Causation, Bias, Confounding & Interaction

Pawin Numthavaj M.D.
(pawin.num@mahidol.ac.th)

Section for Clinical Epidemiology
Faculty of Medicine Ramathibodi Hospital
Mahidol University
http://www.ceb-rama.org

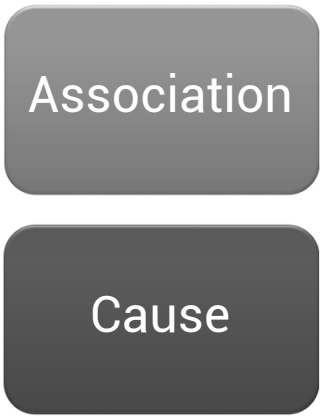
Overview

- Association and cause
  - Causal criteria
- Bias & Confounders
  - Types of Bias
  - Controlling for Bias
- Interaction
  - Types of Interaction

2

ASSOCIATION AND CAUSE

3



4

Association

- An association is a *statistical relationship* between *two or more* events, characteristics or other variables.

5

Association

Alcohol breath test (mg%)	Case (%) N = 43	Control (%) N = 113	OR (95% CI)	P-value
Alcohol breath test				
> 0	11 (44)	9 (8)	8.99 (3.17-25.52)	<0.001
0	14 (56)	103 (92)	1	

6

ตารางที่ 1 ตำแหน่งดาวนับจากลัคนาของผู้ป่วยเบาหวาน โรงพยาบาลบ้านนา ซึ่งทดสอบสมมติฐานว่ามีลัคส่วนมากกว่าค่าเฉลี่ย

ดาว	ตำแหน่งดาว		จำนวนผู้ป่วย		อันดับ	ค่า Z
	นับจากลัคนา	เรียกว่า	(คน)	%		
อาทิตย์	6	-	79	12.36%	1	3.686
	12	-	74	11.58%	2	2.970
จันทร์	9	ตรีโกณ หลัง	66	10.33%	1	1.825
อังคาร	10	จตุโกณ์ ทสเกณฑ์	73	11.42%	1	2.827
พุธ						ไม่ปฏิเสธ Ho
พฤหัสบดี	10	จตุโกณ์ ทสเกณฑ์	77	12.05%	1	3.399
ศุกร์	6	-	72	11.27%	1	2.684
เสาร์						ไม่ปฏิเสธ Ho
ราหู	1	กุมลัคนา	70	10.95%	1	2.397
เกตุ						ไม่ปฏิเสธ Ho
มฤตยู	10	จตุโกณ์ ทสเกณฑ์	69	10.80%	1	2.254
	11	โยคหลัง	67	10.49%	2	1.968

หมายเหตุ : Ho = ไม่แตกต่างจากค่าเฉลี่ย
Ha = แตกต่างจากค่าเฉลี่ย
ค่า Z ที่ = .05 เท่ากับ +/-1.645

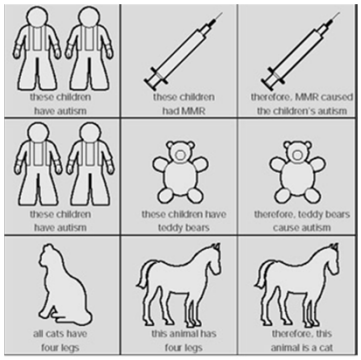
7

Possibility of association

- Artifact from bias ?
- Chance ?
- Confounding ?
- Causation ?

8

Cause ?



9

Association ≠ Causal Relationship

10

Definition

Association

- An association is a *statistical relationship* between *two or more* events, characteristics or other variables.

Cause

- A cause of disease event is an *antecedent event, condition, or characteristic* that was *necessary for the occurrence of the disease* at the moment it occurred, given that other conditions are fixed.

Kenneth J. Rothman & Sander Greenland

11

- Both suspected cause and effect must be associated if they are causally related, but not all associations are causal.
- When statistical associations emerge from clinical research, they do not necessary imply causal associations.

