by Georgia Ede | View Article Comments

The Problem with Epidemiological Studies

Why does nutrition advice seem so confusing, complicated, and controversial? Why are the headlines constantly changing—one day eggs are bad for you, the next day they're perfectly fine? We are supposed to be the most intelligent creatures on the face of the Earth, and yet we are the only species that can't seem to figure out what we're supposed to eat.

The problem is that the lion share of mainstream nutrition advice comes to us not from biochemistry, physiology, or other scientific fields, but rather from the field of nutrition epidemiology, which is not scientific at all. In fact, of all the methods researchers can use to study human nutrition, epidemiology is arguably the least reliable.

Some of the most familiar and influential nutrition studies ever conducted happen to be nutritional epidemiology studies, including:

- the China Study, which argues that meat increases risk for chronic diseases and obesity;
- the vast majority of studies used to support the World Health Organization's 2015 report proclaiming that red meat probably causes colon cancer; and
- much of the research used in the highly publicized 2019 EAT-Lancet report, which seeks to convince the world that plant-based diets are healthiest for humans

- Cancer
- Diabetes
- Other Health Conditions

Read My Latest Posts

Study Finds Serious Mental Illnesses Improve on Ketogenic Diet

read more ...

—including recommendations found in our U.S. Dietary Guidelines—understanding what nutrition epidemiology is, what its limitations are, and how to spot epidemiology-based headlines will empower you to more confidently navigate the turbulent seas of nutrition advice.

Announcement: On Monday June 29th and Tuesday June 30th, 2020 the BMJ and Swiss Re will co-host a very special online event called ***Food for Thought: the Science and Politics of Nutrition***. I will be participating in a panel entitled "Food for Mind and Body: The Impact of Nutrition on Mental Well-Being." Register here to watch the live stream of this special online event and submit your questions; registration is free and open to all. This unique event was designed to foster constructive dialogue between nutrition thought leaders with differing viewpoints to discuss the most important and controversial questions facing nutrition research and public policy today.

The birth of epidemiology

The field of epidemiology (literally, the study of epidemics) was born in the mid-1800s, many crediting its origins to British physician John Snow. In the midst of a deadly cholera outbreak in the Soho district of London, Dr. Snow suspected that contaminated city water might be to blame. To explore this hypothesis, he interviewed townspeople about their water usage habits and meticulously mapped out where infections had occurred. He noticed a striking pattern: most infected households were clustered around a city water pump located on Broad Street. This strong association between proximity to the Broad Street pump and cholera infection risk convinced skeptical city officials to remove the handle from the Broad Street pump. When locals could no longer draw water from that pump, the epidemic came swiftly to an end. [1]

Epidemiology has since proved useful for understanding other diseases also caused by single, quantifiable toxins such as cigarette smoke and

Red Meat Increases Risk for Heart Attacks and Strokes

read more ...

Revolutionizing Anorexia Care Starts with Nutrition

read more ...

Nutrition Can Strengthen the Immune System to Fight COVID-19

read more ...

The Problem with Epidemiological Studies

read more ...

pattern recognition instead of on clinical experiments – after all, it would be unethical to intentionally expose healthy people to deadly bacteria, vicious viruses or tobacco.

More than a century after Dr. Snow's landmark cholera study, Professor Walter Willett, who would come to lead the prestigious Harvard School of Public Health, began using the observational methods of epidemiology to study the connection between human nutrition and chronic disease. Willett is considered by many to be the founding father of *nutrition epidemiology*, having authored a textbook on the subject as well as more than 1,700 research papers on nutrition and public health. Just as Dr. Snow questioned London residents about their water habits, for decades Professor Willett has been using surveys called food frequency questionnaires (FFQs) to inquire about people's eating habits in an effort to understand how their food choices affect their health. His work continues to be tremendously influential around the world.

So, if epidemiology is good enough to help us understand and fight deadly diseases caused by infections and toxins, is it also good enough to help us understand how our food choices affect our health?

This is your brain on berries

When evaluating any nutrition study or headline, I would argue you can save yourself precious time and energy if you begin with these two simple steps:

1. Find out if the study is an epidemiological study
2. If it is an epidemiological study, dismiss its findings

Some may find this approach closed-minded, unfair, or lazy, but I come by this practice honestly. Allow me to explain, using this 2012 study as a typical example.

