior beliefs may be more likely to be published than null or contrary findings. Also, summaries of evidence have sometimes counted an inverse association with one or a few foods as support for a benefit of fruits and vegetables when this may have simply been the result of chance due to multiple comparisons. These potential biases are greatly reduced if large prospective studies are examined systematically; recent findings from such analyses provide a much more tempered view of the benefits to be derived from an overall increase in fruit and vegetable consumption. As described above, a pooled analysis of large prospective studies of lung cancer, which included more than 3000 incident cases in men and women, did not find an overall relationship with consumption of fruits and vegetables, although a small increase in risk in the lowest category of intake was observed (Smith-Warner et al., 2003). A similar lack of overall association with fruit and vegetable consumption was seen in the pooled analysis of breast cancer (7377 cases among 351,825 women) (Smith-Warner et al., 2001b). In the pooled analysis from the Nurses’ Health Study and Health Professionals’ Follow-up Study, no association was seen with total fruit and vegetable consumption and risk of colon cancer (Michels et al., 2000). Also, in a combined analysis of these two cohorts, no association was seen between overall consumption of fruits and vegetables and total cancer incidence (Hung et al., 2004). In this analysis, a statistically significant inverse association was seen between overall fruit and vegetable intake and risk of cardiovascular disease, which indicates that biologically important variation in consumption did exist in these cohorts. Thus, more recent evidence that is less susceptible to methodological bias does not support a major benefit from increasing fruit and vegetable consumption for reducing cancer risk.

Despite the largely null findings for overall fruit and vegetable intake, the possibility remains that higher intake of specific fruits and vegetables, or specific substances in fruits and vegetables, may reduce the risk of specific cancers. For example, evidence noted above suggests that higher intake of folic acid can reduce risk of colon cancer. However, if folic acid is the primary active factor in fruits and vegetables, fruit and vegetable consumption would represent a seriously misclassified measure of folic acid intake because many of these foods contain only small amounts of this vitamin. Also, fortification and supplementation provide a large proportion of the folic acid in the United States food supply, and this would tend to mask the contribution from fruits and vegetables. As another example, higher intake of lycopene, the carotenoid responsible for the red color of tomato products, has been associated with lower risk of prostate cancer in several studies (Giovannucci, 2002c), even though overall fruit and vegetable consumption was not associated with risk of this cancer. It is also possible that biologically active substances in fruits and vegetables exist but in doses that have effects too small to be detected in epidemiological studies, and that purified amounts administered pharmacologically will reduce cancer. Despite these caveats, the best available evidence suggests that the recommendation to increase consumption of fruits and vegetables is good general advice but that this cannot be relied upon to reduce risk of cancer.

Natural Carcinogens and Products of Food Preparation

Ames (Ames et al., 1987) has reviewed the large number of naturally occurring compounds in foods that are known to be mutagenic or carcinogenic in laboratory settings. Many of these toxic compounds have probably been developed by plants during evolution as a form of protection. In addition, a wide variety of mutagens and carcinogens are formed during the process of cooking meat and other foods, even under conditions that do not involve charring (Sugimura, 1986). With few exceptions—such as the bracken fern, which may cause bladder cancer when used regularly as a tea (Pamukcu et al., 1970)—little evidence exists that these compounds contribute substantially to human cancer. However, the role of carcinogens that occur naturally or that are formed during cooking has not been studied adequately with respect to human cancer and deserves further investigation.

PREVENTION OF CANCER BY DIETARY MEANS

Intense interest in the possible prevention of cancer by dietary modification has generated a series of recommendations from numerous private and governmental bodies worldwide and within the United States (Committee on Diet Nutrition and Cancer et al., 1982; Greenland and Sondick, 1986; World Cancer Research Fund & American Institute for Cancer Research, 1997). Those of the National Academy of Sciences (Committee on Diet Nutrition and Cancer et al., 1982) are typical:

1. The consumption of both saturated and unsaturated fats should be reduced, from an average of approximately 40% of total calories to 30%.
2. The intake of fruits, vegetables, and whole grain cereal products in the diet should be emphasized.
3. The consumption of food preserved by salt-curing (including salt-pickling) or smoking should be minimized.
4. Efforts should continue to be made to minimize contamination of foods with carcinogens from any source, including those that are natural or occurring inadvertently during production, processing, and storage.
5. Further efforts should be made to identify mutagens in food and to expedite testing for their carcinogenicity. Where feasible and prudent, mutagens should be removed or minimized.
6. If alcoholic beverages are consumed, this should be done in moderation.

