Re-Thinking treatment effect measure in clinical trials with time-to-event outcomes and competing risks

Tobias Mütze, Novartis Pharma AG, Basel
Stefan Englert, Johnson & Johnson, Germany

17th of June, 2024 – PSI 2024, Amsterdam

Acknowledgment & Disclaimer

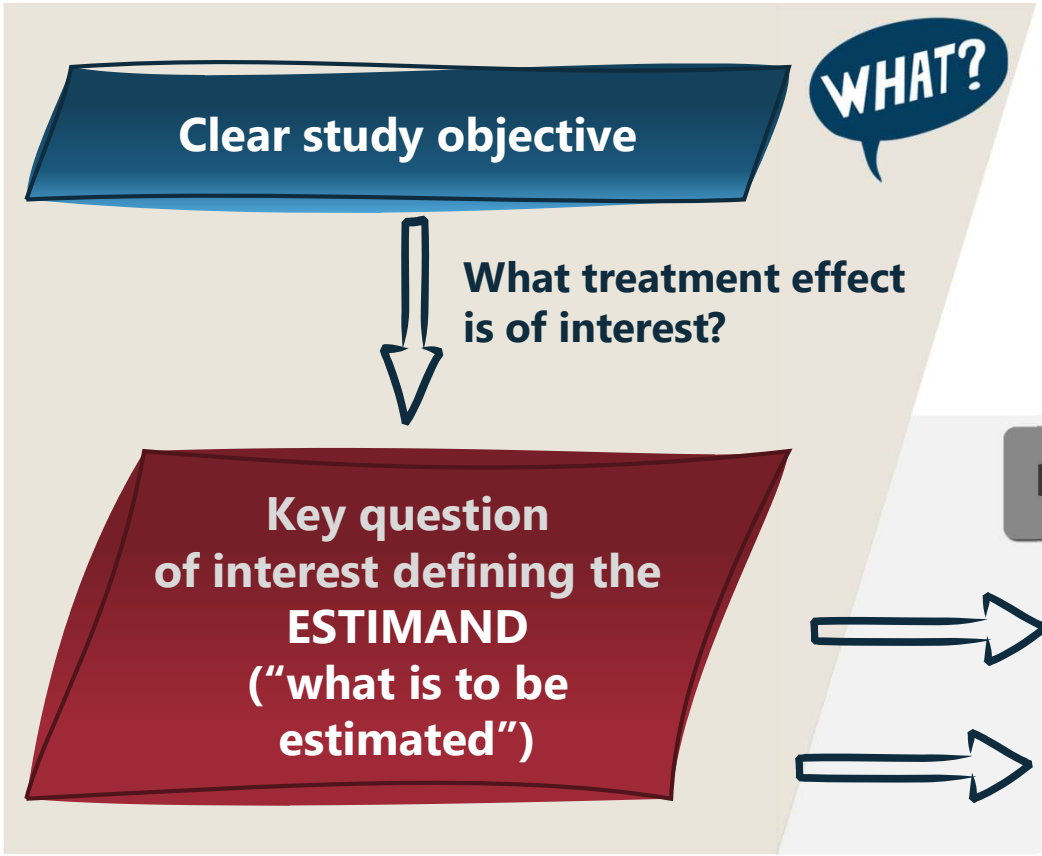
Acknowledgment

We are grateful to the members of the
cross-company working group "*Estimands for time-to-event data*":
Dan Jackson (AstraZeneca), Henrik Ravn (Novo Nordisk), James Bell (Elderbrook Solutions GmbH), Kaspar Rufibach (Roche), Philip Hougaard (Lundbeck), and Vivian Lanius (Bayer AG)
as well as our colleagues for stimulating discussions, contributing thoughts and insights.

