Sensitivity analyses in observational research

OCTOBER 18, 2019

SCOTT QUINLAN, PHD

ASSISTANT TEACHING PROFESSOR

DEPARTMENT OF EPIDEMIOLOGY AND BIOSTATISTICS

Outline

The challenges of observational research

A little history of sensitivity analyses

Unmeasured confounding and the E-value

Other sensitivity analyses to consider:

- Misclassification
- Exposure and outcome definitions
- Reverse causality

Red meat and cardiovascular health

We want to examine the impact of red meat consumption on cardiovascular health.

Randomized trial → not ethically feasible

Observational study → most common, but has limitations

- People who eat red meat are quite different from those who do not.
- Measuring exposure can be tricky.

Is diet soda really bad for you?

Recent observational studies have hinted at an increased risk of mortality associated with high levels of consumption of diet soda.

But…

People who drink >2 sodas per day are really different than people who drink 1 or fewer sodas per month.

What if people change their drinking habits as a result of poor health?

Lung cancer and smoking

In the 1950s evidence was building that linked cigarette smoking to lung cancer.

One study found a relative risk of lung cancer to be 10.73, comparing smokers to non-smokers.

Other scientists (largely associated with the tobacco industry) claimed that this association could be entirely due to an unobserved confounder.

What can we do?

Such results suggest that an error has been made of an old kind, in arguing from correlation to causation, and that the possibility should be explored that the different smoking classes, cigarette smokers, cigar smokers, pipe smokers, etc., have adopted their habits partly by reason of their personal temperaments and dispositions, and are not lightly to be assumed to be equivalent in their genotypic composition. Such differences in genetic make-up between those classes would naturally be associated with differences of disease incidence without the disease being causally connected with smoking. It would then seem not so paradoxical that the stronger fumes of pipes or cigars should be so much less associated with cancer than those of cigarettes, or that the practice of drawing cigarette smoke in bulk into the lung would have apparently a protective effect.

Fisher, R. A., Nature, 108 (1958).

What is a sensitivity analysis?

When conducting observational studies, we need to make many decisions:

- Cohort selection
- Exposure and outcome definitions
- Measuring confounders
- Choosing an analysis

Sensitivity analysis determines how **robust** our results are to small changes in these decisions.

The solution (at the time)

Cornfield (and others) conducted what many consider to be the first sensitivity analysis and showed that:

In order for this confounder to explain away the result:

- The confounder would have to be 10 times more common in smokers than non-smokers

AND

- Those with the confounder would have to be 10 times more likely to develop lung cancer than those without

It seemed implausible that such a confounder could exist and remain unknown.

Current methods

The issue of unmeasured confounding is still an important consideration in observational research.

Two primary methods have been developed to address this.

- **Array approach** → Plug in a range of values for the potential confounder (prevalence in each exposure group, association with outcome) and see how this impacts our result.

- **Rule-out approach** → Identify what conditions the potential confounder would have to meet for the association we observed to be explained away.

The array approach

We can think of the full adjusted RR and the apparent RR (ARR) as being related:

$$ARR = \boxed{RR} * \boxed{\frac{P_{C1}*(RR_{CD}-1)+1}{P_{C0}*(RR_{CD}-1)+1}}$$

Truth Bias

P_{C1}=prevalence of confounder among exposed
P_{C0}=prevalence of confounder among unexposed
RR_{CD}= association between confounder and outcome

