Causal versus Casual Inference

What Happens When I Take This Medication?

Stephen J. Ruberg, PhD
President
Analytix Thinking, LLC
AnalytixThinking@gmail.com
ACKNOWLEDGEMENTS

**Eli Lilly**
Yongming Qu
Haoda Fu

**Sanofi**
Junxiang Luo (work completed while at Lilly)
Outline

Motivation
Tripartite Approach
Estimators for Causal Inference
Example
Discussion
Climbing Mt Kilimanjaro

How long is the hike on Mt. Kilimanjaro?

19,341 feet

On day 6, hikers take on average 4.65 hours.
How long is the hike on Mt. Kilimanjaro?

19,341 feet
“Intent to hike” estimate? (4.65 hrs)
Completers/adherers estimate? (7 hrs)
The whole story? (all three parts)
What does the traveler (i.e. patient) want to know?
WHAT DO YOU WANT TO KNOW?
What would you tell your loved one?
What are the Right Questions?

Patient / Physician
What happens when I take this medication?

Researcher
What are the causal effects of treatment?

Regulator
What are the benefits and risks of treatment?
A.3 ESTIMANDS

A.3.1 Description

“A central question for drug development and licensing is to quantify treatment effects: how the outcome of treatment compares to what would have happened to the same subjects under different treatment conditions (e.g. had they not received the treatment or had they received a different treatment).”
A.3 ESTIMANDS

A.3.1 Description

“A central question for drug development and licensing is to quantify treatment effects: how the outcome of treatment compares to what would have happened to the same subjects under different treatment conditions (e.g. had they not received the treatment or had they received a different treatment).”
A.3 ESTIMANDS

A.3.1 Description

“A central question for drug development and licensing is to quantify treatment effects: how the outcome of treatment compares to what would have happened to the same subjects under different treatment conditions (e.g. had they not received the treatment or had they received a different treatment).”
Treatment Effect

Test Treatment $T(1)$

Control Treatment $T(0)$

Test Response $Y(1)$

Control Response $Y(0)$

Treatment Effect = $Y(1) - Y(0)$
Treatment Effect

$\text{Test Treatment } T(1)$

$\text{Control Treatment } T(0)$

$\text{Test Response } Y(1)$

$\text{Control Response } Y(0)$

Estimator $= \frac{\sum [Y_i(1) - Y_i(0)]}{N}$
Outline

Motivation

Tripartite Approach

Estimators for Causal Inference

Example

Discussion
The Tripartite Approach*

Three Causal (and clinically meaningful) Estimands

1. The proportion of patients that discontinue treatment due to adverse effects
   - Can also assess time to discontinuation

2. The proportion of patients that discontinue treatment due to lack of efficacy
   - Need to assess time to discontinuation

3. For those who could adhere to their treatment, what is the treatment difference for the primary efficacy response
   - Must assess safety in this group as well

Estimands (1) and (2) are pretty easy
- All randomized patients provide a response
- Categorical or survival analysis
- Can include models for important covariates
- Could consider competing risks as a more complex assessment

Estimand (3) is more difficult
- Patients are self-selected (i.e. non-randomized)
- *Akin* to an observational study
- *Casual* inference approach – Completers Analysis
- *Causal* inference approach – Counterfactual / IPW
Outline

Motivation
Tripartite Approach

**Estimators for Causal Inference**

Example
Discussion
### Causal Inference

For T(0) [CONTROL]
- A(0) = 0 for non-adherence
- A(0) = 1 for adherence

For T(1) [EXPERIMENTAL]
- A(1) = 0 for non-adherence
- A(1) = 1 for adherence

<table>
<thead>
<tr>
<th>Control Treatment Adherence</th>
<th>A(1) = 0</th>
<th>A(1) = 1</th>
<th>A(1) ∈ {0, 1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>A(0) = 0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>A(0) = 1</td>
<td>NA</td>
<td>A++</td>
<td>A++</td>
</tr>
<tr>
<td>A(0) ∈ {0, 1}</td>
<td>NA</td>
<td>A**</td>
<td>A**</td>
</tr>
</tbody>
</table>

Population of patients who can adhere to Experimental AND Control treatments.
Causal Inference

