

Confounding

Mabel Carabali

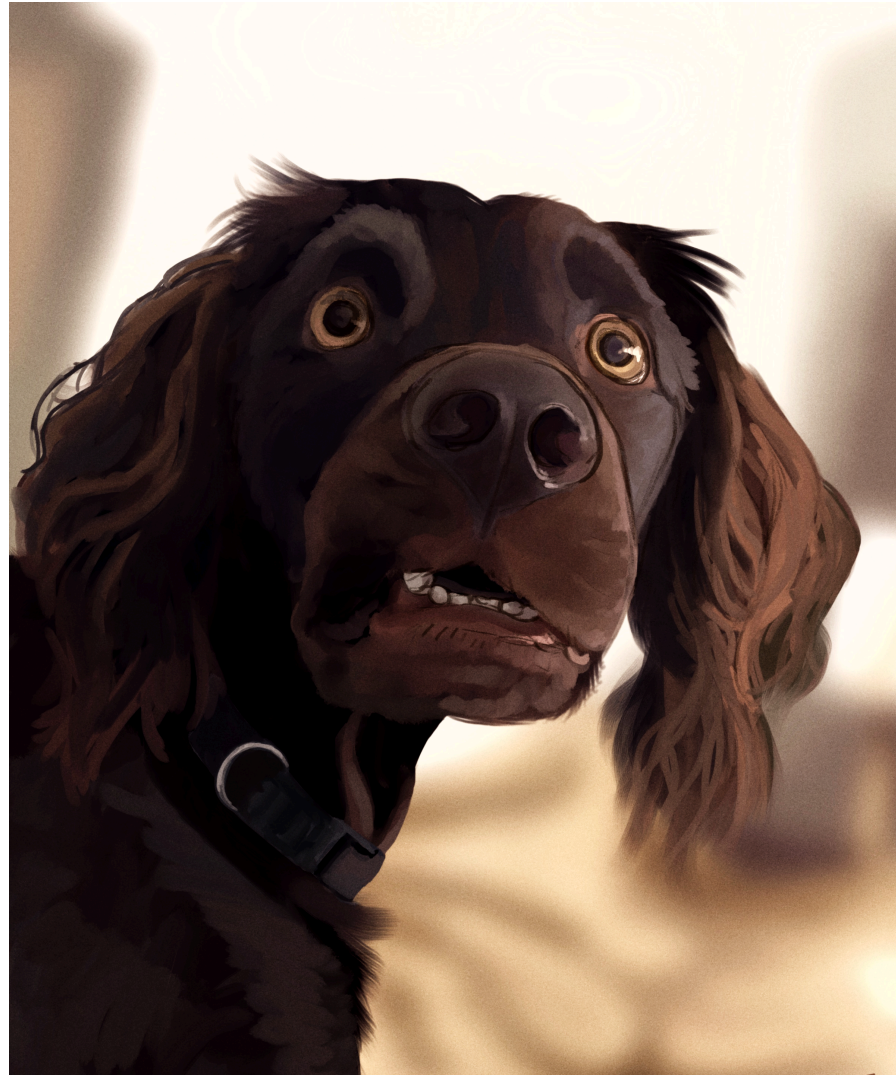
EBOH, McGill University

Updated: 2025-09-19

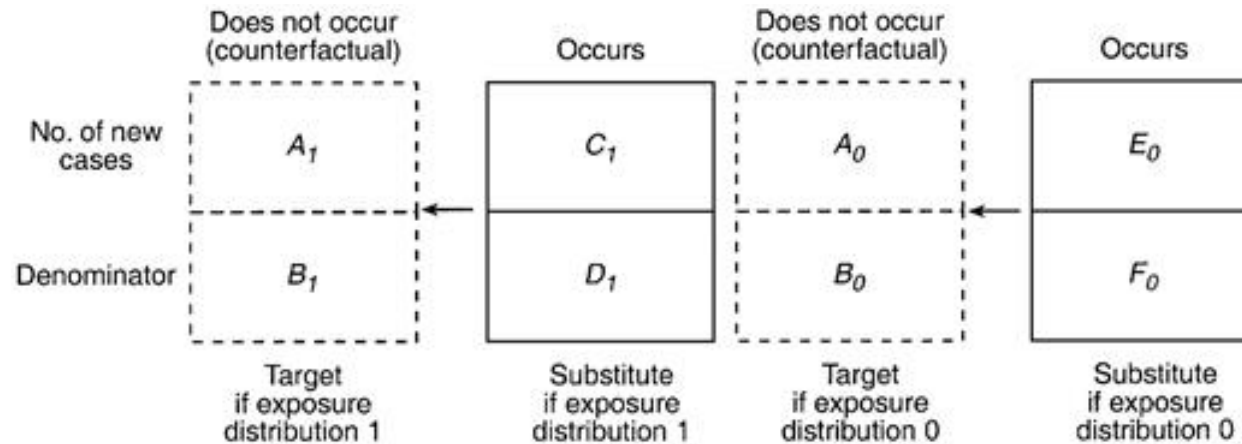
Yes, I measured it, but... is it CAUSAL?

NEW and refurbished material from previous instructors and other resources including material from Epidemiology by design by Daniel Westreich

WHAT??



Fundamental Problem



Maldonado & Greenland 2002

Fundamental Problem

The causal effect:

$$Y^{x=1} - Y^{x=0}$$

- One potential outcome is **observed** & the other potential outcome is **counterfactual**

If we've observed $Y^{x=1}$, then by definition we **have not observed** $Y^{x=0}$

The fundamental problem of causal inference is a problem of missing data.

"Related: we can't formally, with absolute certainty, attribute cause in any individual case except for necessary causes.

- One can't die of AIDS without having HIV – by the definition of AIDS.
 - However, we can sometimes attribute cause with high probability.

You can imagine how it is when an epidemiologist testifies in court about liability due to an exposure.

Four Causal Types

In every population, there exists four types of people, based on their potential outcomes under exposure and non-exposure.

- Doomed
- Immune
- Harmful
- Preventive

Thus, the exposure only has an impact on the outcome on those people that are harmed and those that are protected.

Four Causal Types

Type	E =1	E=0
Doomed	1	1
Harmed ¹	1	0
Preventive	0	1
Immune	0	0

¹ Also called "Causative"