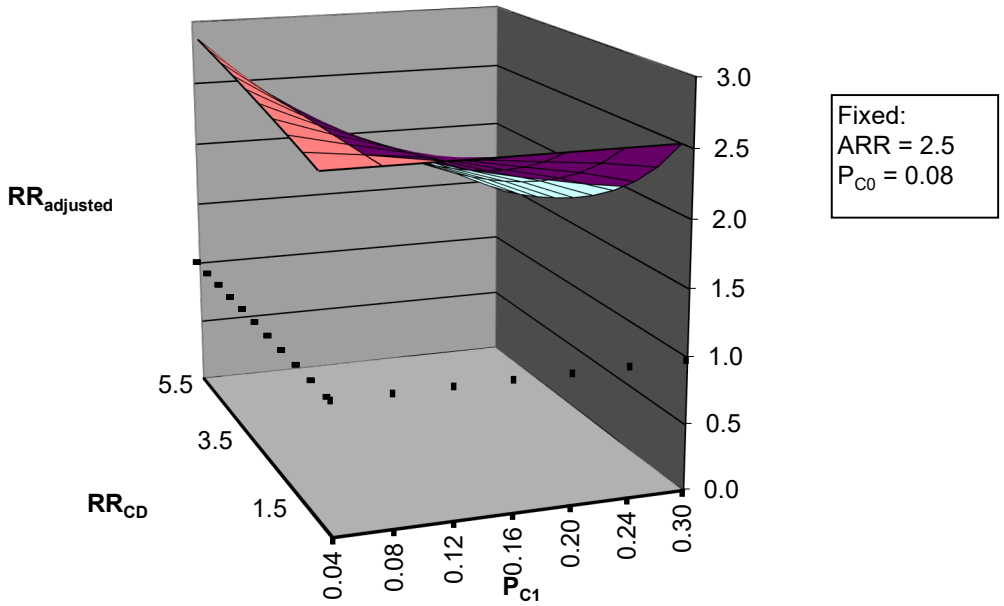
Usually more interesting to solve for RR:

$$RR = \frac{ARR}{\frac{P_{C1}*(RR_{CD}-1)+1}{P_{C0}*(RR_{CD}-1)+1}}$$

If we observe a relative risk of 2.5 linking high red meat consumption to heart disease. We do not have smoking information. What impact might smoking have??

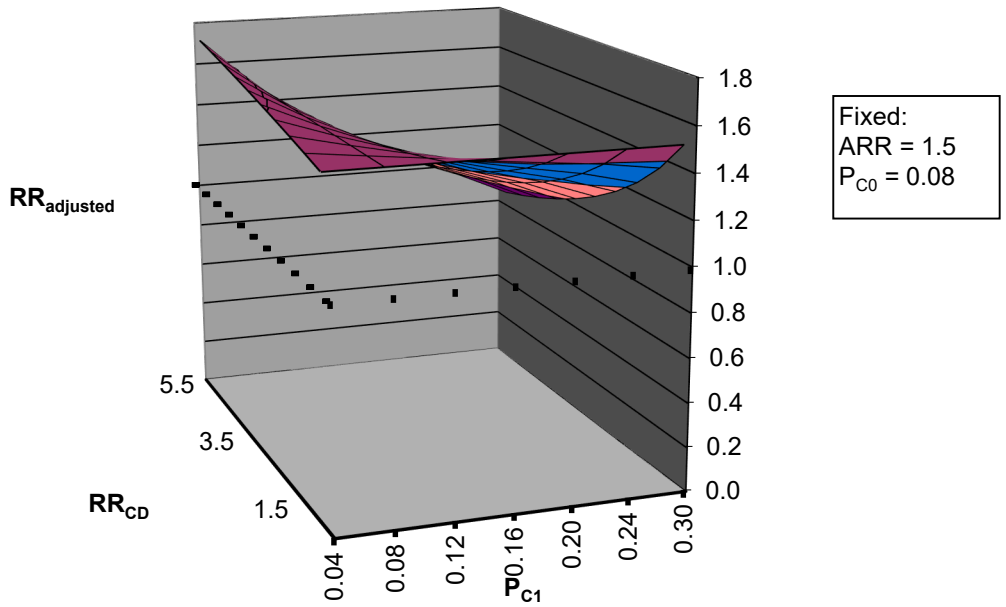
ARR	RR_{CD}	P_{C1}	P_{C0}	$RR_{adjusted}$	% Bias
2.5	5.5	0.12	0.08	2.2	13.24
2.5	5.0	0.12	0.08	2.2	12.12
2.5	4.5	0.12	0.08	2.3	10.94
2.5	4.0	0.12	0.08	2.3	9.68
2.5	3.5	0.12	0.08	2.3	8.33
2.5	3.0	0.12	0.08	2.3	6.90
2.5	2.5	0.12	0.08	2.4	5.36
2.5	2.0	0.12	0.08	2.4	3.70
2.5	1.5	0.12	0.08	2.5	1.92
2.5	1.0	0.12	0.08	2.5	0.00
2.5	5.5	0.16	0.08	2.0	26.47
2.5	5.0	0.16	0.08	2.0	24.24
2.5	4.5	0.16	0.08	2.1	21.88
2.5	4.0	0.16	0.08	2.1	19.35
2.5	3.5	0.16	0.08	2.1	16.67
2.5	3.0	0.16	0.08	2.2	13.79
2.5	2.5	0.16	0.08	2.3	10.71
2.5	2.0	0.16	0.08	2.3	7.41
2.5	1.5	0.16	0.08	2.4	3.85
2.5	1.0	0.16	0.08	2.5	0.00
2.5	5.5	0.20	0.08	1.8	39.71
2.5	5.0	0.20	0.08	1.8	36.36
2.5	4.5	0.20	0.08	1.9	32.81
2.5	4.0	0.20	0.08	1.9	29.03
2.5	3.5	0.20	0.08	2.0	25.00
2.5	3.0	0.20	0.08	2.1	20.69
2.5	2.5	0.20	0.08	2.2	16.07
2.5	2.0	0.20	0.08	2.3	11.11
2.5	1.5	0.20	0.08	2.4	5.77
2.5	1.0	0.20	0.08	2.5	0.00

