Estimands in clinical studies

Kelly Van Lancker

UNIVERSITEIT GENT

FACULTY OF SCIENCES

Acknowledgements

- Saskia le Cessie (Leiden University)
- Stijn Vansteelandt (Ghent University)

Outline

Causal estimands

- The focus of my talk will be on **causal estimands**.

Causal estimands

- The focus of my talk will be on **causal estimands**.
  - These are often **marginal**,

Causal estimands

- The focus of my talk will be on **causal estimands**.
  - These are often **marginal**,
  - which are often necessary (in trials) for **regulatory approval**.

Causal estimands

- The focus of my talk will be on **causal estimands**.
  - These are often **marginal**,
  - which are often necessary (in trials) for **regulatory approval**.
- Should we only focus on them?

Causal estimands

- The focus of my talk will be on **causal estimands**.
  - These are often **marginal**,
  - which are often necessary (in trials) for **regulatory approval**.
- Should we only focus on them?
  - No!

Causal estimands

- The focus of my talk will be on **causal estimands**.
  - These are often **marginal**,
  - which are often necessary (in trials) for **regulatory approval**.
- Should we only focus on them?
  - No!
  - **Less relevant** for exploratory research,

    as well as individual decision-making.
    (see talk Nan; as well as assumption-lean by Stijn)

Estimands first!

- **Defining a causal estimand** is the **first step in a causal analysis**.

Estimands first!

- **Defining a causal estimand** is the **first step in a causal analysis**.
- Causal estimands **translate the scientific question into a quantity that we can (hopefully) communicate well** to clinicians/investigators/....
- Models and algorithms are only tools to learn an estimand, but should never be the primary aim of a causal analysis.

Estimands first!

- **Defining a causal estimand** is the **first step in a causal analysis**.
- Causal estimands **translate the scientific question into a quantity that we can (hopefully) communicate well** to clinicians/investigators/. . . .
- Models and algorithms are only tools to learn an estimand, but should never be the primary aim of a causal analysis.
- This may sound obvious, but it is not.
  - For a statistician / data scientist, the first step is often formulating a model / algorithm.

A common first step

- ICH E9 (FDA and EMA, 1998) and EMA (2015) guidelines are written with the understanding that the **target treatment effect is a model parameter**; e.g.,

$$g\{E(Y|Z,X)\} = \beta_0 + \beta_1 Z + \beta_2 X$$

where $g(\cdot)$ is a pre-specified link function.

A common first step

- ICH E9 (FDA and EMA, 1998) and EMA (2015) guidelines are written with the understanding that the **target treatment effect is a model parameter**; e.g.,

$$g\{E(Y|Z,X)\} = \beta_0 + \beta_1 Z + \beta_2 X$$

  where $g(\cdot)$ is a pre-specified link function.

- This model implies no interaction between Z and X:
  - A **statistical modelling assumption**: **also not implied by e.g. randomization**.
  - When the model is misspecified, the standard likelihood-based estimators of β_1 **may not generally target a causal effect**.

Causal inference road map

- Road map for inferring causal effects
  1. Defining the **estimand**
  2. Stating the **identification assumptions**
  3. **Estimation** method(s) along with statistical assumptions
  4. Evaluate validity of assumptions via **sensitivity analyses**

Causal inference road map

- Road map for inferring causal effects
  1. Defining the **estimand**
  2. Stating the **identification assumptions**
  3. **Estimation** method(s) along with statistical assumptions
  4. Evaluate validity of assumptions via **sensitivity analyses**
- Despite the various positive efforts made, statisticians often tend to go straight to Step 3.
- In my opinion, we should strive to fully follow this road map.
  - **Sensible estimand**: iterate steps 1-2 between statisticians and clinicians.

