Detailed Study Notes

Table of Contents

Introduction to this Episode

When thinking about the effect of eating or not eating a certain food or nutrient, we can't consider this in isolation. Meaning, we need to evaluate the impact within the context of what such an inclusion/exclusion does to an individual's overall diet pattern.

Thinking about this concept, the phrase "compared to what?" has been colloquially used. And while this is an important idea, there has been some misapplication of this principle.

In nutrition science, this is related to the concept of food or nutrient "substitution". And this concept is crucial to understanding the issues that can arise in nutrition studies, particularly when it comes to single food analyses in nutritional epidemiology.

This concept of substitution is quite intuitive in controlled feeding studies. However, it is not as obvious when considering nutrition epidemiology studies. As noted by Ibsen & Dahm (2022):

> *"Whereas studying the effects of eating one food instead of another is typically explicit in interventional study designs, it is often implicit and sometimes hidden in analyses of observational studies."*

However, in nutrition epidemiology substitution is still happening, but it typically emerges as a consequence of adjustment models. In nutritional epidemiology, it is essential to adjust for confounders. E.g., one vital adjustment is often for total calorie intake. However, when our exposure is a specific food/nutrient, we must think about confounding by other foods.

So knowing what, and how, a study is adjusting for variables helps us interpret it better.

In this episode, Dr. Alan Flanagan and Danny Lennon discuss these crucial ideas of food substitution, adjustment models, and "compared to what?".

Connection to Previous Episodes

#378: Nutritional Epidemiology

- In this episode Alan and I discussed some critical aspects to understand about nutritional epidemiology in order to evaluate diet-disease relationships appropriately.
- Of particular focus was:
    - The unique “exposure of interest” in nutrition studies
    - Why it's crucial to understand temporal relationships
    - How to think about relative risk and absolute risk
- We also went on to explain how one can include these findings into an understanding of an overall body of evidence.
- You can find the episode page here.

#343: Understanding Causality in Nutrition Science

- In this episode, Alan and I went through a number of important points of clarification around what causality is, and how it can be demonstrated in nutrition research.
- This included discussing:
    - Inferring causality vs demonstrating causality
    - Hierarchy of evidence vs. standards of proof
    - Reductionism and erroneous application of the biomedical model to nutrition
    - The false causality dichotomy: RCTs vs. epidemiology
    - How nutritional epidemiology can infer causality
- You can find the episode page here.

#386: Deirdre Tobias, ScD – Study Design, Diet Collection Methods and Nutrition Epidemiology

- Dr. Tobias is a nutrition and obesity epidemiologist at Harvard Medical School and is currently the Academic Editor for the American Journal of Clinical Nutrition.
- In this discussion, Alan and I put a number of questions to Dr. Tobias, thus getting into:
    - Trial designs and their application
    - Importance of an appropriate contrast exposure
    - Diet collection methods: use and misuse
    - Benefit of repeated measures of diet
    - How to better understand epidemiology
- You can find the episode page here.

What is Substitution?

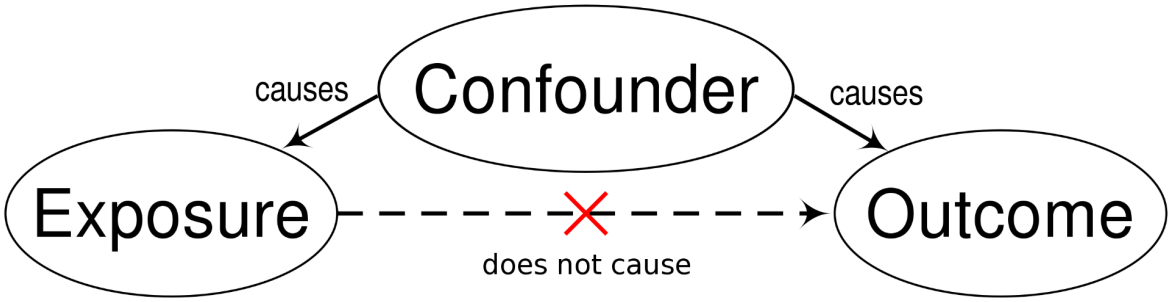
To understand what a substitution analysis is doing in epidemiology, it's perhaps easier to first think about what is happening in a controlled intervention trial…

