EMBLEM
High Quality, High Impact Research

Cancer Research Methods Workshop

Mayur M. Desai, PhD, MPH

Mwanza, Tanzania

September 18-19, 2012

Yale SCHOOL OF PUBLIC HEALTH

Confounding

Confounding

- From the Latin *confundere* – to mix together
- The situation where the apparent effect of the exposure of interest is distorted because the effect of an extraneous (nuisance) factor is mixed with or mistaken for the actual exposure effect (which may be null)
- More likely to occur in observational studies than in RCTs, since randomization works to minimize differences between groups with respect to known and unknown confounders
- As with selection bias and information (misclassification) bias, the distortion introduced by confounding can lead to overestimation, underestimation, or crossover of effect

Is confounding "bias"?

(It's useful to conceptualize confounding as distinct from bias.)

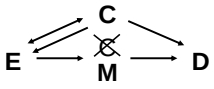
Bias	Confounding
• Distorted measure of association resulting from systematic error in study design or conduct	• Distorted measure of association resulting from interrelatedness of variables as they exist
• Biased associations are bogus (spurious)	• Confounded associations, though not causal, are "real"
• Biased associations are worthless	• Confounded associations can still be useful in terms of identifying "markers" of risk

For an association to be causal, need to rule out bias and confounding, as well as the role of chance (random error)

Confounding

A confounder is a variable that meets all 3 of the following criteria:

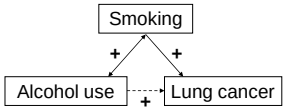
- Confounder [C] is associated with exposure [E]
- Confounder [C] is a risk factor for disease [D]
- Confounder [C] is not in the E → D causal pathway



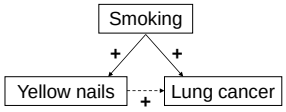
A mediator [M] is a variable that is in the E → D causal pathway (or is hypothesized to be)

Let's look at some hypothetical examples…

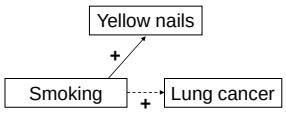
Is smoking a confounder of the association between alcohol use and lung cancer?



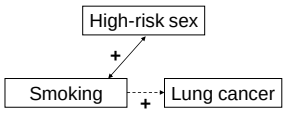
Is smoking a confounder of the association between yellow fingernails and lung cancer?



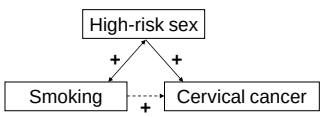
Is having yellow fingernails a confounder of the association between smoking and lung cancer?



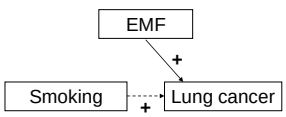
Is having high-risk sex a confounder of the association between smoking and lung cancer?



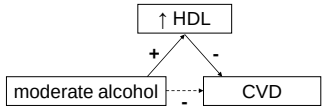
Is having high-risk sex a confounder of the association between smoking and cervical cancer?



Is exposure to electromagnetic fields a confounder of the association between smoking and lung cancer?



Is increased HDL cholesterol level a confounder of the association between moderate alcohol use and CVD?



Assessing the presence of confounding

1. Is the confounder associated with both the exposure and the outcome (and not in the causal pathway)?
2. How do the stratum-specific ORs compare with one another and with the unadjusted OR?
3. How do the unadjusted and adjusted ORs compare with one another?

Note: #3 is the most persuasive approach for determining confounding

Hypothetical example of male gender (exposure) as a risk factor for malaria (outcome)

	Cases (N=150)	Controls (N=150)
Males	88	68
Females	62	82

Unadjusted OR = 1.71

(From Szklo and Nieto)

Is having an outdoor occupation a confounder?

Q: Is the confounder associated with the exposure?

	Male	Female
Outdoor	68	13
Indoor	88	131

$OR_{C-E} = 7.79$

Q: Is the confounder associated with the outcome?

	Cases	Controls
Outdoor	63	18
Indoor	87	132

$OR_{C-D} = 5.31$

Is having an outdoor occupation a confounder?

Q: What's going on in the stratified analysis?

	Cases	Controls
Male	88	68
Female	62	82

Unadj OR = 1.71

How about after stratifying by the confounder?

Adj OR = 1.01

Outdoor

	Case	Control
Male	53	15
Female	10	3

$OR_{outdoor} = 1.06$

Indoor

	Case	Control
Male	35	53
Female	52	79

$OR_{indoor} = 1.00$

Is having an outdoor occupation a confounder of the association between male gender and malaria risk?

- ✓ The confounder was associated with both the exposure and the outcome
- ✓ The stratum-specific ORs were of similar direction and magnitude relative to one another, but differed from the unadjusted OR
- ✓ The unadjusted and adjusted ORs were substantially different – the effect of gender changed after controlling for work environment

→ Answer: Yes

Prevention and control of confounding

- Design phase
  - randomization
  - restriction
  - matching (to be discussed later)
- Analysis phase
  - stratification and adjustment
  - multivariate analysis (statistical modeling)

Randomization

- Used in experimental studies (RCTs)
- Fundamental principle: with a sufficiently large sample size, comparison groups will tend to be similar with respect to known and unknown confounding variables
- Still the possibility of confounding

Restriction

- Restrict the sample to a specific level of a confounder (or combination of confounders) to prevent the confounder(s) from varying between the comparison groups
- Advantages:
  - conceptually straightforward
  - efficient
  - effective

Consider restriction when…

- Studying relatively weak associations involving strong, known confounders
  - e.g., restricting study of passive smoking and liver cancer to nonsmokers
  - e.g., restricting study of mouthwash use and oral cancer to nonsmokers and nondrinkers
- It is clear (or certainly likely) that the number of subjects in a particular level of a confounder will be relatively too small, and thus the study will likely lack sufficient power to examine the effect of this factor
  - e.g., restricting studies by age
  - e.g., restricting studies of veterans to male veterans