Confounding and Stratification

Epidemiological and Statistical Methods Day 1

Mutsa Bwakura Dangarembizi

2019

wellcome

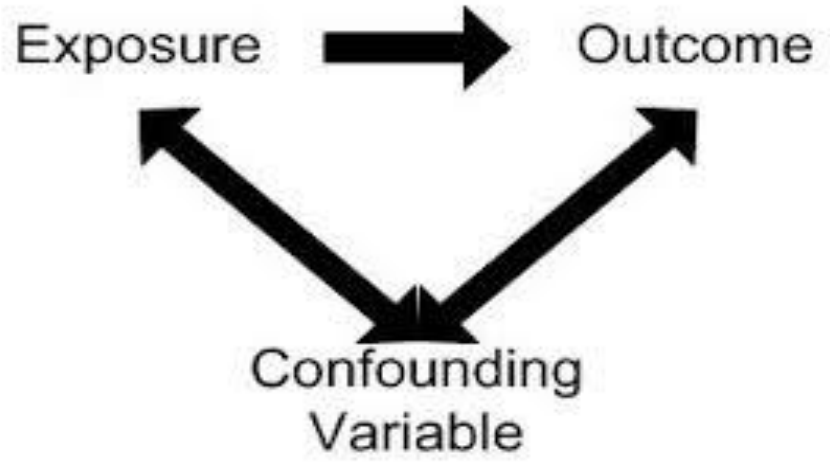
Objectives

- Understand the concept of confounding
- Explain the properties of a confounder
- Describe ways to control and assess for confounding
- Describe residual confounding and Interaction

What is confounding?

What is confounding?

- Distortion (inaccuracy) in the estimated measure of association between an exposure and an outcome
  - Occurs when the exposure of interest is mixed up with some other factor that is associated with the outcome



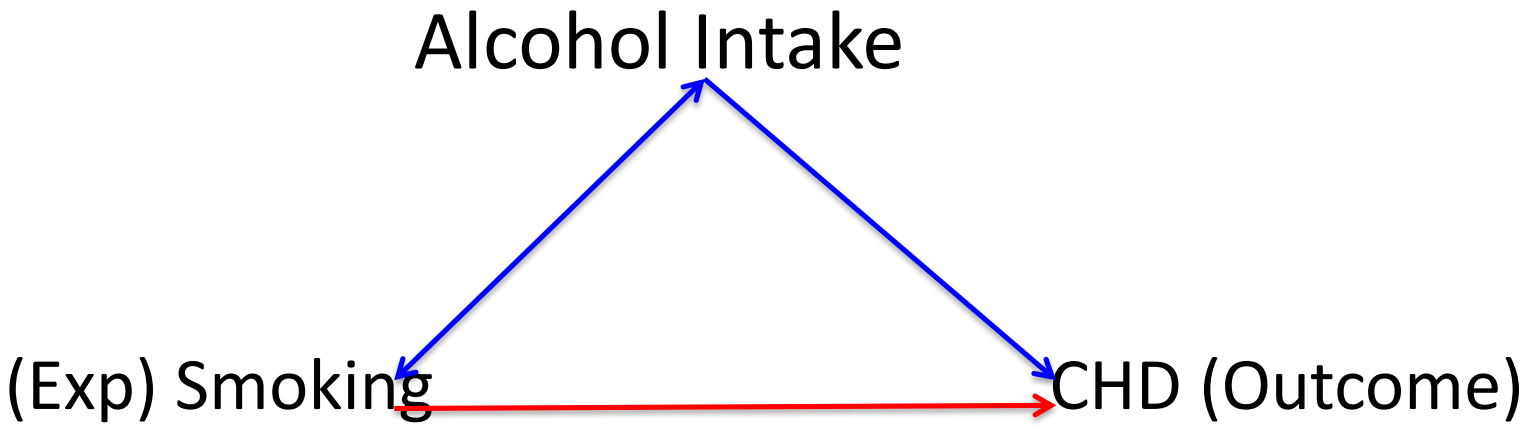
Consequences of confounding

- Can lead to overestimation or under-estimation of the true association between exposure and outcome
- Can change the direction of the observed effect

Example 1.

