Confounding and Bias in Cohort Studies

Chi-Chuan (Emma) Wang, Ph.D.
Assistant Professor
School of Pharmacy, National Taiwan University
Disclosures

• There is no potential conflict of interest relevant to this presentation

• Many materials in this presentation are adopted from the lectures in previous years. Thanks to Drs. Soko Setoguchi and Tobias Gerhard!
Outline

• Bias vs. Chance

• Bias that might occur in cohort studies
  – Confounding Bias
  – Selection Bias
  – Information Bias

• Summary
Bias and Chance

Total Error

Systematic Error (Bias)
- Unaffected by sample size
- Caused by the systematic differences in the case/control or exposed/unexposed groups
- Internal validity

Random Error (Chance)
- Decreases as the sample size increases
- Confidence intervals, p-values
- Precision
Precision and Validity
Precision and Validity

valid, but imprecise

e.g., True RR = 2.0
Estimated RR = 2.0, 95% CI = (0.5 – 4.0)
Precision and Validity

precise, but invalid

- e.g., True RR= 2.0
- Estimated RR= 3.0, 95% CI= (2.8 – 3.2)
Precision and Validity

invalid, and imprecise

e.g., True RR= 2.0
Estimated RR= 3.0, 95% CI= (1.0 – 4.0)
Precision and Validity

precise and valid

e.g., True RR= 2.0
Estimated RR= 2.0, 95% CI= (1.8 – 2.3)
Does the effect detected in your study real?

- Chance?
  - Yes → Not Causal
  - No

- Bias?
  - Yes
  - Not Causal
  - No → Cause?

Statistics  Epidemiology
Bias

• Can occur in all types of studies
  – Particularly in observational studies

• Bias has a direction
  – Bias towards the null
  – Bias away from the null
Types of Bias

• **Confounding**
  – A third factor that distorts the association between exposure and outcome

• **Selection Bias**
  – Due to selection or retention of the study population

• **Information Bias**
  – Measurement errors in exposure, outcome, or confounders
Con founding
Confounding

The quantitative association between exposure and outcome is distorted by a third factor with the following characteristics:

• associated with the exposure
• associated with the outcome
• not an intermediate on the causal pathway between exposure and outcome
Confounding

Confounder

Exposure

Outcome
Confounding - Example

History of heart attack/stroke

Daily low-dose aspirin

Heart attack
Confounding?

On the causal pathway!
This is an intermediary, not a confounder.
Intermediator- Example

Confounding by Indication

• Indication for treatment or disease severity predict the initiation or choice of treatments
• Indication for treatment and disease severity are associated with the outcome of interest
Confounding by Indication - Example

- Depression
  - SSRI
  - Suicide
Addressing Confounding

- Carefully select your comparator!
  - Know your study population and treatment well
- Confounding can be measured or unmeasured
Addressing Confounding

Confounders

- Measured Confounders
  - Design
    - Restriction
    - Matching
  - Analysis
    - Standardization
    - Stratification
    - Multivariate regression
    - Marginal Structural Models
- Unmeasured Confounders
  - Unmeasured, but measurable in substudy
    - 2-stage sample
    - Ext. adjustment
    - Imputation
    - Propensity score calibration
  - Unmeasurable
    - Design
    - Analysis
      - Cross-over
      - Choice of comparison group (active comparison)
      - Instrumental variable?

Schneeweiss PDS 2006, modified by Setoguchi
Selection Bias
Selection Bias

- Distortions that result from procedures used to select subjects and from factors that influence participation/retention in the study

- In cohort studies
  - Selection of exposure and non-exposure group was affected by the risk of the outcome
  - In pharmacoepidemiology study
    - Prevalent user bias
Prevalent User Bias

• Those who develop outcomes stop taking the drug
  – Survival bias; immortal person time

• Prevalent users tend to be healthy adherers and those that benefit from treatment
  – healthy user effect

• Inclusion of prevalent users will oversampling of subjects / person time at low risk
  → underestimation of harms and overestimation of benefits

Solution → New user design
Information Bias
Information Bias

• Measurement of classification errors in exposure, outcome, or confounders
  – Particularly problematic when using secondary data

• Two types of information bias
  – **Non-differential**
    - Misclassification between groups is approximately equal
  – **Differential**
    - Amount of misclassification differs between groups

• More details in the “Confounding and Bias in Case-Control Studies”
Time-Lag Bias
Time-Lag Bias

- Confounding by disease duration and latency time

Suissa and Azoulay, Diabetes Care, 2012
In summary...

• **Best remedy for bias is prevention!**

• **RCTs**
  – Randomization
  – Blinding
  – Primary data collection

• **Observational Studies**
  – Sample selection
  – Choice of comparator
  – Use validated measures
  – Statistical analysis
Thank you

Chi-Chuan (Emma) Wang
chicwang@ntu.edu.tw