The Ninth AACR American Cancer Society Award Lecture on Cancer Epidemiology and Prevention

Diet and Cancer: One View at the Start of the Millennium

Walter C. Willett

Diet and Cancer: One View at the Start of the Millennium

A fundamental observation in cancer epidemiology during the last century was that cancer incidence and mortality rates vary dramatically across the globe (1). In addition, rates of cancer among populations migrating from low- to high-incidence countries change markedly; in most cases, they approximate the rates in the new region within one to three generations. These lines of evidence indicate that the primary determinants of cancer rates are not genetic factors, but rather environmental and lifestyle factors that could, in principle, be modified to reduce cancer rates in high-risk areas (2). During the last two decades, the primary factor of interest, apart from tobacco, has been dietary fat.

The hypothesis that dietary fat is largely responsible for cancers of the colon, breast, and prostate in Western countries derives largely from the strong ecological correlations between national per capita fat consumption and rates of these major cancers (1). Although international comparisons provide fertile soil for the development of etiologic hypotheses, they are a treacherous basis for conclusions because of the major potential for confounding. In the case of breast cancer, the low-risk countries are primarily developing areas and traditional Eastern societies where almost all aspects of lifestyle have been markedly different than those in affluent Western countries. These include differences in reproductive behaviors, physical activity, body composition, and many aspects of diet other than fat consumption. Thus, more detailed studies are needed to control for variables that may seriously confound cross-national comparisons. For the most part, these will necessarily be case-control and prospective cohort studies, which should be interpreted in the light of animal and mechanistic experiments.

Ideally, every dietary hypothesis related to cancer would be tested by multiple large randomized trials in human populations. However, for even a small number of hypotheses, this is likely to be impossible because of the large numbers of individuals required and the long and uncertain follow-up period. For example, the Women’s Health Initiative trial of fat reduction, conducted along with interventions related to hormone replacement therapy and calcium/vitamin D, will cost approximately $1 billion (3). Even so, it is not likely to provide a clear answer to the fat and breast cancer relationship because multiple dietary factors are being changed simultaneously (4) and a 10-year intervention period late in life may well be inadequate even if fat was an etiologic factor.

In this brief overview, I will describe the general process...
by which we have investigated diet and cancer relationships epidemiologically and will provide a short status report on some of the major issues addressed during the last 20 years. In addition to discussing cancer incidence, I will also mention important findings related to non-cancer outcomes, in particular, coronary heart disease. These non-cancer outcomes can be helpful in interpreting findings for cancer, especially when an association is seen with coronary heart disease but the same variable is not associated with cancer. In this case, the lack of association with cancer cannot be attributed to insufficient variation in the dietary factor or inability to measure it. Moreover, coronary heart disease is itself important, and in the end we eat one diet. Thus, any actual dietary decisions should be made in the light of not just cancer, but how these variables relate to cardiovascular disease and other important conditions.

In the late 1970s, when interest in diet as a major determinant of cancer began to emerge, great skepticism existed about the potential for studies of human diets in relation to risk of disease. Two major concerns were raised: that “diets are homogeneous within populations,” and that “people can’t remember what they ate.” Indeed, if either of these conditions was true, it would be impossible to conduct informative analyses of diet and cancer incidence within populations. However, at that time, we had reasons to believe that neither of these assertions was correct. First, even simple surveys within populations indicated that individuals vary widely in food choices so that diets would not likely be homogeneous. Second, the experience of epidemiologists investigating food-borne outbreaks indicated that individuals could indeed remember what they ate, even at some time well removed from the critical event. However, this experience suggested that individuals could more readily answer questions about what they usually would have eaten as opposed to what they actually ate at a specific time on a specific day. Fortunately from the standpoint of cancer incidence, it is usual diet rather than a single meal or intake for 1 day that is likely to be most important etiologically.

