Introduction

Arvid Sjölander

Department of Medical Epidemiology and Biostatistics
Karolinska Institutet

Causal Inference for Epidemiological Research

Definition of epidemiology

*Epidemiology is the science that studies the patterns, **causes, and effects** of health and disease conditions in defined populations.*

Wikipedia, 2017

Statistical association

- Any attempt to discover a causal effect often starts by observing a **statistical association**
- A 'statistical association' between two factors means that they 'tend to appear together'
  - lung cancer is more common among smokers than among non-smokers
  - sickness is more common in hospitals than outside hospitals

Association vs causation

- However, **association does not imply a causation**
- *Apart from a true causal effect, what could possibly explain the association between*
  - *smoking and lung cancer?*
  - *hospitals and sickness?*

The Bradford Hill criteria for causation (1965)

- Strength of association
- Consistency
- Specificity
- Temporality
- Dose-response relationship
- Plausibility
- Coherence
- Experimental evidence
- Analogy

Definition of causal effect

- The Bradford Hill criteria list of (presumably) common features of causal effects
- The criteria do not **define** a causal effect
  - a causal effect may satisfy all or none of the criteria (apart from temporality, which is necessary)
- So what is the definition of a 'causal effect'?

Brief history of causal inference, 70's

- Donald Rubin developed a formal definition of causal effects
  - **potential outcomes**
  - **counterfactuals**

Brief history of causal inference, 80's

- James Robins developed statistical methods for estimating causal effects in longitudinal studies
  - **Marginal Structural Models** (MSMs)
  - **Structural Nested Models** (SNMs)

Brief history of causal inference, 90's

- Judea Pearl developed **Directed Acyclic Graphs** (DAGs)
  - simplify interpretation and communication in causal inference
  - useful for covariate selection in observational studies

Traditional methods vs causal inference methods

- In recent years, many sophisticated analysis methods have been developed for causal inference
  - e.g. Marginal Structural Models, propensity scores, Instrumental variables
- Sometimes, these novel methods are superior to more traditional methods (e.g. logistic regression)
- Often, however, traditional methods works perfectly fine, from a causal inference perspective

Main aims of the course

- Provide a deeper understanding of what a causal effect is, and how it differs from a statistical association
- Provide a deeper understanding of when and why traditional methods succeed/fail to estimate causal effects
- Provide some knowledge about more recent causal inference methods

Outline

- Association vs causation
- Exchangeability
- Directed Acyclic Graphs (DAGs)
- Marginal causal effects
- Instrumental variables and Mendelian randomization