Four Causal Types

Imagine we know that 20% of the population are doomed, 30% are immune, 30% are harmed and 20% are protected by the exposure. What is the causal effect of the exposure on this population (say in 100 people)?

- Under $E=1$, ___ people become diseased
- Under $E=0$, ___ people become diseased
- RD:
- RR:

Four Causal Types

- The proportion that get the diseased when $E = 1$: $p_{doomed} + p_{causative}$
- The proportion that get the diseased when $E = 0$: $p_{doomed} + p_{preventive}$

$$RR_{causal} = \left(\frac{p_{doomed} + p_{causative}}{p_{doomed} + p_{preventive}} \right)$$

- This relationship demonstrates that the size of a causal risk ratio not only **tends to vary** with the proportion of individuals in the target population whose outcome is altered by exposure (**i.e. $P_{causative}$ and $P_{preventive}$**),
- but also tends to vary with the proportion of individuals in the target population for whom disease is **inevitable** by the end of the etiologic time period (**i.e. P_{doomed}**).

Four Causal Types - A note on confounding

$$RR_{causal} = \left(\frac{p_{doomed} + p_{causative}}{p_{doomed} + p_{preventive}} \right)$$

In the real world, the denominator **is not observable**, and so a substitute of some kind is used in its place.

- If this observed substitute population has a risk that **is not equal to the risk in the exposed had they been unexposed**, then the RR estimate is biased
- and we refer to this bias as **“confounding”**.

