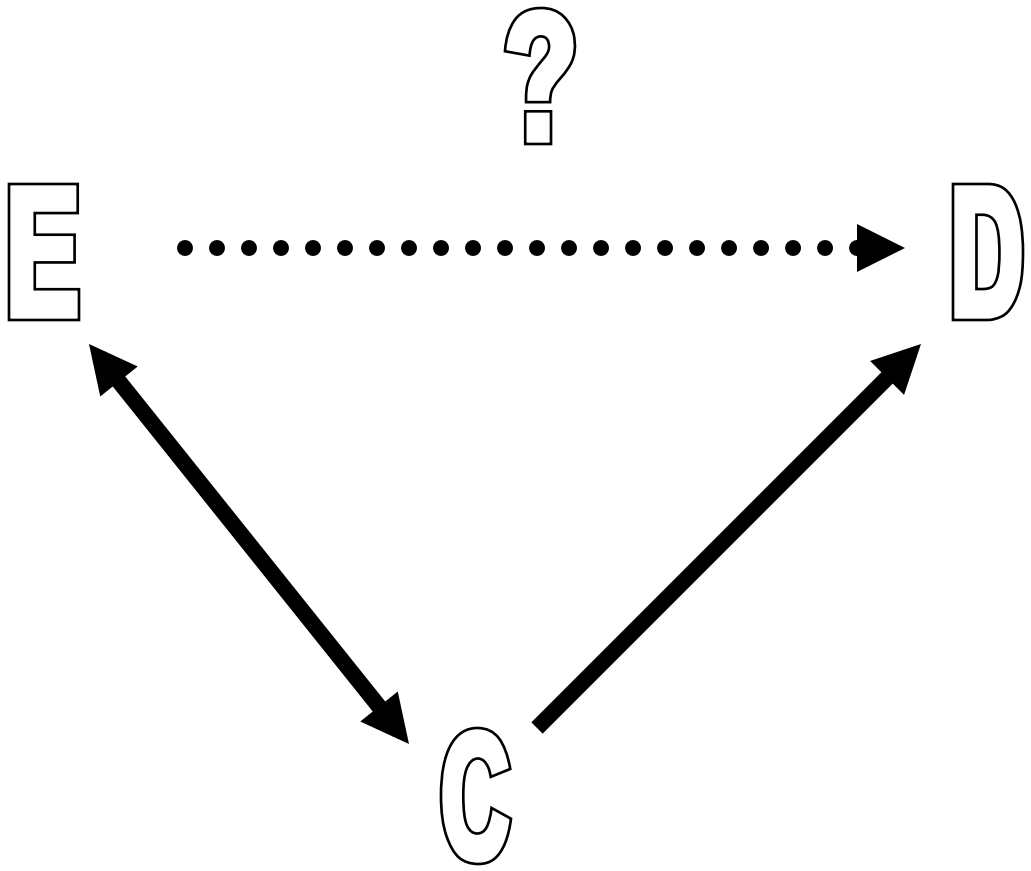
Confounding

A variable that (a) is causally related to the disease under study (or is a proxy for an unknown or unmeasured cause) and (b) is associated with the exposure under study (Kesley)

- Any risk factor for a disease is a potential confounder
- Wholly or partially accounts for apparent effect of exposure on disease (either direction)
- Occurs in nature, not due to study design or execution



Confounding

Examples of Confounding

- Lighters and Lung Cancer
- Breast Cancer Prevention
  - Breast Feeding
  - ? Parity
  - Age at first pregnancy
- Coffee Drinking and Myocardial Infarction

MYOCARDIAL INFARCTION				
COFFEE		**Yes**	**No**	
	Yes	**90**	**60**	**150**
	No	**60**	**90**	**150**
		150	**150**	**300**

OR = ad/bc = (90)(60) / (60)(90) = 2.25

Controlling confounding through stratified analysis

	Smokers		Non-Smokers	
	MI	No MI	MI	No MI
Coffee	80	40	10	20
No Coffee	20	10	40	80
Totals	100	50	50	100
	OR = 1.0		OR = 1.0	