Let's say you are a Harvard nutrition epidemiologist interested in whether the antioxidants in colorful berries could help protect aging women against memory loss.

For 14 long years, you study the eating habits of more than 16,000 middle-aged women. Then, over the course of six more years, you periodically test them for signs of memory problems.

Applying sophisticated statistics to the results, you calculate that women who reported eating two or more servings of strawberries and

Since you are a Harvard-based researcher who studied many thousands of subjects over 20 years, and your results have the potential to improve the lives of countless women around the world, your work is published in the prestigious journal *Annals of Neurology* and enjoys widespread, high-profile media attention, generating headlines like these:

"Eating blueberries and strawberries staves off memory decline, study suggests" —CBS News

"Brain food: berries can slow cognitive decline" —Time Magazine

"Berries keep your brain sharp" —Harvard Gazette

This seemingly impressive study is largely responsible for the common belief that dutifully topping your oatmeal with blueberries every morning will protect your brain. Unfortunately, because this study is an epidemiological study, it cannot tell us anything about how berries affect brain health.

Why should we not place our precious health in the hands of nutrition epidemiologists?

1. Nutrition epidemiology studies are not scientific experiments

The scientific method, which seeks to improve our understanding of the natural world, requires two phases:

- Phase 1: Generate a hypothesis about why something is happening in the world around you. In our example, the hypothesis is that berries may help prevent cognitive decline.
- Phase 2: Test your hypothesis in an experiment to see if you're on the right track. For example: feed some people berries, deprive other people of berries, then test everyone's thinking and memory down the road to see if either group fared better.

Nutrition epidemiology concerns itself entirely with phase one, never venturing into phase two.

Nutrition epidemiology research isn't about changing people's diets to see what happens. Instead, researchers give people questionnaires about their health and eating habits, then look for patterns in the

particular disease may be related.

These guesses easily find their way into the popular media long before they are tested in experiments, generating headlines that often make it sound as if those guesses are scientific facts you can use to make healthier decisions about your diet. Simply put, nutrition epidemiologists jump to conclusions—and journalists, nutrition policymakers, and the general public jump right along with them.

When nutritional epidemiology hypotheses are later put to the test in clinical trials, they have been shown to be wrong at least 80% of the time.

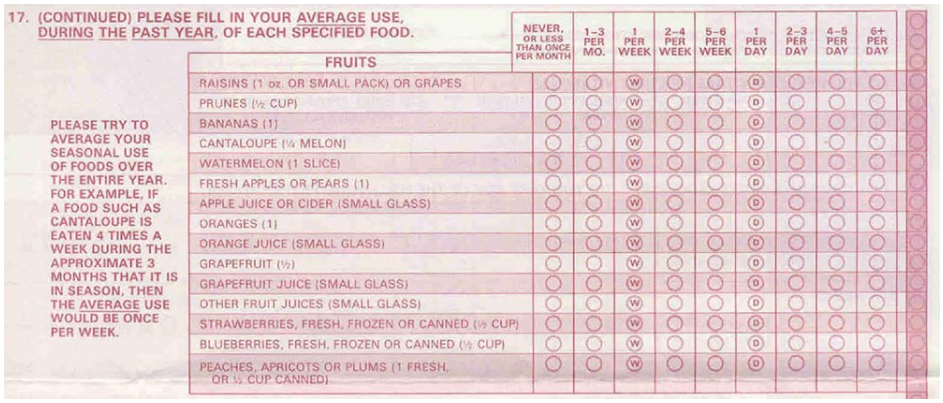
Imagine if the science underpinning air travel were this unreliable:

With a failure rate this high, you would be better off flipping a coin to guess which foods influence which diseases than conducting an epidemiological study.

This is why nutrition science seems so confusing: one day eggs are bad for you (epidemiology), the next day they are perfectly fine (clinical trials). This is not to say that all clinical trials are trustworthy—that is certainly not the case, and that topic deserves its own detailed post. Experimental studies must be carefully scrutinized to figure out whether they add to our understanding of human nutrition, but I would argue that the findings of virtually all nutrition epidemiology studies can be safely dismissed even without reading them.

