

SCIENCE vs. SOUND BITES

*Evaluating the Strengths and Limitations
of Epidemiological Research*



A Guide for Conscientious Journalists and Concerned Consumers by

Katherine A. Beals, PhD, RD, FACSM, CSSD

*Associate Professor, Nutrition & Integrative Physiology
The University of Utah*

COMMISSIONED IN THE INTEREST OF NUTRITION LITERACY BY



SCIENCE vs. SOUND BITES

Evaluating the Strengths and Limitations of Epidemiological Research

Epi-what? A basic understanding	Page 3
Evaluating Epi: Separating Sense from Nonsense	Page 6
Appendices.....	Page 9
1. Glossary (Select words in bold throughout the guide are defined here)	
2. Study Designs	
3. References	

PART ONE

EPI-WHAT? A BASIC UNDERSTANDING

CAN A GOOD SOUND BITE PRESENT GOOD SCIENCE?

You've probably run across headlines like this in the past few days alone:

"Pass the bacon! New study shows eating bacon is not bad for your heart after all!"

"Forget the fiber, research shows fiber-rich diets don't prevent colon cancer."

"Exercise can't save us: Why working out won't make you thin."

Whether you report the news or just read it, this guide will help you to better interpret and understand one of the most widely used study designs in nutrition research today: epidemiology.

Epidemiological investigations have limitations and you can't judge a nutrient, food or dietary pattern by media headline or a single, observational epidemiological study.

They are, indeed, great headlines. But can a good sound bite represent good science? Sometimes, perhaps, but certainly not always. Scientific findings exist in a particular context and if that context isn't communicated or understood properly, chances are the results won't be either.

That's why this guide was created. Whether you report the news or just read it, this guide will help you to better interpret and understand one of the most widely used study designs in nutrition research today (and the basis for the types of attention-grabbing headlines shown above): epidemiology.

Like a bad game of telephone, epidemiological research often gets lost in translation as it goes from the scientific journal that publishes the research to the news media that reports the research to the consumer who ultimately reads or hears about the research. Why? Largely because of misunderstandings or misinterpretations of epidemiological research study design and methodology.

Epidemiological investigations, provide important information regarding associations between dietary **exposures** and health and disease **outcomes**. Nonetheless, they have limitations. Those limitations dramatically affect the validity, generalizability and applicability of the results.

It is important for journalists, health professionals and consumers to read beyond the headlines, to look closely at the study methodology, evaluate the strengths and limitations and then put findings into perspective in order to avoid public health guidance that is not science-based.

When you understand how "epi" research should and shouldn't be conducted, and what information it can and can't provide, you can better understand what coverage of epi research should and shouldn't say – or imply. Just like the saying "you can't judge a book by its cover," where nutrition research is concerned, you can't (nor should you try to) judge a nutrient, food or dietary pattern by a media headline or a single, epidemiological study.

WHAT IS EPIDEMIOLOGY?

Epidemiology is derived from the Greek words *epi*, meaning "upon or among," *demos*, meaning "people or district," and *logos*, meaning "study, word or discourse." Literally translated, it means "*the study of what is upon the people.*"

In more common terms, epidemiology is the study and analysis of the patterns, causes, and effects of health and disease conditions in defined populations.

Epidemiology is the study and analysis of the patterns, causes, and effects of health and disease conditions in defined populations.

When it comes to nutrition, epidemiology can help us understand how particular foods and diets affect health and disease.

There are several different kinds of epidemiological study designs.

In nutritional epidemiology, observational studies are by far the most common type of research.

Observational studies, which make up most of nutritional epidemiological research, can demonstrate “correlation” ONLY

WHY IS EPIDEMIOLOGY IMPORTANT?

Epidemiology is the cornerstone of public health; it shapes health and nutrition policy decisions and evidence-based practice by identifying risk factors for disease and targets for preventive healthcare.

Epidemiological research is essential for identifying associations and assessing hypotheses for further clinical trials. With respect to nutrition, results from epidemiological research can help further our understanding of how particular foods, food patterns and diets affect health and disease.

TYPES OF RESEARCH IN GENERAL

Researchers refer to “study design” when talking about how a particular piece of research is conducted, including how participants are recruited, how data is collected, analyzed, interpreted and so on. Every step along the way offers options, so you can understand how variations in study design multiply quickly.

Generally speaking, there are two broad types of research: **observational** and **experimental**.

