

Causal criteria in nutritional epidemiology^{1,2}

Nancy Potischman and Douglas L Weed

ABSTRACT Making nutrition recommendations involves complex judgments about the balance between benefits and risks associated with a nutrient or food. Causal criteria are central features of such judgments but are not sufficient. Other scientific considerations include study designs, statistical tests, bias, confounding, and measurement issues. At a minimum, the set of criteria includes consistency, strength of association, dose response, plausibility, and temporality. The current practice, methods, and theory of causal inference permit flexibility in the choice of criteria, their relative priority, and the rules of inference assigned to them. Our approach is as follows. Consistency across study designs is compelling when the studies are of high quality and are not subject to biases. A statistically significant risk estimate with a >20% increase or decrease in risk is considered a positive finding. A statistically significant linear or otherwise regularly increasing trend reinforces the judgment in favor of a recommendation. A plausible hypothesis likewise reinforces a recommendation, although the rules of inference for biological evidence are highly variable and depend on the situation. Temporality is, for nutrition recommendations, more a consideration of the extent to which a dietary factor affects disease onset or progression. Evidence supporting these criteria provides a strong basis for making a nutrition recommendation, given due consideration of the balance between presumed benefits and presumed harms. Recommendations should make clear their breadth of application; a narrow recommendation involves a single disease or condition whereas a broad recommendation involves all relevant diseases or conditions. *Am J Clin Nutr* 1999;69(suppl):1309S–14S.

KEY WORDS Causation, diet, epidemiology, nutrition recommendations, disease prevention, cancer prevention, disease etiology, causal criteria, causal inference, dose response, temporality, plausibility, study design, bias, confounding

INTRODUCTION

Causal criteria are central to the inferential methods currently used by epidemiologists to make nutrition recommendations, although these recommendations do not arise solely from the application of criteria. Other scientific considerations, such as study design types, statistical tests, bias and confounding, and the quality of measurements are also important. Ethical considerations also play a role in making nutrition recommendations, as they do for all forms of public health recommendations. The overriding ethical concern is whether a favorable balance of benefits

and harms is expected within a population if recommendations are adopted (1). Whether narrowly conceived as pertaining to a single disease only or broadly conceived as taking into consideration all relevant diseases and conditions affected by the consumption of the nutrient (or food), nutrition recommendations are normative, ethical statements about what ought to be done for the public health.

Nutrition recommendations cannot easily be separated from causal conclusions, which are a form of scientific statement that also emerges from the application of criteria-based inferential methods. We may conclude, for example, that a specific food-related chemical or dietary component causes or prevents a certain cancer. We may subsequently conclude that it is appropriate to warn the public that people should avoid ingesting that factor, or in the case of prevention we may recommend eating more of it. However, the relation between causal conclusions and public health recommendations is considerably more complex than this oversimplified example suggests because of the tangled relation between science and its application. Public health recommendations are possible and even appropriate without firm causal or preventive conclusions but are typically not made without some evidentiary support. Indeed, navigating the space between scientific evidence and public health action in underdetermined circumstances is the basic challenge in making nutrition recommendations.

This article describes how causal criteria are used to help meet the challenge of making nutrition recommendations in the face of complex scientific evidence. Published examples (2) reveal how epidemiologists and others select, interpret, and apply criteria when making nutrition recommendations. We briefly examine the links between the practice of making nutrition recommendations and the methodologic and theoretical literature on inference. Our understanding of the current practice, method, and theory of causal inference in epidemiology provides the rationale for future applications of inferential criteria, with special emphasis on their relevance to nutritional epidemiology.

BACKGROUND: METHODOLOGY, PRACTICE, AND THEORY

Several sources were consulted in developing our view of causal criteria: a historical account of causal inference method-

¹From the National Cancer Institute, Bethesda, Maryland.