P_{C1}=prevalence of confounder among exposed
P_{C0}=prevalence of confounder among unexposed
RR_{CD}= association between confounder and outcome



Helpful spreadsheet at:
http://www.drugepi.org/dope-downloads/

ARR	RR_{CD}	P_{C1}	P_{C0}	$RR_{adjusted}$	% Bias
1.5	5.5	0.12	0.08	1.3	13.24
1.5	5.0	0.12	0.08	1.3	12.12
1.5	4.5	0.12	0.08	1.4	10.94
1.5	4.0	0.12	0.08	1.4	9.68
1.5	3.5	0.12	0.08	1.4	8.33
1.5	3.0	0.12	0.08	1.4	6.90
1.5	2.5	0.12	0.08	1.4	5.36
1.5	2.0	0.12	0.08	1.4	3.70
1.5	1.5	0.12	0.08	1.5	1.92
1.5	1.0	0.12	0.08	1.5	0.00
1.5	5.5	0.16	0.08	1.2	26.47
1.5	5.0	0.16	0.08	1.2	24.24
1.5	4.5	0.16	0.08	1.2	21.88
1.5	4.0	0.16	0.08	1.3	19.35
1.5	3.5	0.16	0.08	1.3	16.67
1.5	3.0	0.16	0.08	1.3	13.79
1.5	2.5	0.16	0.08	1.4	10.71
1.5	2.0	0.16	0.08	1.4	7.41
1.5	1.5	0.16	0.08	1.4	3.85
1.5	1.0	0.16	0.08	1.5	0.00
1.5	5.5	0.20	0.08	1.1	39.71
1.5	5.0	0.20	0.08	1.1	36.36
1.5	4.5	0.20	0.08	1.1	32.81
1.5	4.0	0.20	0.08	1.2	29.03
1.5	3.5	0.20	0.08	1.2	25.00
1.5	3.0	0.20	0.08	1.2	20.69
1.5	2.5	0.20	0.08	1.3	16.07
1.5	2.0	0.20	0.08	1.4	11.11
1.5	1.5	0.20	0.08	1.4	5.77
1.5	1.0	0.20	0.08	1.5	0.00



The problem with the array approach

May or may not be data on the prevalence of the confounder and how strongly it is related to the outcome.

Somewhat of a trial and error approach.

Can be manipulated to show that results are robust.