Disclaimer

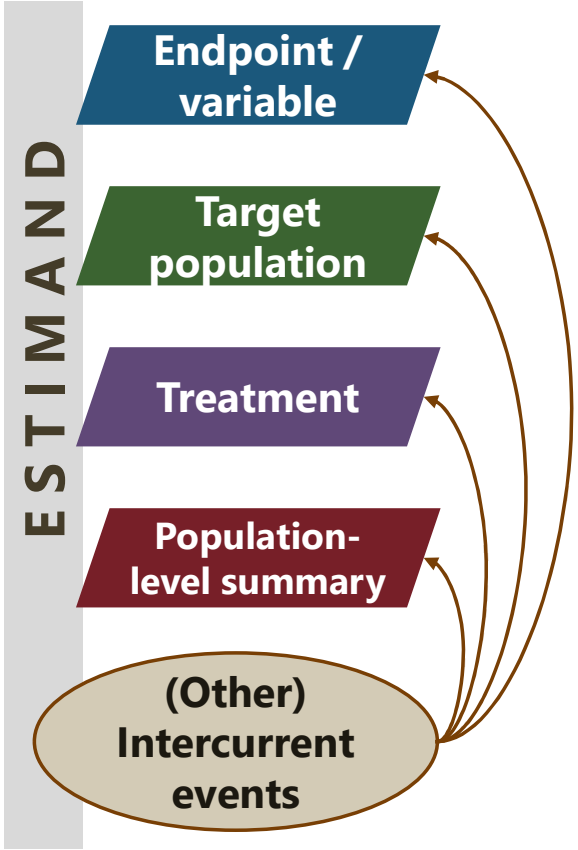
The views and opinions expressed in this presentation are solely those of the presenters and not necessarily those of their employers.

A systematic approach to thinking through study objectives



Estimand: "summarises at a population-level what the outcomes would be in the *same patients under different treatment conditions* being compared". [ICH E9(R1)]

The Estimand as per ICH E9 (R1) – clearly spelled out



The study will compare <**test treatment condition**> with <**reference treatment condition**> in individuals who <**target population**>.

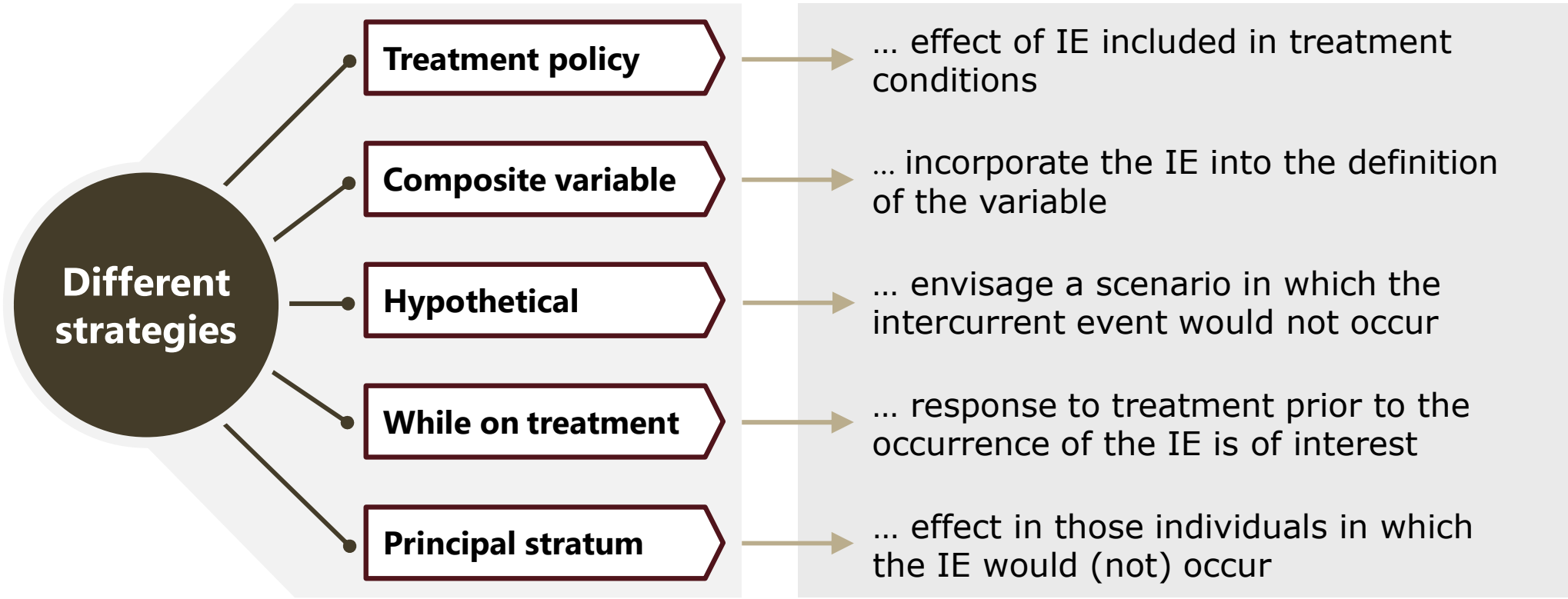
The objective is to <*desired goal/claim**> based on the <**population-level summary measure of treatment effect**> for the <**endpoint/variable**>.