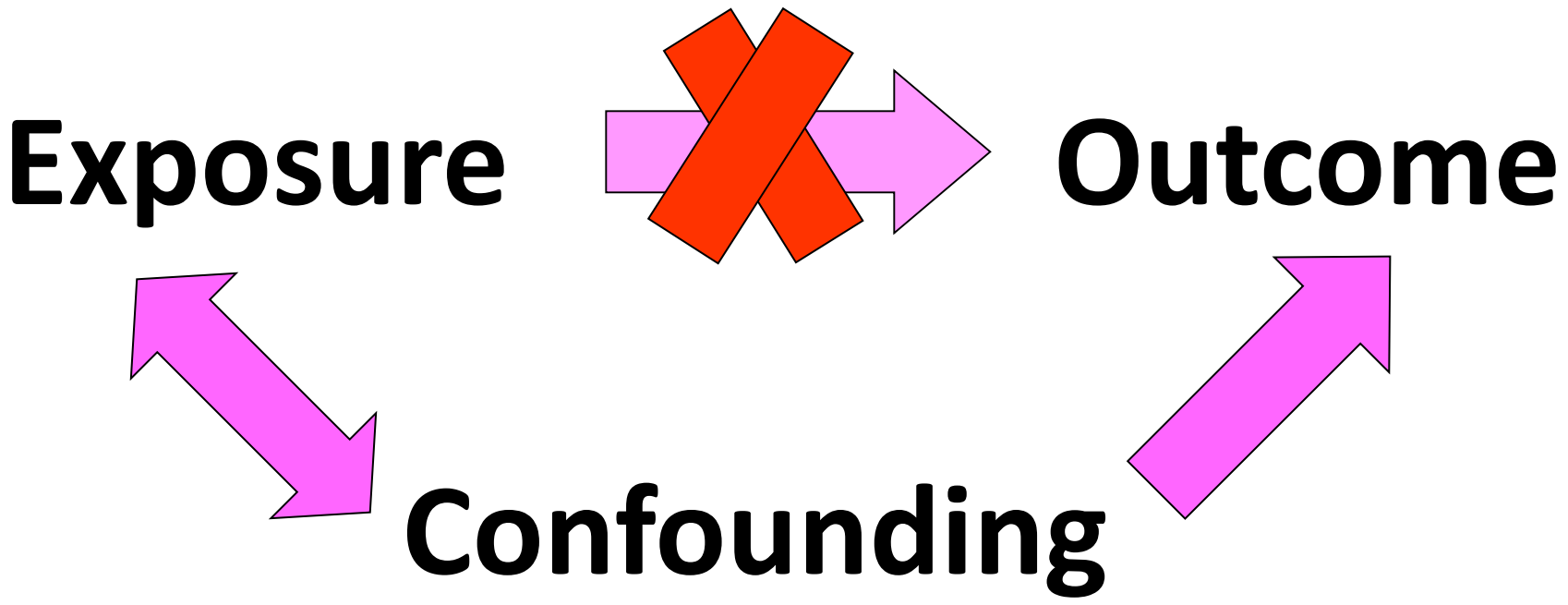
Does smoking influence the risk of coronary heart disease in men aged 18 to 64yr?

Example 1



Effect of alcohol intake mixes up the estimate of the association between smoking and CHD and alcohol intake is also correlated with smoking