For T(0)
\(A(0) = 0\) for non-adherence
\(= 1\) for adherence

For T(1)
\(A(1) = 0\) for non-adherence
\(= 1\) for adherence

<table>
<thead>
<tr>
<th>Control Treatment Adherence</th>
<th>Experimental Treatment Adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A(1) = 0</td>
</tr>
<tr>
<td>A(0) = 0</td>
<td>NA</td>
</tr>
<tr>
<td>A(0) = 1</td>
<td>NA</td>
</tr>
<tr>
<td>A(0) (\in {0, 1})</td>
<td>NA</td>
</tr>
</tbody>
</table>

Population of patients who can adhere to the Experimental treatment *regardless* of adherence to the Control treatment.
### Causal Inference

For $T(0)$

<table>
<thead>
<tr>
<th>Control Treatment Adherence</th>
<th>A(1) = 0</th>
<th>A(1) = 1</th>
<th>A(1) ∈ {0, 1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>A(0) = 0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>A(0) = 1</td>
<td>NA</td>
<td>A++</td>
<td><strong>A</strong>++</td>
</tr>
<tr>
<td>A(0) ∈ {0, 1}</td>
<td>NA</td>
<td>A*+</td>
<td><strong>A</strong>**</td>
</tr>
</tbody>
</table>

Population of patients who can adhere to the Control treatment *regardless* of adherence to the Experimental treatment.

For $T(1)$

<table>
<thead>
<tr>
<th>Control Treatment Adherence</th>
<th>A(1) = 0</th>
<th>A(1) = 1</th>
<th>A(1) ∈ {0, 1}</th>
</tr>
</thead>
<tbody>
<tr>
<td>A(0) = 0</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>A(0) = 1</td>
<td>NA</td>
<td>A++</td>
<td><strong>A</strong>++</td>
</tr>
<tr>
<td>A(0) ∈ {0, 1}</td>
<td>NA</td>
<td>A*+</td>
<td><strong>A</strong>**</td>
</tr>
</tbody>
</table>
Causal Inference

**For T(0)**
- A(0) = 0 for non-adherence
- A(0) = 1 for adherence

**For T(1)**
- A(1) = 0 for non-adherence
- A(1) = 1 for adherence

<table>
<thead>
<tr>
<th>Control Treatment Adherence</th>
<th>Experimental Treatment Adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A(1) = 0</td>
</tr>
<tr>
<td>A(0) = 0</td>
<td>NA</td>
</tr>
<tr>
<td>A(0) = 1</td>
<td>NA</td>
</tr>
<tr>
<td>A(0) ∈ {0, 1}</td>
<td>NA</td>
</tr>
</tbody>
</table>

Population of All Randomized patients.
Causal Inference

General Framework

Let $S \in \{ A++, A*, A+, A** \}$

$$E \left[ Y(1) - Y(0) \mid S \right]$$
Causal Inference

Estimand definition requires POPULATION

It’s more than Inclusion/Exclusion criteria!

**WHAT** \( S \in \{ A++, A^{++}, A^+, A^{**} \} \) are you interested?

\[
E [ Y(1) - Y(0) \mid S ]
\]

**Placebo controlled trials**

The population of patients adherent to Experimental treatment regardless of adherence to Placebo.
Causal Inference

All Patients

Adherers to Experimental
Causal Inference

Estimand definition requires POPULATION

It’s more than Inclusion/Exclusion criteria!

**WHAT** $S \in \{ A++, A^{++}, A^+, A^{**} \}$ are you interested?

$$E[ Y(1) - Y(0) \mid S ]$$

**Active controlled trials**

?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ?? ??
Causal Inference

A**
All Patients

A**
Adherers to Experimental

A**
Adherers to Active Control
Estimand definition requires POPULATION

It’s more than Inclusion/Exclusion criteria!

WHAT $S \in \{A++, A++*, A+, A**\}$ are you interested?

$$E \{ Y(1) - Y(0) \mid S \}$$

Active controlled trials

The population of patients adherent to both the Experimental treatment and the Active Control treatment.
Causal Inference

- **A**
  - All Patients
  - Adherers to Active Control
  - Adherers to Experimental and Active Control

- **A**
  - Adherers to Experimental
Causal Inference

Estimand definition requires POPULATION

It’s more than Inclusion/Exclusion criteria!

