Introduction to Observational Studies

Deborah Friedman, MD, MPH
University of Texas Southwestern Medical Center
Dallas, Texas

Hierarchy of Observational Studies

Case Report
Case Series
Case-Control Study
Cross-Sectional Study
Cohort Study
Anatomy of a Research Study

<table>
<thead>
<tr>
<th>Element</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research questions</td>
<td>What questions will the study address?</td>
</tr>
<tr>
<td>Significance (background)</td>
<td>Why are these questions important?</td>
</tr>
<tr>
<td>Design (time frame, epidemiologic approach)</td>
<td>How is the study structured?</td>
</tr>
<tr>
<td>Subjects (selection criteria, sampling design)</td>
<td>Who are the subjects and how will they be selected?</td>
</tr>
<tr>
<td>Variables (predictor, confounding and outcome variables)</td>
<td>What measurements will be made?</td>
</tr>
<tr>
<td>Statistical issues (hypothesis, sample size, analytic approach)</td>
<td>How large is the study and how will it be analyzed?</td>
</tr>
</tbody>
</table>

Hulley SB et al. Designing Clinical Research, 2nd edition, Lippincott Williams & Wilkins 2001

What is the Research Question?

Every study must have a primary question (hypothesis) that is carefully selected, clearly defined, and stated in advance

- May want to demonstrate a beneficial outcome
- May want to demonstrate no difference between two interventions (non-superiority)

Should be important and relevant scientifically, medically or for public health purposes
Secondary Questions

Specified before data collection begins
Based on reasonable expectations that the intervention will not
Limited in number
Examples:
  ◦ Subgroup Hypotheses (differences at baseline)
  ◦ Adverse effects
  ◦ Ancillary questions, sub-studies
  ◦ Natural history data and prognostic factors

Types of Studies

• Observational
  Cohort
  Cross-sectional
  Case Control

• Clinical Trial
Clinical Research Designs

Question: Does aspirin therapy prevent NAION?

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Key Feature</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observational Designs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort study (Incidence)</td>
<td>A group followed over time to compare those with or without exposure/intervention</td>
<td>Examine a cohort yearly, observing the incidence of NAION in aspirin vs. non-aspirin users</td>
</tr>
<tr>
<td>Cross-sectional study (Prevalence)</td>
<td>A group examined at one point in time</td>
<td>Examine the cohort at once, observing the prevalence of a history of NAION in ASA users and non-users</td>
</tr>
<tr>
<td>Case-control study (Incidence or prevalence)</td>
<td>Two groups, compare exposures based on the outcome that has occurred</td>
<td>Examine people with NAION (cases) and compare them to a group without NAION (controls) and ask about ASA use</td>
</tr>
<tr>
<td><strong>Experimental Design</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Randomized, masked trial</td>
<td>Two groups created by a random process with a masked intervention</td>
<td>Randomly assign patients to ASA or placebo and follow to observe incidence of NAION</td>
</tr>
</tbody>
</table>

Prospective Cohort Study

**Strengths**
- Defines incidence
- Investigates potential causes of a disease
- Opportunity to measure important variables completely and accurately
- Study antecedents of fatal disease

**Weaknesses**
- Expensive and inefficient way to study rare outcomes

Risk factor present or absent
**Retrospective Cohort Study**

**Strengths**
- Same as prospective cohort studies
- Less costly and time consuming
- Cohort is already assembled, measurements made, follow-up has occurred

**Weaknesses**
- Limitations in sampling
- Nature and quality of data
- Key information may be missing, inaccurate or suboptimally measured

---

**Case-Control Studies**

- Generally retrospective
- Look at subjects with a disease and another without it, and look backward at variables that differ between the groups

<table>
<thead>
<tr>
<th>Risk factor present</th>
<th>Risk factor absent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases with disease</td>
<td></td>
</tr>
<tr>
<td>Controls without disease</td>
<td></td>
</tr>
</tbody>
</table>

Largest population with and without disease
Case-Control Studies

Strengths
- High yield for relatively few subjects
- Good for rare outcomes
- Hypothesis generating

Weaknesses
- Can only study one outcome
- Bias
  - Separate sampling of cases and controls
  - Retrospective measurement of predictor variables
  - Cases may not be representative

Why Cases May Not Be Representative of All Cases

- No medical attention
- Seen elsewhere
- Seen but misdiagnosed
- Death or remission before diagnosis
Nested Case-Control Studies

- A case-control study “nested” in a prospective or retrospective cohort study
- Identify outcome of interest
  - Cases: have developed the outcome
  - Controls: have not developed the outcome (selected randomly from the cohort – may include some cases)
- Analyze records, tests, samples to compare risk

Example of Nested Studies: IIHTT
Advantages and Disadvantages of Major Observational Designs

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Establishes sequences of events; Can study several outcomes; # of outcome events grows over time; Yields incidence, relative risk, excess risk</td>
<td>Requires large sample size; Less feasible for rare outcomes</td>
</tr>
<tr>
<td>Prospective</td>
<td>More control over subject selection and measurements; Avoids bias in measuring predictors</td>
<td>More expensive; Longer duration</td>
</tr>
<tr>
<td>Retrospective</td>
<td>Less expensive; Shorter duration</td>
<td>Less control over subject selection and measurements</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross sectional</td>
<td>May study several outcomes; Relatively short duration; Good first step for cohort study; Yields prevalence, relative prevalence</td>
<td>Does not establish sequence of events; Not feasible for rare predictors or outcomes; Does not yield incidence or relative risk</td>
</tr>
<tr>
<td>Case-Control</td>
<td>Study rare conditions; Short duration; Relatively inexpensive; Relatively small; Yields odds ratio (usually a good approximation of RR), rate ratio or incidence proportion ratio depending on sampling methodology</td>
<td>Potential for bias and confounding from sampling two populations; Does not establish sequence of events; Potential survivor bias; Limited to one outcome variable; Does not yield prevalence, incidence or excess risk</td>
</tr>
<tr>
<td>Nested Case-Control</td>
<td>Advantages of retrospective cohort study but more efficient</td>
<td>May require banked samples or data stored until outcomes occur</td>
</tr>
</tbody>
</table>
Observational Studies Pathway

1. Descriptive Studies
   Get a sense for the “lay of the land”
   Distributions of diseases and characteristics in the population, or
   Sensitivity and specificity of a diagnostic test
   
   How common is daily aspirin treatment in patients over age 50 years?

2. Analytic Studies
   Evaluate relationships to determine cause-and-effect relationship
   
   Is taking daily aspirin over age 50 associated with a lower risk of NAION?
3. Clinical Trial

Prospective study to determine whether daily aspirin treatment in individuals over age 50 is associated with a lower risk of developing NAION.
Statistical Issues
The statistical plan should be determined in advance

Investigator initiated studies: consult a biostatistician from the beginning!

For descriptive studies, estimate the number of subjects needed to produce an acceptable level of precision when confidence intervals are calculated for means, proportions and other descriptive statistics.

Physiology of Research
- Research questions cannot be answered with perfect accuracy (impractical)
- Compromise by using a sample of the population
- Variables may be a proxy (e.g., self-report)
- Errors may occur

Truth in the Universe \(\xrightarrow{\text{infer}}\) Findings in the Study
Types of Errors

Random Errors
Due to chance
Larger sample size increases precision

Systematic Errors
Due to bias (source of variation that distort the study findings in one direction)

Need to try to minimize errors in the study design

Plan for Today

- Epidemiology 101
- Case study 1
- Statistics, part 1
- Break
- Case study 2
- Statistics, part 2
- Group sessions – propose a study