- So, imagine we're setting up a diet trial to assess the impact of carbohydrate intake on fasting blood glucose.
- We would have two groups (intervention and comparator/control) consuming two meal plans:
- We could stipulate that group A, the intervention group, is going to consume a diet with an increased carbohydrate intake (or whatever it is we're assessing).
- And we would want total energy intake to be the same between the intervention and control group because that would allow any outcomes to be inferred to be independent of the effects of total energy intake.
    - I.e., if the intervention group consumed 400 calories less than the control group, and we were looking at fasting glucose as an outcome, how do we know the difference between groups was a result of the dietary change or simply because of the difference in total calories?

What is Adjustment?

In nutritional epidemiology, it is essential to ***adjust*** for confounders. This is done via statistical analysis or adjustment models.

In statistics, a ***confounder*** is a variable that influences both the dependent variable and independent variable, but is not a factor in the causal chain. Therefore, if we don't account for the confounder, it may lead to us incorrectly concluding a causal relationship between exposure and outcome, even if one doesn't actually exist.



Note: Whereas a mediator is a factor in the causal chain, a confounder is a spurious factor incorrectly implying causation.

Perhaps the most common adjustment (almost done all the time in epi) is for age. So you'll commonly see at this one "age-adjusted" analysis done. Adjusting for sex is likewise routinely done.

One example of a statistical method to account for such variables is a regression analysis.

In nutritional epidemiology specifically, one common and vital adjustment is for calorie intake.

The importance of this was discussed decades ago. Willett & Stampfer (1986):

We submit, however, that before attributing causality to a *specific* nutrient, the burden is upon the epidemiologist to demonstrate that the effect of this nutrient is independent of caloric intake. For example, perhaps excessive caloric intake increases

SUMMARY

Associations between intake of specific nutrients and disease cannot be considered primary effects of diet if they are simply the result of differences between cases and noncases in body size, physical activity, and metabolic efficiency. Epidemiologic studies of diet and disease should therefore be directed at the effect of nutrient intakes independent of total caloric intake in most instances. This is not accomplished with nutrient density measures of dietary intake but can be achieved by employing nutrient intakes adjusted for caloric intake by regression analysis.

Consider: total calories may be the primary factor associated with increased risk for disease, rather than the nutrient

perhaps excessive caloric intake increases the risk of colon cancer and dietary fat is associated with this disease because of its high caloric content. Before implicating fat per se as a primary cause, however, it would be essential to demonstrate that this effect is not shared by protein or carbohydrate when these are eaten in equicaloric amounts. Otherwise, reduction in the fat content of the diet would have no effect on disease occurrence unless the total caloric intake was also changed. Recognizing the

From: Willett & Stampfer, Am J Epidemiol. 1986 Jul;124(1):17-27.

Nutrient Substitution

- The original evolution of substitution models was focused on macronutrient swaps.
    - *[Remember: total energy intake is always 100%]*
- Macronutrient substitution models are providing a statistical technique to hold total energy intake constant and you are manipulating macronutrient composition by swapping a certain percentage of one nutrient with another.
- By way of analogy to an intervention trial:
    - Intervention trial where participants consume maintenance level of calories.
    - Participants randomized to either:
        - Diet A: 20% protein, 20% fat, and 60% carbohydrate
        - Diet B: 20% protein, 30% fat, and 50% carbohydrate
    - And so conceptually a macronutrient substitution model in epidemiology is seeking to achieve the same effect
- The substitution analysis could help answer questions like:
    - What would be the effect of replacing 5% of the energy in that total energy diet from saturated fat with 5% of energy from polyunsaturated fats?
    - What would be the effect of replacing 5% from saturated fats with 5% from complex carbohydrates?
    - Etc.

Example: Li et al., 2015

- Study: Li et al., J Am Coll Cardiol. 2015 Oct 6; 66(14): 1538–1548.
- Based on one of the big US cohorts: the Nurses Health and Health Professionals follow up study.
- Looking specifically at this substitution analysis of the replacement of saturated fat with other nutrients.
- In the main associations you could see the highest polyunsaturated fat intake was associated with a **20% lower risk** of coronary heart disease.
- *But* when you looked at the substitution analysis, **replacing 5% of energy from saturated fat** with 5% of energy from polyunsaturated fat was associated with a **25% lower risk** of heart disease.
- So the magnitude of the effect in the substitution models can often differ because now you're talking about the effects of these nutrients in relation to each other; the substitution effects in the diet.

