Fundamentals of Program Evaluation
Course 380.611

Experimental, Non-experimental and Quasi-Experimental Designs
Topics to cover:

- Study designs to measure impact
- Non-experimental designs
- Experimental designs
- Quasi-experimental designs
- Observational studies with advanced multivariate analysis
Which study designs (next class) control to threats to validity?

- Experimental (randomized control trials)
  - “gold standard” for assessing impact
  - Controls threat to validity
- Quasi-experimental
  - Controls some threats to validity
- Non experimental
  - Do not control threats to validity
- Observational with statistical controls:
  - Controls some threats to validity
Notation in study designs

- **X** = intervention or program
- **O** = observation (data collection point)
- **RA** = random allocation (assignment)
- **RS** = random selection

How would we interpret?

1) O X O
2) O O O
Source of data for “O”

- Routine health information (service statistics)
- Survey data
- Other types of quantitative data collection
Non-experimental designs

- Post-test only
- Pretest post-test
- Static group comparison

Sources:
- Fisher and Foreit, 2002
- Cook & Campbell, Campbell & Stanley
Posttest-Only Design

Experimental Group

Time

X O₁
Example: % of youth that used condom at last sex

Note: “X” is the campaign

What are the threats to validity?
Pretest-Posttest Design

Experimental Group

Time

O₂  X  O₁
Example: % of youth that used condoms at last sex

Exp. 3% X 15%

Threats to validity?
Static-Group Comparison

Experimental Group

Comparison Group

Time

X  O₁

O₂
Example: % of youth that used condoms at last sex

Exp. group X 15%

Control group 5%

Threats to validity?
Questions

- Do evaluators ever use non-experimental designs?
  - If so, why?
Experimental designs

- Pretest post-test control group design
- Post-test only control group design

Randomized controlled trials (RCTs):
- Are a subset of experimental designs
- Stay tuned for Ron Gray’s lecture (Lecture 12)
Pretest-Posttest Control Group Design

Experimental Group

Control Group

Time

O₁ X O₂

O₃  O₄
Example: % of youth that used condom at last sex

<table>
<thead>
<tr>
<th></th>
<th>Time 1</th>
<th>Time 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp</td>
<td>10%</td>
<td>X</td>
</tr>
<tr>
<td>Control</td>
<td>11%</td>
<td>12%</td>
</tr>
</tbody>
</table>

Note: Youth aren’t randomized to use/not use condoms. This design implies that they are randomly allocated to a group that RECEIVES/DOES NOT RECEIVE the program.
Posttest-Only Control Group Design

- Experimental Group
- Control Group

Time: X \( \text{O}_1 \)

\( \text{O}_2 \)
Example: % of pregnant women who deliver with skilled attendant

Note: Women aren’t randomized to deliver w/w-out. This design implies that they are randomly allocated to a group that RECEIVES/DOES NOT RECEIVE the program.
Although practical difficulties in conducting research increase with design complexity, experiments that examine the multiple treatments are frequently used in OR studies. One advantage is that using multiple treatment designs increases the available alternatives.
Limitations of the experimental design

- Difficult to use with full-coverage programs
- Generalizability (external validity) is low
- Politically unpopular
- (In some cases) unethical
Quasi-experimental designs

- When randomization is unfeasible or unethical
- Treatment must still be manipulable
- Threats to validity must be explicitly controlled
Three most common quasi-experimental designs:

- Time series
- Pretest post-test non-equivalent control group design
- Separate sample pre-test post-test design
Quasi-experimental designs: Time Series

- Multiple observations before and after X
- Can be used prospectively or retrospectively
Time Series Design

Experimental Group

\[ O_1 \ O_2 \ O_3 \ X \ O_4 \ O_5 \ O_6 \]
Example 1 of Time Series

Sudden increase at program intervention

Millions of Condoms Distributed

Program Intervention (X)
Example 2 of Time Series

Steady increase regardless of intervention

Millions of Condoms Distributed

Program Intervention (X)
Example 3 of Time Series

Increases and decreases before & after intervention

![Graph showing millions of condoms distributed over years (2001-2006)]

- Y-axis: Millions of Condoms Distributed
- X-axis: Years (2001-2006)
- Program Intervention (X)

Graph indicates fluctuations in condom distribution before and after an intervention.
Example 4 of Time Series

Temporary impact of intervention

- Millions of Condoms Distributed
- Program Intervention (X)
Using service statistics to demonstrate probable effect

- **Advantages**
  - Lowers costs
  - Takes advantage of existing information
  - Constitutes a natural experiment

- **Disadvantages**
  - Falls short of demonstrating causality
  - Can’t know “what would have happened in the absence of the intervention”
Time Series Analysis: Condom Sales in Ghana

Sales & distribution figures from MOH, GSMF, and PPAG

Program Intervention (February 2000)
Clinic visits in Nepal Before & After Radio Programming

Radio Spots on Air

Radio Serial on Air

No Intervention

Average Monthly Clinic Visits
Example: Mass Media

Vasectomy Promotion in Brazil

- Reasons that it’s a good case study:
  - Topic is almost exclusive to the media campaign
    - Little “naturally occurring” communication on subject
  - Available service statistics
    - Routinely collected
    - High quality
Key media events that promoted Pro-Pater

- 1983:
  - 3 minute broadcast on vasectomy and Pro-Pater
  - Resulted in 50% increase in # vasectomies

- 1985:
  - 10 week newspaper and magazine promotion (with Population Council)
  - 54% increase in # vasectomies
1989 campaign

- Vasectomy promotion in 3 cities (São Paulo, Salvador and Fortaleza)

Objective of the campaign:
- Increase knowledge/awareness (and eliminate misconceptions)
- Increase # vasectomies among lower middle class men aged 25-49
4 phases of the 1989-90 campaign