Four Causal Types

$$RD_{causal} = (p_{doomed} + p_{causative}) - (p_{doomed} + p_{preventive}) = (p_{causative} - p_{preventive})$$

- The causal **RD does not** depend on the proportion of people that are doomed.
- It will only vary with the proportion of individuals in the target population whose outcome is actually altered by exposure

Four Causal Types

- It follows that any factor that affects $P_{causative}$ or $P_{preventive}$ can modify the size of a ratio or difference effect measure,
 - and the factor can modify the size of a ratio effect measure even if it affects only P_{doomed} .
- Not surprisingly, an effect measure could vary from one population to another or from one time period to another
 - Unless we expect other causal factors to have similar distributions across the populations or periods.
 - **Which IRL WE DO NOT!**

Four Causal Types

Type	Outcome E=1	Outcome E=0	Prop. in Exposed	Prop. in Unexposed
Type 1 Doomed	1	1	p1	q1
Type 2 Causative	1	0	p2	q2
Type 3 Preventive	0	1	p3	q3
Type 4 Immune	0	0	p4	q4

RR_{causal} = risk in the exposed / risk in the exposed had they been unexposed =
 $(p1 + p2)/(p1 + p3)$

- As we can't observe the counterfactual we substitute in for $p1 + p3$ using the unexposed group:

$RR_{association} = (p1 + p2)/(q1 + q3)$

- No confounding** when $(p1 + p3) = (q1 + q3)$

Four Causal Types

Imagine we know that:

- 20% of the population are doomed,
- 30% are immune,
- 30% are harmed and
- 20% are protected by the exposure.

What is the causal effect of the exposure on this population (say in 100 people)?

- Under $E = 1$, $(0.2 \text{ (doomed)} + 0.3 \text{ (causative)} = 0.5) \rightarrow 50 \text{ people become diseased}$
- Under $E = 0$, $(0.2 \text{ (doomed)} + 0.2 \text{ (preventive)} = 0.4) \rightarrow 40 \text{ people become diseased}$
- RR: $0.5/0.4 = 1.25 \rightarrow 25\% \text{ more}$
- RD: $0.5 - 0.4 = 0.1 \rightarrow 10 \text{ More}$
 - RD: $0.3 \text{ (causative)} - 0.2 \text{ (preventive)} = 0.1 \rightarrow 10 \text{ More}$

Four Causal Types

- We **never know causal types** of individuals.!!
- We **do not even know the distribution** of those causal types in the population!
- There may be individual causal effects but no average causal effects
 - Consider a drug exposure; equal numbers of people might be harmed and helped by the drug. With more information (genotyping?) we could figure out finer categories of who is helped, and only give the drug to those people. **But those are STILL average causal effects.**
- If there are no individual causal effects, however, there cannot be an average causal effect.
- While a bunch of -1s and +1s might average to 0, the average of a whole bunch of 0s is definitely 0.

Notes on potential Outcomes

- Dichotomous exposures are a simplification
 - If you have more than two levels of exposure, you have more than two potential outcomes
 - One potential outcome for each level of the exposure E.g., if exposure is # cigarettes smoked/day, and that number ranges from 0 to 60, then you have 61 potential outcomes.
 - We still only observe (at best) one potential outcome.
- Deterministic potential outcomes are a simplification.
- Can consider potential outcomes to be stochastic (non-deterministic; random), though we will not.

What are causal effects?

Formally the average causal effect (ACE) in the population, is the **average causal effect** of treatment A on outcome Y if:

$Pr[Y^{a=1} = 1] \neq Pr[Y^{a=0} = 1]$ in the population of interest.

- i.e. the null hypothesis is $Pr[Y^{a=1} = 1] = Pr[Y^{a=0} = 1]$, no average causal effect.

Note that this is different from the absence of an individual causal effect,

because **individuals IRL have different values for $Y^{a=1}$ and $Y^{a=0}$**

What are causal effects?

- Because risk is an average of outcomes, can rewrite the definition of a non-null average causal effect in the population as $E[Y^{a=1}] \neq E[Y^{a=0}]$ so that the definition also applies to non-dichotomous outcomes.
- The average causal effect $E[Y^{a=1}] - E[Y^{a=0}]$ is always necessarily equal to the average $E[Y^{a=1} - Y^{a=0}]$ of the individual causal effects $Y^{a=1} - Y^{a=0}$,
- because a difference of averages is always equal to the average of the differences.

When there is no causal effect for **any** individual in the population, i.e., $Y^{a=1} = Y^{a=0}$ for all individuals, then the “sharp” causal null hypothesis is true.

