Introduction

Arvid Sjölander

Department of Medical Epidemiology and Biostatistics
Karolinska Institutet

A short course on concepts and methods in Causal Inference

The exposure and the outcome

- Many epidemiological research questions are centered around a particular **exposure** and a particular **outcome**
- A few examples:
  - Does your diet (exposure) affect your risk for breast cancer (outcome)?
  - Is the risk for autism (outcome) bigger in IVF pregnancies (exposure), as compared to 'ordinary' pregnancies?
  - Do antibiotics (exposure) cause asthma (outcome)?
  - Is the risk for sudden infant death (outcome) bigger if the baby sleeps on the front (exposure) than if the baby sleeps on the back?

Association

- We often want to learn if there is an **association** between the exposure and the outcome
  - i.e. do the exposure and the outcome tend to 'appear together' in the study population?
- E.g. 'is asthma more common in children who frequently use antibiotics, than in other children?'

Statistics and association

- Statistics is branch of science that deals with association
- Using statistics, we can formally (i.e. mathematically) define and quantify association
- Common statistical measures of association:
  - correlation coefficients
  - regression coefficients
  - risk ratios
  - odds ratios
  - hazard ratios
  - etc

Causation

- However, the goal is often more ambitious
- Ultimately, we often want to learn to what extent the exposure **causes** the outcome
  - E.g. 'Do antibiotics cause asthma?'

Association is not equal to causation

- In observational studies the exposure and outcome may be associated, even in the absence of a causal effect

Statistics and causation

- For most of the 20th century, causation was largely ignored in statistics
- In fact, causation cannot even be defined with 'traditional' statistics language
- For instance, the 'associational' risk ratio

$$\frac{\Pr(Y = 1|A = 1)}{\Pr(Y = 1|A = 0)}$$

  cannot in general be given a causal interpretation
- *But what does the 'causal' risk ratio look like?*

Brief history of causal inference, 70's

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

Brief history of causal inference, 80's

- James Robins discovered - and solved - some important problems with longitudinal studies, from a causal inference perspective
  - **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

Before we start...

- Causal inference has been an intense research field over the last 20 years
  - Countless papers and several books
- This is a brief introduction course
  - We will only have time to scratch the surface