²Address reprint requests to N Potischman, Department of Biostatistics and Epidemiology, Arnold House, University of Massachusetts, Amherst, MA 01003-0430. E-mail: nap@schoolph.umass.edu.

ology and epidemiology textbooks published since the late 1950s (3), 2 accounts of causal inference in nutritional epidemiology (4, 5), and a review of the practice of causal inference in cancer epidemiology (2). On the basis of these assessments, it seems reasonable to make the following summary statements:

- 1) The criteria appearing in articles and textbooks on methodology are typically subsets of the list proposed by the Surgeon General's Committee on Smoking and Lung Cancer in 1964 (6) and expanded on by Hill in 1965 (7).
- 2) The causal criteria used in the practice of causal inference, ie, in review articles, textbook chapters, and editorials are also subsets of the same lists created >30 y ago by the Surgeon General's committee and Hill, although the relative frequency of use is different from the patterns found in the methodologic literature. In practice, the criteria of consistency, strength of association, dose response, and plausibility are used frequently and in that order (2), whereas in the methodologic literature the criteria of strength of association, temporality, consistency, plausibility, dose response, and specificity are most often mentioned (in descending order).
- 3) Across the entire discipline of epidemiology, review articles in which the authors used causal criteria lack methodologic rigor (8). Potential improvements include a careful description of literature search techniques as well as inclusion and exclusion criteria for the literature examined in the review. Although these strategies may not completely control for the exigencies of publication bias, they nevertheless make clear what literature was sought, found, excluded, and reviewed. Improvements in review articles most relevant to causal criteria include providing reference sources for causal criteria and, more importantly, a justification for the criteria selected. Another improvement would be the provision of a clear description of the rules of inference used for each criterion, ie, the types or extent of evidence needed to meet that criterion (9).
- 4) There is only a loose consensus regarding the relative importance and priority of the criteria and the rules of inference assigned to any criterion. The selection, ranking, and definitions of criteria vary from user to user. The criterion of consistency, for example, has been interpreted as an all-or-none phenomenon, a majority-rules phenomenon, or a testable statistical hypothesis about the relative proportions of positive, negative, and null studies (3). Likewise, the criterion of strength of association is interpreted variably, with many different opinions on what size of relative risk constitutes a weak association.

Causal inference involves more than methodology and practice. There is also an underlying theoretical approach to causal thinking that involves issues such as what constitutes the nature of causation, the logic of causal inference, the epistemologic relation of scientific evidence to causal models or theories, and the aforementioned relation of science to ethics. These theoretical commitments are almost never stated by those practitioners who review evidence and make public health recommendations; nevertheless, it is possible that decisions made about public health interventions are influenced by underlying theoretic or philosophic perspectives (10). Although a comprehensive review of the theoretic and philosophic literature on causal inference is beyond the scope of this article, we suggest that the most reasonable theoretic explanation of the conditions within which causal inference

is practiced (with due regard for the long and somewhat contentious discussion about the logic of causal inference) is that the evidence underdetermines (ie, provides neither proof nor disproof of) causation (11). In the absence of proof, judgments about causation and public health recommendations are at least partially value laden, with both scientific and extrascientific values playing roles. Indeed, causal inference appears by this account to be more subjective than objective. Two recent examples are induced abortion in relation to breast cancer and alcohol in relation to breast cancer (12). In both cases, 2 different reviewers examined the same evidence at the same time using similar approaches to causal inferences and yet came to exactly opposite conclusions about causation and public health recommendations. These stark differences can best be explained by the different rules of inference assigned to the criteria (ie, scientific values) and by extrascientific values such as wish bias, moral stances, or political positions regarding the acceptability or appropriateness of the exposure (12). In the case of alcohol and breast cancer, for example, the reviewers assigned different rules of inference for the criterion of consistency; one reviewer used an "all-or-none" rule and the other used a simple majority. This difference may explain why the reviewers' conclusions were so radically different.

The extent to which different causal theories or different types of causal hypotheses (eg, necessary cause, sufficient cause, and components of sufficient causes) might affect the choice of criteria or rules of inference assigned to them has been examined (13). Some criteria (eg, strength of association and consistency) appear to be dependent on the form of the causal hypothesis. Others (eg, biological plausibility, experimentation, and analogy) appear to be independent of such hypotheses. Much more work needs to be done to further elucidate the connections between causal theory and causal criteria. It is reasonable to suppose that a given causal theory may suggest criteria not currently being used.