Key steps to define estimand

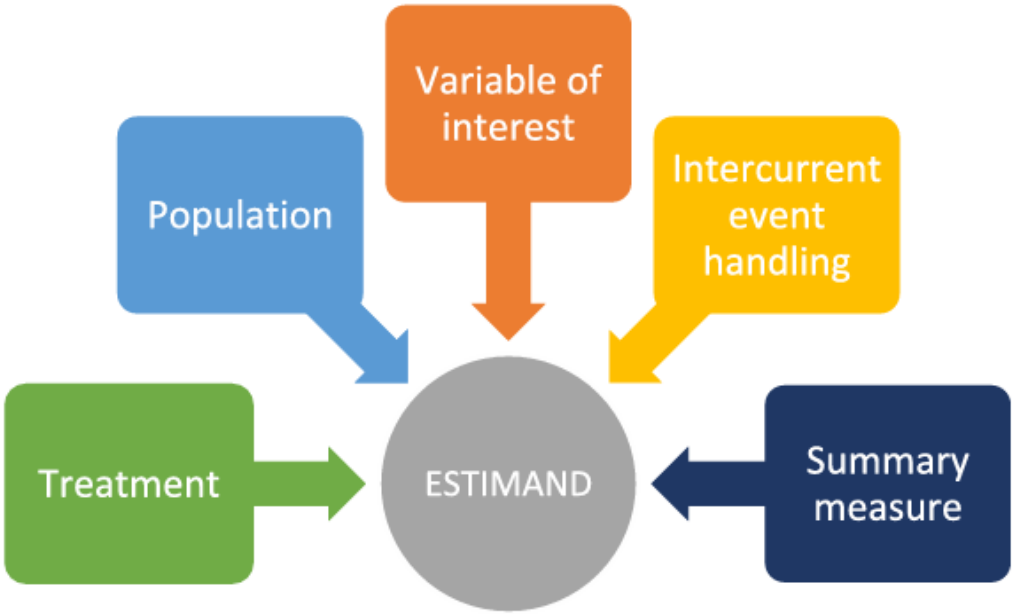


Fig. 1 The five attributes of an estimand according to the ICH E9 (R1) addendum

+ Formalize **potential/counterfactual outcomes**

Outline

Key steps to define estimand

- Relevant time-to-event $\tilde{T}$ as outcome variable.
- A clear starting point (time zero);
- Population of interest:
  - Whole population (marginal estimand)
  - Subpopulation (conditional estimand)
  - Treated only (ATT)
- (Point) exposure/treatment of interest with interventions to compare:
  - Treating versus not treating
  - Hospitalization versus no hospitalization
- **Potential outcomes:** $\tilde{T}^a$ is potential time-to-event under treatment setting $A = a$.

Concept of **emulating Target Trials** can be very helpful (Hernán, 2016).

Estimands: what with censoring?

- Censoring makes the **event of interest 'invisible'**.
- Interest in inference for an **uncensored population**.
- Estimand should **not depend on censoring aspects**.

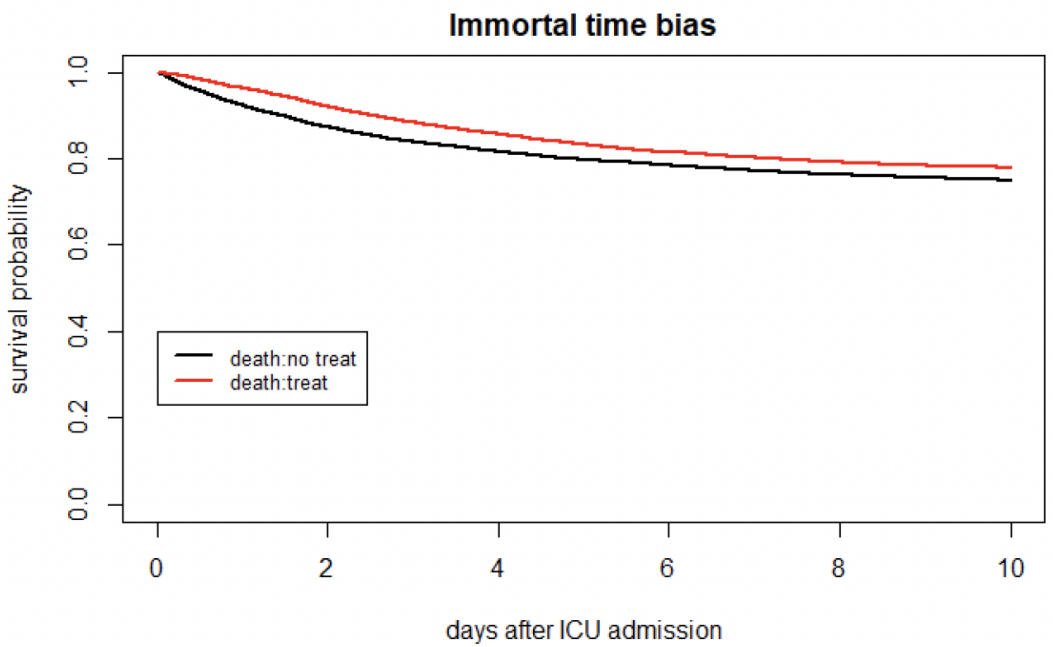
Estimands: how to define time zero?

- In randomized studies: time zero is clear.
  - Start of randomization.
- In observational studies: often less so.
  - Different time scales:
    calendar time vs. start treatment vs. age.
  - Users defined later in follow up $\Rightarrow$ **immortal time bias**.

Immortal time bias: a common made mistake

- Population: patients admitted to ICU
- A treatment without an effect
- Treatment started between day 0 and 5

Ignoring delayed start of treatment yields:

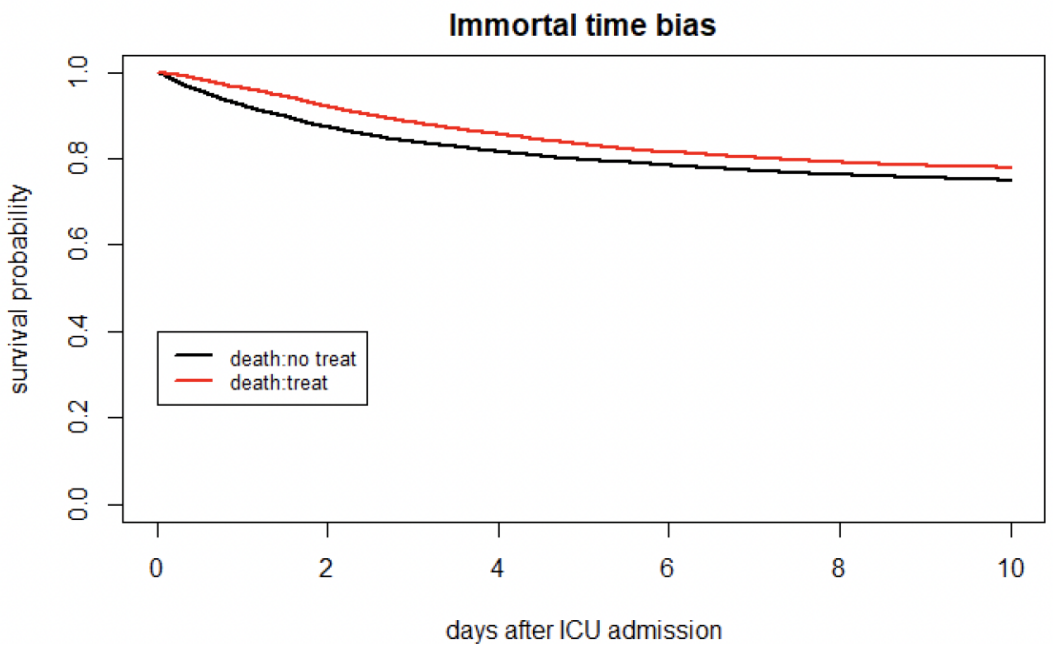


Observed effect (HR=0.88), while no effect was present.

Immortal time bias: a common made mistake

- Population: patients admitted to ICU
- A treatment without an effect
- Treatment started between day 0 and 5

Ignoring delayed start of treatment yields:



Observed effect (HR=0.88), while no effect was present.

Target trials prevent you from making that mistake by mimicking your RCT.

Estimands: scale of contrast

Hazard scale

- Hazard ratio

Estimands: scale of contrast

Hazard scale

- Hazard ratio

Survival/risk scale

- Differences/ratios in survival probabilities at specifc time(s): $P(\tilde{T}^1 > t)$ versus $P(\tilde{T}^0 > t)$.
- Difference in survival curves.
- Difference in median survival time.
- Difference in (restricted) mean survival time: $E(min(\tilde{T}^1, t^*))$ versus $E(min(\tilde{T}^0, t^*))$, with t^* a predefined time horizon.

Estimands: scale of contrast

Hazard scale

- Hazard ratio

Survival/risk scale

- Differences/ratios in survival probabilities at specifc time(s): $P(\tilde{T}^1 > t)$ versus $P(\tilde{T}^0 > t)$.
- Difference in survival curves.
- Difference in median survival time.
- Difference in (restricted) mean survival time: $E(min(\tilde{T}^1, t^*))$ versus $E(min(\tilde{T}^0, t^*))$, with t^* a predefined time horizon.

"Speed" scale

- Estimands based on accelerated failure times

Still a role for hazard ratio? (1)

- It is a **simple summary** when proportional hazards holds.
- In an RCT, it is **protected under the null**.

Still a role for hazard ratio? (1)

- It is a **simple summary** when proportional hazards holds.
- In an RCT, it is **protected under the null**.
- However, over recent years, researchers have **increasingly moved away from reporting hazard ratios**, often choosing restricted mean survival time (RMST) differences as an alternative.

Still a role for hazard ratio? (1)

- It is a **simple summary** when proportional hazards holds.
- In an RCT, it is **protected under the null**.
- However, over recent years, researchers have **increasingly moved away from reporting hazard ratios**, often choosing restricted mean survival time (RMST) differences as an alternative.
  - Issues surrounding **selection bias** due to conditioning on post-baseline variables.(Hernán, 2010)
  - **Non-collpasibility**.
  - More **difficult to interpret**?

Still a role for hazard ratio? (2)

- However, moving to RMST comes with an **information loss**:
  - Prentice and Aragaki (2022) showed that a single averaged hazard ratio is more sensitive to early treatment effect than, for example, the RMST difference.
    - Hazard ratios result in much more powerful tests.

Still a role for hazard ratio? (2)

- However, moving to RMST comes with an **information loss**:
  - Prentice and Aragaki (2022) showed that a single averaged hazard ratio is more sensitive to early treatment effect than, for example, the RMST difference.
    - Hazard ratios result in much more powerful tests.
- Isn't that too much of a sacrifice?
  - It feels like a **large sacrifice for testing**.
  - Estimation?

Thank you for your attention!

E-mail: kelly.vanlancker@ugent.be

References I

EMA (2015). Guideline on adjustment for baseline covariates in clinical trials. Last checked: 2022-05-30.

FDA and EMA (1998). E9 statistical principles for clinical trials. U.S. Food and Drug Administration: CDER/CBER. European Medicines Agency: CPMP/ICH/363/96. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf. Last checked: 2021-02-03.

Hernán, M. A. (2010). The hazards of hazard ratios. *Epidemiology 21*(1), 13–15.

Prentice, R. L. and A. K. Aragaki (2022). Intention-to-treat comparisons in randomized trials. *Statistical Science 37*(3), 380–393.