Causal Inference with Continuous Multiple Time Point Interventions

Michael Schomaker
Ludwig-Maximilians Universität München, Germany

joint work with Iván Díaz, Paolo Denti, Helen McIlleron

Talk at the EuroCIM 2023, Oslo

21 April 2023

Motivation – Continuous Interventions

Data:

- CHAPAS-3 trial (Mulenga et al., *Lancet Infectious Diseases*, 2016):
  children, $\leq$13 years, from Zambia/Uganda, randomized NRTI drug (abacavir, stavudine oder zidovudine) of HIV therapy

Pharmacological Substudy:

- Bienczak et al. (*AIDS*, 2017) evaluated concentration of NNRTI drug regimen component: Nevirapine and Efavirenz
  $\rightarrow$ higher probability of "viral failures" with lower concentrations
- what is the ideal target concentration?; causal question:

> How many percent of children would have had a suppressed viral load at time t if they had had a concentration of "x" mg/L efavirenz at each time point?
>
> In general: how would probability of failure vary for different hypothetical concentration trajectories? $\rightarrow$ **"causal dose-response curve" (CDRC)**

Motivation – Continuous Interventions (II)

We essentially have longitudinal observational data:

- Time-varying confounders: weight, adherence (with treatment-confounder feedback!)
- Note: *Positivity assumption may not be satisfied with continuous interventions*!
- Possible options to answer motivating question:

  Option 1: Change question: "modified treatment policies" (e.g., Diaz et al., *JASA*, 2021)
  Option 2: G-methods $\rightarrow$ simple application (i.e., intervene for many trajectories) today
  Option 3: Find a compromise between interpretability and identifiability today
  Also: for 1 time point, great DR approach developed (Kennedy, JRSS B, 2017)

Notation

- Follow-up time: $t = 0, 1, \ldots, T$
- Outcome: Y_t
- Intervention: A_t
- Confounder, Covariate: L_t
- History: e.g. $\bar{A}_t = (A_0, \ldots, A_t)$
- History up to A_t: H_t
- **Counterfactual:** e.g. $Y_t^{\bar{a}_t}$

Estimand & Estimation with Sequential G-computation

Estimand:

$$m_t : \bar{a}_t \mapsto E(Y_t^{\bar{a}_t} | \mathbf{L_0}^*), \qquad t = 0, 1, \ldots, T$$

Under sequential conditional exchangeability, consistency and **positivity** we have:

$$\mathbb{E}(Y_t^{\bar{a}_t}) \quad = \quad \mathbb{E}(\ldots \mathbb{E}(\mathbb{E}(Y_t | \bar{A}_t = \bar{a}_t, \mathbf{H}_t) | \bar{A}_{t-1} = \bar{a}_{t-1}, \mathbf{H}_{t-1}) \ldots | A_0 = a_0, \mathbf{L}_0)).$$

$\rightarrow$ substitution estimation (sequential g-computation)

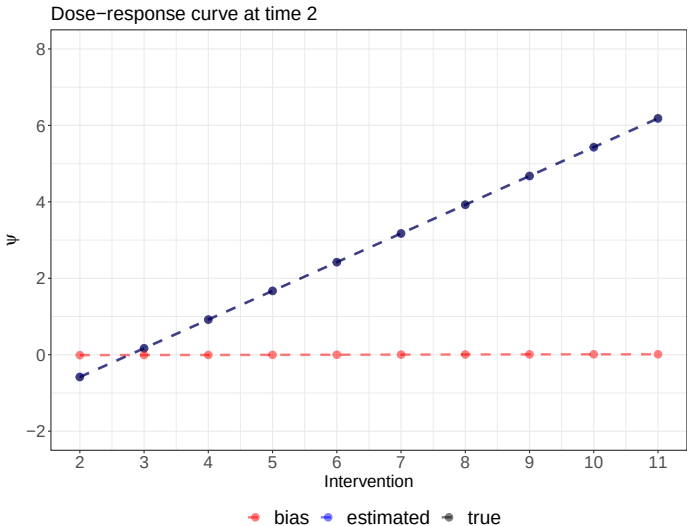
Positivity:

$$\inf_{a_t \in \mathcal{A}_t} g(a_t \mid \mathbf{h}_t) > 0 \quad \text{whenever} \quad p_0(\bar{\mathbf{L}}_t = \bar{\mathbf{l}}_t, \bar{A}_{t-1} = \bar{a}_{t-1}) > 0 \quad \forall t, \bar{a}_t, \bar{\mathbf{l}}_t .$$

