Estimands, target trial emulation, and use of external control data

Kaspar Rufibach
Merck KGaA, Darmstadt
UZH Symposium of Causal Inference in the Health Sciences, 20th March 2025

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 26 January 2024
DOI 10.3389/fphar.2024.1223858

OPEN ACCESS

EDITED BY
Dalia M. Dawoud,
National Institute for Health and Care Excellence, United Kingdom

REVIEWED BY
Michelle Casey,
Pfizer, United States
Enrico Capobianco,
Jackson Laboratory, United States

*CORRESPONDENCE
Letizia Polito,
letizia.polito@roche.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 16 May 2023
ACCEPTED 11 January 2024

Applying the estimand and target trial frameworks to external control analyses using observational data: a case study in the solid tumor setting

Letizia Polito[1*†], Qixing Liang[2†], Navdeep Pal[3], Philani Mpofu[2], Ahmed Sawas[2], Olivier Humblet[2], Kaspar Rufibach[1] and Dominik Heinzmann[1]

[1]Product Development Data Sciences, F Hoffmann-La Roche Ltd., Basel, Switzerland, [2]Flatiron Health, Inc., New York, NY, United States, [3]Genentech, Inc., San Francisco, CA, United States

Polito et al. (2024).

RCT not feasible - emulate it.

Combine estimands and target trial framework.

Structured way of thinking.

Makes assumptions transparent.

Goal: answer causal question about efficacy and safety of a health-related intervention.

Gold standard: RCT.

Answers question under least number of assumptions.

But what if:

RCT is not feasible, ethical, or timely.

Causal question not of primary importance.

Want to accurately analyze existing data.

...

Decisions need to be made even w/o RCT – maintaining status quo is also a decision!

Target trial framework

> *Causal inference from large observational databases (big data) can be viewed as an attempt to emulate a randomized experiment – the target experiment or target trial – that would answer the question of interest.*

Hernan and Robins (2016)

Target trial framework

Target trial framework elements	
Eligibility criteria	
Treatment strategies	
Assignment procedures	
Outcome	
Follow-up period	
Causal contrast of interest	
Analysis plan	

Extends PICO.

Target trial framework

Target trial framework elements	Estimand attributes
Eligibility criteria	Population
Treatment strategies	Treatment
Assignment procedures	
Outcome	Variable of interest
Follow-up period	
	Intercurrent events and their handling
Causal contrast of interest	Population-level summary
Analysis plan	

Extends PICO.

Scientific question:

Can observational data emulate control arm of RCT?

View towards use of external controls.

Case study

Broad scientific question:

> *Is there a difference in OS between patients with metastatic NSCLC receiving front-line platinum-based chemotherapy (pCT) in pivotal trials vs. patients with metastatic NSCLC who received front-line pCT as part of routine care?*

Precise enough? Heterogeneity in 2nd line treatments. Iterate to:

> *Is there a difference in OS between patients with metastatic NSCLC receiving front-line pCT in pivotal trials vs. patients with metastatic NSCLC who received front-line pCT as part of routine care,* **had patients not received subsequent therapy***?*

Risk: Heterogeneity in subsequent therapies across treatment settings may introduce complexities in estimating causal treatment effects for OS and ultimately complicate interpretation.