Most (but not all) epidemiological research is observational. As the name implies, observational studies involve the researcher simply observing the relationship(s) between variables. They do not manipulate variables (as in experimental studies).

Within observational there are three main study designs: **cross-sectional studies, case control studies, and cohort studies**. Within experimental there are two: **randomized control trials** and **quasi-experiments**.

Check out Appendix 2 of this guide for details on each of the three main kinds of observational study designs along with their relative pros and cons.

EPIDEMIOLOGY RESEARCH IN PARTICULAR

Epidemiologists employ a range of study designs — from observational to experimental — in order to examine relationships between various exposures (e.g., a particular nutrient, food or dietary pattern) and outcomes (e.g., health or disease). Randomized control trials (RCTs) represent the strongest research design for studying the effects of an exposure on an outcome and are *the only type of research design that can demonstrate cause-and-effect* relationships between diet and disease and/or health outcomes.

So, why then do researchers conduct observational studies and why are they used to help shape dietary recommendations and nutritional guidance? One reason is that the extensive timeline for the development of many chronic diseases (e.g., cancers, heart disease, diabetes, etc.) and health outcomes renders long-term RCTs impractical or prohibitively expensive.

In addition, there are also ethical considerations inherent to studying diet and disease relationships. For example, it would be unethical to give humans a disease in order to test the effect of a nutritional intervention for its treatment.

Similarly, it would not be ethical to deprive a control group of dietary intervention that could potentially cure or help manage a disease or favorably affect a health outcome. Thus, for the purposes of examining diet and disease relationships, observational studies are by far the most common type of research.

A CRITICAL DISTINCTION

The distinction between randomized control trials and other kinds of study design is absolutely critical to understanding nutrition research. That’s because only randomized control trials — a kind of experimental study — can demonstrate **“cause and effect.”**

ONLY randomized control trials (RCTs) can demonstrate cause and effect.

Observational studies — which make up the lion’s share of all nutritional epidemiological research — can demonstrate “**correlation**” ONLY.

That bears repeating: *Only randomized control trials can demonstrate cause and effect. The typical nutritional epidemiological study can demonstrate correlation only.*

Thus, any (observational) epi study that results in a headline like “Study shows XYZ food causes cancer” has been fundamentally misrepresented as surely as if the headline read “Umbrellas cause rain.” It may be true that the use of umbrellas is strongly correlated with rain. However, to say that umbrellas *cause* rain would be, of course, preposterous.

Confusing correlation and cause and effect is an all-too-common problem when it comes to how epi studies are communicated to the public at large.

PART TWO

EVALUATING EPI: SEPARATING SENSE FROM NONSENSE

Seeking clarity and eliminating confusion should be the goal of all involved.

NOT ALL OBSERVATIONAL STUDIES ARE CREATED EQUAL

As discussed, timeline or ethical considerations have made observational studies more common in nutritional epidemiological research than experimental studies. But, as in all things, the quality of observational studies varies, as does the quality of reporting and critical thinking that media and consumers bring to them. Seeking clarity and eliminating confusion should be the goal of all involved. The rest of this section is devoted to giving you key things to consider when trying to make sense of a nutritional epi study.

ASK “WHO WAS STUDIED?”

The first thing to consider is the population. Whether you’re a member of the media analyzing the data or a consumer reading a news summary, key question to ask are:

- Who were the subjects (e.g., sex, age, education, occupation)?
- How were the subjects identified and recruited?
- Was a “power analysis” done to determine the number of subjects needed for a valid study? (The scope of this statistical process goes beyond the scope of this guide, but suffice it to say, if such an analysis wasn’t done, the study is weak to begin with.)
- And, most important, was the sample of subjects chosen for the study representative of the larger population? This is important because it affects whether the results can be “generalized” to other populations.

The first thing to consider is the population.

NOTE: Not all study subjects are recruited from scratch. Often the data comes from an existing database. For example, one of the most well-known “prospective cohort studies” (collection of data for research purposes) to date is the Nurse’s Health Study. While this study employs a large sample of subjects, one should question whether nurses are representative of the larger population.

ASK “WHAT WAS BEING STUDIED?”

Once you’re sure you understand who the population is, examine what nutritional exposure(s) the researchers were seeking to measure.