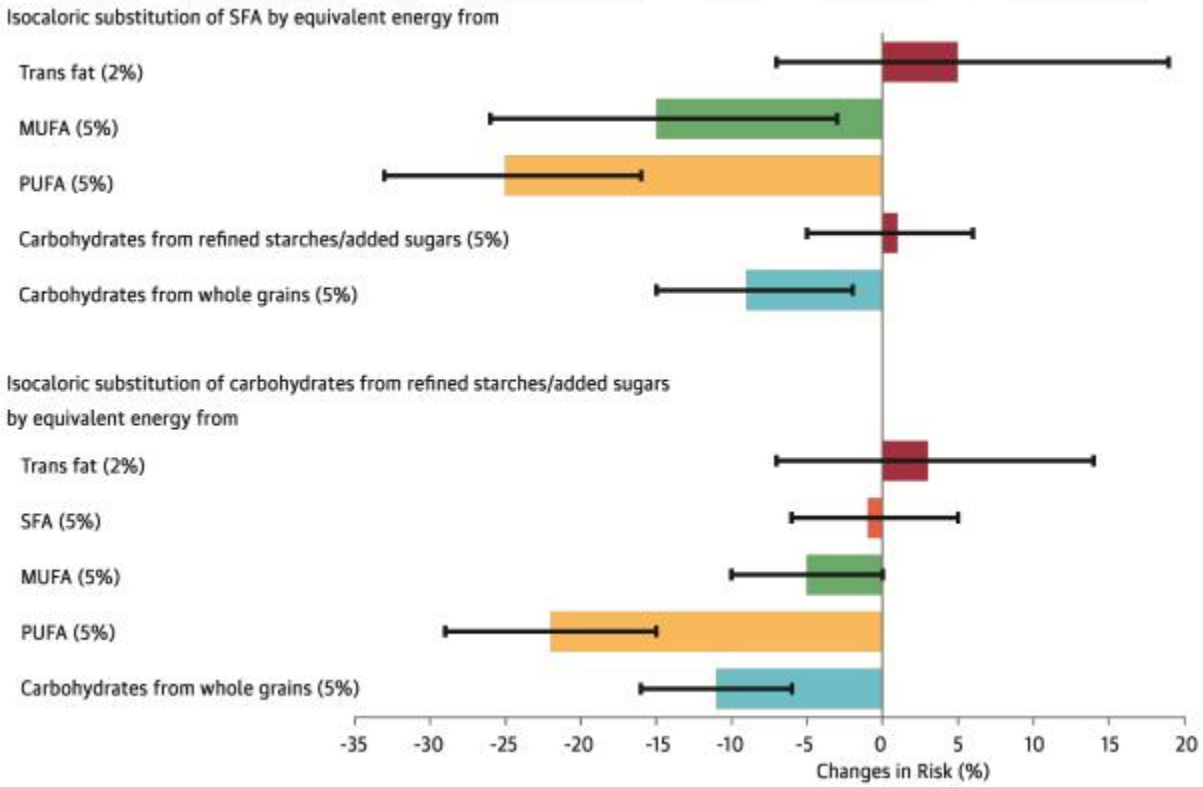


Image from: Li et al., J Am Coll Cardiol. 2015 Oct 6; 66(14): 1538–1548.

Food Substitution: Egg & CVD Example

In the episode I discussed an explanation and analogy that Dr. Deirdre Tobias gave before about how to think about substitution and adjustment in epidemiology, through the example of a hypothetical intervention trial.

This example used a paper by Zhao et al (2022) on egg consumption and cardiovascular disease risk.