2. Nutrition epidemiologists rely on memories, not measurements

what people actually ate over 14 years—in their defense, this would be virtually impossible. Instead, they administered something called a *semi-quantitative food frequency questionnaire* which asks people to remember what they've eaten over the course of the *entire previous year*. Below is the actual berry question used in this study—how accurately can you answer this question?



17. (CONTINUED) PLEASE FILL IN YOUR AVERAGE USE, DURING THE PAST YEAR, OF EACH SPECIFIED FOOD.

PLEASE TRY TO AVERAGE YOUR SEASONAL USE OF FOODS OVER THE ENTIRE YEAR. FOR EXAMPLE, IF A FOOD SUCH AS CANTALOUPE IS EATEN 4 TIMES A WEEK DURING THE APPROXIMATE 3 MONTHS THAT IT IS IN SEASON, THEN THE AVERAGE USE WOULD BE ONCE PER WEEK.

FRUITS	NEVER, OR LESS THAN ONCE PER MONTH	1–3 PER MO.	1 PER WEEK	2–4 PER WEEK	5–6 PER WEEK	1 PER DAY	2–3 PER DAY	4–5 PER DAY	6+ PER DAY
RAISINS (1 oz. OR SMALL PACK) OR GRAPES	○	○	(W)	○	○	(D)	○	○	○
PRUNES (½ CUP)	○	○	(W)	○	○	(D)	○	○	○
BANANAS (1)	○	○	(W)	○	○	(D)	○	○	○
CANTALOUPE (¼ MELON)	○	○	(W)	○	○	(D)	○	○	○
WATERMELON (1 SLICE)	○	○	(W)	○	○	(D)	○	○	○
FRESH APPLES OR PEARS (1)	○	○	(W)	○	○	(D)	○	○	○
APPLE JUICE OR CIDER (SMALL GLASS)	○	○	(W)	○	○	(D)	○	○	○
ORANGES (1)	○	○	(W)	○	○	(D)	○	○	○
ORANGE JUICE (SMALL GLASS)	○	○	(W)	○	○	(D)	○	○	○
GRAPEFRUIT (½)	○	○	(W)	○	○	(D)	○	○	○
GRAPEFRUIT JUICE (SMALL GLASS)	○	○	(W)	○	○	(D)	○	○	○
OTHER FRUIT JUICES (SMALL GLASS)	○	○	(W)	○	○	(D)	○	○	○
STRAWBERRIES, FRESH, FROZEN OR CANNED (½ CUP)	○	○	(W)	○	○	(D)	○	○	○
BLUEBERRIES, FRESH, FROZEN OR CANNED (½ CUP)	○	○	(W)	○	○	(D)	○	○	○
PEACHES, APRICOTS OR PLUMS (1 FRESH, OR ½ CUP CANNED)	○	○	(W)	○	○	(D)	○	○	○

How is anyone supposed to recall what was eaten up to 12 months ago? Most people can't remember what they ate three days ago. Note that you are not given the option of saying "I don't know", "I can't remember", or "You can't be serious"—you are forced to enter a specific amount, even if you're not sure. This particular question even requires that you do math to convert the number of servings of fruit you consumed seasonally into an annual average!

Your actual annual fruit intake is not weighed, measured, or recorded in any way. Notice also the vague descriptions of serving sizes—one slice of this, one small glass of that. These are *meaningless, unscientific quantities*. Imagine a laboratory chemist trying to follow an experimental protocol calling for "one small glass" of hydrochloric acid.

Human memory is also subject to conscious and unconscious distortion. Some people may believe they eat healthier than they actually do, or feel shame around how they eat, and these feelings may influence their answers. Therefore, instead of neutral, objective, quantifiable measurements, we have forced, subjective, inaccurate estimates. These wild guesses become the "data" that form the backbone of the entire study. No matter how sophisticated the statistics and analysis you apply to your research may be, your results will only be as good as the "data" you are analyzing. As the saying goes: garbage in, garbage out.

3. Modern diets are too complicated for

lend itself well to the study of human nutrition. The chances are small that any single dietary component is solely responsible for our salvation or our demise, but even if a single ingredient were at fault, estimating our consumption of that ingredient is extraordinarily challenging. When asked about their nicotine habit, most people can give a fairly accurate answer about a) whether or not they've smoked in the past year and b) how many cigarettes they typically smoke per day. By contrast, it is virtually impossible to design a food frequency questionnaire capable of capturing the complexity of most modern diets, which typically contain many hundreds of individual ingredients.