- Was it a single nutrient (and if so, was it a dietary source or a supplemental source or both) or several nutrients?
- Was it a particular food (or foods) or dietary pattern?
- Did the authors determine the nutritional exposure “*a priori*” (meaning they knew in advance what dietary factor or factors they wanted to look at) or did they let the exposure(s) emerge as the study unfolded?

Next, consider the exposure.

NOTE: Studies that target a particular nutrient, food or pattern *a priori* may introduce a degree of bias (“Let’s look for this” vs. “Let’s see what we find”) and, thus, those that allow the data and results to unfold are sometimes considered preferable.

ASK “WHAT HAPPENED?”

- What was the outcome of interest?
- Was it a risk factor (or factors) for a single disease and or multiple diseases?
- Would the disease(s) be considered chronic or acute?

Also, consider the outcome ... and consider “all-cause mortality” a major sign of a potentially meaningless study.

NOTE: Observational studies that examine “**all-cause mortality**” as their primary outcome *should be examined with extreme caution*. “All-cause mortality” literally means those participants who died during the study died from any cause. The death of a participant as a result of an auto accident will count the same in the data as the death of a participant as

a result of heart disease. But if you're studying the effects of nutrition on heart health, one of those deaths is far more relevant than the other. *When death for any reason is the outcome, the data is not meaningful.*

ASK "WHERE WAS THE STUDY?"

Similar to knowing who the subjects are, knowing where the study took place is important for evaluating the generalizability of results.

The more unique the location of the study, the less the results can be generalized to a broader population.

- Was the study conducted in a single location (school, city, state, country, etc.) or multiple locations?
- Was it conducted in the United States or another country (or other countries)?

NOTE: Dietary intakes and patterns as well as disease incidence differ from city to city, state to state and country to country. There are also heritability effects that could potentially confound the data and skew the results.

ASK "HOW LONG DID THE STUDY LAST AND WHEN WERE THINGS MEASURED?"

The "when" of an observational study includes the time course of the study (roughly speaking, did it occur in a relatively restricted moment in time or over the course of many years?). "When" also includes the frequency of measuring the nutritional exposure(s) as well as the outcomes.

How long a study lasts, how frequently exposures are made, and how often data is collected all determine the quality of a study.

- Was it a cross-sectional study (a weaker design) or a **longitudinal study**, which is a stronger design?
- If it was a longitudinal study, what was the length of the follow-up: 1 year, 2 years, 5 years? 10 years or more?

NOTE: Because many chronic diseases have long progressions, a study with sufficient follow-up is necessary for valid results.

- How often were exposures and outcomes measured?
- When did those measurements begin? When did they end?

NOTE: A single dietary measurement, completed at the beginning of a study to establish a baseline is considered significantly less valid, reliable and robust than multiple dietary measurements taken throughout the course of the follow-up period.

ASK "HOW WERE THINGS MEASURED?"

The main question here is simply:

- How was food intake/exposure measured?

If the answer is a "**food frequency questionnaire**" (FFQ), which is the most common method, then you need to temper any conclusions drawn from the study with the known limitations of FFQs.

If data is collected via a "food frequency questionnaire," be especially cautious about over-generalizing the results of the study.

Food frequency questionnaires are used by subjects to indicate the frequency of consumption (over a discrete period of time e.g., the past month, 6 months, a year) of a select number of foods (typically in the range of 80 to 160 foods).

NOTE: FFQs are vulnerable to flawed memories and misreporting, and to lack of precision in the data itself. Within the nutrition science community, it is widely acknowledged that FFQs are associated with substantial error and therefore represent imperfect measures of exposure. In some instances, this represents random error, which produces imprecise estimates for nutritional exposure(s) and disease outcomes and can lead to biased results (Maki et al. 2014).

In other cases, the error is systematic, such as the known tendency for participants to under-report food (particularly unhealthy foods) and thus energy intake, which can affect estimates for intakes of other nutrients (Maki et al. 2014).

Also problematic is that foods and/or food categories found within or constructed from FFQ data (e.g., vegetables, fruits, red meat, plant foods) may include many types of foods and prepared in many ways, which can significantly affect the reliability of the data and generalizability of the results (Maki et al 2014).

Finally, FFQs cannot quantify amounts (i.e., grams, milligrams, micrograms) of specific nutrients.

ASK “WERE OTHER FACTORS CONTROLLED FOR?”

Most diseases are multifactorial in terms of causation. In other words, there can be many causes; thus, isolating a single dietary exposure responsible for a single disease state can be difficult.