Circulation

AHA Journals | Journal Information | All Issues | Subjects | Features | Resources & E

Home > Circulation > Vol. 145, No. 20 > Associations of Dietary Cholesterol, Serum Cholesterol, and Egg Consumption With Overall and Cause-Specific Mortality

FREE ACCESS
RESEARCH ARTICLE

PDF/EPUB

Tools Share

Associations of Dietary Cholesterol, Serum Cholesterol, and Egg Consumption With Overall and Cause-Specific Mortality: Systematic Review and Updated Meta-Analysis

Bin Zhao, Lu Gan, Barry I. Graubard, Satu Männistö, Demetrius Albanes and Jiaqi Huang

Originally published 1 Apr 2022 | https://doi.org/10.1161/CIRCULATIONAHA.121.057642 | Circulation. 2022;145:1506–1520

This article is commented on by the following:

Other version(s) of this article

This paper did an analysis of a prospective cohort study and used those results to update a meta-analysis. The authors concluded that greater egg consumption increased risk of CVD mortality and all-cause mortality:

Conclusions:

In this prospective cohort study and updated meta-analysis, greater dietary cholesterol and egg consumption were associated with increased risk of overall and CVD-related mortality. Our findings support restricted consumption of dietary cholesterol as a means to improve long-term health and longevity.

Some quick details on this paper:

- They did an original prospective analysis in a Finnish cohort
    - Males, smokers, in Finland
- They compared risk between men with higher vs. lower intakes of cholesterol and egg (at study baseline).
- Then used these results in an updated meta-analysis of 41 prospective cohort studies.
- The prospective analysis was from the ATBC Study cohort…

ATBC Study

- ATBC = Alpha-Tocopherol, Beta-Carotene Cancer Prevention
- A controlled primary prevention trial
- Originally conducted to evaluate whether cancer incidence could be decreased by supplementation with:
    - α-tocopherol (50 mg/d)
    - β-carotene (20 mg/d)
    - or both
- >29,000 Finnish male smokers between 1985 and 1988, age 50 to 69 years
- Followed for >30 years
    - 22,035 deaths
    - 9,110 deaths from CVD

They found: *"… for each **additional 50-g egg** consumed daily, hazard ratios were **1.06** and **1.09,** respectively, for overall and CVD-related mortality"*

Remember: A hazard ratio (HR) of 1.06 means a 6% increase in risk

Multivariate Adjustments

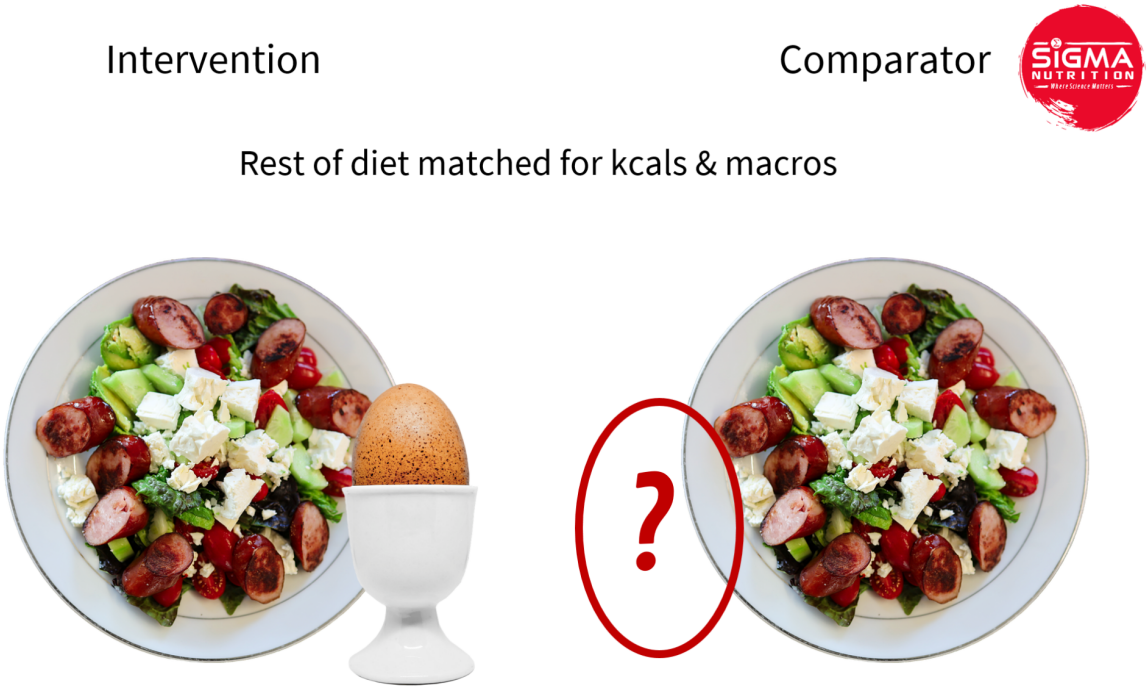
- In nutritional epidemiology, it is essential to adjust for confounders
- In this study many potential confounders were adjusted for.
- The study had three different adjustment models:
    - **Model 1** - Age-adjusted only
    - **Model 2**
        - Characteristics: age; body mass index; cigarettes smoked per day; years of smoking; education; physical activity;
        - Health markers: serum total and HDL cholesterol; intervention assignment; systolic and diastolic blood pressure; history of CVD; diabetes;
        - Energy intake
        - Nutrient intake: levels of serum α-tocopherol, β-carotene, and retinol; alcohol intake, percentage of energy from protein, carbohydrates, saturated fatty acids, monounsaturated fatty acids, and polyunsaturated fatty acids.
    - **Model 3**
        - All above in model 2 + further ***adjusted for dietary cholesterol***

