Designing The Right Study

.....Observational & Experimental Designs

Susan G. Fisher, M.S., Ph.D.
Chair, Department of Clinical Sciences

October, 2015
Asking the Right Question

A Good Research Question is:

• Relevant and interesting
• Feasible
• Ethical
• Novel…maybe!
• Well-built
Start with a research question:

Is statin use associated with an increased risk of type II diabetes?

Formulate a hypothesis:

Null Hypothesis \((H_0)\):

There is no association between statin use and risk of diabetes.

Alternate Hypothesis \((H_A)\):

There is an association between statin use and diabetes.
Refining the Research Parameters

Population
- High cholesterol
- Previous cardiac event
- Presence of risk factors of diabetes
- Normal glycated hemoglobin (HbA$_{1C}$)

Independent Variable: statin use
- Specific type/drug
- Dose
- Length of time on drug

Dependent Variable: diabetes
- Onset of newly diagnosed diabetes
- Changes in HbA$_{1C}$
Strength of Evidence

<table>
<thead>
<tr>
<th>Strength</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>Case Report</td>
</tr>
<tr>
<td></td>
<td>Case Series</td>
</tr>
<tr>
<td></td>
<td>Ecological Study</td>
</tr>
<tr>
<td></td>
<td>Cross-sectional Survey</td>
</tr>
<tr>
<td></td>
<td>Case-control Study</td>
</tr>
<tr>
<td>Strong</td>
<td>Cohort Study</td>
</tr>
<tr>
<td></td>
<td>Clinical Trial</td>
</tr>
</tbody>
</table>

- Descriptive
- Analytic
Case Report / Case Series

• Anecdotal Reports of Interesting Observations
  – Unusual cluster of symptoms
  – Departure from a normal pattern of known disease
  – Repetitive disease occurrence among people with a specific exposure

• Cluster of observations in short time period or small geographic area
  – New epidemic of known disease
  – New disease occurrence
  – New cause of existing disease

[Example: Three well-controlled diabetic patients prescribed statins over the last 6 months have unexpected elevations in HbA1C]
Ecologic Studies

• Evaluation of associations between exposures and outcomes in populations rather than individuals

• "Ecological Fallacy"
  - results from making causal inferences about individual phenomena based on observations of group
Cross-Sectional Studies

- Provide “snapshots” of the health of a specified population at one moment in time.
- Usually descriptive in nature
- Often used to determine ‘prevalence’ of a condition or correlation between 2 variables
- Temporality cannot be determined → ‘chicken or egg problem’
- Low cost and no loss to follow-up

[Example: Identify 200 males over age 40; obtain history of statin use and measure their HbA1C level.]
Analytic (Observational) Studies

- Case Control study
- Cohort Study

Exposure, Intervention, or Treatment

Disease or Outcome
Case Control Study

- Select subjects with outcome/disease of interest (Cases)
- Select similar group of individuals without disease/outcome of interest (Controls)
- Determine exposure status of all subjects

<table>
<thead>
<tr>
<th></th>
<th>Cases (Diabetes)</th>
<th>Controls (No Diabetes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposed (Statins)</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Unexposed (No Statins)</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>Total</td>
<td>a + c</td>
<td>b + d</td>
</tr>
</tbody>
</table>
Case Control Study Advantages

• Quick and easy
• Able to study multiple risk factors simultaneously
• Efficient for rare diseases
• Requires ‘small-ish’ sample sizes
Case Control Study Disadvantages

• Cannot address causality
• Only investigates 1 disease outcome
• Can only compare odds of exposure; not incidence of outcome
• High, **HIGH** likelihood of bias
Control Sources

- General population controls
- Hospitalized individuals
- Neighborhood residents
- Spouses / relatives/ friends of case
In a case control study, we use the **ODDS RATIO** to estimate the odds of a case being exposed versus the odds of a control being exposed.

**ODDS RATIO (OR) = \frac{AD}{BC}**

<table>
<thead>
<tr>
<th>Case (Disease)</th>
<th>Control (No Disease)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exposure</strong></td>
<td><strong>No Exposure</strong></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

\[
\text{OR} = \frac{\text{Odds of case exposed}}{\text{Odds of control exposed}} = \frac{\frac{A}{C}}{\frac{B}{D}} \quad \text{or} = \frac{AD}{BC}
\]
Interpreting an Odds Ratio

If OR = 1

- Odds of exposure is equal between groups (no association)

If OR > 1

- Odds of exposure is greater in cases than in controls (positive association);

If OR < 1

- Odds of exposure in cases is less than odds of exposure in controls (negative association; possibly protective)
Example of an Odds Ratio