The potential impact of these changes in diet is difficult to quantify. Based on similar recommendations, the U.S. National Cancer Institute suggested that within 10 years, cancer of the colon and rectum would be reduced by 50%; cancer of the breast by 25%; cancers of the prostate, endometrium, and gallbladder by 15%; and the cancers of the stomach, esophagus, pancreas, ovaries, liver, lung, and bladder by a possible but not precisely quantifiable amount (Greenwald and Sondick, 1986). In this same publication, the large uncertainty of these estimates was noted, including the possibility that no reductions for some of these cancers might occur.

Dietary recommendations for the United States and other Western countries have primarily focused on ways to reduce cancers of the breast, large bowel, and prostate, and to a lesser extent pancreas and ovary because these are the most important cancers not caused by smoking or alcohol that are plausibly related to diet. In other parts of the world different cancer sites, such as stomach and esophagus, are dominant. These upper gastrointestinal cancers in particular are likely influenced by dietary factors; although uncertainty exists about the specific alterations of diet that will be effective, a change toward the contemporary United States diet is likely to be beneficial. However, a more precise understanding of etiologic factors is needed to avoid exchanging one pattern of cancer for another that includes high rates of breast and colon cancer.

The actual benefit in cancer reduction that might be realized by dietary change is, of course, a function of both the biological relationships and the success of an intervention. The most certain effects of diet on cancer incidence are those related to total energy balance, especially with cancers of the (postmenopausal) breast, colon, kidney, and endometrium. In the United States, 20% of cancer mortality in women and 15% of cancer mortality in men are estimated to be attributable to overweight and obesity (Callie et al., 2003). In addition to reducing incidence, avoidance of midlife weight gain is likely to contribute to lower breast cancer mortality, in part due to a lower case-fatality rate among lean women. Although reduction in obesity among premenopausal women could lead to an increase in breast cancer incidence, breast cancer mortality in this group is not related to body weight due to delayed diagnosis among overweight women (Tretli, 1989; Willett et al., 1985). Thus, weight control deserves to be high
on the agenda for cancer prevention, not far behind efforts to reduce cigarette smoking. Weight control programs need to be inextricably intertwined with promotion of physical activity, which also can reduce risk of several cancers independent of its effect on body weight (International Agency for Research on Cancer, 2002). Because rates of overweight and obesity are rapidly increasing in the United States and many countries (Flegal et al., 2002; Seidell, 2002), intensified weight control efforts are needed to avoid further increase in cancer incidence. Although the overall trends are discouraging, most motivated individuals can control their weight; this is manifested by the 50% lower prevalence of obesity among those with higher education compared to groups with low educational levels (Flegal et al., 2002). Epidemiologic data also provide strong evidence that reduction in alcohol consumption will reduce cancers of the upper gastrointestinal tract, liver, colon, and breast (International Agency for Research on Cancer, 1988: Smith-Warner et al., 1998). Although heavy consumption (more than two alcoholic drinks per day) has long been known to be associated with cancer incidence, modest increases in risk of breast cancer are seen with 1–2 drinks per day.

Considerable effort has been given to increasing fruit and vegetable consumption with the expectation that this could have a major impact on cancer incidence. This has been the rationale for the “Five-a-Day” program of the U.S. National Cancer Institute. However, the likely benefits appear to have been seriously overstated, largely due to biases in case-control studies. Whether any relationship exists with overall cancer incidence in the United States is unclear. Association may exist between intakes of specific foods and specific cancers that are obscured by combining all fruits and vegetables, and some evidence suggests that benefits exist for cancers of the esophagus and stomach, which are uncommon in the United States but extremely important elsewhere. Also, some studies suggest that only very low intakes of fruits and vegetables are associated with excess risk, and that increasing intake beyond about three servings/day may have little further effect. Specific components of fruits and vegetables may be protective, as folic acid appears to be for colon cancer, but overall fruit and vegetable intake represents a poor indicator of intake. Despite the diminished prospects for reducing cancer incidence by increasing fruit and vegetable consumption, evidence is strong that this will reduce risk of coronary heart disease, the most important cause of death in the United States.

The primary focus of most dietary recommendations thus far has been on intake of dietary fat; it has been suggested that intake should be reduced to an average of about 30% of energy (Committee on Diet Nutrition and Cancer et al., 1982; National Research Council (U.S.) & Committee on Diet and Health, 1989; World Cancer Research Fund & American Institute for Cancer Research, 1997). Anticipated benefits have related primarily to cancers of breast and colon and possibly prostate, endometrium, and ovary. The recommendation to reduce fat intake has often been in part justified on the basis of being harmless and also being likely to result in a reduced risk of coronary heart disease. However, this recommendation is overly simplistic and not without potential harm. Total fat intake is not related to risk of coronary heart disease (Hu and Willett, 2002; National Research Council (U.S.) & Committee on Diet and Health, 1989); and unsaturated fats, contained primarily in vegetable oils, are likely to reduce risk of coronary heart disease, whereas saturated and trans fats are likely to increase the risk (Hu and Willett, 2002). The relation of type of fat to risk of human cancer remains unclear; what evidence there is suggests that any positive association pertains mainly to fat from animal sources. Without a careful, selective reduction in the type of fat, it is quite conceivable that some individuals might reduce their intake of vegetable fat and thereby increase their risk of cardiovascular disease. Rather than focusing on fat intake, a reduction in red meat and dairy fat consumption is better supported; associations with red meat have been seen with cancers of the colon and prostate in multiple studies. Whether these relationships are due to the fat content of meat is uncertain, and some data for colon cancer suggest other components contribute to risk.

