Can We Really Derive Etiology from Human-based Studies?1

James R. Marshall2

Department of Social and Preventive Medicine, State University of New York at Buffalo, Buffalo, New York 14214

Abstract

This paper discusses the need for and limitations of human-based studies of diet and cancer. Four problems of such studies remain especially acute: assessment and monitoring of dietary practice; assessment of compliance in diet change studies; validation of dietary questionnaires; and the use of intermediate biological end points. These limit human-based studies' utilization of recent advances in the cellular and molecular biology of cancer. Those researchers who emphasize human-based studies cannot proceed without acknowledgment of and interface with these advances, which promise more precise definition of both exposure and disease and may facilitate evaluation of associations of dietary practice and carcinogenesis in the context of human-based studies. To take advantage of these advances will require critical attention to the causal mechanisms by which intermediate biological end points link dietary exposures with cancer risk.

Introduction

Because this symposium was intended to consider the state of our knowledge with respect to the human-, animal-, and culture-based study of diet and cancer, I begin by considering the distinct contributions and limitations of human-based studies. I then deal with the questions of how to monitor food intake, how to assess compliance in human-based studies, how to validate dietary questionnaires, and how to choose intermediate end points. I also discuss where biochemistry and molecular biology fit in the consideration of these questions by epidemiologists. Finally, I address what might be done to make human-based studies contribute more effectively to our understanding of the etiology of diet and nutrition in cancer.

Contributions and Limitations of Human-based Studies

It is generally not difficult to generalize human-based studies to people. No small matter. With our concern over representative sampling in studies, we worry about whether the clinical or cooperative populations we have studied are typical and whether the experience of these populations applies to that of others. We have lately seen concern expressed over whether cancer studies conducted on men generalize to the other half of the human population. Note that all these studies involve extrapolating from one category of human to another. Generalizability is in general a qualitatively less thorny issue with human-based studies than it is with others.

We study humans because we wish to know the actual and relative strength of various exposures as risk factors for human cancer. We are concerned with confounding. We are concerned with effect modification: with whether phenotypic characteristics alter the effects of epidemiological exposures.

Among humans, we study the effects of exposure on cancer in terms of the remarkable heterogeneity of people. Many of the findings which at first blush strike us as exciting and promising, with time, present a more sobering and complicated picture. The statistical variability of human populations is not kind to our relatively simplistic generalizations. For example, we have seen the causal scenario derived from Vogelstein's (1) exciting findings: A process of hypomethylation engenders deletion on the 5th chromosome, various ras mutations, deletions on the 18th chromosome, deletions on the 17th chromosome, and so on. These are often perceived as an orderly cascade. It is likely, however, that the cascade is not orderly; as Vogelstein has recognized, exceptions are nearly as common as manifestations, and we will in all likelihood get a much more variable picture when we study large human populations. One of the distinct contributions of human-based studies is that they force us to confront human heterogeneity in our efforts to understand cancer.

With human-based studies, we evaluate the entirety of the process of cancer etiology. We attempt to draw statistical links between exposure as it occurs in the natural world and the risk of cancer. We deal less with doses which produce disease in 50% of exposed animals, or with cultures in which substantial proportions of the cells manifest changes, than with the slight changes in risk that accompany the probable presence or absence of this or that nutrient in the diet.

We pay dearly for these advantages. Human-based studies are expensive, and they require large amounts of time to complete. If we could experiment on human populations, we would need to watch them for decades to observe their subsequent experience. We cannot, however, experimentally induce or even reduce disease in human populations except in very limited circumstances. We cannot invade people's bodies simply for research purposes. We cannot interrupt the process of disease formation at various points in the process, as we can with rodents at 8, 16, and 32 weeks. We must work in large part with the natural experiments afforded us by the play of chance in human life.

How to Monitor Food Intake

Any influence of diet on cancer risk probably involves exposure to etiologically relevant elements over an extended time span. In addition, there is probably, even for cancers with shorter gestational periods, a lag of years between exposure and disease. Thus, our dietary instruments must summarize diet over an extended period. For some studies, such as case-control studies, our diet summary must describe diet over an extended period at an earlier point in time.