As discussed earlier, adjusting for total calories is important. But… when our exposure is a specific food/nutrient, we must think about **confounding by other foods.**

- So when we make our comparisons isocaloric, then any additional dietary factors we control for are crucial for the interpretation of our exposure.

To look at demonstrate the adjustments were done in the Zhao prospective cohort study, Dr. Tobias gave a hypothetical example of what this would look like in an intervention trial:

- So in the intervention trial we would be comparing two diets, but one has additional egg in it.
- The background diets for each should be similar and we want to control for some variables…
    - First, to make sure any differences are not just down to a greater calorie intake, we need to control for calories back by making each diet the same.
        - N.B.: In epidemiology, the equivalent is achieved through a statistical analysis that adjusts for calories.



- So in the case of an intervention trial, when we add egg to the intervention group's diet, we have to add the same number of calories to the other group, but from non-egg foods.

- If we wish to control for other variables (e.g. protein), then this would need to be the same too.
- In nutritional epidemiology we have to do the same thing. Except using statistical analysis.
    - *"Whereas studying the effects of eating 1 food instead of another is typically explicit in interventional study designs, it is often implicit and sometimes hidden in analyses of observational studies."* – Ibsen & Dahm, 2022
    - As Dr. Tobias puts it: *"This notion of substitution is implicitly happening in a regression model whether it is intentional or not."*
- And in epidemiology, what we choose to adjust for has implications for the findings…

Looking at the results from the cohort analyzed by Zhao et al. (below), note a few things:

1. Results are for egg consumption, per 50 g/d
2. Outcomes are all-cause mortality and CVD mortality
3. Risk is shown via hazard ratios (HR) for three different adjustment models (described already).

Table 3. Associations Between Daily Egg Consumption and Overall and Cause-Specific Mortality in the ATBC Study

Cause of death	Egg consumption (per 50 g/d)*		
	ARD, % (95% CI)	HR (95% CI)	*P* value
All causes			
Model 1: Age-adjusted	1.18 (0.81, 1.56)	1.06 (1.04, 1.08)	<0.0001
Model 2: Multivariable†	1.19 (0.75, 1.65)	1.06 (1.04, 1.09)	<0.0001
Model 3: Multivariable‡	−1.83 (−3.50, −0.14)	0.91 (0.84, 0.99)	0.029
CVD			
Model 1: Age-adjusted	1.00 (0.55, 1.44)	1.07 (1.03, 1.10)	<0.0001
Model 2: Multivariable†	1.25 (0.72, 1.74)	1.09 (1.05, 1.12)	<0.0001
Model 3: Multivariable‡	−1.33 (−3.51, 0.62)	0.92 (0.81, 1.04)	0.18

Table from: Zhao et al., Circulation. 2022 May 17;145(20):1506-1520.

Now let's focus on Model's 2 and 3. Note that we see:

- **Model 2:**
  - All-cause mortality: HR = 1.06 (i.e., 6% *increase* in risk)
  - CVD mortality: HR = 1.09 (i.e., 9% *increase* in risk)
- **Model 3:**
  - All-cause mortality: HR = 0.91 (i.e., 9% *decrease* in risk)
  - CVD mortality: HR = 0.92 (i.e., 9% *decrease* in risk)

So now let's think about what each of these adjustment models has done, by viewing them from the perspective of the hypothetical intervention trial we mentioned.