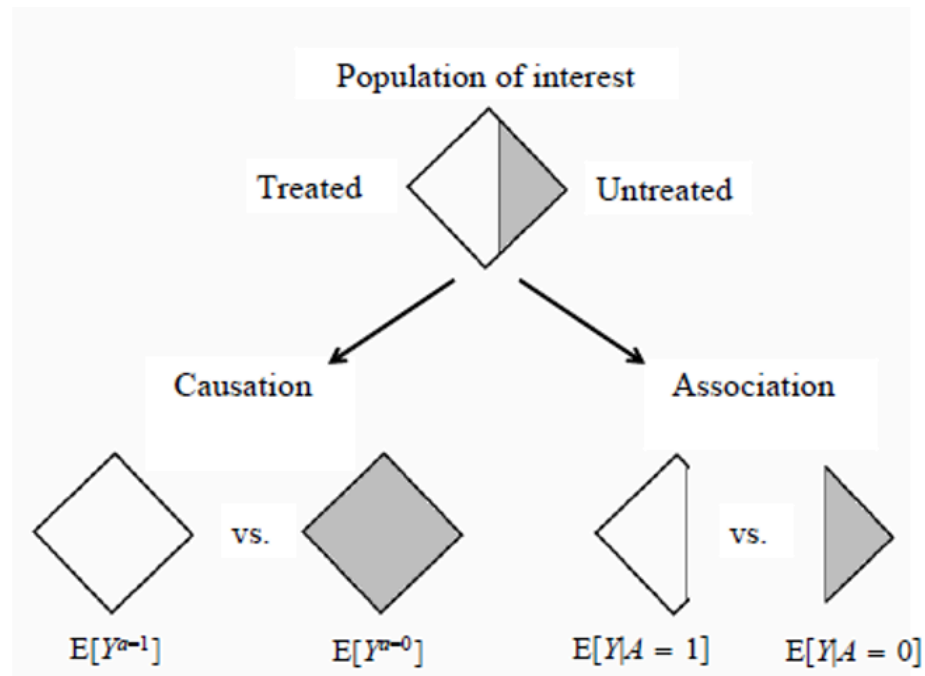
Average causal effects can be identified from data, even if individual causal effects cannot.

How do we estimate causal effects?

- At the individual level, we generally do not.
- At the population level, we find substitute populations.
 - If a population P1 is exposed to smoking, we find a substitute population P2
 - which (*we believe*) represents the experiences of population P1 if,
 - counter to fact, P1 had not been exposed to smoking. Maldonado & Greenland 2002

Validity of inference depends on the appropriateness of the substitute population.

From Hernán & Robins Figure 1.1 as a causal contrast:



What if? Hernán M & Robins J

Counterfactuals

BBT- Counterfactuals- The game or "The Zazzy substitution"

Oversimplified counterfactuals

A paper came out several years ago estimating population attributable fractions for social epidemiology exposures.

- For example, population attributable fraction for the outcome of death, and the exposure of living in high-poverty neighborhoods, for the US in (say) 2010.
- What contrast the population-attributable fraction estimating?

Think about the comparison:

- a US in 2010 with no high-poverty neighborhoods, but everything else remains constant.
 - Is that plausible?

Identification assumptions

What if? Hernán M & Robins J

We say that an average causal effect is (non-parametrically) identifiable when the distribution of the observed data is compatible with a single value of the effect measure. Conversely, we say that an average causal effect is non-identifiable when the distribution of the observed data is compatible with several values of the effect measure.

What is required for causal inference?

- Tricky question. See Greenland EJE 2017

What is required for causal inference?

One sufficient (though not necessary!) set of conditions¹ for causal effect estimation is:

- **Consistency** (well-defined interventions; treatment variation irrelevance)
- **Exchangeability** , or
 - Conditional exchangeability with positivity
- **Positivity**
- We assume temporality in all cases: that causes precede effects.
- We expect absence of Systematic Error!

¹ You will get more of this in 705 and 710

Consistency



Consistency

What if? Hernán M & Robins J,

Consistency simply means that the outcome for every treated individual equals his outcome if he had received treatment, and the outcome for every untreated individual equals his outcome if he had remained untreated.

- This is an axiom rather than an assumption.
- The more critical assumption is that to the extent that there are multiple versions of treatment, they do not matter.