Three methods offer attractive approaches to the description we need (2). The brief period recall allows the respondent to focus on a brief time span during which recall is relatively straightforward. Such a recall is intended to provide a very accurate snapshot of a small sample of the long-term diet. The snapshot, however, is subject to a great deal of sampling error. Moreover, in the absence of evidence that the period of data collection represents earlier points in time, one would be chary about using this brief recall method to refer to earlier dietary practices.

Inducing subjects to record their food intake over a few days provides a clear, specific data focus. It is likely that one's recorded dietary practice is a relatively faithful representation of actual diet during that period. The use of the method for studies of diet and cancer requires that the method not alter actual dietary practice, that the number of days of records provides an adequate sample size, and that the reference period be well represented by the period during which data are recorded.

The third method, food frequency, requires the subject to describe his or her typical diet during a reference period. He or she is asked to describe the frequency and often the quantity of his/her consumption of a number of basic food items. The focus of the data request is...
ETIOLOGY AND HUMAN-BASED STUDIES

neither distinct nor well identified. However, the reference period can readily be specified, and that reference period is by nature diet over an extended time span.

There appears to be a growing consensus that, in general, food frequency instruments provide a modestly effective means of evaluating food intake. The food frequency provides information integrated by the respondent over time: a broad brush picture of general dietary practices. In general, validity studies suggest that the food frequency is able to account for a substantial proportion of variance in nutrient intake (2). However, the effectiveness of the food frequency varies; it is substantial for describing total and saturated fat, cholesterol, carbohydrate, carotene, dietary fiber, vitamin C, and vitamin E intake, less so for total calorie, protein, and folate intake (2).

What is important about any method is that we can improve our data by repeat measurement. To the extent that we can design studies which enable us to measure and measure again, and we have not yet begun to explore the situations in which this is readily possible, we can increase the accuracy of our measures a good deal (3).

What is important about our use of repeat measures is that the focus of the measures be clearly specified and carefully considered. For example, one report may refer to time \( t \), a repeat measure to time \( t + n \). Differences in the two reports would partly result from the fact that they refer to different periods: distinguishing the proportion of non-reproducibility due to time change from that due to reporting inaccuracy is quite difficult.

What is also important is that errors in repeat measures not be repeated. Respondents may on one measure have over- or under-reported their true diet; the second measure should be designed to minimize the likelihood that they will make the same mistake again. Thus, for example, in the Polyp Prevention Trial, in our efforts to make our measurement as effective as possible, we measure diet by a food frequency method, but we also use a 4-day food record. Our goal in this use of two procedures is to lessen the likelihood that errors accrued as part of the food frequency method will be reported as part of the 4-day food record.

Our inability to accurately monitor food intake imposes a major limitation on our ability to use human-based studies to explain the role of diet in cancer. We know that, under the best, most predictable of circumstances, error can severely attenuate our estimates of the effects of the 4-day food record. Inasmuch as a central goal of our human-based studies is to compare the relative magnitudes of nutrients \( a \) and \( b \), we need to continually remind ourselves that, if nutrient \( a \) is measured a good deal more effectively than nutrient \( b \), our measure of the effect of nutrient \( a \) will be less biased than will that of nutrient \( b \). Unfortunately, we see in epidemiology at best only ceremonial attention to measurement error as it might obscure our knowledge. A greater level of sophistication is needed if human-based studies are to reach their potential to inform our cancer prevention efforts. As an example, consider the impressive efforts of Howe et al. (4), in pooling the major case-control studies of diet and breast cancer. On the basis of his analysis, Howe concluded that fat increases risk, that vitamin C intake is protective, and that \( \beta \)-carotene has no effect. This study, however, does not consider the possibility that the reliabilities of fat, vitamin C, and \( \beta \)-carotene in all likelihood differ substantially. Procedures for addressing this possibility have been developed by Rosner and Willett (5). It is possible that vitamin C exposure is measured better than that of \( \beta \)-carotene; in general, considering the possibility that the reliabilities of certain nutrients exceed those of others might substantially improve our understanding.