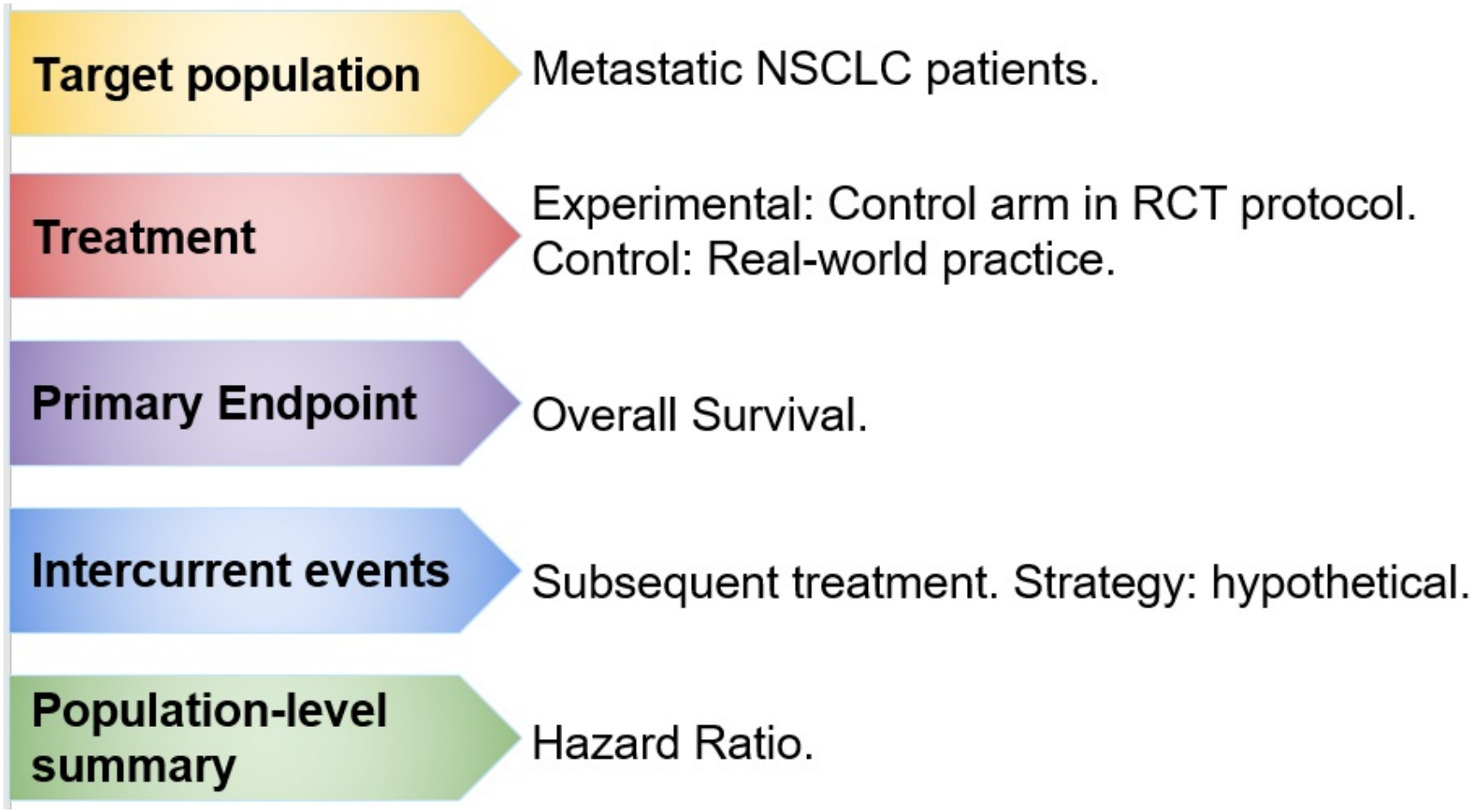
Data:

Control arms of three RCTs.

Flatiron EHR-derived data.

Target trial assumes we would have randomized to the two sources.

Estimand


Target population
Metastatic NSCLC patients.
Treatment
Experimental: Control arm in RCT protocol.
Control: Real-world practice.
Primary Endpoint
Overall Survival.
Intercurrent events
Subsequent treatment. Strategy: hypothetical.
Population-level summary
Hazard Ratio.

Population

Population target trial: $\approx$ I/E criteria of RCTs.

Emulation:

- Align Flatiron cohort as much as possible to that.
- Flatiron: US only. RCTs: global. **Risk**.
- Exclude patients with missing covariate values. **Risk of selection bias**.

Backbone CT

Backbone CT target trial: Nab-paclitaxel and paclitaxel.

Emulation:

- RCTs offered both.
- Decision to include paclitaxel-treated patients. **Limit treatment-assignment bias**.

Risk of bias:

- RCT patients received care according to protocol.
- Flatiron cohort patients received routine clinical care.

Start / end of follow-up

Target trial:

- Start of follow-up: when eligibility met, i.e. treatment is assigned.
- End of follow-up: relevant clinical cutoff date.

Emulation for Flatiron cohort:

- Start of follow-up: first enrollment start date of the three RCTs.
- End of follow-up: latest clinical cutoff date of the three RCTs.
- All patients with cycle 1 dose 1 within this range.
- Assumption such that this approach does not introduce **immortal bias**:
  - No reasons other than death for a patient to not initiate treatment once assigned to treatment.
  - Death unlikely between assignment and start of treatment: short interval & mNSCLC no rapid course in first line.
- Flatiron cohort patient follow-up truncated at 21m $\Rightarrow \approx$ RCT maximal follow-up.

Endpoint, ICEs, summary

Endpoint:

- Target trial: OS.
- Emulation: validity of rwOS established.

ICEs:

- Target trial: ICE subsequent therapy, **hypothetical** strategy.
- Emulation: same.

Summary:

- Target trial: hazard ratio.
- Emulation: hazard ratio.

Treatment effect of interest

Average treatment effect on the treated (ATT).

Treatment effect difference

- of using front-line chemotherapy in a clinical trial
- versus in clinical practice,
- where target population is defined by population targeted by three RCTs.

Assignment

Target trial: Participants randomly assigned to RCTs or Flatiron cohort.

Emulation:

- Weighting observations by **inverse probability of treatment** (IPTW).
- Assumptions:
  - Assignment explained through Age, gender, race, metastatic tumor type, time from initial diagnosis to index date, smoking history, histology, and treatment type.
  - Consistency, conditional exchangeability, positivity, and correct model specification.
- Positivity: non-zero probability to end up in RCTs or Flatiron cohort. Not met, but alignment of I/E criteria + propensity scoring.

Estimation of **average effect on the treated** (ATT):

- Propensity scores estimated using multiple logistic regression = P(assigned to RCTs | confounders).
- RCT patients get weight 1.
- Weights Flatiron cohort: odds of being treated in the clinical setting $\Rightarrow$ IPTW-ATT weights.