The “rule out” approach

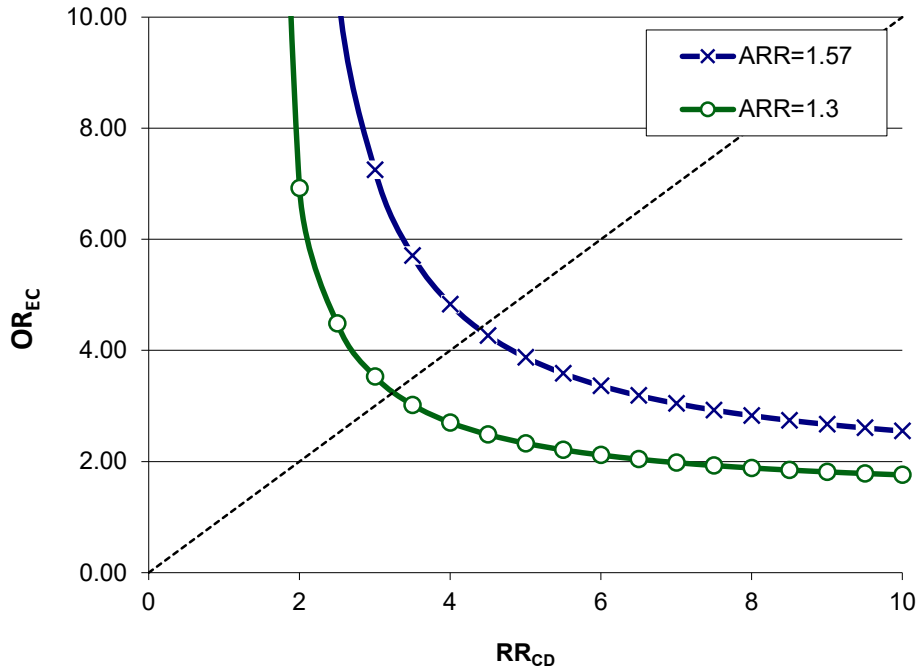
Asks a slightly different question:

- How strong would an unmeasured confounder have to be to completely explain away the observed association (i.e. change to RR =1)?

We conduct a study that shows that those who eat eggs regularly have 1.57 times higher risk of developing coronary artery disease than those who do not. How strong would an unobserved confounder have to be to completely explain away this association?

- What if we found a relative risk of 1.33?
- Are these realistic?

RR_{CD}	P_C	P_E	ARR=1.57	OR_{EC}	ARR=1.3	OR_{EC}
2	0.1	0.01	1.57	24.51	1.3	6.92
2.5	0.1	0.01	1.57	10.67	1.3	4.49
3	0.1	0.01	1.57	7.25	1.3	3.53
3.5	0.1	0.01	1.57	5.71	1.3	3.02
4	0.1	0.01	1.57	4.83	1.3	2.70
4.5	0.1	0.01	1.57	4.27	1.3	2.49
5	0.1	0.01	1.57	3.88	1.3	2.33
5.5	0.1	0.01	1.57	3.59	1.3	2.21
6	0.1	0.01	1.57	3.36	1.3	2.12
6.5	0.1	0.01	1.57	3.19	1.3	2.04
7	0.1	0.01	1.57	3.05	1.3	1.98
7.5	0.1	0.01	1.57	2.93	1.3	1.93
8	0.1	0.01	1.57	2.83	1.3	1.88
8.5	0.1	0.01	1.57	2.74	1.3	1.85
9	0.1	0.01	1.57	2.67	1.3	1.81
9.5	0.1	0.01	1.57	2.61	1.3	1.78
10	0.1	0.01	1.57	2.55	1.3	1.76



Helpful spreadsheet at:
http://www.drugepi.org/dope-downloads/

$RR_{CD} = how\ strongly\ is\ confounder\ related\ to\ outcome?$
$p_c = how\ common\ is\ the\ confounder?$
$p_E = how\ common\ is\ the\ exposure\ of\ interest?$
$OR_{EC} = how\ strongly\ are\ confounder\ and\ exposure\ related?$

These methods are great, but…

All require us to make assumptions about the nature of the confounder.