How to Validate Dietary Questionnaires

A central goal of our human-based studies of nutrition and cancer is to compare the relative magnitudes of effects, comparing one exposure to another. Thus, we wish to consider individuals along a spectrum of exposures to various nutrients and evaluate those individuals with respect to their relative intake of nutrient \( a \) as opposed to nutrient \( b \). Our concern is in comparing individuals, comparing their relative intake. We are less interested in the exact actual amount of substance or nutrient to which that subject is exposed than in the subject’s relative intake of that substance or nutrient. As we showed (7), the replicability of our instrument will provide an index of reliability and that, in most cases, constitutes the information we need. There has been much worry about validating questionnaires but perhaps too little attention to understanding the replicability or reliability of our measures of different nutrient exposures. We worry about what a “gold standard” would reveal. We should. Nonetheless, with dietary practice after intake lessens the likelihood that the subject’s dietary practice will be altered by the fact that it is being observed. It will undoubtedly be colored by that subject’s understanding of what he or she is supposed to have been doing. One approach to this problem in the Polyp Prevention Trial has been to adopt and use a 24-h recall. The subject is called with no advance notice and asked to recall what he or she has consumed in the past 24 h. One’s diet in the past 24 h provides a direct, distinct, and objective focus that minimizes the possibility of perception altering recall. In addition, focusing on dietary practice after intake lessens the likelihood that the subject’s dietary practice will be altered by the fact that it is being observed. It is more likely to be reflective of true diet. The Polyp Prevention Trial also uses a series of biological markers derived from blood samples to provide what is hoped will be an additional, potentially more objective indicator of compliance with dietary prescriptions.

How to Assess Compliance in Human-based Studies

The problems of monitoring food intake in general intensifies as we attempt to assess compliance in human-based studies. Our concern is response to an intervention. We want to understand the effect of the intervention on the dietary practices of experimental and control subjects. We have been encouraging our experimental subjects (cajoling and pressuring them) to change their dietary practice. We want to know whether they have, and so we ask them. We hope that their responses to our questions will be objective and truthful; that those responses will be reliable measures of the changes that have actually taken place. Unfortunately, people’s responses are filtered by their perceptions, and in their perceptions they are aware not only of what they have been doing but of what the experimenter wants them to have been doing. Our subjects will be tempted to see themselves as they think they should be, and they will be tempted to describe themselves as they think they should be. Thus, in the Women’s Health Trial pilot project, the experimental subjects reported substantial alterations in dietary practice (6). In fact, the experimental subjects’ caloric intakes as measured by a 4-day food record were so low that those experimental subjects should have been losing substantial amounts of weight relative to controls. They were not. This does not mean that the subjects were intentionally dishonest. It suggests, however, that their reporting suggested different behavior during the evaluation period than their weight change indicated. The subjects tended to see themselves as compliant and they were, especially when they were being monitored. They were perhaps more compliant when they were being watched than when they were not.

We want our markers to be passive indicators of actual dietary practice—of change in dietary practice—and we want errors to be independent of program goals. This is less likely than desirable with the conventional food frequency instrument; the subject’s perception will undoubtedly be colored by that subject’s understanding of what he or she is supposed to have been doing. One approach to this problem in the Polyp Prevention Trial has been to adopt and use a 24-h recall. The subject is called with no advance notice and asked to recall what he or she has consumed in the past 24 h. One’s diet in the past 24 h provides a direct, distinct, and objective focus that minimizes the possibility of perception altering recall. In addition, focusing on dietary practice after intake lessens the likelihood that the subject’s dietary practice will be altered by the fact that it is being observed. It is more likely to be reflective of true diet. The Polyp Prevention Trial also uses a series of biological markers derived from blood samples to provide what is hoped will be an additional, potentially more objective indicator of compliance with dietary prescriptions.
suspected. Validity studies tend to provide information similar to what is found in a focus on reliability. In large part, our instruments appear to generate data in which between 40 and 60% of the observed variance is error variance. The rest is variance associated with actual exposure. For their expense and logistical encumbrances, validity studies provide little that is not extractable from reliability studies. The fact that validity studies are not easy to carry out should not deter us from reliability studies. Every diet and cancer study should involve careful evaluation of the reliability of diet data, considered among cases separately from controls.