The level of scientific certainty that is appropriate for launching public health interventions aimed at changing diets to prevent cancer is an issue of legitimate debate. As has been pointed out elsewhere (Greenwald and Sondick, 1986), waiting to change dietary behavior until the scientific evidence is almost certain could mean the loss of thousands of lives. On the other hand, promulgating policies that turn out to be wrong has costs in terms of diverting attention and resources from interventions that are effective, such as smoking cessation and mammography, and in the loss of credibility, which is essential for any public health program. Fortunately, the dietary recommendations noted above, with the qualification regarding type of fat, are generally consistent with those designed to reduce the incidence of coronary heart disease (Consensus Conference, 1985; Hu and Willett, 2002), and for which the evidence is more compelling.

Specific actions for implementing dietary recommendations have been developed by the U.S. National Cancer Institute (Greenwald and Sondick, 1986). These include encouraging federal agencies and industries to include cancer prevention dietary recommendations in federal food production, marketing, and distribution policies; encouraging the production of leaner meat products and lower fat content of dairy products; informing the public about the relationship between diet and cancer; expanding nutrition labeling to cover the full range of mass-marketed foods so consumers can be better informed and make wiser shopping decisions; and developing diet and cancer programs that make use of the mass media and other high-technology communication approaches. State and local agencies are also encouraged to review school curricula to reflect newer knowledge of diet and cancer risks and strategies for risk reduction; to review school menus in relation to the cancer control objectives; to promote diet and cancer information programs; to encourage restaurants to provide sufficient information to allow patrons to choose nutritious foods; to coordinate governmental planning activities to ensure that attention is given to reducing dietary risk factors for cancer; to promote dissemination of information about proper food selection to protect against cancer risk; and to include information on diet and cancer in existing food, nutrition, and health programs with the use of innovative approaches to reach high-risk groups. These approaches remain valid, but the emphasis will need to be shifted from reducing total fat intake if there is to be an appreciable impact on cancer incidence. Because of the importance of weight control, excessive energy intake from all sources and increasing physical activity will need to be emphasized.

To accomplish this, health care providers will need to be more engaged in weight counseling, physical education programs will need to be increased and improved, and the physical infrastructure of America will need to be modified to encourage walking and bicycle riding for transportation and a wide range of enjoyable forms of recreational activity.

**SUMMARY**

A wide variety of evidence based on comparisons of cancer rates in different geographic areas, migrating populations, religious orders, and rapid changes in rates over time strongly suggests that the high rates of breast, colon, and other important cancers in the affluent countries are due to environmental rather than genetic factors. Furthermore, aspects of diet are likely to be important etiological factors for some of these cancers. Similarly, the incidence rates of many cancers that are still of great importance in other parts of the world, such as those of the oral cavity, esophagus, and stomach, are also likely to be influenced by dietary factors. Despite the substantial evidence that dietary factors are likely to be important, the specific aspects of food and nutrient intake that are either causative or preventive remain inconclusively defined for most of these cancers. The most consistent evidence is that excessive energy intake in relation to level of physical activity increases the risk of many of these cancers and that this is the second most important avoidable cause of cancer mortality in many countries after cigarette smoking. Also, convincing evidence indicates that excessive alcohol consumption increases the risk of several important cancers, including those of the upper gastrointestinal tract, liver, colon, and breast. Substantial evidence, although not entirely consistent, indicates that red meat or animal fat intake is associated with risks of colon...
and prostate cancers. Evidence that overall increases in fruit and vegetable consumption will reduce cancer incidence appreciably has become much weaker than believed earlier, although modest benefits for some specific cancers cannot be excluded. Better evidence supports a protective effect of folic acid against risk of colon cancer, but this is more reliably achieved by taking a multiple vitamin than by increasing fruit and vegetable intake. Individual and policy decisions regarding dietary changes should consider not only the possible benefits for cancer reduction but also the effects on coronary heart disease, because this remains the dominant cause of mortality in the United States and its relation with diet is better established.

References


