Discussion on hypothetical strategies: a causal inference perspective

UNIVERSITEIT GENT

Kelly Van Lancker

FACULTY OF SCIENCES

Causal inference road map

- Positive evolution towards ‘specification of an **estimand as the starting point** of a data analysis’.
    - This is central to the **causal inference tradition**. (Hernán and Robins, 2020)

Causal inference road map

- Positive evolution towards ‘specification of an **estimand as the starting point** of a data analysis’.
  - This is central to the **causal inference tradition**. (Hernán and Robins, 2020)
- Road map for inferring causal effects
  1. Defining the **estimand**
  2. Stating the **identification assumptions**
  3. **Estimation** method(s) along with statistical assumptions

Causal inference road map

- Positive evolution towards 'specification of an **estimand as the starting point** of a data analysis'.
  - This is central to the **causal inference tradition**. (Hernán and Robins, 2020)
- Road map for inferring causal effects
  1. Defining the **estimand**
  2. Stating the **identification assumptions**
  3. **Estimation** method(s) along with statistical assumptions
- Despite the many positive steps, statisticians often tend to go straight to Step 3.
- In my opinion, we should strive to fully follow this road map to achieve the most benefits.

Defining *a* hypothetical estimand

- Broad range of hypothetical scenarios can be considered: important to be **precise enough** [Mouna and Frank]
  - Different (hypothetical) strategies for different intercurrent events can be combined. [Ian]

Defining *a* hypothetical estimand

- Broad range of hypothetical scenarios can be considered: important to be **precise enough** [Mouna and Frank]
  - Different (hypothetical) strategies for different intercurrent events can be combined. [Ian]
- By using **counterfactual outcomes** notation, one can uniquely define the estimand. [Jonathan]
  - Controlled direct effect: $E\left[Y_2(Z=1, Sym=0) - Y_2(Z=0, Sym=0)\right]$ [Florian]
  - Natural direct effect (Michiels et al., 2021): $E\left[Y_2(Z=1, Sym(0)) - Y_2(Z=0, Sym(0))\right]$

Defining *a* hypothetical estimand

- Broad range of hypothetical scenarios can be considered: important to be **precise enough** [Mouna and Frank]
  - Different (hypothetical) strategies for different intercurrent events can be combined. [Ian]
- By using **counterfactual outcomes** notation, one can uniquely define the estimand. [Jonathan]
  - Controlled direct effect: $E\left[Y_2(Z=1, Sym=0) - Y_2(Z=0, Sym=0)\right]$ [Florian]
  - Natural direct effect (Michiels et al., 2021): $E\left[Y_2(Z=1, Sym(0)) - Y_2(Z=0, Sym(0))\right]$
- Importantly, the estimand should target the **scientific question** and be **model-free**.

Identification assumptions for hypothetical estimands

- Identification refers to the **translation** of a causal estimand into a quantity involving only the observed data.
    - "Making connection between estimand, study design, data collection, and analysis."
- E.g., consistency, positivity, and exchangeability. [Jonathan]

Identification assumptions for hypothetical estimands

- Identification refers to the **translation** of a causal estimand into a quantity involving only the observed data.
  - "Making connection between estimand, study design, data collection, and analysis."
- E.g., consistency, positivity, and exchangeability. [Jonathan]
- Different assumptions exist; e.g., **instrumental variable** conditions.
  - What is the role for instrumental variable assumptions/methods to correct for non-compliance?

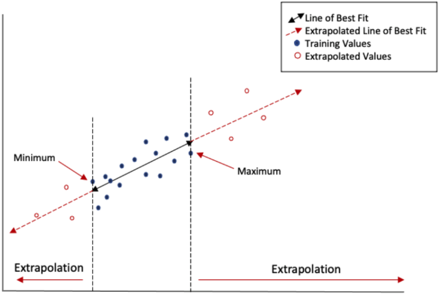
Estimation of hypothetical estimands (1)

- **Positivity violations** are a real concern with intercurrent events
  - Study protocols may often indicate when e.g., rescue treatment should be started.

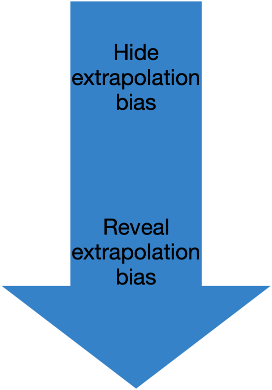
Estimation of hypothetical estimands (1)

- **Positivity violations** are a real concern with intercurrent events
  - Study protocols may often indicate when e.g., rescue treatment should be started.

- This brings serious **concers for extrapolation**, and means that trade-offs need to be made between relevant and feasible estimands. [Jinglin]

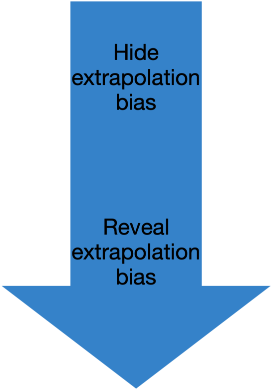


Estimation of hypothetical estimands (2)



- **Multiple imputation** and **maximum-likelihood**
- **G-estimation**[Florian]
  - Less clear how to use them for other type of endpoints (e.g., binary).
- **Inverse weighting**
  - Can be inefficient.
  - More sensitive to near-positivity violations than G-estimation.

Estimation of hypothetical estimands (2)



- **Multiple imputation** and **maximum-likelihood**
- **G-estimation**[Florian]
  - Less clear how to use them for other type of endpoints (e.g., binary).
- **Inverse weighting**
  - Can be inefficient.
  - More sensitive to near-positivity violations than G-estimation.

- **Augmented inverse weighting** / **targeted learning** deserves much more attention. (Van der Laan et al., 2011)
  - More efficient than inverse weighting.
  - Less assumptions than G-estimation, thus usually less efficient.

Friendly criticism towards causal inference community

- Sometimes **overly ambitious** when defining the estimand. (Hernán and Scharfstein, 2018)
- Assumptions are often defined in a pure mathematical way, without attention to the **clinical relevance**.

Friendly criticism towards causal inference community

- Sometimes **overly ambitious** when defining the estimand. (Hernán and Scharfstein, 2018)
- Assumptions are often defined in a pure mathematical way, without attention to the **clinical relevance**.

I believe that, together, the two communities can push each other forward!

kelly.vanlancker@ugent.be

References I

Hernán, M. and J. M. Robins (2020). *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC. Taylor & Francis.

Hernán, M. A. and D. Scharfstein (2018). Cautions as regulators move to end exclusive reliance on intention to treat.

Michiels, H., C. Sotto, A. Vandebosch, and S. Vansteelandt (2021). A novel estimand to adjust for rescue treatment in randomized clinical trials. *Statistics in Medicine 40*(9), 2257–2271.

Van der Laan, M. J., S. Rose, et al. (2011). *Targeted learning: causal inference for observational and experimental data*, Volume 10. Springer.