How to Choose Intermediate End Points

In their promising book, Hulka et al (8) have pointed to distinctions among types of intermediate end points. Some are markers primarily of exposure, and others are markers primarily of very early disease. Those concerned with screening will be interested most in early markers of disease. Epidemiologists are interested in markers of exposure as well as in markers of disease.

These intermediate end points usually involve the assay of some bit of biological material: blood, skin, intestinal mucosa, urine, feces. We seek them because we hope by their use to increase the accuracy (the objectivity) of our measures of exposure or of disease presence.

We want the link of exposure markers to exposure to be as direct as possible. We are hindered in our search for such passive markers to the degree that our markers are a function not just of exposure but also of response. The presence of metabolized exogenous agents, or cellular or molecular changes, may mark not just exposure but also response. Other exposures may cloud our use of markers. The level of β-carotene in blood may be a function not just of β-carotene intake but also of other environmental factors that may alter absorption or retention of β-carotene. Smoking may be one such factor.

With markers of disease, we may come to find ourselves on the horns of a dilemma; earlier markers may be desirable because they appear well before the onset of disease, but they are less likely than those more proximal to symptomatic disease to have high predictive value with respect to actual disease.

It is also worth considering markers in the context of the central task of the nutritional epidemiology of cancer: to evaluate the association of diet and nutrition and cancer. Thus, our first need will be for markers primarily of exposure and also for markers of disease. It might well be that, in light of this central task, biological markers which are more strictly reflective of exposure or of disease are more valuable. Such markers facilitate our ability to refine our understanding of associations between exposure and disease. Markers which reflect both exposure and disease may prove of less value, because they obscure our ability to evaluate the statistical association of exposure and disease.

One of the great untapped reservoirs of material for epidemiology may be early, presymptomatic disease. To the degree that markers are highly predictive of actual disease, their epidemiology should help to inform the epidemiology of the disease they predict. An excellent example is our recent experience with respect to adenomatous polyps. Most adenomas never progress to become colon cancers: they are removed when first observed. In any case, most would never have an opportunity to mature and reach the point where they could cause actual cancer. They are small, relatively inactive, and not highly dysplastic (9). Thus, the positive predictive value of the presence of adenomatous polyps is relatively low. Many of those with polyps will never develop cancer. On the other hand, it is clear that almost no cancer of the colon ever develops except out of an adenomatous polyp. Thus, the negative predictive value of the absence of polyps for the nondevelopment of cancer is virtually 100%.

Therefore, although polyps do not commonly become colon cancers, few cancers emerge but from polyps. Because they are generally asymptomatic and much more common than colon cancer, screening studies which note their presence can be used as case-control studies. Such studies avoid the problems of recall bias generally encountered among sick patients.

Biological markers may represent a tremendous, untapped resource. Nevertheless, for all their attractiveness, they must be carefully used. A marker of the maturity of a form of pathology may be unlikely to inform us. Epidemiologists may be tempted to propose to evaluate colonic polyps with chromosome 5, 17, or 18 deletions, comparing the dietary practices of those whose polyps have them to those whose polyps do not. It is likely that we will, in the very near future, see studies of the alteration of the p53 gene, comparing "cases" with a mutation to "controls" without it. To the extent that these are a function mainly of stage or maturity of the lesion, whether polyp or cancer, these studies may not be very informative.