Is also the assumption that we do in fact observe one of the potential outcomes for each individual.

Consistency

- Among people who **were exposed** their outcome was not different than it would have been if they had being **assigned** the exposure - the same for unexposed.
 - Recall Potential outcomes Observed vs substitute population observed outcomes!

So Westreich calls this assumption, ***"treatment variation irrelevance"***

- because it relates to whether there is **meaningful** variation in the **observed exposure or treatment**

Treatment variation irrelevance? (a.k.a., Consistency)

Example: say we're interested in addressing the issue of **hypertension on mortality**. So we compare, *observationally*, a group of individuals with hypertension to a group of individuals without.

But what is the implied causal mechanism? Many possible ways to change a hypertensive person to a non-hypertensive person:

- Medication
- More exercise
- Meditation / stress management
- Less salt in the diet
- Quit smoking
- Combinations of all the above.

Some of these have quite different implications for overall health than others.

Treatment variation irrelevance? (a.k.a., Consistency)

Often, the route you take to changing a factor has a strong implication for the effects of changing that factor

- So you have to assume that exposure contrasts **vary only in ways which are functionally irrelevant to the effect of that cause.**
 - E.g., daily aspirin. Let's assume we don't care whether you take it in the morning or at night, and we don't care how much water you drink when you take it.
- Does this seem like something we could assume in the hypertension case?
- What if (unknowingly) the individuals:
 - exercise more; start meditating; stop or reduce salt consumption and quit smoking

But we measure the effect of Aspirin - Alone 🤪

Interference & Consistency

“Consistency” means that if $A_i = a$, then $Y_i^a = Y^{Ai} = Y_i$

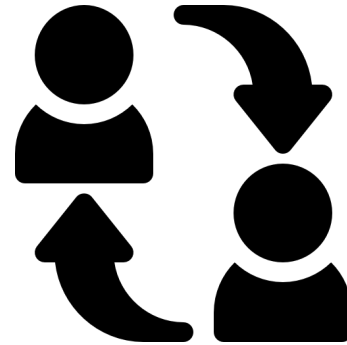
An implicit assumption in the definition of counterfactual outcome is that an individual's counterfactual outcome under treatment value $A=a$ does not depend on other individuals' treatment values.

Interference:

- The counterfactual OUTCOME depending on another person's treatment.
- Interference between individuals is common in studies of infectious or transmissible exposures (viruses, information, etc).
- The counterfactual Y_i^a for an individual i is not well defined because an individual's outcome depends also on other individuals' treatment values.
 - We can only make inferences conditional on this additional exposure.

The assumption of no interference is included in the “stable-unit-treatment-value assumption (SUTVA)” described by Rubin.

Exchangeability



Exchangeability

- The condition that the potential outcomes under **assigned treatment are independent** of the actual treatment received.
 - Possibly, conditional on a set of confounders.
- ~The baseline risk (pre-exposure/treatment) for the outcome among the exposed and unexposed is independent of the exposure they got.
 - E.g: In randomized trial in which codes got switched and everyone randomized to exposure got placebo and vice versa: we would expect the same result, because exchangeability is expected to hold.

Formally: Statistical independence between the potential outcomes and the exposure or treatment **actually received** by the study participants.

$$Y^x \perp\!\!\!\perp X \text{ or } Y^x \perp\!\!\!\perp X \text{ or } Pr(Y^x|X=1) = Pr(Y^x|X=0)$$

Exchangeability

Generally speaking, this is the assumption that there is **no confounding, and no selection bias**.

- Or more likely, ***no unmeasured/uncontrolled confounding, selection bias***.
 - I know! you started this lecture to talk about **confounding** and we have not done so yet.
- We will get there ... **But! Confounding is the main thing that people are worried about when they tell you that correlation is not causation.**

But there IS SO MUCH MORE THAN THAT!.