12

Web of Causation

Fig. 3. Relationships between obesity, OSA, disease mechanisms, and disease conditions. Solid lines show recognized associations, and dashed lines show possible associations.

13

Association to Cause?

14

Koch's Postulate

Robert Koch, Discovered the Tubercle Bacillus, 1882

1. The specific organism should be shown to be *present in all cases of animals suffering* from a specific disease but should *not be found in healthy* animals.
2. The specific microorganism should be isolated from the diseased animal and grown in pure culture on artificial laboratory media.
3. This freshly isolated microorganism, when *inoculated into a healthy* laboratory animal, should *cause the same disease* seen in the original animal.
4. The microorganism should be reisolated in pure culture from the *experimental infection*.

15

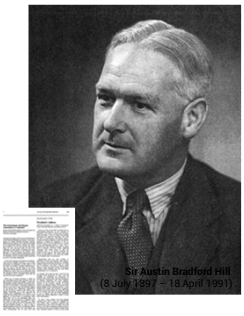
But Koch's postulate is for infectious diseases.

What about non-infectious diseases?

16

Hill's Causal Criteria

- Proposed by Hill, AB
- 1965 paper
- 9 elements that support the strength of association to be "cause"



Sir Austin Bradford Hill (8 July 1897 – 18 April 1991)

17

1. Strength

- More strength – more possibility of cause
- Relative Risk / Odds Ratio
- Cholera epidemic (Snow 1855)
  - Polluted water supply by Southwark and Vauxhall Company: 71 death/10,000 house
  - Lambeth Company: 5 death/10,000 house

18

2. Consistency

- Has the observed association been repeatedly observed by different persons, in different places, circumstances and times?

19

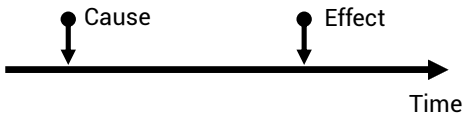
3. Specificity

- If the association is limited to specific workers (patients) and to particular sites (area) and types of disease and there is no association between the work and other modes of dying, then clearly there is a strong argument in favor of causation
- *Does exposure lead only to outcome?*
- Weak criterion: some exposure (Ex. smoking) leads to multiple outcomes

20

4. Temporality

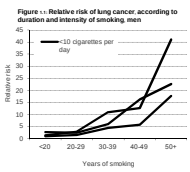
- *Time Order*
- Exposure always precede outcome
- Essential criterion
- Study design: Cohort, RCT, Case-control



21

5. Biological Gradient

- Dose-response curve
- Rate of lung cancer rises linearly with number and duration of of cigarettes smoked daily



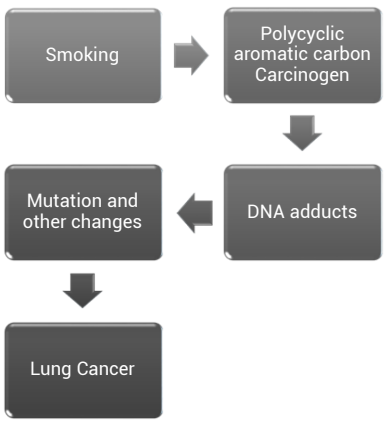
http://www.cancerresearchuk.org/cancer-info/cancerstats/causes/lifestyle/tobacco/tobacco-and-cancer-risk

22

6. Biological Plausibility

- Biologically plausible
- *Does the association make sense?*
- Depend on our lack of knowledge in the field

23



24

7. Coherence

- "Internal consistency"
- Is the association consistent with other available evidence?
- Consistently evident within subgroups
- If smoking cause lung cancer
  - It should cause in men, women, those finished high school, those finished university

25

8. Experiment

- Has a RCT been done?