Thus, during the late 1970s, we and several other groups embarked on prospective studies of diet in relation to cancer incidence. At that time, virtually all investigators converged on the concept that a food-frequency questionnaire would be the only practical means of investigating the effects of long-term dietary intake in large numbers of individuals. The major alternative, 24-h recall, has practical limitations because it requires a professional dietitian and is expensive. More importantly, detailed studies had shown that diet varies greatly from day to day for most individuals so that even perfect information about a 24-h period would not reasonably represent a person’s longer-term intake (5, 6). The food-frequency questionnaire that we initially developed included 61 foods that were carefully selected to provide the maximum amount of information on intakes of nutrients hypothesized to influence cancer and cardiovascular disease risk (7). For each of these 61 foods, participants were asked about their average use over the past year and were offered a multiple choice of answers ranging from never to six or more times a day. In addition to collecting data over an extended period, from a practical standpoint the food-frequency questionnaire has tremendous advantages because it could be self-administered and thus used in large populations. In 1982, we developed and tested an optically scanned version of our food-frequency questionnaire, which was found to be highly acceptable. Importantly, this substantially reduced data-entry costs and errors, which allowed us to expand the questionnaire and to collect dietary data repeatedly over time in a large population. This approach has become standard in almost all large prospective studies of diet and cancer.

A central concern in the study of diet and cancer is validity of the dietary assessment. We therefore conducted a series of detailed studies to address this issue using three methods to assess question of validity: comparison with a detailed quantitative method, comparison with biochemical indicators of diet, and ability to predict known disease relationships (7). In the first validation study, using our 1980 61-item food-frequency questionnaire, we identified ~200 participants in our large cohort study, the Nurses’ Health Study. Each participant was instructed to weigh and record each food on a meal-by-meal basis for four 1-week periods over the course of the year. We then compared nutrient intakes from this detailed quantitative method with intakes calculated from our self-administered food-frequency questionnaire. After adjustment for total energy intake, most correlations were in the range from 0.5 to 0.6 (8). Although this degree of validity was less than perfect, it was sufficient to ensure that important associations would not be missed if the study population was sufficiently large. Moreover, with Drs. Rosner and Spiegelman, we developed statistical methods to take these data on measurement error into account by adjusting relative risks and confidence intervals for the degree of measurement errors (9–11).

The primary applications of our dietary assessment methodology have been in three large cohort studies. The Nurses’ Health Study began in 1976 when 121,700 women across the United States were enrolled in a major study primarily to investigate the relationship between oral contraceptive use and cancer outcomes. In 1980, we first mailed our dietary questionnaire to this cohort, and ~95,000 women returned completed questionnaires. Follow-up in the study is conducted primarily by additional questionnaires mailed at 2-year intervals to update information on medical history, diet, and lifestyle, and to ascertain diagnoses of major diseases. These diagnoses are documented with medical records. In this cohort, we have repeated dietary assessments at 2- to 4-year intervals so that by year 2000 we have obtained six repeated measurements of diet. Because the Nurses’ Health Study included only women, in 1986 we enrolled 52,000 men into a parallel cohort study, the Health Professionals Follow-up Study. In 1989, we enrolled an additional 116,000 women into the Nurses’ Health Study II; this cohort provided the next generation of younger nurses, who were maximally exposed to oral contraceptive use during their early reproductive lives.

By the early 1980s, the belief that dietary fat was a major cause of breast cancer had become extremely strong. Based largely on the international comparisons and two case-control studies [one of which was said to be positive although there was no statistically significant association between fat intake prior to diagnosis and breast cancer risk (12)] the Committee on Diet Nutrition and Cancer of the National Academy of Sciences in 1982 came to the following conclusion: “There is sufficient evidence that high fat consumption is linked to increased incidence of certain common cancers (notably breast and colon cancer) and that low fat intake is associated with lower incidences in these cancers. The committee recommends that the consumption of both saturated and unsaturated fat be reduced . . . to 30% of total calories in the diet” (13). This statement was highly influential because it served as the primary basis for national recommendations to reduce fat intake of all types, which has been the dominant nutritional advice for the last 20 years. Not surprisingly, our first publication on dietary fat and breast cancer, which included 4 years of follow-up and 601 incident breast cancer cases, was controversial because it did not support a positive association (14). Indeed, the relative risk for women in the top quintile of fat intake compared with
the lowest quintile was 0.85 (95% confidence interval, 0.66–
1.09). This report, although the most statistically powerful
available data at that time, had many limitations. These in-
cluded the relatively short follow-up period and the fact that we
did not address the association with fat intake below 30% of
calories, although the lowest group at 32% of energy from fat
was close to the recommended intake. Moreover, an effect of
fat intake before adulthood could not be excluded.