Finally, consider if or how other factors were accounted for. Isolating a single dietary exposure responsible for a single disease state can be difficult.

Moreover, the dietary causes can be related — a concept known as collinearity. For example, foods rich in fiber tend to be those in the fruit and vegetable categories, which are also rich in a host of other nutrients (B vitamins, vitamin C and magnesium), thus making it difficult to decipher the *independent* effects of fiber from fruits and vegetables and/or individual vitamins and minerals on disease outcomes (Maki et al. 2014).

Through statistical techniques, researchers can try to account for these variables (known as “confounding” variables), but it is impossible to control for all of them and, in doing so, researchers may end up over adjusting, which can also result in invalid results. Indeed, reports from observational studies examining a particular diet-disease relationship can vary markedly in the potential confounders included in statistical models, even within the same cohort and, thus produce very different results (Maki et al 2014).

BOTTOM LINE: CHECK THESE BOXES

When it comes to observational studies, the strongest evidence will be derived from ...

If you encounter a nutritional epidemiological study that doesn't check these boxes, approach it with a healthy dose of skepticism

- A longitudinal, prospective cohort study ...
- in which the subjects are representative of the general population ...
- the dietary exposure of interest is chosen without bias (or is allowed to emerge as the study unfolds) ...
- and is measured *multiple times* over the course of the follow-up period using a valid and reliable assessment tool (e.g., a validated FFQ) ...
- and the outcome is specific and could realistically be determined in the time period allotted.

APPENDIX 1: GLOSSARY

A

all-cause mortality

In the context of epidemiological research, “all-cause mortality” literally means those participants who died during the study died from *any cause*. Consider that a red flag; observational studies that examine all-cause mortality as their primary outcome should be examined with extreme caution. When death for any reason is the outcome, nothing meaningful can be gleaned from data.

See also **outcome(s)**

C

case-control studies

The “case-control” is a type of epidemiological observational study in which participants are divided into those who have a particular attribute (the “cases”) and those who don’t (the “controls.”) For example, in a nutrition case-control study examining the role of diet on blood pressure, the “case” group would include those with high blood pressure and the “control” group would include those without.

See also **cohort studies, observational research**

cause and effect

“Cause and effect” is a type of relationship in which an event occurs at least in part *because* of an event the occurred before it. The confusion of cause and effect with “correlation” is the basis of much misunderstanding of observational research, which *cannot* demonstrate cause and effect.

See also **correlation, epidemiological investigation**

cohort studies

In research, a “cohort” refers to a group of people with a particular characteristic in common (e.g., they all have a certain medical condition or were born in the same city in the same year, etc.) A cohort study uses a cohort as its population. “Retrospective” cohort studies use existing data (e.g. medical records) whereas “prospective” cohort studies require collecting new data. Cohort studies are a form of longitudinal research.

See also **case-control studies, longitudinal study, population(s)**

correlation

Think of “correlation” as a synonym for “association” In non-experimental research, researchers want to find out if there is an association, or correlation, between two variables. Correlation is not the same as causality.

See also **cause and effect, epidemiological investigation**

cross-sectional studies

“Cross-sectional studies” are a type of observational study that analyzes data from a population at a single, specific point in time. This is in contrast to longitudinal studies, which look at data over a period of time.

See also **cohort studies, longitudinal studies**

D

dietary patterns

“Dietary patterns” refers to the quantities, proportions, variety or combination of different foods, drinks, and nutrients in diets, and the frequency with which they are habitually consumed. If you are a vegetarian you have a dietary pattern that is different than that of a carnivore.

E

exposure(s)

In epidemiological research an “exposure” is the particular variable a researcher is concerned with. In nutrition research an exposure might be a nutrient, food or dietary pattern.

epidemiological investigation(s)

Often referred to casually as “epi” by researchers, epidemiology is the study and analysis of the patterns, causes, and effects of health and disease conditions in defined populations. “Epidemiological investigations” can take a variety of forms, the most common of which is “observational” research.

See also **observational research**

experimental research

“Experimental research” refers to studies where the researcher manipulates a variable or variables (known as the “independent” variable(s)) and examines its effect on another variable or variables (termed the “dependent” variable(s)). The most well-known, and typically most definitive form of experimental research is the “randomized control trial.

