Cross-sectional Study Design

ADA Research Toolkit

ADA Research Committee
Learning Objectives

At the end of this presentation the participant will be able to:

• Define a cross-sectional study and its components
• Describe the study design
• Describe the strengths and limitations of cross-sectional studies
• Identify the main statistical tools used to analyze cross-sectional studies
Introduction

Definition of a cross-sectional study

- A study where exposures and outcomes are observed or measured simultaneously in a population. In this design, a researcher examines the association between the exposure and the outcome, but cannot infer cause and effect. (ADA terms on EAL)
Distinguishing Features

- Prevalence studies
- Exposure and disease status are measured simultaneously
- Measurement occurs at a single point in time or over a short period of time
- In general, they are fast and less resources are needed because there is no follow-up
- Useful for identifying associations
- Repeated cross-sectional studies can be conducted to measure change in a population
- Hypothesis generating
- This study design is exploratory; therefore, associations can be identified, but not cause and effect or the sequence of events
Basic Characteristics

• It is more commonly used than a longitudinal study design
• Subjects stratified into different groups at one time
• Draws conclusions about the population from the sample studied providing the sample is randomly chosen
• Compares the characteristics of those strata and provides an inference
Confounding Effects

- There are the extraneous variables that interfere with the actual results of any cross-sectional study.
Examples of Subject Characteristics

- Age
- Household income
- Education level
- Exposure to information about health
- Historical events that influenced life choices
- Behavioral choices, e.g., diet, smoking etc.
Basic Design

In a cross-sectional study, the researcher

• Selects a sample from the population
• Measures exposure and outcome variables
• For example, identifying the presence or the absence of a risk factor and disease

Sample Study Design

- Exposure; disease
- Exposure; no disease
- No exposure; disease
- No exposure; no disease
Questions Answered

1. Are exposure variables associated with outcome variables?
2. What exposure variables are correlated?
3. Do the findings suggest a more rigorous study is needed?
Generalizability

Depends on:
1. Sample representative of reference population
2. Biologically plausible that outcome variables are related to exposure variables
External Validity

Definition

- The extent to which study findings can be generalized beyond the sample used in the study (Burns and Grove)

External validity is usually limited in regard to generalizing from sample to population is enhanced by good sampling strategies

Internal Validity

Definition

• The extent to which the effects detected in the study are a true reflection of reality, rather than being the result of the effects of extraneous variables (Burns and Grove)
  – Factors that affect internal validity are
    • Selection and measurement bias
    • Blinded data collectors
    • Quality control procedures

Recall Bias

Recall from participants can be faulty due to:

- Desire to report socially desirable behaviors, i.e., eating fruits and vegetables or exercising
- Reluctance to report embarrassing or socially undesirable behaviors (binge eating/alcohol/drug abuse)
- Individuals who have a disease are more likely to remember past exposures than those who do not
- Forgetfulness
- Self-selection of the importance or relevance of a potential exposure
Internal Validity

The internal validity of a cross-sectional survey design is critically determined by

- The researcher’s understanding and inclusion of relevant predictor and confounding variables
- The validity of the measurements
- Sampling
Cross-Sectional Study Design

1. Define the population of interest
2. Sample randomly from the population of interest?
3. The random sample
4. Unexposed (to?)
   - Individuals are presented to an observer
5. Exposed (to?)
6. Identify health event status
7. Evaluate the association between exposure and disease

Bayona M and Olsen C. Young Epidemiology Scholars Program, Observational studies and bias in epidemiology
Advantages

1. Everything done in the present, taking a random sample from the population of interest
2. Can be a pilot for future studies
3. Generates hypotheses for future studies
4. Study multiple outcomes and exposures
5. Can measure prevalence
Disadvantages

1. Cannot make conclusions about cause and effect or sequence of events
2. Everything is measured at one specific time point
3. Not good for rare diseases or those that have short duration or short survival because of the limitation in finding a large number of cases in the population
4. Prone to selection and measurement bias
5. Subject to selective survival bias.
   - For example, if the people that are alive have a different exposure for the disease than those that die early, the sample may inadvertently be enriched for a particular sub-type of the disease.
Typical Statistics

- Cross-sectional study design employs various statistical measures. Some of the important statistical tools are:
  - Odds ratio
  - Multivariate analysis
- Relative prevalence and excess prevalence are the cross-sectional analogs of relative risk and excess risk
  - Chi-square test: compares observed frequencies within categories to frequencies expected by chance

*normal distribution: a symmetrical bell-shaped theoretical distribution that has defined properties
Logistic Regression

• Logistic regression is a statistical method which can be used to estimate the probability of disease for a given covariate pattern.

• The results of logistic regression lead naturally to an estimate of the odds ratio but may also be used to estimate the prevalence ratio, simply by taking the ratio of the estimated probabilities.
Key Points

• The study employs a cross-sectional, sample survey design. The “cross-sectional” component of the design denotes that subjects will be surveyed at one point in time.

• The “sample” component of the design indicates that only a portion of the target population will be surveyed.

• Survey designs in general share three common characteristics
  – A relatively large number of respondents
  – Formal observation involving interviews or questionnaires
  – Statistical analysis of data

• Examines prevalence of disease and exposure at one point in time
Key Points

- Sequential studies can be used to measure change in characteristics (e.g., behavior) over time
- No cause and effect, but associations can be identified
- More commonly used than a longitudinal study design