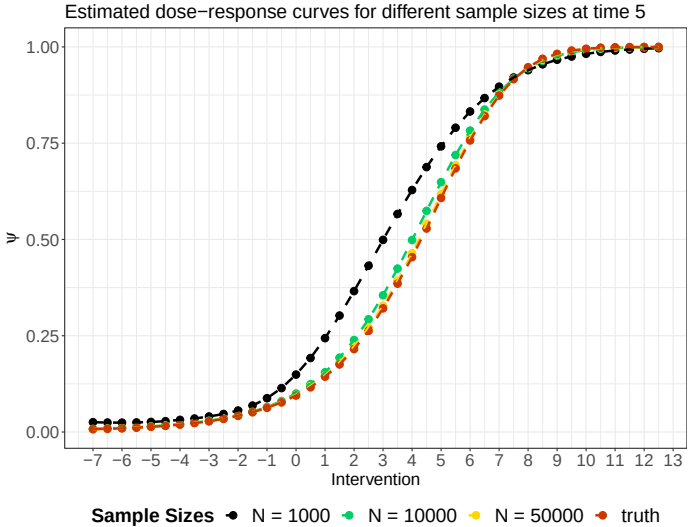
where $\bar{\mathcal{A}}_t$ denotes the set of all relevant strategies $\bar{a}_t = (a_0, \ldots, a_t)$

$\rightarrow$ What if we simply assume positivity and apply g-computation for many $\bar{a}_t$?

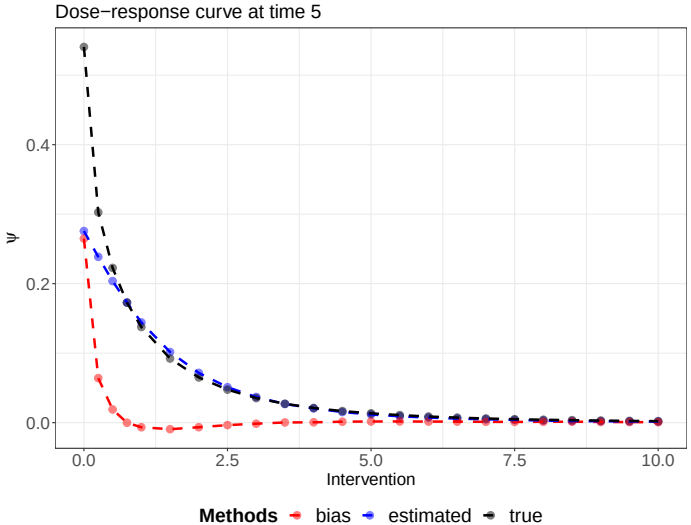
Simulation (simple)



Simulation (survival)



Simulation (complex, as in data)



Overall Consideration

The **tradeoff** to make is between

estimating the CDRC as closely as possible, at the risk of bias due to positivity violations because of the continuous intervention

and

minimizing the risk of bias due to positivity violations, at the cost of redefining the estimand (e.g. by using modified treatment policies)

Alternatively: compromise!

Proposal: Weighted Estimand (1 Time Point)

The general dose-response curve $m : a \mapsto E(Y^a)$ can be identified[1] with the g-formula as

$$m(a) = \int E(Y \mid A = a, \mathbf{L} = \mathbf{l}) p_0(\mathbf{l}) d\nu(\mathbf{l}) \,,$$

Proposal: instead, rather use

$$m_w(a) = \int E(Y \mid A = a, \mathbf{L} = \mathbf{l}) w(a, \mathbf{l}) p_0(\mathbf{l}) d\nu(\mathbf{l})$$

with

$$w(a, \mathbf{l}) = \begin{cases} 1 & \text{if } g(a \mid \mathbf{l}) > c \\ \frac{g(a|l)}{g(a)} & \text{otherwise.} \end{cases}$$

[1] under consistency, positivity and conditional exchangeability

Weighted Estimand – Implications

- yields the desired dose-response curve under enough support (i.e., $g(a \mid \mathbf{l}) > c$)