Another major problem with food questionnaires is that epidemiologists are unlikely to include questions about potential culprits they haven't thought of or aren't interested in. For decades, nutrition epidemiologists ignored refined carbohydrate intake in their studies, focusing instead on saturated fat, because saturated fat was hypothesized to be the dangerous ingredient in modern diets. While many researchers have recently begun taking refined carbohydrate into consideration when designing and analyzing their questionnaires—and this is certainly a step in the right direction—one can imagine countless other processed food ingredients that aren't examined for their effects on human health. How many FFQs inquire about soy lecithin, carrageenan, or hydrolyzed vegetable protein? Do you know how many grams of soy lecithin you consumed last year?

The FFQ used in this Harvard berry study considered only 130 food items. Notice in the fruit question above, only 15 fruits are represented. The number of kiwis, papayas, figs, mangos, cherries, dates, pineapples, honeydew melons, plantains, blackberries, raspberries, or cranberries these women ate was apparently considered unimportant. Asking about some fruits and not others makes no scientific sense—and smacks of blatant fruitism. #nofruitleftbehind

4. Food frequency questionnaires are too infrequent

If you think using a single questionnaire to represent an entire year's worth of food choices is absurd, then be sure you're sitting down for this next bit. Most epidemiologists don't even administer FFQs every year. In our berry study example, questionnaires were administered only five times over the course of 15 years; researchers then simply averaged the five sets of answers together to arrive at total berry intake. To make matters worse, researchers didn't monitor food intake at all during the entire six-year period that memory was being monitored. Even if the

and 2001—and that women's eating habits between 1995 and 2001 had no impact on memory—is hard to swallow.

Patient: "Doctor, I seem to be having more trouble remembering things lately."

Doctor: "Well, Barbara, research shows that memory problems can be due to berry deficiency. How many ½-cup servings of strawberries and blueberries did you typically eat per week ten years ago?"

Patient: "I have no idea, but this year my husband got a new job at Krispy Kreme and gets to bring me home all the defective donuts for free. Is that anything to worry about?"

Doctor: "Not according to this prestigious study by the Harvard School of Public Health."

5. Association is not necessarily causation

Even the most thoughtfully designed nutrition epidemiology study is only capable of documenting possible patterns of association between a particular food and a particular health problem, not of establishing cause-and-effect relationships between the two. If people who eat more berries really do suffer less cognitive decline than people who don't, this doesn't necessarily mean that berries have anything to do with it—the relationship between the two could be pure coincidence.

For example: if people who report eating more pretzels are also more likely to suffer from alcoholism, that doesn't necessarily mean pretzels cause alcoholism—it could simply mean that alcoholics spend more time in bars that serve free pretzels. Yet it would be completely acceptable in the field of nutrition epidemiology to publish that untested association using rather prescriptive language—phrased as medical advice—in a manner such as this: "Eating Pretzels Increases Risk of Alcoholism."

I like to use this imaginary example because the idea that pretzels might

contrast, if an association is found between two things we've been conditioned to believe are really connected—such as red meat and cancer—we are far more likely to take the headline at face value. This is called confirmation bias. In other words, believing is seeing.

[Associations between two things such as proximity to a water pump and cholera infections or smoking and lung cancer can sometimes indicate a likely cause-and-effect relationship, but for this to be the case, the associations must be very strong and ideally meet certain other standards as well. The Bradford-Hill criteria is a list of conditions that researchers can use to help understand whether the relationship between two things is more likely to be causal or simply coincidental. I'll publish a short companion post on this topic shortly.]*

6. Risky business: Absolute risk vs. relative risk

Very strong, consistent associations may be worth our attention, but we tend not to see these in nutrition epidemiology. In an effort to generate meaningful information from thousands of FFQ responses, epidemiologists often use sophisticated statistics to crunch the numbers into a single value called relative risk, which is supposed to help people understand how beneficial or dangerous a particular food is. Unfortunately, relative risk conceals true risk, because without knowing what that risk is relative to, we don't know how big the risk actually is.