### Role of Statins in Risk of New Onset of Diabetes

<table>
<thead>
<tr>
<th></th>
<th>CASES Diabetes</th>
<th>CONTROLS No Diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statin Use for &gt; 2 yrs (before dx)</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>No Hx of Statin Use</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>90</td>
</tr>
</tbody>
</table>

\[
\text{OR} = \frac{ad}{bc} = \frac{25 \times 80}{50 \times 10} = \frac{2000}{500} = 4.0
\]
Cohort Studies

• Designed to address a specific hypothesis;

• Select a group of subjects exposed to factor of interest and a group not exposed

• OR select a group of subjects and then categorize them by presence or absence of risk / exposure / treatment

• Prospectively follow both the exposed and unexposed group to determine occurrence of outcome of interest
Prospective & Retrospective Cohort Studies

Defined Population

Non-Randomized

Exposed

Non-Exposed

Disease

No Disease

Disease

No Disease

2015

2024

2006

2015
## Cohort Study

Role of Statins in Risk of New Onset of Diabetes

<table>
<thead>
<tr>
<th></th>
<th>Diabetes</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Exposed</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>(Statin Use)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexposed</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>(No Statin Use)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Relative Risk (RR)** = incidence of disease in exposed divided by incidence of disease in the unexposed

\[
RR = \left( \frac{a}{a+b} \right) / \left( \frac{c}{c+d} \right)
\]
Interpreting the Relative Risk of a Disease

If $RR = 1$

- Risk in exposed equal to risk in unexposed (no association)

If $RR > 1$

- Risk in exposed greater than risk in unexposed (positive association);

If $RR < 1$

- Risk in exposed less than risk in unexposed (negative association; possibly protective)
# Cohort Study

## Role of Statins in Risk of New Onset of Diabetes

<table>
<thead>
<tr>
<th>Diabetes</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Exposed</td>
<td>(Statin Use)</td>
</tr>
<tr>
<td>Unexposed</td>
<td>(No Statin Use)</td>
</tr>
</tbody>
</table>

$$RR = \frac{a}{a+b} / \frac{c}{c+d}$$

$$RR = 0.10 / 0.05 = 2.00$$
Advantages of Cohort Studies

• Cases are incident cases and may be more representative of all cases of the disease
• Provides more information on the natural history of a disease
• Incidence rates are available
• Fewer sources of bias
• Temporal relationship between exposure and disease can be established
• Able to study a rare exposure and a common disease
Disadvantages of Cohort Studies

• Duration may be long with difficulty maintaining consistent study methods and staff
• Expensive
• Large population required
• Exposure may not have been measured at baseline or may change
• Rare diseases cannot be studied
When is a Prospective Observational Study the RIGHT Design?

• Good evidence of an association between an exposure and a disease exists;

• Attrition of study population can be minimized;

• Ample funds are available;

• The investigator has a long life-expectancy
Randomized Clinical Trial: What?

- Experimental design to test a specific hypothesis involving a new intervention(s);
- Controlled and randomized;
- Assign a group of subjects to one of two or more interventions;
- Follow subjects prospectively to determine outcome of interest.
Randomized Clinical Trial: When?

- Exposure or treatment of interest is modifiable;
- Individuals are willing to relinquish control;
- Legitimate uncertainty exists about benefit of treatment;
- Health condition and/or outcome is reasonably common or detrimental.
Randomized Clinical Trial: WHY?

• Best method for providing evidence related to direct treatment benefit

• “Clinical equipoise”
Randomized Clinical Trial

- Target Population
- Randomize
- Treatment A
- Treatment B

Outcome:
- Yes
- No
Hallmark #1: Randomization

- Randomization is the process of assigning subjects to different treatments by using a predetermined, random scheme;
- Eliminates bias in treatment assignments;
- Balances known and unknown prognostic factors between treatment groups;
Hallmark #2: Blinding

- Process in which the identity of the treatment being received is unknown to certain individuals.
  - Single blind ➔ patient
  - Double blind ➔ patient & physician
  - Triple blind ➔ patient, physician, & reviewer
Hallmark #3: Validity of Results

• Inclusion criteria provide defined, homogeneous population;
• Treatments/interventions administered with a systematic, planned approach;
• Treatment groups provided similar care and follow-up;
• Outcomes/endpoints are defined and objectively assessed;
• Statistical analyses carefully planned *a priori.*
Numerous exclusion criteria leads to decreased generalizability;

Lack of treatment choice, inflexible schedule lead to decreased accrual;

Expensive & lengthy;

Measurement of medical endpoints rather than patient-centered outcomes.
Randomized Clinical Trials

- Designed to provide best available care to patients;
- Maximize patient safety;
- Optimize data integrity;
- Minimize study bias;
- Provide compelling evidence of treatment efficacy.
If Research Were So Easy, EVERYONE would do it!