CAUSAL CRITERIA IN NUTRITIONAL EPIDEMIOLOGY

We find that the literature on the practice, methods, and theory of causal inference provides those who would make nutrition recommendations with an almost unfettered set of choices about which criteria to use and how to prioritize them, which rules of inference to apply to them, and their interrelations. Although research on these topics is important and is underway (14), we do not have the luxury of waiting for progress. Rather, as public health professionals (15) and clinicians involved in preventive medicine, we have a moral obligation to consider the need for nutrition recommendations given the currently available evidence. It follows that we must provide our own view of the role of causal criteria in making nutrition recommendations, however muddy the current state of understanding.

In view of the emphasis in current practice on the criteria of consistency, strength of association, dose response, and biological plausibility, we examined these first and then added temporality. For the purpose of considering the need for nutrition recommendations, these represent the minimum set of criteria that should be considered, although not every one of these need be met to make a recommendation in a given case. For each criterion we proposed a rule or rules of inference to be used; we then discussed the extent to which evidence contradicting a given rule counts against making nutrition recommendations and the extent to which supporting evidence moves our judgment in the opposite direction.



Four of Hill's original criteria—experimentation, analogy, specificity, and coherence—were not discussed. We presumed that the typical situation in nutritional epidemiology involves making recommendations in the absence of evidence from randomized prevention trials. We further presumed that the criteria of analogy and specificity are secondary to the criterion of plausibility. On the other hand, we considered coherence to be a "meta" criterion, inasmuch as it applies not only to the evidence assembled for a given factor–disease association but also to the criteria themselves. In other words, we strove for an integrated—ie, coherent—understanding of the use of causal criteria in nutritional epidemiology and stressed that the goal of the practice of causal inference is to examine the evidence for the purpose of making a judgment regarding the need for public health recommendations, whether or not causation is concluded.

Although we intended to examine a specified list of criteria with accompanying rules of evidence, many other considerations are important in making nutrition recommendations, as described in the Introduction; these include study design, statistical testing, and confounding to name a few. These we left for others to explore, although some issues, such as the quality of dietary measurements, are so important that we did discuss them in the context of the causal criteria.

An additional methodologic point regarding the nature of nutrition recommendations should be considered. We believe there is a certain particularity to making nutrition recommendations. Each specific nutrient or dietary factor–disease association has some unique characteristics. Our judgment about the need for recommendations stems at least in part from this particularity, ie, on the specific circumstances arising from the evidence and from the hypothesis in question.

Consistency of association

Consistency across studies is the most commonly used criterion in epidemiology, reflecting the basic scientific notion of replicability. Nutritional epidemiology does not differ from other epidemiologic subdisciplines in this regard. In general, consistency across populations, study designs, and statistical methods bears much weight in making nutrition recommendations and must be considered in light of the potential effect of publication bias. What counts as consistent (ie, the rule of evidence assigned to this criterion) often depends on the individual reviewer, although it would seem difficult to make a case for consistency without at least a majority of studies (ie, studies not excluded on methodologic grounds) supporting the hypothesis in question. In other words, unless some methodologic feature of a study supporting the minority view is both unique and overwhelmingly relevant—such as a randomized prevention trial—the majority view should rule, with due consideration for other aspects of this criterion. Consistency across study designs is particularly compelling, but only if the studies are judged to be of high quality and not subject to obvious biases. An alternative rule of evidence, which is only sometimes available, still controversial, and yet promising at least for specific study design types, is to evaluate the criterion of consistency in terms of the results of meta-analysis (16). Whether or not a meta-analysis is undertaken, evaluation of consistency may require the exclusion of studies because of bias or other methodologic problems. These exclusions should be clearly stated (9).

