

Introduction to Causal Inference and Causal Data Science

Day 5: Group Work, Discussion, Q&A

Oisín Ryan, Bas Penning de Vries, Wouter van Amsterdam

Department of Data Science and Biostatistics
Julius Center
UMC Utrecht

July 11, 2025

Group Assignment

- 1 Define a causal research question.
 - Can this research question be answered with a hypothetical **target trial**? If so, what would that look like?
 - Tip: Start with a broad research question, then operationalize as more specific

Group Assignment

- 1 Define a causal research question.
 - Can this research question be answered with a hypothetical **target trial**? If so, what would that look like?
 - Tip: Start with a broad research question, then operationalize as more specific
- 2 What type of (non-RCT) **data** would you need to answer this question?
 - What would the ideal dataset look like?
 - What type of data might you reasonably have in practice?
 - Think about setting where there is variation in the **treatment**. What other variables would you need to observe?

Group Assignment

- 1 Define a causal research question.
 - Can this research question be answered with a hypothetical **target trial**? If so, what would that look like?
 - Tip: Start with a broad research question, then operationalize as more specific
- 2 What type of (non-RCT) **data** would you need to answer this question?
 - What would the ideal dataset look like?
 - What type of data might you reasonably have in practice?
 - Think about setting where there is variation in the **treatment**. What other variables would you need to observe?
- 3 Describe the **assumptions** you would need to identify this causal effect from your chosen data setting.
 - If possible, **draw a DAG**
 - Discuss **exchangeability**, **SUTVA**, **positivity**
 - Which assumptions are most challenging / problematic?

Group Assignment

- 1 Define a causal research question.
 - Can this research question be answered with a hypothetical **target trial**? If so, what would that look like?
 - Tip: Start with a broad research question, then operationalize as more specific
- 2 What type of (non-RCT) **data** would you need to answer this question?
 - What would the ideal dataset look like?
 - What type of data might you reasonably have in practice?
 - Think about setting where there is variation in the **treatment**. What other variables would you need to observe?
- 3 Describe the **assumptions** you would need to identify this causal effect from your chosen data setting.
 - If possible, **draw a DAG**
 - Discuss **exchangeability**, **SUTVA**, **positivity**
 - Which assumptions are most challenging / problematic?
- 4 Describe how you would go about **estimating** this effect

Group Assignment

Form groups of 4/5, choose a flipboard (4 groups)

Session 1: Design your causal analysis

Session 2: Groups pair up, present their causal research project, get feedback.

Session 3: Plenary. Each group shares 1/2 insights they had

Practical Matters

Before we go to lunch, let's take a **group picture**

Before starting again after lunch, please fill out the evaluation form (link on course website)

Closing

This Course

Goals

- **Introduce** you to the foundational concepts of causal modeling with **observational** or **non-experimental** data.
- Show you how to **view data problems** through the lens(es) of causal inference
- **Equip you** with the **skills** to perform causal inference and causal modeling
- **Get you started** on your own journey of causal learning

Other Materials / Events

What if (Hernan and Robins): https://content.sph.harvard.edu/wwwhsph/sites/1268/2024/04/hernanrobins_WhatIf_26apr24.pdf

Causal policy evaluation methods / Diff-in-diff, synthetic control, etc:
<https://causalpolicy.nl/>

UU/UMCU/NL: Causal Data Science SIG <https://www.uu.nl/en/research/applied-data-science/sig-causal-data-science>.
Quarterly meetings with Leiden, Delft, Rotterdam MC's

Julius Center: Monthly Causal Infernece seminars (Wouter & Oisín)

Thank You