A more useful approach might be to attempt to focus on biological markers of process rather than of the stage of disease. Hulka has emphasized the importance of a focus on process. For example, proliferation of the colonic mucosa appears substantially predictive of subsequent risk of disease (10). Proliferation, a process, appears to generate a number of changes in the colon. Similarly, the process of methylation of the DNA, however that takes place, also appears to be related to the risk of polyp and cancer formation (11). Thus, we see two disease processes; biological markers of the activity of those processes might be much more useful than the presence of intermediate markers of stage of disease in informing our efforts.

Where Do Biochemistry and Molecular Biology Fit in Human-based Studies of Diet and Cancer?

These open new doors for use of biological markers. It is important that the question as phrased is not "Do these fit" but "Where do these fit?" It is clear that health and disease are increasingly being seen as a continuous process. Dietary exposures initiate processes which lead eventually to cancer.

The interest in mechanisms of biochemistry and molecular biology has forced epidemiologists to sharpen their traditional focus on the simple statistical association of exposure and disease. We can, for example, distinguish with a great deal more sophistication various forms of disease. We can also by the use of markers of disease risk focus on individuals who are at elevated risk of disease but who have not yet even begun to exhibit frank symptoms of disease. As mentioned, our ability to mount case-control studies, comparing individuals at elevated risk with those who are not, has gained greater and greater actuality. In the Polyp Prevention Trial, we use biological markers of colorectal cancer risk, rather than cancer, as our dependent variables, focusing both on adenoma formation and on proliferation.

However, two notes of caution regarding our forays into molecular biology are appropriate. First, if an association between exposure and disease is weak, it may be tempting to evaluate a series of biological or molecular biological intermediate end points by which that exposure weakly transforms the risk of disease. Filling in details, in this case, though, is not in all likelihood going to be helpful; it will simply provide information on details of a weak process. If we hope to explain much about alterations in the risk of disease, we will need to use advances in molecular biology to help us identify genotypes and phenotypes in which exposure-disease associations vary in strength. It is readily possible that, in subsets of the population identified by the use of genetic markers, exposure may substantially increase the risk of disease. In other subsets of the population, however, the effects of
exposure on disease will be weaker or even negligible. It may well be in this context that advances in molecular genetics hold the most promise.

Second, reference to mechanism must be used cautiously as a means of explaining, or of explaining away, scientific findings. An interesting example of this is seen in Meyskens’ thoughtful comment on the finding of Greenberg et al. that β-carotene supplements do not decrease the risk of non-melanoma skin cancer recurrence (12, 13). Meyskens suggested that Greenberg et al. saw no effect because β-carotene might depress vitamin E levels, that it is actually that decreased vitamin E is etiologically significant rather than that increased β-carotene is not. He may well be right. Nonetheless, this should not distract attention from the fact that the epidemiological association between β-carotene and skin cancer recurrence risk was shown convincingly to be null by Greenberg and Baron’s excellent work. Their study showed convincingly, on a treatment and biological level, that there is no such association. Whether there could be some interaction between β-carotene and vitamin E levels, such that, if β-carotene and vitamin E levels were substantially increased there would be a lessened risk of recurrence, is another question.

What human-based studies can contribute to our understanding of nutrition and cancer is the connection between diet-exposure and nutrition and the subsequent risk of disease as it occurs in natural human populations. The experience of nutritionists in understanding the etiology of coronary heart disease may be instructive. With substantial attention devoted to the association of lipid status and subsequent CHD^1 risk, there is no question that certain lipid profiles substantially increase risk. On the other hand, however, the connection between dietary practice and CHD risk is much less clear. Willett (14) recently claimed that, in spite of the monumental efforts devoted to etiology based on human studies of CHD over the past 50 years, our knowledge of the connection between dietary practice and CHD risk is surprisingly limited.

Thus, while molecular biology allows us to collect substantial amounts of data, we must remain careful about how we use that information. Biomarkers, including those made possible by molecular biology, certainly provide the opportunity for us to better evaluate indices of exposure and to better document the effects of exposure on changes at the tissue level. Molecular biology also affords us the opportunity to better understand presymptomatic disease and its determinants. This is especially true when we can identify the substrate of changes from which disease springs.