**WHAT** \( S \in \{ A++, A*+, A+, A** \} \) are you interested?

\[
E [ Y(1) - Y(0) | S ]
\]

The Intent-to-Treat Population (all randomized patients).
Method A – Counterfactual

- Build a model of response (Y) based on baseline covariates (X) and post-randomization data (Z)
- Create a “virtual twin”
- Predict response on the unobserved treatment
- Compare observed response with predicted response of “twin”

This is consistent with the ICH-E9(R1) statement:

- “quantify treatment effects: how the outcome of treatment compares to what would have happened to the same subjects under different treatment conditions (e.g. had they not received the treatment or had they received a different treatment).”
Method B – Inverse Probability Weighting (IPW)

- Estimate the probability that a patient from the unobserved treatment (e.g. Control) would adhere to the observed treatment (e.g. Experimental)
- Use IPW to estimate the average response for the population ($A_{ij}$) of interest.

This is consistent with the ICH-E9(R1) statement:
- But takes a different approach
# Estimators Based on Adherence

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimator Based on Distribution of $(X, Z, Y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A**</td>
<td>$\frac{1}{n_1} \sum_{j \in {j: T_j = 1}} \hat{\psi}<em>1(X_j, Z_j) - \frac{1}{n_0} \sum</em>{j \in {j: T_j = 0}} \hat{\psi}_1(X_j, Z_j)$</td>
</tr>
<tr>
<td>A*+</td>
<td>$\frac{1}{n_{11}} \sum_{j} T_j A_j Y_j - \frac{1}{n_{11}} \sum_{j} \hat{\phi}_0(X_j)$</td>
</tr>
<tr>
<td>A++</td>
<td>$\frac{1}{n_{01}} \sum_{j} \hat{\phi}<em>1(X_j) - \frac{1}{n</em>{01}} \sum_{j} (1 - T_j) A_j Y_j$</td>
</tr>
<tr>
<td>A+++</td>
<td>$\frac{1}{\sum_{j} \hat{\phi}<em>1(X_j)(1 - T_j) A_j} - \frac{1}{\sum</em>{j} \hat{\phi}_0(X_j) T_j A_j}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimator Based on Distribution of $(X, Z, A)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A**</td>
<td>$\frac{1}{n_1} \sum_{j \in {j: T_j = 1, A_j = 1}} \frac{Y_j}{\hat{g}(X_j, Z_j)} - \frac{1}{n_2} \sum_{j \in {j: T_j = 0, A_j = 1}} \frac{Y_j}{\hat{g}(X_j, Z_j)}$</td>
</tr>
<tr>
<td>A*+</td>
<td>$\frac{1}{n_{11}} \sum_{j} T_j A_j Y_j - \frac{n_{11}}{n_0 n_{11}} \sum_{j} \hat{h}_1(X_j)(1 - T_j) A_j Y_j$</td>
</tr>
<tr>
<td>A++</td>
<td>$\frac{n_0}{n_1 n_{01}} \sum_{j} \frac{\hat{h}<em>0(X_j) T_j A_j Y_j}{\hat{g}(X_j, Z_j)} - \frac{1}{n</em>{01}} \sum_{j} (1 - T_j) A_j Y_j$</td>
</tr>
<tr>
<td>A+++</td>
<td>$\frac{1}{\sum_{j} \hat{h}<em>1(X_j)(1 - T_j) A_j Y_j} - \frac{1}{\sum</em>{j} \hat{h}_1(X_j)(1 - T_j) A_j}$</td>
</tr>
</tbody>
</table>
### Estimators Based on Adherence