- otherwise the estimand is $E(Y|A = a)$

  $\rightarrow$ not a causal quantity but does not require positivity assumption

Weighted Estimand – Multiple Time Points

$$w_t(a_{t+1}, \mathbf{h}_{t+1}, c) = \begin{cases} 1 & \text{if } g_t(a_{t+1} \mid \mathbf{h}_{t+1}) > c\,, \\ \frac{g_t(a_{t+1}|\mathbf{h}_{t+1})}{g_t(a_{t+1}|a_t,\mathbf{h}_t)} & \text{if } g_t(a_{t+1} \mid \mathbf{h}_{t+1}) \leq c \text{ and } g_t(a_{t+1} \mid a_t, \mathbf{h}_t) > c\,, \\ \frac{g_t(a_{t+1}|\mathbf{h}_{t+1})}{g_t(a_{t+1}|a_{t-1},\mathbf{h}_{t-1})} & \text{if } g_t(a_{t+1} \mid \mathbf{h}_{t+1}) \leq c \text{ and } g_t(a_{t+1} \mid a_t, \mathbf{h}_t) \leq c \\ & \text{and } g_t(a_{t+1} \mid a_{t-1}, \mathbf{h}_{t-1}) > c\,, \\ \vdots & \vdots \\ \frac{g_t(a_{t+1}|\mathbf{h}_{t+1})}{g_t(a_{t+1})} & \text{otherwise}\,. \end{cases}$$

Weighted Estimand – Implications

- returns the CDRC if there is enough conditional support in terms of $g_t(a_t \mid \mathbf{h}_t) > c$
- if there is not enough conditional support (and the weight denominator is $> c$)

  we can show that the estimand equates to $E(Y_t | A_t = a_t, \ldots, A_0 = a_0)$

  $\rightarrow$ not a causal quantity but does not require positivity assumption
- if there is not enough conditional support and the weight denominator is too small the estimand entails a compromise

Weighted Estimand – Estimation (I)

For example, substitution estimator based on the following expression:

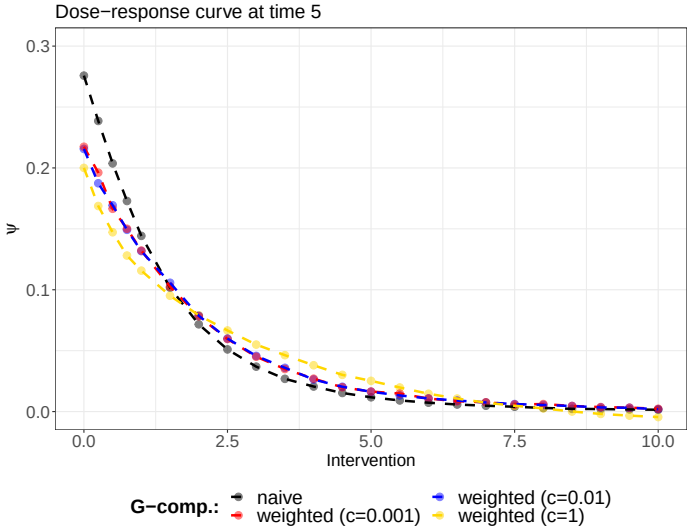
$$\mathbb{E}_{w_0}(\ldots \mathbb{E}_{w_{t-1}}(\mathbb{E}_{w_t}(Y_t|\bar{A}_t = \bar{a}_t, \mathbf{H}_t)|\bar{A}_{t-1} = \bar{a}_{t-1}, \mathbf{H}_{t-1})\ldots|A_0 = a_0, \mathbf{L}_0)),$$

where we define $\mathbb{E}_{w_t}(Y_t|\bar{A}_t = \bar{a}_t, \mathbf{H}_t) = \mathbb{E}(w_t Y_t|\bar{A}_t = \bar{a}_t, \mathbf{H}_t)$.

$\rightarrow$ can also re-expressed into parametric g-formula-type expression, but then requires estimation of conditional *densities*, not only expectations

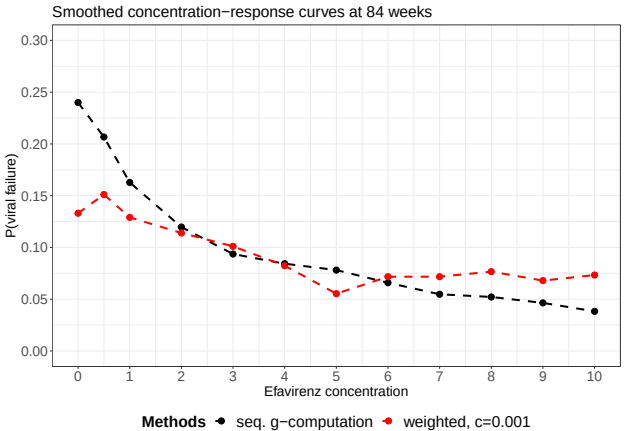
Note: even if Y_t is normal, $w_t Y_t$ may not be normal; so we may need a data-adaptive approach