**General Recommendations**

What can we improve in our human studies to make them more useful components of our effort to understand the etiology of diet and nutrition in cancer? I wish, in summary, to address these on a global level, considering two summary recommendations about our attempts to evaluate the many possible links among diet, nutrition, and cancer.

First, we need to continue to attempt to better appreciate the strength and weaknesses of our dietary measures. We need to assess the reliability of every measure within our armamentarium and to evaluate that reliability every time we undertake an epidemiological study. We simply cannot in general assume that the reliability of a measure documented in one region or county with one population will be useful for every population we choose to study. Neither can we simply throw up our hands at the weakness of our measures and hope things work out. We must do all we can to improve our measures, and we must continue to evaluate how good they are.

Second, we need to seek better and more extensive use of our data. Many of the niceties of experimental research are not possible with human populations. If study A reveals a strong nutrition-cancer association and study B reveals a much weaker one, we need better means of reconciling those findings (15). The reliability-sensitivity evaluations suggested are merely a starting point. We need to work less on "discovering" associations and more on evaluation (critical evaluation) of the robustness of our discoveries. We need to continue to seek to transform our studies so that they are as amenable as experiments to replication and to scrutiny by other investigators. Experiments can be replicated. Human-based study, in general, is not subject to that degree of scrutiny. It is often not possible for several investigators to precisely replicate a study which has uncovered a surprising finding. Even if the study were replicated, slight changes in questionnaire wording, subject recruitment, or societal dietary practice could change the nature of the findings and even the association of exposure and disease. Even with its drawbacks, Howe’s pooled analysis is a good model.

It is possible, however, for investigators to provide better access to their data. The stuff of the experiment for the human-based scientist is the data set which he or she has collected. It would be reasonable, it seems, to expect that any investigator who has published a finding should provide access to the original data which generated that finding. Other investigators should have the opportunity to evaluate the nuances of the data: decisions about coding, data analysis, computer programs used; how outliers were handled; the procedures used for controlling or not controlling for covariates. These would, of course, make it somewhat more difficult for epidemiologists to publish findings as rapidly and as blithely as we presently do. If we knew others were looking over our shoulders, we would be more careful about establishing the robustness of our findings.

This is not to suggest that we should overanalyze our data, torturing them until the data surrender the desired result, merely that we should individually and together make more use of those data than we presently do. At present we often see a finding analyzed with control for a series of variables as well as with control for no other variables. What is often troublesome is that the criteria used by the investigator for deciding whether to control or not control for any of a host of variables are not made very clear. Thus, the reader is not permitted to evaluate the extent to which the findings are an artifact of the analysis. It might be far more useful if investigators would focus on revealing the robustness of their findings and the fact that any one finding does or does not go away with control for dozens if not hundreds of covariates.

In the social sciences, books are often written on scantier data sets than go into the standard case-control epidemiological study with 300 cases and 300 controls which produces one paper. Books may not always be justified by such data sets. On the other hand, however, it is not difficult to imagine that much more could be made of the data. Often, in fact, it seems as if the motto or goal of the epidemiologist is to make as little of the data as possible rather than as much. Funding agencies must be encouraged to support data analysis.

We have considered, in the light of what is unique about human-based studies, how to monitor food intake in general, how to assess compliance in intervention studies, how we might best validate dietary questionnaires, how to choose intermediate end points for diet and cancer and chemoprevention studies, and where biochemistry and molecular biology fit in the consideration of these issues by epidemiologists. We must recognize the distinct role of human-based studies as we design our studies, our data collection, and our use of the data to maximize our contributions to the understanding of diet and cancer.

**References**

ETIOLOGY AND HUMAN-BASED STUDIES

Can We Really Derive Etiology from Human-based Studies?

James R. Marshall

_Cancer Res_ 1993;53:2437s-2441s.

Updated version

Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/53/10_Supplement/2437s

E-mail alerts

Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions

To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/53/10_Supplement/2437s. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.