<table>
<thead>
<tr>
<th>Population</th>
<th>Method A: Estimator Based on Distribution of ((X, Z, Y))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_{**})</td>
<td>[ \hat{\psi}_1(X_j, Z_j) ]</td>
</tr>
<tr>
<td>(A_{*+})</td>
<td>[ \frac{\sum_j h_1(X_j)(1-T_j)A_j}{\sum_j h_0(X_j)T_jA_j} ]</td>
</tr>
<tr>
<td>(A_{++})</td>
<td>[ \frac{\sum_j h_1(X_j)A_j}{\sum_j h_0(X_j)A_j} ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Population</th>
<th>Method B: Estimator Based on Distribution of ((X, Z, A))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_{**})</td>
<td>[ \frac{1}{Y_j} \sum_j \frac{Y_j}{(X_j, Z_j)} ]</td>
</tr>
<tr>
<td>(A_{*+})</td>
<td>[ \frac{1}{\sum_j \hat{h}_0(X_j)T_jA_j} ]</td>
</tr>
<tr>
<td>(A_{++})</td>
<td>[ \frac{1}{\sum_j \hat{h}_1(X_j)(1-T_j)A_j} ]</td>
</tr>
</tbody>
</table>

**An extension of MMRM**

**An extension of Marginal Structural Models**
Estimators Based on Adherence

Under certain reasonable assumptions, these estimators are consistent

Our simulations studies show they are unbiased
- Even with modest sample sizes ($N = 150 / \text{treatment}$)
- For different discontinuation patterns
  - Differential discontinuation on Experimental and Control Treatments
  - Discontinuations ranging from $\sim10\%$ to $\sim50\%$

Method A and Method B perform similarly
- Method A is easier computationally when the outcome is normally distributed.
Outline

Motivation
Tripartite Approach
Estimators for Causal Inference
Example
Discussion
Discontinuation due to administrative reasons occurred in ~11% of each treatment group and were randomly distributed over the course of the study.
In **WHAT** Population are we interested?

\[ E \{ Y(1) - Y(0) \mid S = A^{++} \} \]

(adherent to experimental **and** active control)

---

We could examine ...

\[ E \{ Y(1) - Y(0) \mid S = A^{*+} \} \]

(adherent to experimental **regardless** of active control adherence)
Treatment Difference between Experimental and Active Control in HbA1c at 52 Weeks

Abbreviations: NA = not applicable; ITT = Intent-to-Treat; MMRM = Mixed Model Repeated Measures; J2R – Jump to Reference; CR = Copy Reference
Outline

Motivation
Tripartite Approach
Estimators for Causal Inference
Example
Discussion
Casual Inference – Completers Analysis

$$E[Y(1) \mid A(1)=1] - E[Y(0) \mid A(0)=1]$$

Causal Inference

$$E [Y(1) - Y(0) \mid S]$$

More work on variance of these estimators
Missing at Random assumption

- Quite reasonable to believe that discontinuations of treatment are related to the efficacy and safety of the treatment
- Some data MCAR (administrative drop-outs)

Causal Inference

- It’s more complicated, but ... worth it given the cost of clinical trials.
What are the Right Questions?

Patient / Physician
What happens when I take this medication?

Researcher
What are the causal effects of treatment?

Regulator
What are the benefits and risks of treatment?

Tripartite Estimands
What are the Right Questions?

Probability of Adverse Event

Probability of Lack of Efficacy

Benefit (and Risk) Given Adherence

Tripartite Estimands

Probability of Adverse Event
Why can’t Estimand 3 be the gatekeeper to regulatory review (i.e. $p < 0.05$)?

(PS: It can!)

(PPS: Physicians think this is what we give them.)

... Then assess risks in this context.
“... I hope we also recognize when what’s meaningful to our patients trumps anything medical that we can offer.”

Mikkael A. Sekeres, M.D.
“The Best Medicine? What’s Meaningful to Our Patients”
New York Times
3 May 2018

Stephen J. Ruberg
ASA Bipharm Workshop
13 Sep 2018
“... I hope we also recognize when what’s meaningful to our patients trumps anything statistical that we can offer.”

Mikkael A. Sekeres, M.D.
“The Best Medicine? What’s Meaningful to Our Patients”
New York Times
3 May 2018
Stephen J. Ruberg
ASA Bipharm Workshop
13 Sep 2018
Presentation Estimands

1. Proportion of those who had an adverse reaction to these concepts/recommendations
2. Proportion of those who tuned out due to lack of interest
3. For those who followed this presentation to the end ...

THANK YOU.

I hope the expected change in your thinking is scientifically meaningful!!