Positivity



Positivity

Formally: $Pr(X = x) > 0$

- There must be observed individuals at all levels of exposure (or treatment) for all levels of all covariates (confounders) in the study.
 - E.g., aspirin: Say we're studying the effect of daily aspirin on risk of heart attack.
 - We would want to account for age, because older people more likely to take daily aspirin;
 - Older people have more heart attacks. (Draw the DAG.)
 - But IF in our study, everyone over 40 takes daily aspirin;
 - No one under 40 takes daily aspirin.
 - We can't separate the effects of aspirin and age. **No positivity.**

Positivity

An extreme version of this is “exposure opportunity”

- in a study in which the exposure is prostatectomy:
 - we should not be studying people without a prostate

This condition can be loosened (carefully) with model interpolation, and (still more carefully) with model extrapolation. **⚠ not always advisable ⚠**

Positivity

Note:

- Positivity only matters with regard to variables on which you need exchangeability.
 - So if you have exchangeability conditional on certain variables, then you also need positivity on those variables.

Aspirin and Age Example:

we were explicit that we lacked exchangeability because of age (“because older people more likely to take daily aspirin; older people have more heart attacks”), so we had to account for it in the model.

- Because of this, we also needed positivity on age.

So positivity is itself contingent on exchangeability, called “unconditional exchangeability”

- Westreich calls it “conditional exchangeability with positivity”

Other conditions

- In reality, other conditions have to be met in order to assert causality.
- In real data, most of the time, you have to model;
 - So you have to assume the statistical model is correct.

We generally have to assume that there are no dependent happenings (no interference):

- That is, that my exposure does not affect your outcome
- When is this violated? When is this violated and good for public health?

No measurement error is also a necessary condition for causal effect estimation

A case:

Suppose we have a database with information collected among 800 alcohol drinkers and abstainers over 10 years.

Is this risk difference for death at 10 years a causal effect of alcohol?



A case:

Consistency: alcohol consumption is a big category.

- Do you want to "*combine*" people who drinks one beer a day in with those who drinks a six pack a day?
- Only beers, or do wine and cocktails and whiskey counts? And does abstainers mean the same thing to everyone:

If you only drink a glass of wine a week – when you're out at the bar – is that "non-drinking"?
(Consider: how do you imagine moving drinkers into the abstaineres category?)

Exchangeability: are **other** causes of death over represented among drinkers?

E.g., do drinkers exercise less, or have worse LDL cholesterol? Smoking?
If so, then how much of the observed risk difference is due to drinking, versus the other factors?

Positivity: following on exchangeability, can you ever separate out baseline risks? What else?

Finally! ... Confounding

Internal validity versus external validity:

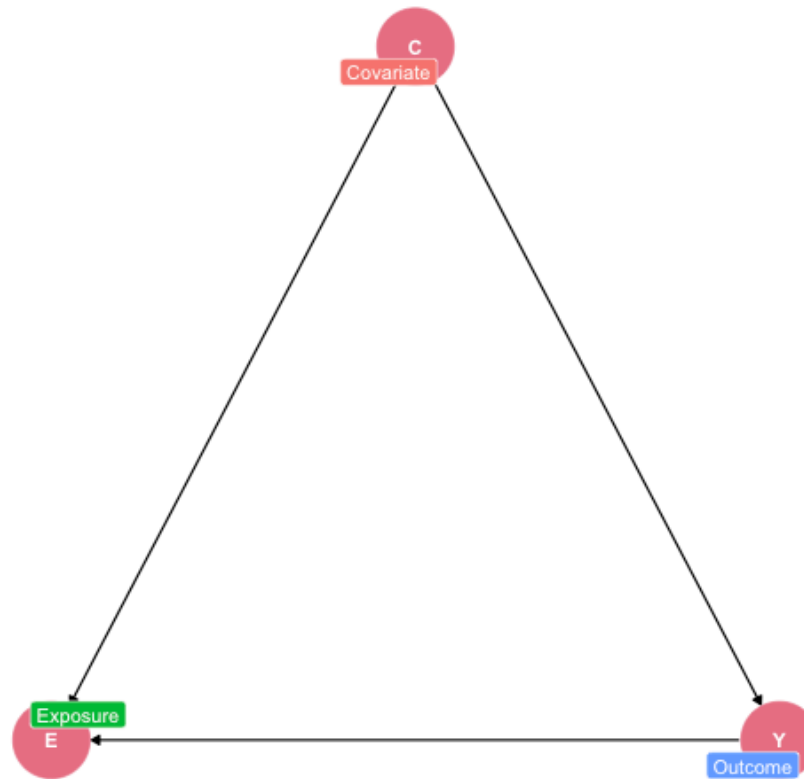
External Validity also involves successful generalization outside of the sample.