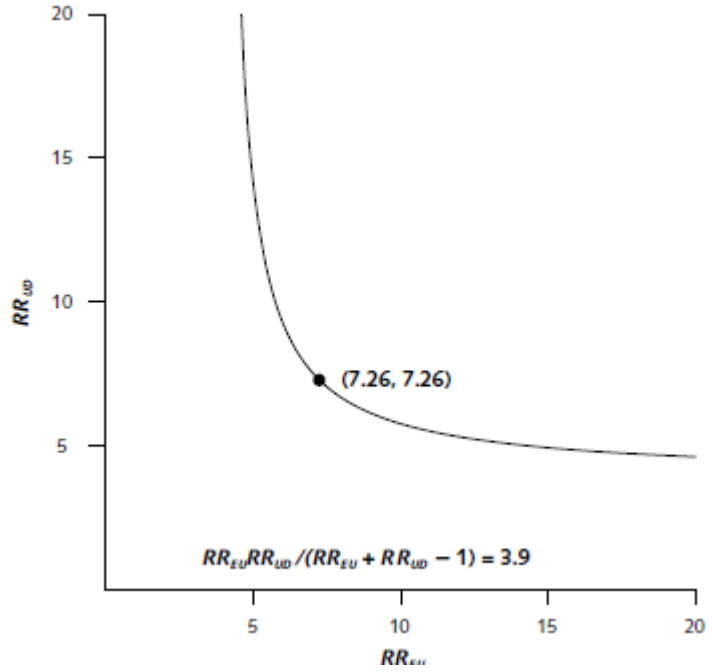
Provide a range of possible impacts on our results, but this can be difficult to summarize succinctly.

We would like a single summary measure of the potential impact of a confounder on our results (similar to p-value).

Introducing the E-value

The E-value is the **minimum** strength of association a potential unmeasured confounder would need to have with **both** the exposure of interest and the outcome of interest to fully explain away an observed association.

- Measured on the **risk ratio** scale.
- Conditional on the measured and controlled confounders.



Example

A study is conducted and finds that the chance of infants dying from a respiratory infection is 3.9 times (95% CI 1.8 to 8.7) higher in infants who were formula fed compared to infants who were exclusively breastfed.

- Already adjusted for age, birthweight, social status, maternal education, and family income.
- But, not able to adjust for smoking status.

- How strongly would smoking have to be associated with both the exposure (breastfeeding or not) and the outcome (infant mortality from respiratory infection) to completely explain this finding?

$$E - value = RR + \sqrt{RR * (RR - 1)}$$

$$E - value = 3.9 + \sqrt{3.9 * (RR - 1)} = 7.26$$

VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med*. Am Coll Physicians; 2017;167:268–274

Example

A study finds the chance of infants dying from a respiratory infection is 3.9 times (95% CI 1.8 to 8.7) higher in infants who were formula fed compared to infants who were exclusively breastfed.

$E - value = 7.26$

This means that to explain away the observed association:

- Mothers who use formula would have to be at least 7.26 times as likely to smoke as those who exclusively breastfeed

AND

- Infants of mothers who smoke would have to be at least 7.26 times as likely to die from a respiratory infection than infants whose mothers do not smoke.
- If one association is weaker than this, the other must be stronger.

Is this realistic?

How confident are we in our confidence?

A study finds the chance of infants dying from a respiratory infection is 3.9 times (95% CI 1.8 to 8.7) higher in infants who were formula fed compared to infants who were exclusively breastfed.

- How strong of a confounder would smoking have to be to make the lower limit of the 95% CI cross/include the null value?

- $E - value = LL + \sqrt{LL * (LL - 1)}$
- $E - value = 1.8 + \sqrt{1.8 * (1.8 - 1)} = 3$

Protective effects of coffee drinking?

A recent study found that people who drink 4 or more cups of coffee per day had notably lower risk of death than people who consumed no coffee (RR: 0.82; 95% CI: 0.78 to 0.87).

- Concerns over whether income might be a confounder here.