Eggs, adjusted for calories, is like designing your trial's comparator meal to omit egg but allowing participants to eat whatever they want in its place:

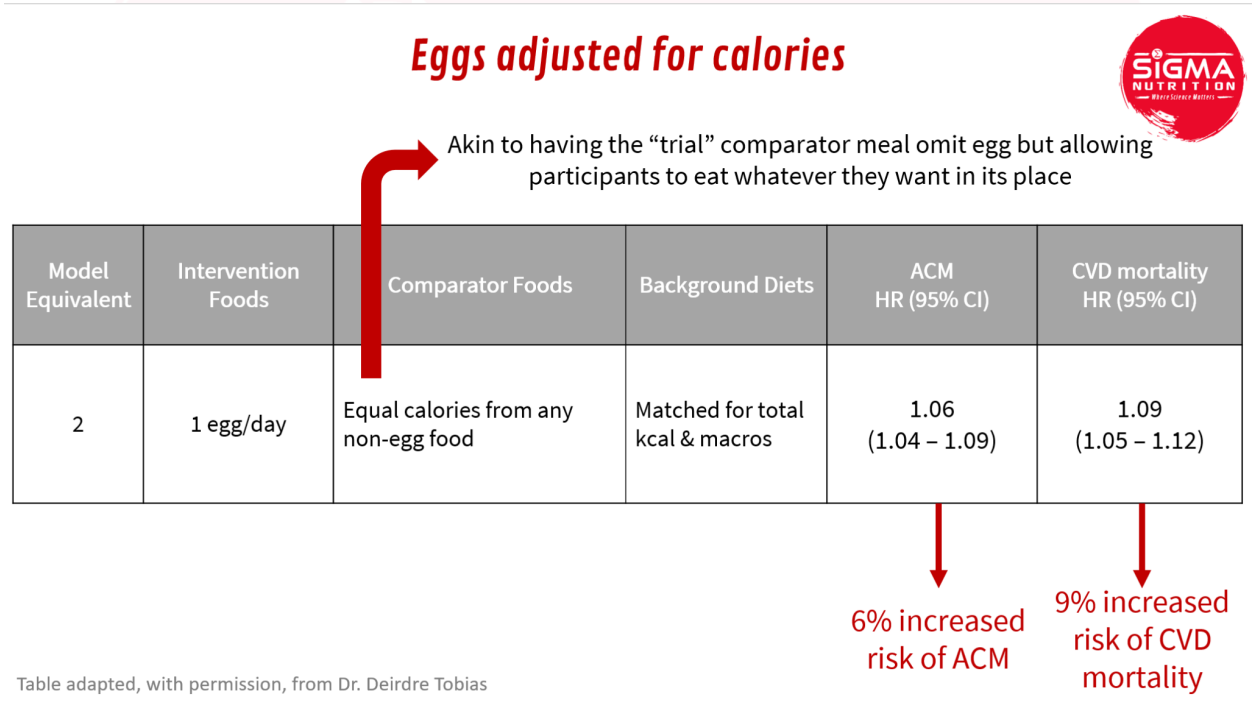


Image by: Sigma Nutrition. Adapted from: Tobias, D.

Point to note: at this point, when interpreting these results, we would be asking: if those not eating the eggs are at lower risk, then what are they eating instead? Because what replaces the eggs will matter.

Now let's consider Model 3. Recall that this adjusted for all the same variables as Model 2, except it additionally adjusted for dietary cholesterol. In doing so, the results then showed a risk reduction from greater egg intake. So what happened?

Again, let's translate that to a hypothetical intervention trial:

Adjusting for dietary cholesterol means…

Why?