- (1) Pre-campaign P.R.
  - Press releases, press conference

- (2) Campaign- “vasectomy is an act of love”
  - TV spots (May, June 1989)
  - Radio
  - Pamphlets, billboard, magazine ads
4 phases of the 1989-90 campaign

(3) Rebroadcasting of TV spots
- September 1989
- 2-5 times daily in evenings

(4) Mini-campaign: (Jan-March 1990)
- Ad in *Veja* magazine
- Electronic billboard in San Paulo
- Direct mailing of pamphlet to *Veja* subscribers
Mean number of daily calls to PRO-Pater clinic: 1989-90
Results of Brazil Study

Number of vasectomies

- Poisson regression: measured slopes based on monthly points
- Immediate/substantial increase after start of the campaign
- Gradual decline through 1990
- After last campaign, downward trend became even greater
Possible hypotheses

- Cost of operation increased
- Competition from other doctors (that Pro-Pater had trained) increased
Take-home messages re: time series

- “What was the impact of the program?”
- Combining service stats/special studies
  - Useful to understanding program dynamics
- Time series (quasi-experimental design) can’t definitively demonstrate causality
- “What would have happened in the absence of the program”???
Non-Equivalent Control Group

Experimental Group

Control Group

Time

$O_1 \times O_2$

$O_3 \quad O_4$
Example: vasectomy promotion in Guatemala (1983-84)

- 3 communication strategies
  - Radio and promoter
  - Radio only
  - Promoter only

- One community per strategy (all on Southern coast)

- Before/after survey
  - 400 men per community, each survey
## Results of Vasectomy Promotion:
% interested in vasectomy

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radio &amp; Promoter</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>Radio only</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Promoter only</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>Comparison</td>
<td>23</td>
<td>27</td>
</tr>
</tbody>
</table>
Results of Vasectomy promotion: % operated

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radio &amp; Promoter</td>
<td>0.5</td>
<td>0.8</td>
</tr>
<tr>
<td>Radio only</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Promoter only</td>
<td>1.0</td>
<td>3.0*</td>
</tr>
<tr>
<td>Comparison</td>
<td>0.3</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Separate Sample Pretest-Posttest Design

- Target population
  - RS
  - $O_1 \times X$

- Target population
  - RS
  - $X \times O_2$

Time

- $O_1$
- $O_2$
Findings from Mayan Birthspacing Project
Alternative to experimental designs

Application of multi-level multi-variate statistical analysis to “observational” (cross sectional) data.
Example: Measuring the effects of exposure (dose effect)

- Did increased exposure to Stop Aids Love Life in Ghana relate to desired outcomes?
  - Controlling for socio-economic factors and access to media
% aware of at least one way to avoid AIDS by level of exposure

![Bar chart showing percentage awareness of ways to avoid AIDS by level of exposure for males and females.](image-url)
% believe their friends approve condoms (HIV/Al DS)
ABC’s: Abstinence
Mean age at first sex

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>19.5</td>
<td>17.5</td>
</tr>
<tr>
<td>Low</td>
<td>20.5</td>
<td>18.5</td>
</tr>
<tr>
<td>High</td>
<td>21</td>
<td>19</td>
</tr>
</tbody>
</table>
ABC’s: Being faithful % with 1 partner in past year
ABC’s: Condom use
% used condom at last sex

None
Low
High
Example of estimating the effects of exposure to media

- Example from Tsha Tsha (South Africa)

- Using the data related to exposure to the communication program as a determinant of outcome

- (Credits to CADRE in S. Africa, and to the authors: Kevin Kelly, Warren Parker, and Larry Kincaid)
Perceived Qualities of Boniswa among Females

- Concerned: 69
- Honest: 62
- Courageous: 62
- Confident: 68
- Loyal: 61

N=269
Example: relating outcome to identification with characters

- Does identification with specific characters in a soap opera affect outcomes?
- Controlling for SES
Multiple regression to estimate the independent effect of exposure, controlling for other influences *

* Including lagged attitude means that the impact of other variables is on change in attitude.
Fundamentals of Program Evaluation

Path model: Effects of recall of drama and identification with Boniswa on AIDS Attitudes

- Watches “Days of Our Lives”
- Abstained for a month or more
- Learned about AIDS on TV
- Female
- Lagged AIDS attitude
- Education
- Frequency of TV viewing
- Kwazulu province

Identification with Boniswa

Recall of the Drama

AIDS Attitude

Wave 3; South Africa, 2004
Observational studies with statistical controls

- Widespread use, esp. among academics

- Advantages
  - Doesn’t require an experimental design
  - Allows greater understanding of the pathways to change (tests conceptual framework)

- Limitations:
  - Difficult to rule out confounding factors
Why are the Victoria et al and Habicht et al articles important?

- Challenge RCT as the gold standard for evaluating ALL interventions
- Suggest different levels of evidence:
  - Probability (RCTs)
  - Plausibility (includes comparison group and addresses confounding factors)
  - Adequacy (based on trends in expected direction following intervention)
Reasons why RCTs may have low external validity (Victora)

- Causal link between intervention and outcome can vary depending on external factors:
  - Actual dose delivered to the target population varies
    - Institutional, provider, recipient behaviors
  - Dose-response relationship varies by site:
    - Nutritional interventions (health of target pop)
Trade-off between internal and external validity

- Example: evaluating communication campaign

- Pilot projects: don’t deliver the full impact of a national full coverage program
  - Results not generalizable

- Mass media: lower internal validity
  - Can’t control for all confounding factors
Take home message from Victora et al./Habicht et al.

- RCTs are the gold standard:
  - But may not be appropriate for evaluating large scale public health interventions

- Evidence-based public health must draw on studies with designs other than RCTs
  - Plausibility
  - Adequacy