- Let $RR^* = \frac{1}{RR} = \frac{1}{0.82} = 1.22$

- Same formula as before, but use RR*:

- $E - value = RR^* + \sqrt{RR^* * (RR^* - 1)} = 1.22 + \sqrt{1.22 * (1.22 - 1)}$
- E-value = 1.73

$$E - value = RR + \sqrt{RR * (RR - 1)}$$

What about other measures?

Measure	Approach
OR or HR for rare outcomes (<15%)	Plug in OR or HR for RR in formula above
Rate ratios	Use rate ratio instead of risk ratio
OR for common outcomes	Use $RR \approx \sqrt{OR}$
HR for common outcomes	Use $RR \approx \frac{1-0.5^{\sqrt{HR}}}{1-0.5^{\sqrt{1/HR}}}$
Continuous outcomes	See article
Risk difference	See article

Ann Intern Med. 2017;167:268-274

E-value in action

A study found that the risk of macrovascular complications following bariatric surgery was notably lower than no surgery (HR: 0.60; 95% CI: 0.42 – 0.86) in people with type 2 diabetes.

Computed an E-value of 2.72 for risk estimate and 1.60 for upper limit of confidence interval.

Is an unobserved confounder with a RR of 2.72 feasible?

- Hypertension: 1.09
- Dyslipidemia: 1.88
- Current smoker: 1.48

JAMA. 2018;320(15):1570-1582

E-value resources available

https://mmathur.shinyapps.io/evalue/

Allows for E-values to be calculated across a range of different effect measures.

Also provides plots and R code.

mmathur.shinyapps.io/evalue/

E-value calculator | Instructions | Compute an E-value | More resources

Outcome type

Relative risk / rate ratio

Point estimate

3.9

Confidence interval lower limit

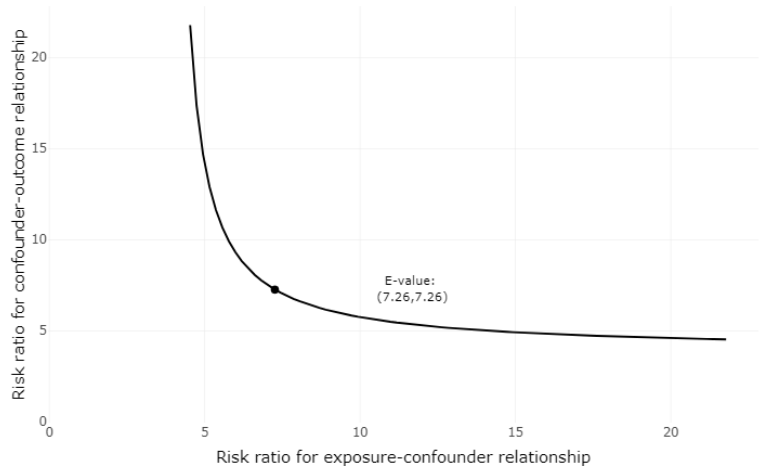
1.8

Confidence interval upper limit

8.7

True causal effect to which to shift estimate (default: null)

1



E-value for point estimate: 7.26 and for confidence interval: 3

Use with caution

The E-value is an important tool, but be wary of misuse and misinterpretation:

- The E-value is **conditional** on other covariates measured and adjusted.
- The E-value only addresses **unmeasured confounding** and there are many other potential biases we need to consider.
- Cannot prove causality, even with a really large E-value.
- Small sample size with large effect size (even if imprecise) can lead to an extremely large E-value.
  - Report E-value for confidence limits too!

The obesity paradox

Several large prospective cohort studies have demonstrated an interesting finding, those who were classified as obese had comparatively *lower* mortality than those who were classified as normal weight among people with certain conditions.

What is going on here?

Could this be a real effect?

Or, the result of how we selected the study sample (selection bias)?

Considering the obesity paradox

Several large prospective cohort studies have demonstrated an interesting finding, those who were classified as obese had comparatively *lower* mortality than those who were classified as normal weight.

How censoring might explain this finding.