26

9. Analogy

- Is the association similar to others?
- Indirect information
- Ex:
  - Oral contraceptive cause MI in smoker
  - Oral contraceptive cause stroke

27

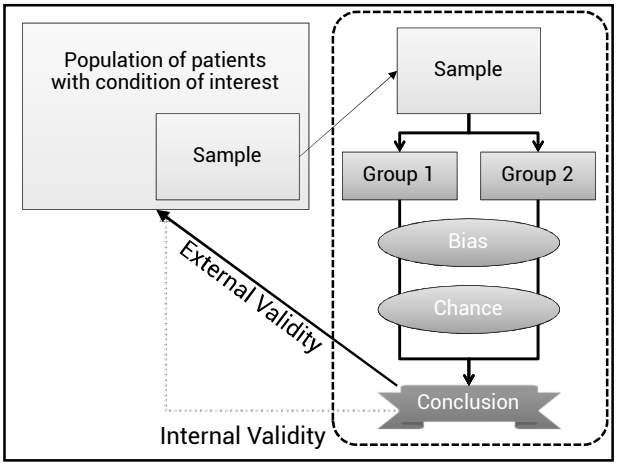
BIAS AND CONFOUNDER

28

Bias

- "Deviation from the truth"
- Undermines ***internal validity*** of research

29



Classifications of Bias

- Sackett: 35 different classifications
- Feinstein:
  - *Susceptibility bias* (difference in baseline)
  - *Performance bias* (different proficiencies of treatment)
  - *Detection bias* (different measurements of outcome)
  - *Transfer bias* (differential losses to follow-up)

31

3 Types of Bias

1. Selection bias
2. Information bias
3. Confounding

32

1. Selection Bias

- Are the groups similar in all important respects?
- **Membership bias**
  - Member of group differ from others
  - Ex: Jogging to prevent MI: Joggers may be different in smoking history, diet

33

- **Berkson bias**
  - Different *rates of admission* for case and control
  - Ex: case control study in hospital setting with higher rate of admission of salphingitis with IUD device use
- **Neyman bias**
  - Disease that is *quickly fatal*
  - Ex: MI and snow shoveling: more cases died at the site and never reach hospital – so less case in the hospital

34

- **Unmasking bias**
  - Exposure unmasked the outcome
  - Ex: estrogen therapy cause symptomless cancer to bleed, thus showing sign of cancer even if it is already there
- **Non-respondent bias**
  - Nonresponders are different than responders
  - Ex: Smokers are less likely to return questionnaires

35

Controlling Selection Bias

- Define criteria of selection of case and control independent of exposure in case-control study
- Define criteria of selection of exposed and non-exposed independent of disease outcome in cohort study
- Use RCT

36

2. Information bias

- Has information been gathered in the same way?
- Synonym: "*Observation*" "*Classification*" "*Measurement*" bias
- Outcomes should be obtained in the same way for exposed and unexposed

37

- **Ascertainment bias**
  - Gathering information in different ways
  - Ex: bedside interview for case but telephone interview for control
- **Diagnostic suspicion bias**
  - Knowledge of exposure leads to intensive search for disease
  - Ex: IV drug users are likely to have HIV tested

38

- **Recall bias**
  - Cases tends to remember better than healthy controls
  - Ex: breast cancer patients recall their abortion history better than controls

39

Non-differential misclassification

- Information bias might not leads to one direction (increase/decrease RR) but may also leads to obscured difference
- Ex: Ambiguous questionnaires leads to error in data collection in both case and control → found no association

40

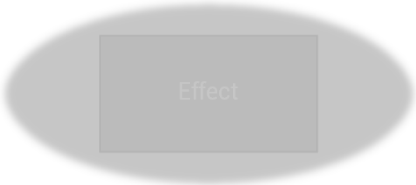
Controlling Information Bias

- Standardized protocol for data collection
- Methods of data collection similar
- Interviewers and study personnel are unaware of exposure/disease status
- Strategy to assess potential information bias

41

3. Confounding

- Is extraneous factor blurring the effect?