Bioethics professor David Shaw of Maastricht University in the Netherlands published a clear, succinct explanation and critique of nutrition epidemiology's relative risk problem in the BMJ's *Journal of Medical Ethics*. In his excellent piece, he examines a 2019 epidemiology study that found eating two extra slices of bacon per day increases colon cancer risk by 20%. This sounds like a scary increase . . . but 20% compared to what?

If I'm a jellybean merchant and offer to give you 20% more jellybeans for a dollar, it makes a difference how many you're starting with. If you're buying five gallons of jellybeans, then 20% more is a whole gallon more, which is a LOT of jellybeans. If you're buying five jellybeans, then 20% more is just one measly additional outrageously expensive jellybean.

It turns out that the 20% increased risk of colon cancer in people eating two extra slices of bacon per day was in comparison to people who reported eating about one slice of bacon per day, who had a colon cancer risk of 0.4%. Since 20% of 0.4% is 0.08, a 20% higher

were estimated to have a colon cancer risk of 0.4% (40 cases per 10,000 people) and people who report eating three slices of bacon per day were estimated to have a colon cancer risk of 0.48% (48 cases per 10,000 people).

It is clearly more intellectually honest, transparent, and helpful to report absolute risk along with relative risk, so why don't nutrition epidemiologists typically do this? Perhaps because a conclusion that reads **"Two extra slices of bacon per day increase risk of colon cancer by 0.08%"** would make their findings appear far less important . . dare I say trivial.

[For more information, please read "Relative risk vs. absolute risk: one cannot be interpreted without the other," written by bioinformatics scientists at the University of Amsterdam.]

It is common to read critiques of nutrition epidemiology studies that focus on these typically tiny absolute risks to reassure people that eating bacon isn't as dangerous as studies would have us believe. However, the danger in this practice is that it takes those tiny risks seriously, thereby legitimizing the methodology. The truth is that the risk estimated by any nutrition epidemiology study, no matter how large or small, is calculated using profoundly inaccurate raw "data" from FFQs and therefore cannot possibly represent true risk. Calculating risk by applying fancy math to junk data gives the *illusion of precision*.

7. The field of nutrition epidemiology Is in denial

Nutrition science is rife with bias on all sides of the debate. Making biased observations is part and parcel of being human (and is actually important to the hypothesis-generating phase of the scientific process), but nutrition science is particularly vulnerable to entrenched, malignant bias. There are many psychological reasons for this—cultural and emotional attachments to food, strong feelings about whether or not we should eat animal foods, personal dietary experiences that influence our beliefs about which foods are healthiest, concern about how our food choices impact our environment, and many others.

Unfortunately, the big nutrition questions people want answered don't easily lend themselves to experimentation, and this is where the weak, pseudoscientific field of nutrition epidemiology finds a toehold. Here are just a few examples of questions nutrition epidemiologists feel uniquely qualified to answer:

- Which dietary patterns promote longevity?
- Do plant-based diets protect against chronic diseases?

These questions seek to understand risk over long periods of time—decades, or even lifetimes.

Since human clinical trials of extremely long duration simply are not feasible, nutrition epidemiologists argue that they are uniquely positioned to answer these kinds of questions because they can conduct their studies over decades and take multiple variables into consideration. However, rather than succumbing to the "epidemiology may be flawed, but it's better than nothing, so let's run with it" philosophy, why not call a spade a spade? Let's face the fact that we do not currently have meaningful scientific ways of addressing these big questions—is a car without wheels really better than no car at all? Nutrition epidemiologists understandably have great difficulty acknowledging the powerlessness and irrelevance of their methodologies.

8. Epidemiology is epidemi-illogical

Critical thinking skills are central to the scientific process, yet often seem to be lacking in nutrition epidemiology studies. It is one thing to earnestly generate a hypothesis that turns out to be unsupported by your findings. It is another thing to repeatedly ignore evidence against your hypothesis and continue to plow forward. When epidemiologists observed that French people tended to have lower death rates from cardiovascular disease despite reporting relatively high intakes of saturated fat, they were surprised because this pattern flew in the face of their hypothesis that saturated fat causes heart disease. Instead of opening their minds to the possibility that saturated fat might not cause heart disease, they instead stuck to their guns and famously called their findings a paradox.