Three main threats to internal validity:

- Confounding, Selection bias, and Information bias
 - Confounding means “confusion” of effects. A formal definition of confounding is still actively debated. VanderWeele TJ, Shpitser I. On the definition of a confounder. Ann Stat. 2013 Feb;41(1):196–220. PMID: 25544784
 - Confounding is an imbalance, and can occur by chance (in a randomized trial) or by the influence of a common cause of exposure and outcome. Greenland S. Randomization, statistics, and causal inference. Epidemiology. 1990;1(6):421–9.

What is a confounder?

Relationship between Exposure, Outcome and Covariates



What is a confounder?

General conditions that allow a variable to be a confounder:

Modern Epidemiology 3 p. 133:

- **1) Independently predictive of the disease (i.e. within strata of exposure)** When assessing whether the putative confounder is predictive of outcome, it should be done within strata of exposure, to guard against the possibility that it predicts outcome only because it is correlated with exposure.

Suppose this is the DAG: $C \rightarrow E \rightarrow Y$

- In this case, C is not a confounder because it does not have an independent effect on Y.
 - But there will be an observed association between C and Y, by virtue of their common association with E.
 - But it is not an independent association.

That's why we should assess this criterion within levels of exposure.

- Stratified by E, the association between C and Y is null if there is no direct effect (as shown in the DAG).

General conditions that allow a variable to be a confounder:

Modern Epidemiology 3 p. 133:

- **2) Associated with the exposure (often in the non-diseased)** The association between putative confounder and exposure can be made in the entire study population, but traditionally has been assessed in the controls (or non-diseased) as a stand-in for the population. $A \leftarrow L \rightarrow Y$
 - If the study is a cohort, then presumably the assessment at base-line is made among a population without any disease.
 - If the study is a case-control study then the controls are meant to represent the source population that gave rise to the cases.
 - If the outcome is common, however, do NOT stratify on the outcome when making this assessment, as this would introduce collider stratification bias (since the disease node is a collider, being affected by both confounder and exposure).

General conditions that allow a variable to be a confounder:

Modern Epidemiology 3 p. 133:

- **3) Not on the causal pathway between exposure and outcome.**

This third criterion was overthrown by Robins in 1986-1987: Robins J. The control of confounding by intermediate variables. Stat Med. 1989;8(6):679-701.

But it remains true for traditional confounder control methods in some textbook.

- This can't be known from the data.
- There is no test that will verify this condition, and so it must be based on subject matter knowledge.

Hernán MA, et al. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002;155:176-84.

- We'll start to learn to go beyond this *"But is not an intermediate in the causal pathway between exposure and outcome"* in few slides and if you take EPIB 710

Conditions that allow a variable to be a confounder:

 **Modern Epidemiology 4th, page 268** 

The developments in causal inference over the past decades, summarized in Chapter 3, have made clear that this definition [...the traditional criteria described from ME3...] of a “confounder” is inadequate. It is inadequate because there can be a pre-exposure variable associated with the exposure and the outcome, the control of which introduces, rather than eliminates, bias

Confounding

More formally knowing about the Potential Outcomes/counterfactual thinking:

- Confounding is an **inadequacy of the substitute population** to “stand in” effectively for the experience of the target population under the specified exposure condition.
- Confounding is present if **the substitute imperfectly represents** what the target would have been like under the counterfactual condition.
- An association measure is confounded (or biased due to confounding) for a causal contrast if it does not equal that causal contrast because of an imperfect substitution.
- A confounder is a variable that at least partly explains why confounding is present.

Confounding and the substitute population

The equalities that must be met to control confounding are:

- $E0/F0 = A0/B0$ in scenario 1 (target observed under exposure 1) and
- $C1/D1 = A1/B1$ in scenario 2 (target observed under exposure 0).
- If the entire target is not observed in one or the other exposure states,
- Then both substitutions are necessary (except in the unlikely case of incidental balancing out of the two substitution biases).