Simulation (complex, as in data)



Data Analysis

Based on a complete case analysis of $n = 58$ kids



Weighted curve deviates from estimated CDRC in areas of low support

Interpretation

The weighted curve shows the probability of failure

- if patients whose covariate trajectory that makes the intervention value of interest at t not unlikely to occur (under the desired intervention before t), receive the intervention at t
- all other patients get different interventions that produce, on average, outcomes as we would expect among those who actually follow the trajectory of interest.

The weighted curve acts like a magnifying class and sensitivity tool if we don't want to rely on parametric extrapolation in regions of low support, where fixing the concentration to a specific level seems unrealistic.

Summary

- Standard g-computation can be used for continuous interventions
  - + targets the CDRC of interest
  - – relies on positivity assumption
- Simulations show that strategy may work, but can be problematic in regions of low support or with limited sample size
- Weighted curves offer a compromise and do not require the positivity assumption
- The toolkit for causal effect estimation with longitudinal continuous interventions should ideally be broad

Literature

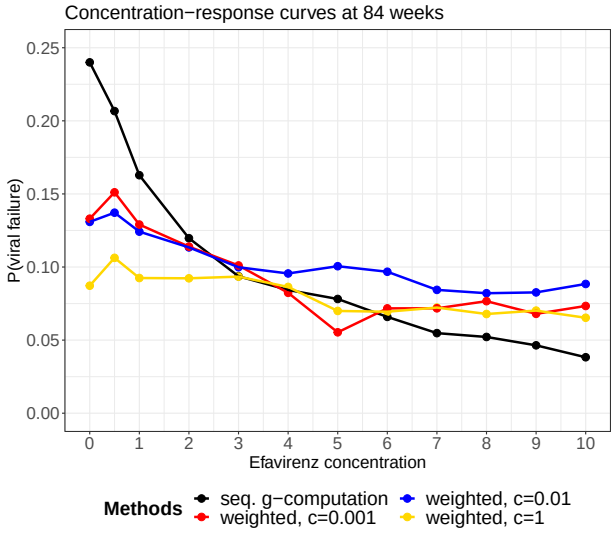
[1] Andrzej Bienczak, Paolo Denti, Adrian Cook, Lubbe Wiesner, Veronica Mulenga, Cissy Kityo, Addy Kekitiinwa, Diana M. Gibb, David Burger, Ann S. Walker, and Helen McIlleron.
Determinants of virological outcome and adverse events in african children treated with paediatric nevirapine fixed-dose-combination tablets.
AIDS, 31(7):905–915, 2017.

[2] Iván Díaz, Nicholas Williams, Katherine L. Hoffman, and Edward J. Schenck.
Non-parametric causal effects based on longitudinal modified treatment policies.
arXiv e-prints, 2020.

[3] Edward H. Kennedy, Zongming Ma, Matthew McHugh, and Dylan S. Small.
Nonparametric methods for doubly robust estimation of continuous treatment effects.
Journal of the Royal Statistical Society. Series B, Statistical methodology, 79:1229–1245, 2017.

[4] Veronica Mulenga, Victor Musiime, Adeodata Kekitiinwa, Adrian D. Cook, George Abongomera, Julia Kenny, Chisala Chabala, Grace Mirembe, Alice Asiimwe, Ellen Owen-Powell, David Burger, Helen McIlleron, Nigel Klein, Chifumbe Chintu, Margaret J. Thomason, Cissy Kityo, A. Sarah Walker, and Diana Gibb.
Abacavir, zidovudine, or stavudine as paediatric tablets for african hiv-infected children (chapas-3): an open-label, parallel-group, randomised controlled trial.
The Lancet Infectious Diseases, 16(2):169–79, 2016.

APPENDIX


Concentration−response curves at 84 weeks
P(viral failure)
0.25
0.20
0.15
0.10
0.05
0.00
0
1
2
3
4
5
6
7
8
9
10
Efavirenz concentration
Methods
seq. g−computation
weighted, c=0.001
weighted, c=0.01
weighted, c=1

Treatment-Confounder Feedback