Model Equivalent	Intervention Foods	Comparator Foods	Background Diets	ACM HR (95% CI)	CVD mortality HR (95% CI)
3	1 egg/day	Equal calories of non-egg, cholesterol-containing food	Matched for total kcal & macros	0.91 (0.81 – 1.04)	0.93 (0.81 – 1.08)

Men with higher intake of eggs had a lower risk of mortality… compared to men eat more meat and butter :-/

Men with higher intake of eggs actually had a **lower** risk of mortality…

Table adapted, with permission, from Dr. Deirdre Tobias

Adjusting for dietary cholesterol

- So if we are adjusting for dietary cholesterol, then what replaces the eggs would be something that keeps cholesterol content the same.
- So in this Finnish cohort, the most likely foods would be butter and fatty cuts of meat.
- And so what the risk reduction from increased egg intake in this model is really saying, is that those who ate more eggs are at lower risk than those who ate more meat and butter.
- Therefore, what we adjust for matters for the interpretation.

Compared To What?

- The crucial concept of being aware what substitution effects are at play has been communicated colloquially with the phrase “compared to what?”
- This is a really useful reminder for us to always bear in mind what exactly is being compared to what in an analysis.
- However, as we discussed in the episode, sometimes people are misusing or misapplying this idea.
- Sometimes people interpret this idea as if it's a straight substitution; that if we model food A versus food B, food A replacing food B, and there is a lower risk that results, that means that food A is better than food B, and we should be looking to remove food B from the diet.
- But that's not necessarily the inference that should be derived from these models because it depends on the nature of the foods being modeled and exchanged.
- So if, for example, we're talking about protein from red meat for example.
    - But we're talking about a population that consumes red meat twice a week.
    - And there's an analysis and it models the substitution of a serving of red meat with a serving of lentils.
    - First, the actual serving size for those foods might actually be different and so they should in an ideal scenario for a food substitution, those foods should be consumed within reasonably comparable intake amounts.
    - Otherwise you get a substitution that can reflect the actual consumption pattern of a food that's highly consumed and frequently consumed, could actually end up having a stronger effect.
    - In this analysis then, as we would probably predict this model of replacing some red meat with some lentils shows that there's a lower risk of heart disease, and everyone says *"that's lentils better than red meat. We should replace red meat with lentils"*.
    - But that's not necessarily the case, because a valid interpretation could be that in that person consuming red meat once/twice a week, they should just probably add a few servings of lentils to their weekly level of intake.
    - So there are a number of interpretations that we can make, and the actual serving of the foods is going to be important to how we would interpret it.
- In some cases it might simply be that someone is adding foods to their diet not necessarily entirely removing a food from the diet. And it shouldn't necessarily be interpreted to mean that outright one food is better than another.
- Example from the Daniel Ibsen paper:

    - They modeled the effects of replacing foods and they looked at poultry, white meat, chicken and red meat.
    - And they modeled the comparison of substituting a serving a week of red meat or poultry, but **not specifying** the replacement foods.
        - In that context, the analysis for poultry didn't show any significant effect on risk of type two diabetes.
        - The hazard ratio was 1.01 and the confidence intervals were 0.98 to 1.04.
        - So you could be left thinking there's no beneficial effect of poultry.
    - But then when they specified the substitution and the poultry is replacing a serving of processed red meat, then it was associated with a lower risk of diabetes: a 7% lower risk of diabetes.

- Substitution is not independent of the effects of the wider dietary pattern.
    - So the magnitude of any effect of, for example in our lentil swapping for red meat example, then the characteristics of the wider dietary pattern associated with those consuming red meat versus those consuming lentils is going to be important in that analysis.

Conclusions

1. In nutrition, when thinking about the impact of a nutrient or food, we need to ask "compared to what?"
2. Eliminating or adding a food to your diet will inherently displace other foods.
3. In research, adjusting for confounders is important
4. But knowing what, and how, a study is adjusting for things helps us interpret it better
5. Be mindful of the difference between a macronutrient based substitution model and food-based substitution models.
    a. Macronutrient based models maybe somewhat more straightforward in their interpretation
6. Be mindful that where it's **food-based substitution**:
    a. This is still an emerging area of nutritional epidemiology, relatively speaking in terms of the methodology being employed, the nitty gritty of the statistical models being used and their and what they show.
    b. When we use that phrase, **"compared to what"** for food-based substitution models, don't consider that as food A versus food B.
7. Once you've identified what those foods being modeled in the substitution analysis are, you should take that "compared to what" question and extend it to think "what's the comparison in the characteristics of people who consume food A versus food B? What's the comparison and how do they compare in their wider dietary characteristics? What are the implications for this analysis?"