See also **observational research, randomized control trials**

F

food frequency questionnaire

“Food frequency questionnaires” are a common method of recording nutrition exposures in epidemiologic research. Often abbreviated as “FFQ,” the questionnaires rely on the subjects themselves to recall and record the frequency of consumption of a select number of foods. While FFQs are frequently the best available method for collecting data, they are vulnerable to flawed memories and misreporting, and to lack of precision in the data itself. Those weaknesses should be taken into account when examining the outcomes of any study that utilizes FFQs.

L

longitudinal study

A “longitudinal study” is a research design that involves repeated observations of the same exposures and population over a period of time. It is often a type of observational study, although they can also be structured as longitudinal randomized experiments.

See also **cohort studies, cross-sectional studies**

O

observational research

“Observational studies” involve the researcher identifying a nutritional exposure and outcome of interest and then simply observing the relationship between the two. The subjects are not randomized by the researcher or separated into exposed and unexposed groups. Rather, the nutritional exposure is measured and then the frequency, incidence and/or pattern of the outcome(s) is/are observed and the association between exposure and outcome(s) is estimated using statistical techniques.

See also **experimental research, epidemiological investigation, randomized control trials**

outcome(s)

A research “outcome” is the nature of the result of a particular sequence of exposures. For example, there may be a health outcome or a disease outcome associated with a particular research study.

See also **exposure(s)**

P

population(s)

In research a “population” simply refers to the group of people of interest who have received an exposure and for whom an outcome is examined.

See also **exposure(s), outcome(s)**

R

randomized control trials

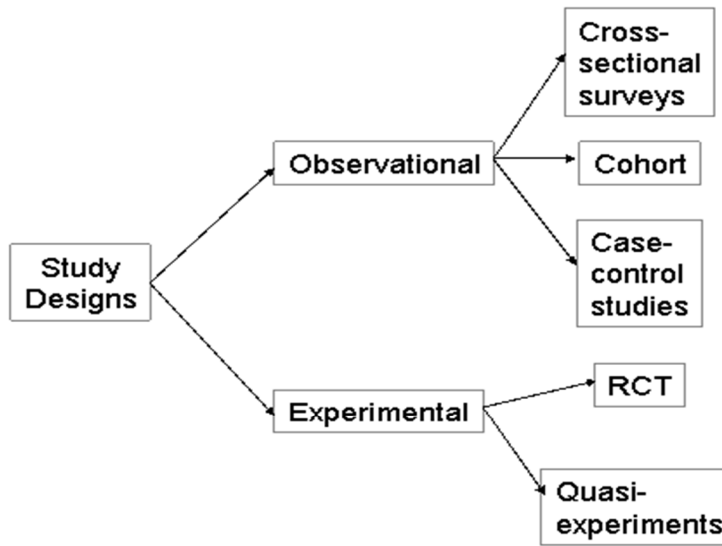
Often abbreviated “RCT,” “randomized control trials” are a type of experimental research that aim to reduce bias by randomly placing trial participants in either the group receiving the exposure under investigation or to a group receiving no exposure as the control. RCTs represent the strongest research design for studying the effects of an exposure on an outcome and are the only type of research design that can demonstrate cause-and-effect relationships between diet and disease and/or health outcomes.

See also **cause and effect, correlation, experimental research, exposure(s), outcome(s)**

APPENDIX 2: STUDY DESIGNS

FIGURE 1

At-a-glance breakdown of research study designs in general



Advantages	Disadvantages
Quick; can cover whole population, giving representative information whether or not people are seeking care	Based mainly on self-report (biases?); diagnostic information usually inaccurate; can't establish causal sequence
Prospective, so can establish causal sequence; can estimate incidence	Time-consuming; costly; attrition of cohort?
Relatively cheap way of focusing on causal factors	Requires recall of past events (inaccurate?); controls not equivalent to cases
Controls for all main forms of bias; good for both etiological and evaluative research	Ethical concerns in etiological applications; Often uses selected populations: issue of generalizability?
May be more practical than RCT: can use "natural experiments"	Allocation bias often significant (exp'tal and control groups not equivalent)

FIGURE 2

For a closer look at the three general categories of observational studies used in epidemiological research, this chart describes each along with its strengths, weaknesses, and bottom-line analysis.