The treatment effect of interest is <high-level description of accounting for ***other*** **intercurrent events**>.

**e.g. 'show superiority'*

Detailed clinical objective template from Bell J, Hamilton A, et al. The detailed clinical objectives approach to designing clinical trials and choosing estimands. Pharmaceutical Statistics. 2021;1–13

How to address intercurrent events (IEs) according to ICH E9(R1)



Pre-Estimands Traditional Thinking in Drug Development

Clinical & project lead

"Interest is in the occurrence of *outcome event* through time."

"We target a relative risk reduction of 25%."

"Given tight timelines, maximum study duration should not exceed 3 years and we'll need ~1.5 years for recruitment."

Statistician

"A time-to-event analysis is needed to account for censoring."

"They believe the hazard ratio is HR=0.75."

"If hazards are not proportional, maybe an RMST analysis should be considered."

Health Authority

"Perform a competing risk analysis."

Current practice often results in ambiguous questions of interest with unclear clinical relevance

- Reverse transfer of standard analysis practices into the estimand framework,
  eg, if analysis censors competing risks then declare it a hypothetical or while-alive strategy
  → Analysis defines the question of interest, not vice-versa

- Timeframe of interest is left
  → Censoring distribution (of mostly event-driven studies) implicitly part of the estimand

- Unclear terminology regarding summary measure especially in clinical trial papers,
  eg, confusing 'risk' and 'hazard', or "event was taken into account as a competing risk"
  → Imprecise communication of trial results

Hazard ratio is treated as solely relevant summary measure

- Prevalent use of the hazard ratio as population-level summary measure
  → Concerns around interpretability & causality
- Assumptions (*eg*, proportional hazards) made on the estimand rather than on the analysis level
  → Clinical question of interest tied to assumptions that cannot be assessed at planning stage
- Strong focus on calculating single effect estimates
  → Isn't primary interest almost always the complete survival/incidence function?

Estimand framework offers clarity around the clinical question of interest, instead discussions (within teams and between sponsor and HAs) focus on analysis and we are missing out on an opportunity for precise treatment effect definitions

Clinical question of interest depends on clinical context & indication

Compared to 'control', how much does assigning 'experimental treatment' …

- decrease the **probability** for event at *or* up to time T^*?
- decrease the **hazard** of event up to time T^*?
- increase the
  - **median time** to event?
  - **expected mean time** to event up to time T^*?
- **accelerate** the occurrence of event up to time T^*?

Summary measure is not one size fits all

- Choice depends on disease/therapeutic area
- Different 'scales' might be relevant for, *eg*, patients, physicians, or payers
  - *Example*: use of time scale via Restricted Mean Survival Time (RMST)

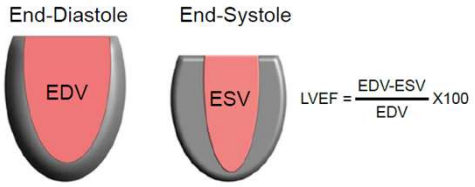
Key questions after …

… *cancer diagnosis*	… *cardiovascular (CV) event (with very low incidence)*
"How long have I got?" → time of high relevance	"What's the risk for a CV death within 1 year?" → the *if* is often more relevant than the *when*

- Time horizon $[0, T^*]$ has to be chosen based on medical justification, (eg, acute setting) rather than design or estimation considerations

Example: Heart failure with reduced ejection fraction

- Heart failure (HF) means that the heart is unable to pump blood around the body properly
- Symptoms typically include shortness of breath, excessive fatigue, and leg swelling
- HF segmented by left ventricular ejection fraction (LVEF)
- HF with reduced ejection fraction (HFrEF): LVEF<40%



- Primary outcome: Composite of cardiovascular (CV) death and HF hospitalization (HHF)
- Key intercurrent/competing event: Non-Cardiovascular (non-CV) death

What clinical questions could be of interest in the HFrEF example?

Compared to 'control', how much does assigning 'experimental treatment' decrease the …

- … probability for CV death at *or* up to time T^* …
  - … in a setting where patients can die of non-CV death?
  - … in patients who would not die of non-CV death regardless of their treatment assignment?
  - … if patients would not die of non-CV death?
- … (cause-specific) hazard for CV death up to time T^* in those who are alive?
- … probability for all-cause death at *or* up to time T^*?