Evaluation of consistency in nutritional epidemiology is a challenge. Nutritional studies often have null findings for a vari-

ety of reasons including measurement error, lack of variation of intake in the population, or a distribution of intakes unrelated to the disease process. Careful evaluation of inconsistencies between positive and null studies can be informative. This effort is hampered, however, by the noncomparability of dietary instruments, especially with regard to the level of nutrient intakes as measured by food-frequency questionnaires. An assessment of the level of intake needed for an effect to be observed across studies is difficult given that food-frequency questionnaires are adequate for comparisons within a study but are not accurate in terms of absolute nutrient values. Thus, cutpoints in one study usually cannot be compared with cutpoints in another study. In some circumstances, a lack of consistency across vastly different study populations may provide insights rather than suggesting a lack of effect. For example, cutpoints for fruit intake in a null study of invasive cervical cancer in the United States were 7.3 and ≥ 19 servings/wk for the lowest and highest quartiles, respectively (17), whereas they were 4.3 and ≥ 30 times/wk, respectively, in a Latin American study that showed protective effects (18). This example suggests that, although the findings were inconsistent between studies, very high fruit intakes may be necessary to observe an effect on cervical cancer. Extrapolation of absolute values for cutpoint intakes is not usually possible unless a study has validation data that can be used to estimate what the true cutpoints might be. Put another way, the presence of consistency across studies with different cutpoints is hampered unless information on true (ie, absolute) values of intake is available.

Inconsistency across studies, populations, and study designs requires careful evaluation and may suggest erroneous findings; it may also suggest new directions for research. Clarification of a discordance of results across food groups, associated nutrients, or associated blood variables is also important. Sex-based inconsistency is often overlooked by reviewers of the literature. Unless the disease in question is related to hormones, effects related to dietary intake should be consistent across sexes. It was noted, however, that differences between men and women may be related to methodologic as well as dietary differences (19, 20).

Strength of association

Nutritional epidemiology is fraught with evidence of weak associations. It is far more common to find risk estimates of 0.8–1.2 than to find a 2-fold (much less a 4-fold) estimate of risk. Indeed, strong risk estimates (≥ 4) arising in a single study are so uncommon that they may be viewed as the result of bias if they are not reproduced in other, similarly designed studies. Generally, weak associations are also viewed with caution because they too can often be explained by bias. Indeed, the criterion of strength of association is more likely to be problematic in nutrition studies as a result of the frequent occurrence of measurement error, although this fact could be used to claim that small risks have likely been underestimated and are therefore stronger than observed. Weak associations in dietary studies may have large public health effects if the dietary factor is common and the disease presents an important public health concern.

There is a considerable range of opinion—call it an example of methodologic subjectivity—in setting a threshold for what counts as a weak association, and therefore what counts as a strong association. It may be reasonable for this threshold to vary according to the prior hypotheses and whether the exposure consists of a food group, a diet-derived nutrient, or a blood marker. For our purposes, a statistically significant risk estimate that is a



>20% increase or decrease in risk is considered a positive finding. A change of 40–50% may be considered strong, especially for protective effects. Large risk estimates are not necessary, although they are desirable for strong inferences. The use of serum or other markers may lead to more stable and potentially larger risk estimates if the measured constituent is truly related to the disease (eg, LDL cholesterol and heart disease) or is a good marker of the food or dietary pattern (eg, β -carotene for high vegetable intake) related to the disease.

Dose response

Our view of the criterion of dose response is that the presence of a statistically significant linear or otherwise regularly increasing trend clearly reinforces the evidence in favor of causality. However, such an ideal situation may not be achieved easily when dietary data are being evaluated. In nutritional epidemiology, it is often the case that only the extreme categories of exposure are related to risk. Under these circumstances, a test for trend may be significant. Although such a finding may represent a statistical artifact, it also does not preclude a trend nor does it preclude the possibility of a threshold effect. Nevertheless, studies that reveal no obvious linear trend (eg, relative risk estimates from lowest to highest quintile of 1.0, 1.2, 0.8, 1.0, and 1.4) but that also show a statistically significant trend test should be viewed with caution.