When Dr. Snow created his iconic cholera map, it revealed that most infected residents lived near the Broad street water pump, but there were exceptions—stricken people who lived in neighborhoods serviced by a different water pump, and people served by the Broad Street pump who remained perfectly healthy. Rather than lazily writing these exceptions off as paradoxes, he pursued these apparent outliers, and discovered that healthy locals never drank water from Broad Street (some preferring beer instead), and those living elsewhere who fell ill with cholera had indeed been drinking Broad Street water (some preferring its taste). The harder he looked, the more support for his

Nutritional Mythology," and in my 2019 presentation and article about the EAT-Lancet report, I give examples of the extraordinary lengths that some epidemiologists go to in order to defend their nutrition hypotheses, even in the face of overwhelming evidence to the contrary.

Science requires intellectual curiosity, humility, open-mindedness, a willingness to learn from others, the ability to wrestle with facts that don't fit your beliefs, and the courage to acknowledge the limitations of your hypotheses and your methodologies when they no longer serve your goal. Shouldn't the goal be the pursuit of knowledge in the service of public health rather than the pursuit of rationalizations in the service of nutrition ideology?

Tips for recognizing nutrition epidemiology studies

Fortunately, identifying nutrition epidemiology studies can be quick and easy if you know what to look for. Nutrition epidemiology studies:

- Often have at least one of these phrases in the title:
    - associated with
    - linked to
    - increases (or decreases) risk
    - more likely (or less likely)
- Tend to be of very long duration, sometimes even decades long
- May include very large numbers of subjects—tens or even hundreds of thousands of people.
- Often generate simplistic, appealing or appalling headlines such as:

    "People who eat dark chocolate less likely to be depressed"

    "Millions of cardiovascular deaths attributed to not eating enough fruits and vegetables"

    "Your Fries May Be Deadly"

In short, if it sounds too good (or too bad) to be true, it probably is.

To be absolutely sure a study is an epidemiological study: If you have access to the original journal article, scroll immediately to the

The bottom line about nutrition epidemiology

Unfortunately, unlike the clear relationship between contaminated water and cholera infections, the relationship between modern diets in all their staggering complexity and chronic diseases like obesity, cancer, and heart disease do not lend themselves well to questionnaire-based methodologies. If Dr. Snow were alive today to map out the occurrence of contradictory nutrition headlines, convoluted dietary guidelines, nutrition confusion and declining public health, he could trace them all to prestigious wells of scientific information which have been contaminated by the antiscientific practice of nutrition epidemiology.

Curious to learn more about nutrition epidemiology?

*If you'd like to be notified when my short companion piece discussing the **Bradford-Hill criteria** is published, please sign up for my free newsletter by clicking on the gold button at the bottom of this post. The Bradford-Hill criteria help scientists understand whether the associations observed in an epidemiological study suggest a possible cause-and-effect relationship between two things like smoking and cancer or saturated fat and heart disease.

Suggested Readings and Presentations

Stanford Professor John Ioannidis' 2018 critique of epidemiology published in JAMA: "The Challenge of Reforming Nutritional Epidemiologic Research"

University of Colorado Professor James O. Hill's 2018 critique of epidemiology published in *Frontiers in Nutrition:* "The Failure to Measure Dietary Intake Engendered a Fictional Discourse on Diet-Disease Relations"

Science journalist Gary Taubes' classic 1995 piece in *Science:* "Epidemiology Faces its Limits"

Humorist and filmmaker Tom Naughton's thoroughly entertaining (and educational!) 2011 video presentation about epidemiology for the general public: "Science for Smart People"

"What evidence can we trust?" [video of expert panel discussion from Swiss Re/BMJ conference Food for Thought: The Science and Politics of Nutrition] with Dr. John Schoonbee, Dr. Zoe Harcombe, Prof. Walter Willett, Dr. Aseem Malhotra, Gary Taubes, and Prof. Rita Redberg

Re/BMJ conference Food for Thought: The Science and Politics of Nutrition

Additional articles and videos I've created on this topic:

"Brainwashed: The Mainstreaming of Nutritional Mythology" [video presentation]

"Latest Low-Carb Study All Politics, No Science"

"WHO Says Meat Causes Cancer?" blog post and video presentation

"EAT-Lancet's Plant-Based Planet: 10 Things You Need to Know"

"EAT-Lancet's Plant-Based Planet: Food in the (Mis)Anthropocene" [video presentation]

Go back