Controlling Confounding

- A. Controlling by Design
  - a) randomization – assures same # with and without any potential confounder in both groups
  - b) restriction – only allow into study if fall into specific groups
  - c) matching – for every person with a factor in case have person without in controls
- B. Controlling by Analysis
  - a) ***stratified analysis – make groups homogenous***
  - b) multivariate analysis – most popular

Effect Modification and Interaction

- 1. Definitions
  - a) traditional (statistical)
    - Risk of Lung Cancer
  - b) biological
  - c) public health
- Additive vs. Multiplicative (lack of one implies the other)

		Asbestos	
		Yes	No
Smoking	Yes	50	10
	No	5	1

New thinking about interaction

- synergy – parallelism = positive additive interaction = R(AB) – R(B) – R(A) + R(ab)

Bias

- A systematic error in the *collection or interpretation of data* in an epidemiologic study. (Henneken) Any systematic error in the *design, conduct or analysis* of a study resulting in a mistaken estimate of an exposure effect. (Schlesselman)
- Found in the design or conduct of study, as opposed to confounding which is found in nature

Types of Bias

- 1. Recall Bias
  - Particular problem in case-control studies
- 2. Diagnosis Bias
  - knowledge of E may influence Dx (e.g. BCPs and PE)
- 3. Hawthorne Effect
  - General Electric plant in Hawthorne, NY
  - Productivity tied to ↑ (and ↓) in lighting
  - Called ‘placebo’ effect in medicine; participants and researchers ‘blinded’ to actual treatment status

Selection Biases

"a distortion in the estimate of effect resulting from the manner in which subjects are selected for the study"(KKM)

- Detection Bias – differential surveillance based on exposure status
  - Surveillance Bias (Schlesselman) – BCPs and endometrial CA (Feinstien)
  - Greater in 'milder' diseases picked up on routine visits

Selection Biases

- Loss to Follow up (Non-response Bias)
  - Cohort Studies
  - Compliant participants tend to be healthier
- Healthy Worker Bias
  - Even 23 years after d/c soldiers healthier
  - Caution comparing work cohorts to general population
- Volunteer bias
  - ↓ smokers, ↑ exercise,

HRT ➔ CAD Controversy

- Observational Studies
  - HRT Protective for CAD
    - Tended to be studies of volunteer worker cohorts
- Randomized Trials
  - Slight increase in risk

Incidence-Prevalence Bias

- Incidence – all **new** cases of disease in a time period
  - Tend to be acute
- Prevalence – **existing** cases of disease at one point in time
  - Tend to be chronic
- Cross-sectional studies tend to pick up chronic cases

Direction of Incidence-Prevalence Bias Depends on Population

- Hospital-based study of depression
  - Systematically miss patients who improved (or committed suicide)
- In-patient study of MI patients
  - Systematically miss sudden deaths and those successfully thrombolysed and released
- Studies of schizophrenia
  - Bias can be in either direction. Prognosis fairly bright (60-80% go on to productive lives) if based on outpatient population; fairly grim if based on in-patient population (DSM)

Does public assistance breed dependency?

	1-2 yrs	3-7 yrs	>7 yrs
% who have ever received AFDC	30%	40%	30%
% receiving AFDC at particular time	7%	28%	65%

Long-term recipients more likely to be picked up in a cross-sectional survey

Berkson's Bias: A selection bias due to differing rates of hospitalization

TYPE OF CANCER				
VAGINAL BLEEDING		Endometrial	Other	
	Yes	100	100	200
	No	900	900	1800
		1000	1000	2000

OR = (100)(900) / (100) (900) = 1.0

In the general population, there is no association between vaginal bleeding and endometrial cancer.

Numbers from Hospital-Based Study

TYPE OF CANCER				
VAGINAL BLEEDING		**Endometrial**	**Other**	
	Yes	**73**	**85**	**158**
	No	**90**	**450**	**540**
		163	**535**	**698**

Probability of admission varies: vag bleed = 70%, endometrial CA = 10%, other Cancer = 50%

Now, OR = (73)9450)/(85)(90) = 4.3

Spurious association

How to address selection biases?

- if a, b, c, d represent selection probabilities for the cells in 2x2 table, ensure ad/bc = 1
- Overestimate:

  ad/bc > 1

	D	d
E	a	b
e	c	d