The effect of misclassification errors in quantile designation from dietary data may be underappreciated by readers and reviewers and may have a profound effect on the assessment of dose-response trends. For example, validation studies typically report correlation coefficients of 0.30–0.60 for nutrient intakes assessed from a food-frequency questionnaire compared with a validation instrument such as a diet diary (21–23). In one such study, subjects were grouped into quintiles on the basis of the food-frequency questionnaire and again by nutrient estimates from 2 7-d diaries (23). Only 30–45% of subjects were correctly classified into the lowest quintile and similar results were observed for the highest quintile. Percentages were better for the lowest 2 or highest 2 quintiles (50–80% were correctly classified). Nevertheless, because less than half of the sample was correctly designated in the extreme categories and certainly correct designation into the intermediate quintiles would be lower, it seems unreasonable to expect a dose-response relation to emerge.

Measurement error has another facet worthy of discussion. Increasing the sample size diminishes some of the problems associated with misclassification. As a result, more stable estimates may emerge along with trends. However, if nutritional epidemiologists (along with the rest of the discipline) become increasingly interested in genetic susceptibility, then very large studies indeed will be needed to evaluate subgroups that may be more or less affected by dietary factors (24). Evidence of dose-response relations in small subgroups may be an unrealistic expectation. Similarly, dose-response relations may be missed if the effect is restricted to one genetically susceptible group.

Threshold effects are often encountered in nutritional epidemiology; their interpretation is problematic. Consider an example in which the risk is elevated 2-fold in all quartiles above the reference (25). One interpretation is that there were too few people in the low category to evaluate the trend. In other words, if the range of intake in the population were shifted down so that there were more subjects with lower intakes, then a trend might emerge. Nonetheless, this example remains consistent with a dietary threshold.

The assessment of dose response may be facilitated when the exposure is represented by a blood marker for a nutrient. Serologic measures may be more relevant to (ie, proximal to) the disease process because they have already incorporated dietary intake, absorption, distribution, and to some extent, utilization. Nevertheless, serologic markers have their own share of problems. They represent only one metabolic site for a nutrient in the organism and may not reflect the functional role of the nutrient with regard to the disease process. The timing of the presence of the marker and the occurrence of disease (and nutrient exposure for that matter) is another issue, perhaps better handled under the criterion of temporality, discussed below.

Biological plausibility

Biological plausibility is one of the most challenging and promising of all the causal criteria. In nutritional sciences, the biological evidence is collected from animal models, *in vitro* cell systems, and human metabolic and clinical studies. The relevance of each type of evidence is controversial. Decisions on usefulness tend to be rather subjective, as is frequently the case in causal inference. The incorporation of genetic and other biological markers as exposures (and sometimes as endpoints) in epidemiologic studies suggests that biological plausibility will become more important to causal inference in the future (26).

Nevertheless, the relevance of biological plausibility in any given situation depends on the disease outcome of interest. For example, in a defined clinical syndrome such as type 1 diabetes, a biological mechanism for an association between early infant feeding practices and onset of the disease years later (27, 28) provides a better explanation of the evidence than does confounding or another methodologic artifact. For chronic diseases with multifaceted causal pathways and years of latency before clinical manifestations, biological plausibility is also desirable. The essential question, however, is the extent to which it is reasonable to expect that this criterion be met in nutritional epidemiology; clearly, part of the problem is the difficulty of assigning a rule of inference to this criterion.

In situations in which an *a priori* hypothesis of nutrient-disease association is linked with a known (ie, established) biological mechanism, evidence of the association in an epidemiologic study can be called biologically plausible. However, evidence of associations that were not anticipated often emerges from epidemiologic studies. Given the number of nutrients and food groups evaluated in most nutritional studies, some new findings may be due to chance. Furthermore, it may not be difficult to contemplate biologically relevant functions of nutrients, but *post hoc* justifications should not hold the same evidentiary status as *a priori* hypotheses.

Consider the general case of a nutrient and a cancer, wherein multiple biological functions of the nutrient could be relevant to the development of the cancer. For example, an association between β -carotene and cervical cancer could be mediated by antioxidant function (29), immune enhancement (30), and other mechanisms. All of these potential mechanisms could operate together to inhibit carcinogenesis. However, although the natural history of cervical cancer—from preneoplasia to neoplasia—is well studied, it remains unclear at which point or points nutrients play a role. Further, with the complex interrelations of nutrients in the diet coupled with the potential for metabolic nutrient-nutrient interactions, a simplistic approach can be misleading.