Since our original report on fat intake and breast cancer,
we have continued to work on the refinement of our dietary
measurement methodologies. As noted above, we expanded the
dietary questionnaire and collected more detailed information
about many aspects of diet and have taken into account changes
in the fat content of foods over time. Importantly, the use of
repeated measures of intake provides an enhanced measure of
long-term diet because random variation can be minimized by
taking the average of multiple replicates. For example, in a
second validation study that also included ~100 women in our
first validation study, the correlations for saturated fat were
0.54 between the 1980 food-frequency questionnaire and the
1980 diet record, 0.76 between the 1986 questionnaire and the
1986 diet record, and 0.80 between the average of the 1980,
1984, and 1986 food-frequency questionnaires with the average
of 1980 and 1986 diet records (7). Thus, the attenuation of
relative risks due to errors in dietary assessment are substi-
tually reduced by the increased precision of our methods and use
of multiple replicates. In addition, we have recently used fasting
triglyceride levels as an indicator to assess the ability of our
questionnaire to measure dietary fat. In controlled metabolic
studies, reductions in dietary fat as a percentage of calories
increased levels of fasting plasma triglycerides. In the Nurses'
Health Study, we found that fasting triglyceride levels among
women with fat intake <20% of energy as assessed by a
food-frequency questionnaire were approximately twice as high
compared with women with 45% of energy or more from fat
(15). This was confirmed in a similar analysis among members
of the Health Professionals Follow-up Study. This provides
additional evidence that our dietary assessment method is ca-
ble of identifying physiologically relevant differences in fat
intake.

The continued follow-up in the Nurses’ Health Study has
provided additional opportunities to examine the relationship
during dietary fat and breast cancer risk. In the most recent
update (16), we examined the relationship between fat intake
and breast cancer incidence over a 14-year period. This allowed
the use of repeated questionnaire measurements to provide a
better estimate of long-term average diet, the examination of
breast cancer incidence among postmenopausal women (who
were relatively few in our first analysis), and the evaluation of
a wide range of fat intake because of both the larger numbers
of cases and decreases in fat intake over time. During the
14-year follow-up, nearly 3000 women in the Nurses’ Health
Study developed breast cancer. As before, after adjustment for
standard risk factors we found a weak overall inverse associ-
dation between dietary fat intake, which ranged from <20% of
energy to >55% of energy. With the larger number of cases,
this inverse of trend was statistically significant.

During this period of time, six other large prospective
studies have examined the relationship between fat intake and
breast cancer incidence, with all finding weak or no association.
We and the investigators of the other large prospective studies
have conducted a pooled analysis combining data from all these
studies, which included nearly 5000 incidence cases of breast
cancer (17). Overall, there was no significant relationship be-
betwenn fat intake and breast cancer risk. However, among the
small number of women who reported <15% energy from fat,
we observed a significant 2-fold increase in risk of breast
cancer. This elevation in risk intake provides additional evi-
dence that a major reduction in incidence at very low intake is
unlikely. Furthermore, hyperinsulinemia, which has been hy-
pothesized to increase breast cancer risk (18), is exaggerated
with high-carbohydrate/low-fat intake among persons with
underlying insulin resistance (19). Because much of the post-
menopausal US population has insulin resistance due to inac-
tivity and overweight, it is conceivable that high carbohydrate
intake in the context of American lifestyles may increase breast
cancer risk.