- Among the obese patients those in the poorest health drop out of the study making this group look artificially healthier.
- No such difference occurs among those with healthy weight.

What can we do?

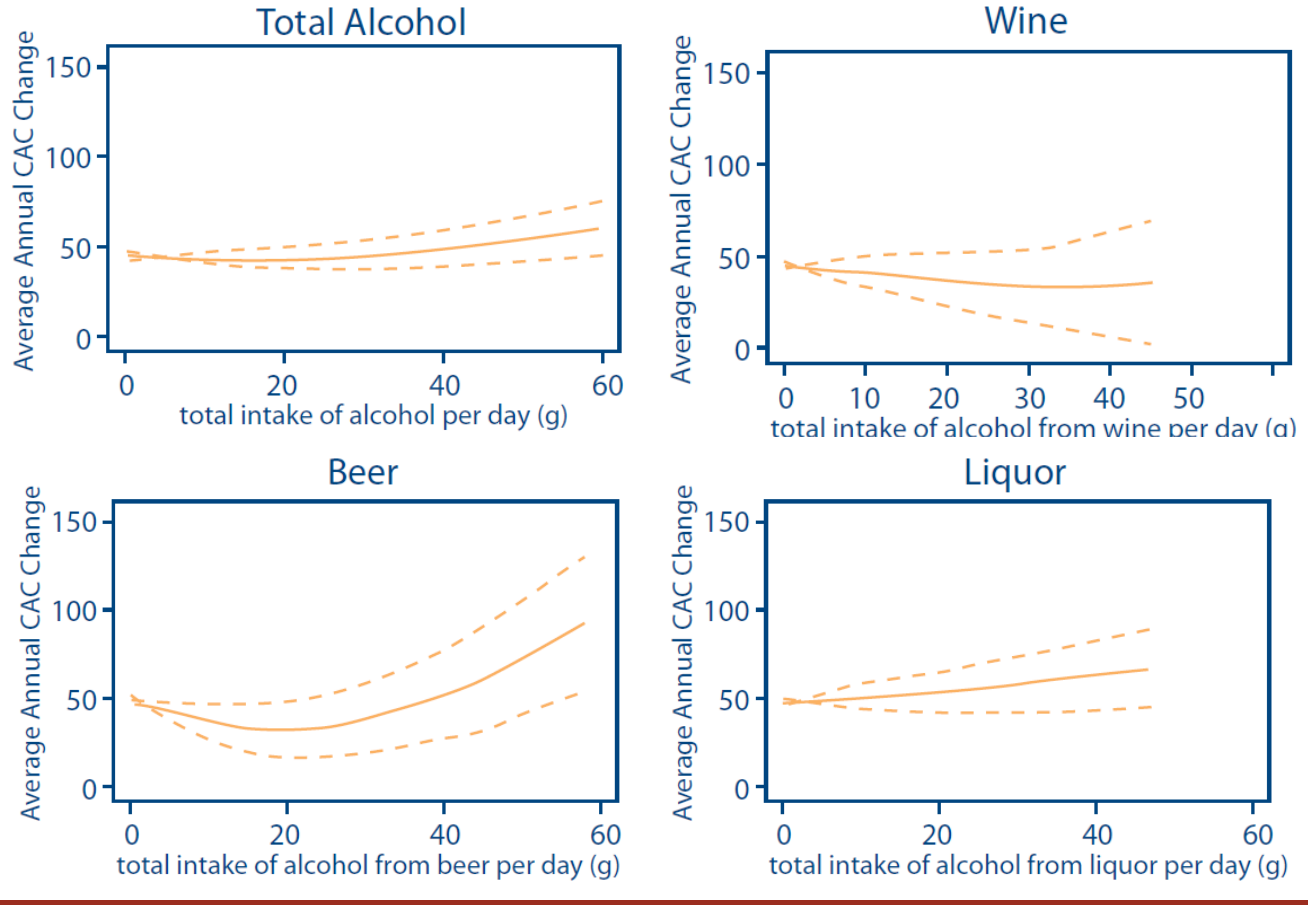
- Make certain assumptions about those lost to follow-up and see how they impact our results.