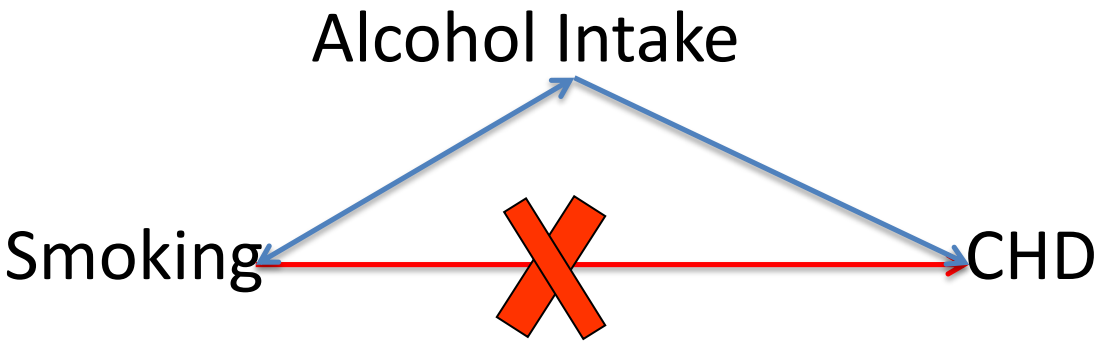
Criteria for Confounding



Criteria for confounding

- Must be associated with the exposure of interest in the source population
- Must be a risk factor for the outcome of interest in those who are not exposed
- Must not be on the causal pathway between the exposure and outcome of interest

For example 1

- 

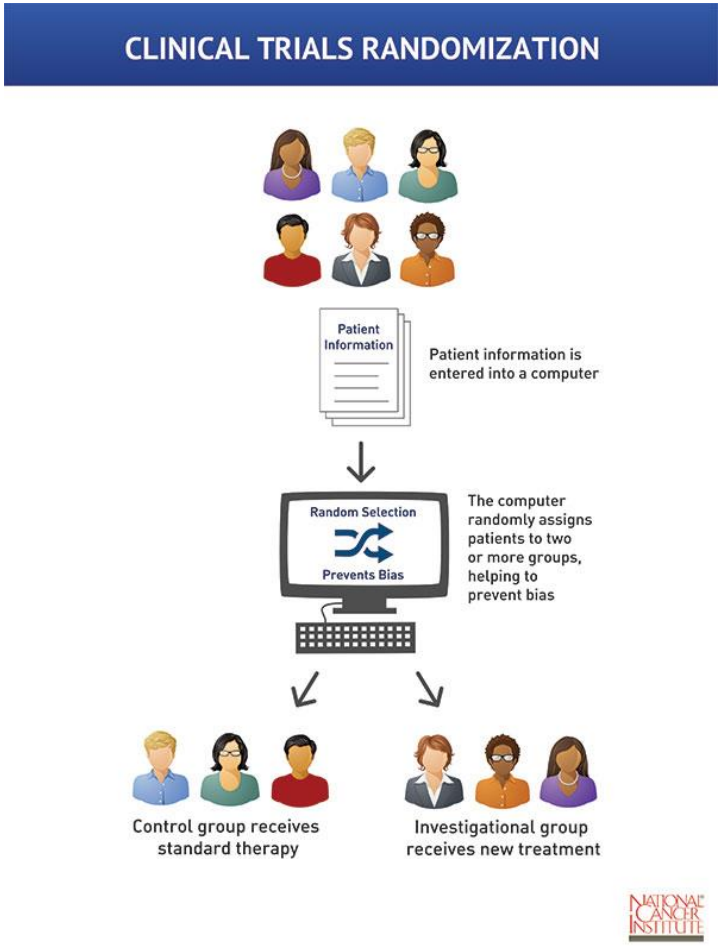
1. Smokers generally have a higher alcohol intake
2. Alcohol intake is a risk factor for CHD in non-smokers
3. Alcohol intake is not on the causal pathway between smoking and CHD

An example of being in the causal pathway

Controlling for confounding

Study design stage

- Randomization
  - Ensures that all known and unknown confounding factors are evenly distributed in the study groups
  - Need a large sample size



Restriction

- Limit participation to individuals who are similar in relation to the confounder
  - eg if smoking limit to non-smokers
  - Convenient and cheap
  - But
  - Reduces eligible participants
  - Limits generalisability

Matching

- Individuals are selected in such a way that the potential confounders are distributed in similarly among the study groups
  - Eg case control study of exercise and risk of MI
    - If age, sex and smoking are potential confounders, match the cases and controls
  - Difficult and expensive
  - Useful in case control studies

Controlling for confounding during analysis

- You must have measured the confounder
- Stratification
  - Stratify the data according to the potential confounder
  - Then measure the association between exposure and outcome in the different strata
  - If they are different from the "crude" measure of effect but similar to each other, this is evidence of confounding