Although fat and breast cancer have been the focus of
methodological developments in diet and cancer, similar find-
ings exist for colon cancer. A strong positive association was
suggested by international comparisons, but case-control stud-
ies have not shown an association with fat intake after ade-
quately adjusting for total energy intake (20). Prospective stud-
ies, although more limited in number and size, have also not
supported an association (21). Although not entirely consistent,
an association with red meat consumption has been seen in a
number of prospective studies (21), which is not true for breast
cancer.

Findings on dietary fat in relation to coronary heart disease
in the Nurses’ Health Study provide further evidence that our
dietary methodology can detect important associations. In the
14-year follow-up, which included nearly 1000 incidence cases
of coronary heart disease, we found a significant positive asso-
ciation with intake of trans fat, a weak positive association
with saturated fat, and significant inverse associations with
monounsaturated and polyunsaturated fat intake (22). These
correspond to the associations predicted on the basis of meta-
bolic studies examining the effects of specific dietary fats on
LDL and HDL cholesterol. Among women with high intake of
trans fat and low intake of polyunsaturated fat, compared with
those with low intake of trans fat and high intake of polyunsat-
urated fat, there was a 3-fold of higher risk of coronary heart
disease. These data indicate that the food-frequency question-
aires can detect important disease associations and that the type
of fat, not total fat intake, should guide dietary choices.

Another widely accepted hypothesis has been that higher
dietary fiber, particularly from grains, reduces risk of colon
cancer (23). This was largely based on cross-national compar-
sions and some animal studies. Although case-control studies
had suggested that higher intake of fiber from fruits and veg-
etables might be associated with lower risk of colon cancer,
such studies had quite consistently failed to demonstrate lower
risk with higher cereal fiber consumption (24), thus, raising
questions as to whether fiber per se was the responsible factor.
In our prospective studies of both men (25) and women (26), we
found no relationship between overall dietary fiber or fiber
from cereal products and risk of colon cancer. Within the same
cohorts, we also found no evidence that dietary fiber was
related to occurrence of colon adenomas in men (27) or women
(26), thus reducing the likelihood that an association with colon
cancer had been missed simply because we were observing the
process too late. A lack of association between fiber intake and
colon cancer has also been seen in other prospective cohort
studies (28–31). In contrast, in both men (32) and women (33)
we found significant inverse associations between cereal fiber
intake and risk of coronary heart disease, with an ~30% lower
risk among those in the highest compared with the lowest
quintiles of intake. Similarly, we have seen inverse associations
between fiber consumption and risk of adult onset diabetes in
both cohorts (34, 35) and diverticular disease (36). Again, the
non-cancer outcomes provide strong evidence that we do have informative variation in fiber intake and the ability to measure it within our cohort studies. Recently, intervention trials among individuals with a previous diagnosis of adenoma have also provided evidence that there is little or no important effect of fiber intake up to \( \sim 30 \) g per day on recurrence of adenomas (37, 38).