Carefully considering our exposure definition

Sensitivity analyses can examine how changes in how we define exposure might impact our key results.

McClelland RL, Bild DE, Burke GL, et al. Alcohol and coronary artery calcium prevalence, incidence, and progression: results from the Multi-Ethnic Study of Atherosclerosis (MESA). Am J Clin Nutr 2008 Dec;88(6):1593- 601

Figure 11.1. Smoothed plot of alcohol consumption versus annualized progression of CAC with 95% CIs



Misclassification

A certain amount of misclassification occurs in all studies, usually indicated as non-differential.

We can make certain assumptions about misclassification and see how these impact our results.

- What if 25% of both cases and controls underreport smoking?
- What if 15% of cases but 25% of controls underreport smoking?

Should we be concerned?

- Compare our estimates of prevalence of certain characteristics to other sources to see if they are similar.
- Any validation or calibration studies available?

Red meat consumption and health

The recent article indicated that much of what we know about red meat consumption and health comes from observational studies, which have many limitations.

One challenge in nutritional studies such as this is how do we accurately measure exposure?

Impact of misclassification

We observe:

OR = 1.30

	Case	Control
90 g or more	554	1869
<60 g	376	1648

Cases: 90% Se, 90% Sp
Controls: 90% Se, 90% Sp

	Case	Control
90 g or more	576	1897
<60 g	354	1620

OR = 1.39

Cases: 90% Se, 90% Sp
Controls: 70% Se, 70% Sp

	Case	Control
90 g or more	576	2035
<60 g	354	1482

OR = 1.19

Cases: 90% Se, 70% Sp
Controls: 90% Se, 70% Sp

	Case	Control
90 g or more	458	1357
<60 g	472	2161

OR = 1.54

Reality

Se = sensitivity, likelihood someone who is exposed is classified as exposed
Sp = specificity, likelihood someone who is not exposed is classified as not exposed

Want to live longer? You may want to ditch these drinks

By Sandee LaMotte, CNN

Updated 12:57 AM ET, Wed September 4, 2019

This study included multiple European countries and one of the outcomes was mortality. The countries differed in how they assessed mortality:

- Active follow-up
- Linkage to death registries

The researchers conducted a sensitivity analysis in which they compared the strength of association between soft drink consumption and mortality stratified by the method for death ascertainment.

- Results were **stronger** in countries that used active follow-up.

Measles, Mumps, Rubella Vaccination and Autism

A Nationwide Cohort Study

Anders Hviid, DrMedSci; Jørgen Vinsløv Hansen, PhD; Morten Frisch, DrMedSci; and Mads Melbye, DrMedSci

---

Original definition for an autism diagnosis required **at least one** diagnosis of any of the following:

- Autistic disorder
- Atypical autism
- Asperger syndrome
- Other/unspecified pervasive developmental disorder

How would the results change if they required **at least two diagnoses** on separate dates?

Original result: aHR of 0.93 (95% CI: 0.85 to 1.02)

Updated result: aHR of 0.99 (95% CI: 0.88 to 1.11)

Pass the diet soda?

The same study showed that high levels of consumption of artificially sweetened soda was associated with increased risk of a variety of ailments compared to low consumption.

The study conducted a sensitivity analysis in which they excluded any events that happened within the first 8 years of follow-up. Why?

Trying to rule out the possibility of **reverse causality**.

Summary

Sensitivity analyses are a useful tool to allow us to formally consider some of the limitations of observational research.

Sensitivity analyses can address:

- Unmeasured confounders (E-value)
- Misclassification
- Selection bias
- Reverse causality

Even with sensitivity analyses we cannot prove that an exposure caused to outcome.

We can show the robustness of our evidence.

Learn more

Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiology & Drug Safety* 2006 May. 15(5):291-303

Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiol.* Wolters Kluwer Health; 2016;27:368

VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med*. Am Coll Physicians; 2017;167:268–274

SQUINLAN@GWU.EDU