	Cross-Sectional Studies	Case Control Studies	Cohort Studies
Description	<ul style="list-style-type: none"> Measure the exposure (e.g., nutrient, food, dietary pattern) and outcome (e.g., a disease) at a single point in time Asks, <i>“What is the association between currently having the disease and the nutritional exposure?”</i> Subjects are selected for inclusion in the study without regard for disease status or nutritional exposure Used in the well-known National Health and Nutrition Examination Surveys (NHANES) 	<ul style="list-style-type: none"> Compares a group of subjects with a disease (“cases”) to a similar group of subjects who don’t have the disease (“controls”) by looking back through time at potential exposures that both groups (cases and controls) may have encountered Asks, <i>“Did individuals with the disease consume foods, nutrients, diets, etc. that were different from those consumed by persons who do not have the disease?”</i> Considered “retrospective” as subjects are asked to recount their dietary habits before the onset of the disease with the goal of determining the usual nutritional exposure before the onset of the disease Food-frequency or dietary history questionnaires are typically used to assess usual intake in the past (at least one year but often as much as 10 years before onset of symptoms) 	<ul style="list-style-type: none"> Measure the effect of an exposure (e.g., nutrient, food, dietary pattern) on and outcome (e.g., disease) over time. Also referred to as prospective, follow-up or longitudinal studies Asks, <i>“Do persons with a nutritional exposure develop (or die) from the disease more (or less) frequently or sooner than those without the nutritional exposure?”</i> For example, are people who eat the recommended servings of fruits and vegetables less likely to develop coronary heart disease (CHD) than persons who do not? In some cohort studies the nutritional exposure is measured only once at baseline, (and typically asks the subjects to recall their dietary intake over the previous year), while in others, nutritional exposures are measured periodically over the course of the follow-up period
Strengths	<ul style="list-style-type: none"> Generate results fairly quickly Cost-effective Low subject burden Can employ large sample sizes, increasing the generalizability of the results 	<ul style="list-style-type: none"> Relatively inexpensive to conduct Provide results fairly quickly Require smaller sample sizes than cross-sectional or cohort studies 	<ul style="list-style-type: none"> The time sequence in which the nutritional exposure precedes disease development can be established and nutritional exposures can be measured without the influence of the disease (Sempos et al. 1999) Yield absolute estimates of risk, whereas case-control studies yield only relative estimates; thus, they can hint at, although not directly indicate causation
Weaknesses	<ul style="list-style-type: none"> Measures both nutritional exposure and disease status at the same time, thus cannot determine causation, nor can they indicate the direction of the relationship (impossible to determine whether the nutritional exposure is a cause of the disease, or if the disease process affected the nutritional exposure) For example, the disease may lead individuals to change their dietary intake patterns, or the disease process may produce changes in serum concentrations of nutrients 	<ul style="list-style-type: none"> Relies on self-reporting and subject recollections from the remote past, which can significantly impact the validity and reliability of the nutritional exposure data A key assumption is that the measure of nutritional exposure has not been influenced by the disease process itself or by a change in dietary habits 	<ul style="list-style-type: none"> Extremely costly and time-consuming as they require a large number of subjects to be followed for long periods of time In many (but not all) cases nutritional exposure(s) is/are measured only once at the beginning of the study with the assumption that the exposure(s) did not change over the entire course of the follow-up period which sometimes spans 5, 10 and even 15 years This is an extremely tenuous assumption as it is unlikely that a nutritional exposure measured in the past accurately reflects long-term exposure (Sempos et al. 1999)
Bottom Line	<ul style="list-style-type: none"> Cross-sectional studies are considered to be a relatively weak method for studying diet-disease relationships, and results from them should be interpreted extremely cautiously (Tarasuk 1997) 	<ul style="list-style-type: none"> Similar to cross-sectional studies, results from case control studies should be interpreted very cautiously 	<ul style="list-style-type: none"> Generally considered the “best” (methodologically speaking) of the observational studies because the time sequence in which the nutritional exposure precedes disease development can be established and nutritional exposures can be measured without the influence of the disease (Sempos

APPENDIX 3: References

Sempos CT, Kiang L, Ernst ND. Food and nutrient exposures: what to consider when evaluating epidemiologic evidence. *Am J Clin Nutr.* 1997; 1999;69(suppl):1330S–8S.

Tarasuk VS, Brooker AS. Interpreting epidemiologic studies of diet-disease relationships. *J Nutr.* 1997;127:1847-52.

Maki KC, Slavin JL, Rains TM, Kris-Etherton PM. Limitations of observational evidence: Implications for evidence-based dietary recommendations. *Adv Nutr.* 2014;5:7-15.