Confounding and “study design”?

- Because it is the goal of all etiologic designs to estimate causal contrasts, different study designs are just different ways of choosing a target that corresponds to the study question, and choosing substitutes and sampling subjects from target and substitutes into the study.
- Advantages and disadvantages of different designs are just intended to balance trade offs among bias, variance, study costs and study time.
- Some studies, by design, are more prone to confounding ¹
 - but **all** are potentially susceptible to confounding if the substitute *miss-represents the target population*

¹ More on this in EPIB 703.

Confounding

Recall:

No confounding when $(p_1 + p_2) = (q_1 + q_2)$.

- This is a different criteria for confounding that previously.
- Thus, confounding may be present for one target population but absent for another target population.

Example using the causal types

E=1			E=0		
Type 1:	Pdoomed (p1)	= 0.3	Type 1:	Pdoomed (q1)	= 0.2
Type 2:	Pcausal (p2)	= 0.1	Type 2:	Pcausal (q2)	= 0.3
Type 3:	Ppreventive (p3)	= 0.2	Type 3:	Ppreventive (q3)	= 0.2
Type 4:	Pimmune (p4)	= 0.4	Type 4:	Pimmune (q4)	= 0.3

Target: E=1

- $RD_{causal} = (p_1^{E=1} - p_2^{E=1}) = 0.1 - 0.2 = -0.1$
- $RD_{association} = (p_1 + p_2) - (q_1 + q_3) = (0.3 + 0.1) - (0.2 + 0.2) = 0$

Exposure appears null but is actually preventive

E=1			E=0		
Type 1:	Pdoomed (p1)	= 0.3	Type 1:	Pdoomed (q1)	= 0.2
Type 2:	Pcausal (p2)	= 0.1	Type 2:	Pcausal (q2)	= 0.3
Type 3:	Ppreventive (p3)	= 0.2	Type 3:	Ppreventive (q3)	= 0.2
Type 4:	Pimmune (p4)	= 0.4	Type 4:	Pimmune (q4)	= 0.3

Target: E=0

- $RD_{causal} = (p_1^{E=0} - p_2^{E=0}) = 0.3 - 0.2 = 0.1$
- $RD_{association} = (p_1 + p_2) - (q_3 + q_1) = (0.3 + 0.1) - (0.2 + 0.2) = 0$

Exposure appears null but is actually harmful

Direction of confounding depends on the target population of interest

QUESTIONS?

COMMENTS?

RECOMMENDATIONS?

Some extra slides

A Classic Example to Read: GREENLAND and ROBINS Int J Epidemiology 1986; 15 (3): 412-418

- Suppose (unknown to us) exposure has no effect, i.e., there are no Type 2 or 3 individuals, and that we measure a risk factor having the following joint distribution with exposure and type:

Type	Factor present		Factor absent		Crude	
	Exposed	Unexposed	Exposed	Unexposed	Exposed	Unexposed
1 ('doomed')	60	70	40	180	100	250
4 ('immune')	40	30	60	220	100	250
Total	100	100	100	400	200	500
Incidence	0.60	0.70	0.40	0.45	0.50	0.50

Note that only the last 2 lines of the table would be observable in a real study.

- From these last two lines (the observable data) we can see that the factor is associated with exposure (100 of 200 exposed have the factor, as opposed to 100 of 500 unexposed), and is predictive of outcome among the unexposed (among the unexposed, 0.45 of those without the factor will get the disease, as opposed to 0.70 of those with the factor).

A Classic Example to Read:

GREENLAND and ROBINS Int J Epidemiology 1986; 15 (3): 412-418

- If we fail to control the factor, we will observe disease in an equal proportion (1/2) of the exposed and unexposed, correctly indicating no exposure effect,
- while if we control the factor, exposure will incorrectly appear to be preventive of disease in both strata.
- More precisely, $p_3=0$ and thus $p_1+p_3 = p_1 = 0.50 = q_1 = q_1+q_3$ so that the crude estimate is unconfounded ,
- while $0.70 (100) + 0.45 (100) = 115 \neq 0.50 * (200)$, so that the adjusted estimate is confounded.

Mindbogglingly!