Example: Case control study

- Coffee drinking and risk of pancreatic cancer
  - Cigarette smoking potential confounder

	Coffee drinker	Non coffee drinker
Ca pancreas	450	300
No Ca Pancreas	200	250

Estimated OR = 450/300 ÷200/250=1.9

Data analysis according to smoking habit

	Non smokers	Non smokers	Smokers	Smokers
	Coffee	No coffee	coffee	No coffee
Ca pancreas	50	100	400	200
No Ca pancreas	100	200	100	50

Estimated OR= 50/100÷100/200=1.0 400/200÷100/50=1.0

Smoking confounded the association between coffee consumption and Ca pancreas
No effect of coffee on Ca pancreas for smokers or non smokers

Example: cohort study

Is there an association between exposure to chemical X and the risk of developing lung cancer?

		Exposure to chemical X	
		Yes	**No**
Lung cancer	Yes	480	360
	No	83,520	95,640
	Total	84,000	96,000

Occupational cohort study

Risk in the exposed=480/84000

Risk in unexposed=360/96,000

Crude risk ratio
=(480/84,000)/(360/96,000)

=1.52 (95% CI 1.33-1.75)

$p<0.001$

What is the interpretation?

Stratified analysis

Stratified analysis

SMOKERS		Exposure to chemical X	
		Yes	No
Lung cancer	Yes	80	160
	No	3,920	15,840
	Total	4,000	16,000

- Risk ratio among smokers =

  (80 / 4,000) / (160 / 16,000)

  = **2.00** (95% CI 1.53-2.61)

NON-SMOKERS		Exposure to chemical X	
		Yes	No
Lung cancer	Yes	400	200
	No	79,600	79,800
	Total	80,000	80,000

- Risk ratio among non-smokers =

  (400 / 80,000) / (200 / 80,000)

  = **2.00** (95% CI 1.69-2.37)

An example of negative confounding Crude RR = 1.52

Hypothetical examples of confounding

	Crude effect	Effect in stratum 1	Effect in stratum 1	Adjusted effect	Interpretation
1	2.0	2.0	2.0	2.0	No confounding
2	4.0	3.0	3.0	3.0	confounding
3	1.53	2.0	2.0	2.0	confounding

Summary of stratification analysis

- Confounding is present when there is a difference in the crude and adjusted estimates
  - Evidence of confounding if >10% difference
- A weighted average is used when strata have varying number of individuals
  - Mantel-Haensel method
- There is no statistical test for confounding

Controlling for confounding during analysis

- Standardization
  - Techniques used to remove the effects of differences in age or other confounding variables when comparing 2 or more populations
- Multivariate analysis
  - Allows adjustments for many confounders simultaneously
  - Involves construction of a mathematical model to describe the association between exposure and disease as well as confounders

Residual confounding

- Results from the incomplete adjustment for a confounder
  - Eg adjusting for social status, using broad categories (age)
  - Usually difficult to measure

Interaction (Effect Modification)

- Applied to situations in which the magnitude of the effect of an exposure of interest differs depending on the level of a third variable.
- A variable that differentially modifies the observed effect of a risk factor on a disease status
- Is associated with the outcome but not the exposure
- It must be detected and described

Example of interaction

- Reye's syndrome is a rare but severe condition characterized by brain damage and liver dysfunction after a viral illness
  - Most common in 4-14 yr age-group treated with aspirin
  - Can also occur in adults but is rare
- Effect of aspirin treatment for a viral illness is clearly modified by age

Conclusion

- Confounding is concerned with “alternative” explanations for the effect seen between an exposure and the outcome
- A confounder is associated with the exposure and a risk factor for the outcome but is not in the causal pathway
- To prevent confounding, choose an appropriate study design and control for it using statistical methods
- When confounding is present, the effect of the exposure of interest is the same in all strata of the confounding factor, but is different from the “crude” effect
- Effect modification must be detected and described

Sources

- Epidemiology in Medicine Hennekens and Buring
- ICEMS - LSHTM 2018

Thank you