High consumption of fruits and vegetables has been thought to be particularly important for cancer prevention and has served as the basis for the national “5-a-day” campaign. A recent major review concluded, “There is convincing evidence that diets high in vegetable and/or fruits protect against cancers of the mouth and throat, esophagus, lung, stomach, and colon and rectum. Such diets probably also protect against cancers of the larynx, pancreas, breast, and bladder . . .” (23). However, the vast majority of literature up to this point has been based on case-control studies, and in recent cohort studies the associations usually have been considerably weaker. For example, in 32 case-control studies of stomach cancer, there was a remarkably consistent inverse association between consumption of fruits and vegetables (23), whereas in the first detailed prospective study, there was no overall association after excluding the first year of follow-up (39). In addition, we have recently combined the data from the Nurses’ Health Study and Health Professionals Follow-up Study to examine the relationship between consumption of fruits and vegetables and risk of colon cancer, which provided a particularly powerful analysis including nearly 1000 incidence cases (40). No overall association was seen, and there was no suggestion of an inverse association with any specific grouping of fruits and vegetables. A number of possible reasons may explain why the inverse relationships reported from case-control studies may have been overstated. First, recall bias is possible in the context of a case-control study. This bias is almost guaranteed to exist in a situation where control participation is less than complete because those who participate are likely to be more health conscious and consume greater amounts of fruits and vegetables. In addition, selective reporting and publication is likely in part because many fruits and vegetables are examined and typically only a few are reported. Frequently, studies are said to be positive when only one or a few of the fruits and vegetables examined were significantly related to risk of cancer, which may be due to chance (41, 42). Also, the broad category of “fruits and vegetables” may be too nonspecific; if biological effects are related to specific nutrients, then perhaps only one or a few vegetables may be related to a specific cancer. Furthermore, within the US, use of vitamin supplements and fortified food is now widespread so that a beneficial effect of fruits and vegetables could well be missed if it was due to one of the nutrients contained in supplements. Furthermore, it is also possible that the risk of cancer becomes elevated only at a very low intake of fruits and vegetables, i.e., the relationship is nonlinear, and few individuals in the populations being studied consume such a low amount.

Although increasing fruits and vegetables overall appears to be less promising as a way to substantially reduce cancer risk, multiple lines of evidence suggest that higher intake of folic acid may be beneficial. Two hypotheses exist whereby folic acid could reduce cancer incidence. First, folic acid provides a methyl group needed for the synthesis of methionine, which then is used for DNA methylation and regulation of gene expression. In addition, folic acid provides a methyl group necessary for the conversion of uracil to thymine, and inadequate folic acid can result in the replacement of uracil for thymine in DNA. Thus, among individuals in the lowest quartile of red blood cell folate, a 10-fold increase in uracil incorporation into DNA was observed, which was reversible with folic acid supplementation (43).

Because of the potential for reduction in cancer risk with higher folic acid intakes, we examined the associations with colon and breast cancer in our prospective studies. We have good evidence that our dietary questionnaire measures folic acid well. In the Framingham Heart Study population, plasma folic acid levels were positively related to intake assessed by our food-frequency questionnaire, and plasma homocysteine levels were strongly inversely related to our calculated folic acid intake (44). In a combined analysis of the Nurses’ Health Study and Health Professionals Follow-up Study, Giovannucci et al. (45) found that folic acid intake was inversely related to risk of colon adenoma. Furthermore, use of multiple vitamins containing folic acid for 15 or more years was associated with a substantial reduction in risk of colon cancer in the Nurses’ Health Study, but shorter-term use had only a weak relationship with risk (46). These data suggest that folic acid consumption during the early stages of colon carcinogenesis may be particularly important. Further evidence for a role of folic acid in colon carcinogenesis is provided by a relationship with a functionally important polymorphism in the methylenetetrahydrofolate reductase gene. Specifically, the val/val genotype, which results in lower enzymatic activity, has been associated with reduced risk of colon cancer (47, 48). Although additional confirmatory data are needed, this association provides support for a role of folic acid because it is unlikely that a polymorphism in an enzyme involved in folic acid metabolism would be associated with colon cancer if the substrate for the enzyme did not influence the carcinogenic process.

Alcohol consumption, a known risk factor for colon cancer, can interfere with folic acid availability (49–52), and an interaction between folate intake and alcohol consumption has been seen in relation to colon cancer risk (53). Specifically, among men consuming the highest amounts of folate acid, alcohol did not appear to increase risk of colon cancer. A similar interaction between folic acid intake and alcohol consumption has also been seen for breast cancer; alcohol consumption appears to increase breast cancer primarily among women who also have low folic acid consumption (54). These findings, which have practical implications, have recently been reproduced by others (55, 56).