Collinearity of nutrients in the same foods and in associated foods also provides an opportunity for multiple mechanisms. For example, some foods high in folate (orange juice, greens, and



legumes) are also high in vitamin C (orange juice and greens), β -cryptoxanthin (orange juice and greens), and fiber (legumes). These foods may also be consumed in the context of diets high in fiber and other beneficial micronutrients. When considering cancer, there is inherent appeal in discussing the DNA damage and repair, hypomethylation, and associated functions linked to low folate status (31, 32) in combination with the enhanced cytokine production, antiimmunosuppressive activities (33), and antioxidant functions of vitamin C (34, 35) and the enhanced immune function and free radical-scavenging activities of carotenoids (29, 30). Given that current knowledge of cancer etiology reveals a poorly defined disease process, these known functions of nutrients are likely to be describing disjointed components of a complex process. Use of broad categories of food groups, such as fruits and vegetables, makes assigning a specific biological mechanism nearly impossible.

Because there are strong precedents in epidemiology and nutrition for making public health recommendations without evidence of biologically plausible mechanisms, we suggest that it may be reasonable to continue with this practice. We are aware, however, that such a strategy introduces another problem when relatively new dietary constituents (such as phytoestrogens, lignans, and individual carotenoids) for which the biological function is not well known are studied. In such cases it may be wise to be wary of claims about new or single-constituent associations that have no known role in biological systems. We must also be prudent about recommending increases or decreases in the consumption of foods (in contrast to single nutrients) without understanding the biological mechanism. In these situations, a balance must be sought between the known benefits and harms (at a population level) of the foods and the uncertainty added because of the unknown effects of other nutrients found in these same foods. When biological evidence is unavailable and yet recommendations based on epidemiologic evidence seem prudent, it is wise to make clear that some recommendations are more tentative than others. Emerging scientific evidence may require a reassessment of recommendations, which may in some instances lead to a change in the recommendations.

Temporality

For the purpose of reaching causal conclusions, it is desirable—even necessary—for the exposure assessment to precede the onset of disease. For the purpose of making public health recommendations, it is desirable to determine the extent to which dietary factors may influence either onset or progression of disease. In cohort studies, diet is assessed before disease is diagnosed, but for diseases of long latency the disease process is usually already present and progressing at the time diet is assessed. Investigators often evaluate the effect of excluding those subjects whose disease is diagnosed close to the time of dietary assessment. This is done in part because dietary intake after the disease is established may not provide the appropriate exposure information; the disease can affect the diet or the reporting of the diet. In case-control studies, investigators often attempt to ask about a time period preceding the diagnosis; this is problematic because the disease may have been already present even if undiagnosed, as can occur in cohort studies. An additional problem with case-control studies is the potential for recall bias by case subjects (4).

THE INTERRELATIONS AND LIMITS OF CRITERIA

It is our view that when criteria are used in causal inference for the purpose of making nutrition recommendations, no single crite-


riion (of these 5: consistency, strength of association, dose response, biological plausibility, and temporality) is absolutely critical or absolutely irrelevant. For each criterion we argue that nutrition recommendations are not necessarily precluded even when epidemiologic evidence appears to conflict with our rule of inference for that criterion. For example, a lack of consistency may be explained by a lack of variation in intake, a weak association may actually reflect a larger relative risk obscured by misclassification bias, a dose-response relation may only appear for extreme categories of intake, and our assessment of biological plausibility depends on the knowledge gained to date on the specific disease or disease process. However, for any potential nutrient-disease association, we are not willing to forego considering any of these criteria.