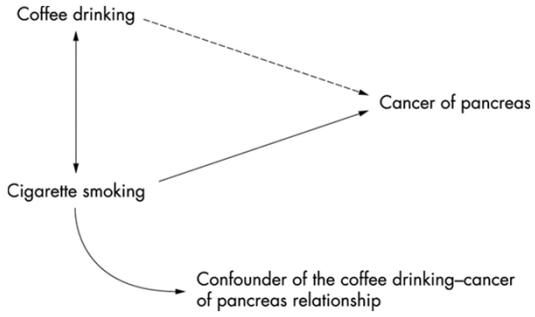
42

Confounding Variable

- Associated with the exposure
- Affects the outcome
- But not an intermediate link in the chain of causation between exposure and outcome

43

Cigarette smoking as a confounder of the coffee drinking-cancer of pancreas relationship.



Zaccai J H Postgrad Med J 2004;80:140-147

PMJ 44

Control for confounding

- Confounding can be corrected if it was anticipated and requisite information gathered
- Possible approaches:
  - Restriction
  - Matching
  - Stratification

45

- **Restriction**
  - Enroll only people without confounder
  - Ex: Enroll only non-smokers
  - Cons:
    - Hinders recruitment
    - Cannot extrapolate to those with confounder e.g. cannot extrapolate to smokers (poorer external validity

46

- **Matching**
  - Match case and control with smoking status
  - Cons
    - Recruitment
    - Cannot examine effect of matched variables
- **Stratification**
  - Post hoc analysis
  - Results are calculated separately for smokers and non-smokers to see if the same effect arises independent of smoking

47

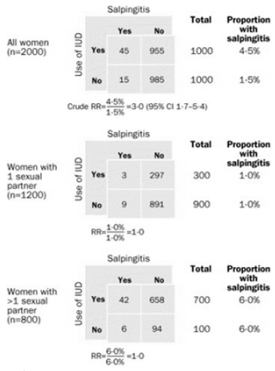
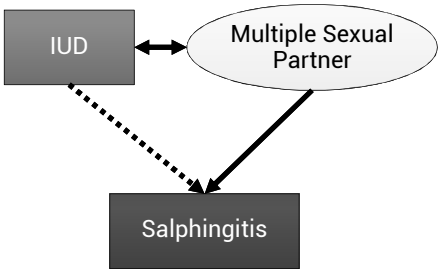


Figure 1. Example of confounding in a hypothetical cohort study of intrauterine device use and salpingitis. When the crude relative risk is controlled for the confounding effect of number of sexual partners, the raised risk disappears.

David A Grimes , Kenneth F Schulz **Bias and causal associations in observational research** The Lancet Volume 359, Issue 9302 2002 248 – 252 http://dx.doi.org/10.1016/S0140-6736(02)07451-2

48



49

- **Multivariate techniques**
  - Mathematical modeling
  - Similar to stratification but can control for multiple variables

50

INTERACTION

51

Interaction

- Effect between *2 risk factors* and *outcome*
- Effect of one risk factor is different within strata defined by the other

52

- Effect of aflatoxin in chronic hepatitis B on the development of liver cancer
  - RR for CA liver from HBV = 7.3
  - RR for CA liver from aflatoxin exposure = 3.4
  - RR for CA liver from both = 59.4!

53

Types of Interaction

- If the observed risk for having both A and B is equal to the expected, then there is **no interaction**
- If the observed risk for having both A and B is greater than expected risk, then there is **synergistic interaction**
- If the observed risk for having both A and B is less than expected risk, then there is **antagonistic interaction**

54

Relative Risk		Factor A	
		-	+
Factor B	-	RR_{00}	RR_{10}
	+	RR_{01}	RR_{11}

No interaction :	$RR_{11} = RR_{10} \times RR_{01}$
Synergistic Interaction :	$RR_{11} > RR_{10} \times RR_{01}$
Antagonistic interaction :	$RR_{11} < RR_{10} \times RR_{01}$

55

Ex: RR of CA Oral cavity from smoking and alcohol

Relative Risk		Smoking	
		No	Yes
Alcohol	No	1.00	1.53
	Yes	1.23	5.71

Expected RR for both smoking and alcohol
= 1.23×1.53 = 1.88
In reality it is 5.71 (>1.88)
So it is "Synergistic Interaction"

56

Questions?

57