If neither fat, fiber, nor fruits and vegetables are important risk factors for breast and colon cancer, then what can explain the large international differences in cancer risk? A consideration of the animal literature is worthwhile and clearly suggests that the most powerful and consistent dietary influence on carcinogenesis is simply energy restriction (23). This has been documented for numerous tumor sites and across multiple animal models. The effects are often profound; for example, a 30% restriction in energy intake can reduce mammary tumors by as much as 80% (57). The study of energy intake in human cancer is complex because energy intake is largely determined by physical activity; thus, lean individuals and populations often have higher energy intake because of high levels of physical activity. An assessment of the difference between energy intake and energy expenditure is not feasible because neither variable is measured with adequate precision to assess the small differences that could cumulatively be important over many years. Fortunately, indirect indicators of energy balance, such as body weight and changes in weight, are readily and precisely measured (7).

Epidemiologic data accumulated during the last decade
have provided strong support for an important effect of energy balance on human cancer risk. A positive association between adiposity and risk of colon cancer has been noted in men for a number of years, but more recent data indicate that the relationship exists for women as well (58, 59). Waist circumference may be an even stronger predictor for colon cancer risk in men (60). Although the relationship between body weight or weight change and breast cancer risk has generally been weak in most prospective studies, this appears to be due in part to a masking effect of hormone replacement use, which leads to elevated levels of circulating estrogens regardless of body weight. Among women who never use hormone replacement therapy, a strong dose-response relationship between weight change during adult life and breast cancer risk has been observed (61–66).

Regular physical activity contributes to a lower prevalence of overweight and obesity, and will thus help reduce risk of breast and colon cancer by this mechanism alone. Studies of physical activity and cancer risk suggest that regular activity also has an independent effect. The evidence is quite consistent for colon cancer (58, 60, 67). For breast cancer, the association has been somewhat less consistent, but the overall evidence suggests that a modest benefit exists (68, 69). More data are needed to further our understanding regarding the form of activity and period in life when the benefits are greatest. Because of imperfect assessments of physical activity, both within countries and between populations, it is presently difficult to estimate quantitatively the contribution of physical activity to international differences. However, because the levels of physical activity in traditional rural societies are probably much higher than present US levels, the contribution is likely to be large.

We have recently examined within our cohort of men the proportion of colon cancer that is potentially preventable by simultaneous reduction of six modifiable risk factors (70). For this analysis, we defined a low-risk group as body mass index <25 kg/m², ≥30 min a day of vigorous to moderate physical activity, alcohol consumption <15 g/day or 15–30 g/day with supplemental folic acid, folate supplement of 100 μg/day or more, less than three pack-years of lifetime smoking, and two or fewer servings of red meat per week. Although none of these definitions involve extreme changes in lifestyle, only 3.1% of the population fell into the joint low-risk group. On the basis of the data within this cohort of men, the population attributable risk was 71% (95% confidence interval, 33–92%), suggesting that the large majority of colon cancer in this group is potentially avoidable. Because this group had already adopted many health-conscious behaviors compared with the general United States population, the overall percentage in the US population should be even higher.

Summary

Available evidence suggests several conclusions:

(a) Weight control and regular physical activity should be high priorities for prevention of cancer as well as cardiovascular disease.

(b) Decisions about dietary fat should be made primarily on its role in human cancer risk. A positive association between physical activity and period in life when the benefits are greatest. Because of imperfect assessments of physical activity, both within countries and between populations, it is presently difficult to estimate quantitatively the contribution of physical activity to international differences. However, because the levels of physical activity in traditional rural societies are probably much higher than present US levels, the contribution is likely to be large.

(c) Consuming an abundance of fruits and vegetables and eating grains in a minimally processed, high-fiber form is desirable, but the benefits appear greater for cardiovascular disease than for cancer.

(d) Adequate folic acid intake appears important in reducing cancer risk, particularly if alcohol is consumed regularly. Taking a multiple vitamin is the most reliable way.

(e) Much remains to be learned about diet and cancer, in particular about long-term effects. A wealth of new data will emerge from the more than 30 large prospective studies that are already underway (7).

References


Diet and Cancer: One View at the Start of the Millennium

Walter C. Willett