If the evidence at hand clearly conflicted with all 5 of these criteria, it is highly unlikely that we would conclude that public health recommendations (to increase or decrease intake of the nutrient, depending on the situation) were warranted. Conversely, if the evidence strongly supported all 5 criteria, we would likely be in a very strong position to make a public health recommendation, as long as other (eg, ethical) considerations were also met. For example, the presumed benefit of the change in diet should not be overshadowed by the presumed harm. However, situations in which all 5 criteria are clearly and unequivocally met or not met are the exception rather than the rule in the current practice of causal inference. Put another way, it is difficult to lay out clear rules to be followed in the majority of cases considered by decision makers. Public health decision-making is typically a complex affair: these criteria are not independent of each other, nor are they easily separable from a host of other considerations such as measurement error, confounding, and other sources of uncertainty. Dose-response curves, after all, not only depend on our ability to measure but also comprise a series of relative risk estimates that ideally progress from weak to strong. Similarly, the criterion of consistency depends on the extent to which it makes biological sense to expect the same effect in different (eg, genetically or culturally diverse) populations in which potential confounders are not usually fully known.

We conclude, therefore, that the traditional causal criteria are important for making public health recommendations despite their insufficiencies. Nevertheless, we also conclude that, in the vast majority of situations, it is not possible to define a single set of rules for public health decision-making from the criteria alone. Until a new approach (to replace the current approach) is proposed and tested, we do not advocate dispensing with these familiar and still useful criteria. Indeed, we suggest that any new approach maintain some aspects of the traditional approach; evolution seems more likely than revolution.

Applying causal criteria and other considerations to the epidemiologic and biological evidence available for assessing the benefits and risks of a nutrient or food may (and indeed typically does) involve more than one disease or condition. In current practice and theory, causal criteria are applied to one association at a time. The user of this inferential method is then faced with making a balanced judgment about overall benefits and risks across different outcomes.

Thus, the best we can hope for in these circumstances is a coherent and ethically defensible judgment about the need for public health recommendations regarding, in a narrow sense, a specific nutrient-disease or food-disease association, or in a broader sense, a nutrient or food, taking into account all possible diseases and conditions affected. Depending on the situation, we

may reasonably conclude that no change in dietary recommendations should occur but that we may change our recommendation in the face of new evidence in the future. When we, as reviewers of the evidence, make recommendations others should expect that we have searched the literature carefully and described the causal criteria we used, their rules of inference, their relative importance, and how in this circumstance we put the complex mix of factors, issues, criteria, and evidence together to come to a decision about causation and about the need, or lack of need, for a specific nutrition recommendation. Although others may disagree, they are similarly charged with sorting out the complex interrelations for the same purposes: to advance scientific knowledge and to use that knowledge to improve the public health. 