Some advice on defining estimands for time-to-event outcomes

- Definition of **target population** and **treatment conditions** as per clinical context, not specific to time-to-event outcomes
- **Variable**: status of the individual over the course of time (from treatment assignment) w.r.t. to the outcome event of interest and competing risks, if applicable
- Important to align choice of **summary measure** – *univariate or multivariate* – appropriately, characterizing the distribution of the time-to-event data with (clinical) question(s) of interest, including timeframe of clinical relevance and interest
- Regarding **intercurrent events** distinguish
  - the occurrence of competing risks (or terminal events) and
  - other intercurrent events where the individual is still at risk of the event of interest

Example revisited:
Estimands for heart failure with reduced ejection fraction

Question of interest in a competing risk setting

Compared to ‘control’, how much does ‘drug’ decrease the probability for CV death up to 2 years in HFrEF patients who can also die of non-CV causes?

Example revisited:
Estimands for heart failure with reduced ejection fraction

Question of interest in a competing risk setting

Compared to 'control', how much does 'drug' decrease the probability for CV death up to 2 years in HFrEF patients who can also die of non-CV causes?

Estimand

- Population: Patients diagnosed with heart failure with reduced ejection fraction
- Treatment: 'drug' vs 'control'
- Variable: State: 'event free' ($X_t = 0$), 'dead of CV cause' ($X_t = 1$), 'dead of non-CV cause' ($X_t = 2$) at each time t since baseline
- Summary measure: Probability to have died of CV cause at time $t \in [0, 2]$, *eg*, $P(X_t = 1)$ with between-group comparison based on difference/ratio
- Intercurrent events: Non-CV related death is a competing risk and included in the variable definition.

ICH E9 (R1) Strategies & competing event setting

- In drug development: precisely describe the clinical question of interest
- *One* possibility to structure such a description is using the attributes (population, treatment, variable, summary measure) as outlined in ICH E9(R1)
- Should we use any of the five intercurrent event strategies from ICH E9(R1) to describe how to incorporate the competing events into the attributes?
  - Treatment policy, hypothetical and while-on-treatment have been used in literature without consensus, with some authors suggesting that it requires a new strategy
  - Probably not sensible and might lead to more confusion
  - Can be described in the variable definition

Conclusions and Discussion

- Current practices in many clinical trials with time-to-event data and competing risks are not aligned with estimand thinking
- Issues outlined (*eg*, regarding summary measure) also occur when there is no competing risk (*eg*, in trials focusing on overall survival)
- Applying the estimand thinking process to time-to-event and competing risk data adds value → separate the question of interest from the analysis (estimand first, analysis second)
- Estimands might have impact on design choice (minimum follow-up time, …)
- Strategies outlined in the ICH E9(R1) addendum have limitations in the competing risk setting and competing risks should be included and acknowledged in the variable definition.

References

- ICH E9 (R1). ICH Harmonised Guideline. Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials E9(R1). International Conference on Harmonisation [online]. Available at https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

- Bell J, Hamilton A, et al. The detailed clinical objectives approach to designing clinical trials and choosing estimands. Pharmaceutical Statistics. 2021;1–13.

Thank you!

Contact

Tobias Mütze, Novartis Pharma AG, Basel — tobias.muetze@novartis.com
Stefan Englert, Johnson & Johnson, Germany — senglert@its.jnj.com

Abstract

In randomized clinical trials with a time to event outcome, the hazard ratio is still the most common effect measure. Post-randomization (i.e., intercurrent) events are often addressed through censoring without explicitly discussing or stating the underlying clinical question of interest. Alternative summary measures, especially on a probability scale or time scale, are rarely considered in clinical trials despite being seemingly easier to interpret and potentially more meaningful to patients and practitioners.

In this talk we will present the status of ongoing discussions on estimands for clinical trials with time-to-event outcomes and competing risks. In detail, we will discuss what key clinically meaningful questions of interest are when measuring the effect of an intervention through a time-to-event endpoint. We will reflect on the interpretation of various summary measures, the role of causality when defining an estimand in a clinical trial, and on how the choice of the estimand affects the design of a trial with a time-to-event endpoint. We will elaborate on the practicalities of summarizing the effect of treatment through a single number in a time to event setting and discuss separating testing and estimation. We will also propose a new approach to embed competing risks within the framework that we believe is helpful for describing estimands in a competing risk setting.