REFERENCES

- Weed DL. Epistemology and ethics in epidemiology. In: Coughlin SS, Beauchamp TL, eds. *Ethics and epidemiology*. New York: Oxford University Press, 1996:76–94.
- Weed DL, Gorelic LS. The practice of causal inference in cancer epidemiology. *Cancer Epidemiol Biomarkers Prev* 1996;5:303–11.
- Weed DL. Causal and preventive inference. In: Greenwald P, Kramer B, Weed DL, eds. *Cancer prevention and control*. New York: Marcel Dekker, 1995:285–302.
- Willett W. *Nutritional epidemiology*. New York: Oxford University Press, 1990.
- Committee on Diet and Health, National Research Council. *Diet and health: implications for reducing chronic disease risk*. Washington, DC: National Academy Press, 1989.
- Surgeon General's Advisory Committee on Smoking and Health. *Smoking and health*. Rockville, MD: Public Health Service, 1964. [DHEW publication no. (PHS) 1103.]
- Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295–300.
- Breslow RA, Ross SA, Weed DL. The quality of reviews in epidemiology. *Am J Public Health* 1998;88:475–7.
- Weed DL. Methodologic guidelines for review papers. *J Natl Cancer Inst* 1997;89:6–7.
- Nijhuis HG, Van der Maesen LJJG. The philosophical foundations of public health: an invitation to debate. *J Epidemiol Community Health* 1994;48:1–3.
- McMullin E. Underdetermination. *J Med Philos* 1995;20:233–52.
- Weed DL. Underdetermination and incommensurability in contemporary epidemiology. *Kennedy Inst Ethics J* 1997;7:107–27.
- Weed DL. Causal criteria and Popperian refutation. In: Rothman KJ, ed. *Causal inference*. Chestnut Hill, MA: Epidemiology Resources Inc, 1988:15–32.
- Weed DL. On the use of causal criteria. *Int J Epidemiol* 1997;26:1137–41.
- Weed DL. Science, ethics guidelines, and advocacy in epidemiology. *Ann Epidemiol* 1994;4:166–71.
- Weed DL. Meta-analysis under the microscope. *J Natl Cancer Inst* 1997;89:904–5.
- Ziegler RG, Brinton LA, Hamman RF, et al. Diet and risk of invasive cervical cancer among white women in the United States. *Am J Epidemiol* 1990;132:432–45.
- Herrero R, Potischman N, Brinton LA, et al. A case-control study of nutrient status and invasive cervical cancer. I. Dietary indicators. *Am J Epidemiol* 1991;134:1335–46.
- Jacques PF, Sulsky SI, Sadowski JA, Phillips JC, Rush D, Willett WC. Comparison of micronutrient intake measured by a dietary questionnaire and biochemical indicators of micronutrient status. *Am J Clin Nutr* 1993;57:182–9.
- Mares-Perlman JA, Klein BE, Klein R, Ritter LL, Fisher MR, Freudenheim JL. A diet history questionnaire ranks nutrient intakes in middle-aged and older men and women similarly to multiple food records. *J Nutr* 1993;123:489–501.
- Block G, Thompson FE, Hartman AM, Larkin FA, Guire KE. Comparison of two dietary questionnaires validated against multiple dietary records collected during a 1-year period. *J Am Diet Assoc* 1992;92:686–93.
- Jain M, Howe GR, Rohan T. Dietary assessment in epidemiology: comparison of a food frequency and a diet history questionnaire with a 7-day food record. *Am J Epidemiol* 1996;143:953–60.
- Decarli A, Franceschi S, Ferraroni M, et al. Validation of a food frequency questionnaire to assess dietary intakes in cancer studies in Italy: results for specific nutrients. *Ann Epidemiol* 1996;6:110–8.
- Goldstein AM, Falk RT, Korczak JF, Lubin JH. Detecting gene-environment interactions using a case-control design. *Genet Epidemiol* 1997;14:1085–9.
- Potischman N, Swanson CA, Brinton LA, et al. Dietary associations in a case-control study of endometrial cancer. *Cancer Causes Control* 1993;4:239–50.
- Weed DL, Hursting SD. Biologic plausibility in causal inference: current method and practice. *Am J Epidemiol* 1998;147:415–25.
- Gerstein HC. Cow's milk exposure and type I diabetes mellitus. A critical overview of the clinical literature. *Diabetes Care* 1994;17:13–9.
- Kostraba JN. What can epidemiology tell us about the role of infant diet in the etiology of IDDM? *Diabetes Care* 1994;17:87–91.
- Di Mascio P, Murphy ME, Sies H. Antioxidant defense systems: the role of carotenoids, tocopherols, and thiols. *Am J Clin Nutr* 1991;53(suppl):194S–200S.
- Bendich A. Carotenoids and immunity. *Clin Appl Nutr* 1991;1:45–51.
- Krumdieck CL. Role of folate deficiency in carcinogenesis. In: Butterworth CE, Hutchinson ML, eds. *Nutrition factors in the induction and maintenance of malignancy*. New York: Academic Press, 1983:225–45.
- Eto I, Krumdieck CL. Role of vitamin B12 and folate deficiencies in carcinogenesis. *Adv Exp Med Biol* 1986;206:313–30.
- Blumberg JB. Vitamins. In: Forse RA, Bell SJ, Blackburn GL, Kab-bash LG, eds. *Diet, nutrition and immunity*. Boca Raton, FL: CRC Press, 1994:237–46.
- Henson DE, Block G, Levine M. Ascorbic acid: biologic functions and relation to cancer. *J Natl Cancer Inst* 1991;83:547–50.
- Ascorbic acid: biologic functions and relation to cancer. Proceedings of a conference held at the National Institutes of Health, Bethesda, MD, September 10–12, 1990. *Am J Clin Nutr* 1991;54(suppl):1113S–327S.