Non-random censoring

Non-random censoring at ICE:

- **Inverse probability of censoring weighting** (IPCW).
- Patients **artificially censored** at time of receipt of first second-line treatment.
- Use IPCW to estimate weights for follow-up information for remaining patients using both baseline and time-varying variables.
- Fit Cox model within each arm to estimate probability of not being censored by time t.
- IPCW weights = inverse of conditional probability of not being censored.

Primary result: HR = 0.94 with 95% CI from 0.77 to 1.13

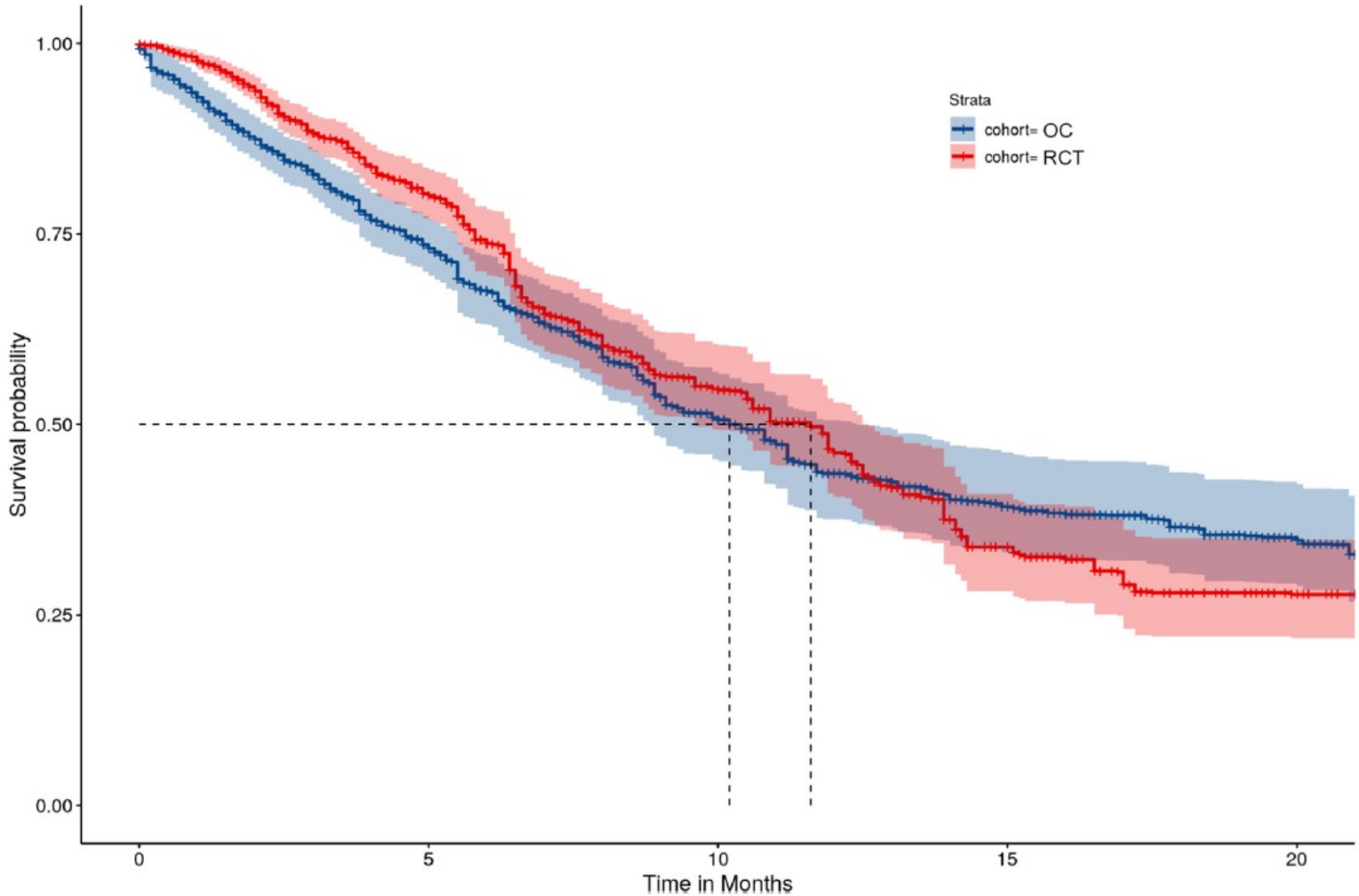


FIGURE 2
IPTW-ATT-IPCW weighted Kaplan-Meier curves.

RCT not feasible - emulate it.

Combine estimands and target trial framework.

Structured way of thinking.

Makes assumptions transparent.

Where does causal inference appear in drug development?

Causal inference in drug development

- Clear **definition** of intervention effect of interest. ICH E9(R1) addendum, target trial framework.
- Formal **justification of randomization**.
- Trade-off randomization $\Leftrightarrow$ assumptions.
- **Transparency of assumptions** needed to make causal claim, especially in presence of **post-baseline** events.
- **Postbaseline** confounders – e.g. principal stratification.
- Structured way to think about:
  - Validity of RWD, external controls to answer causal question.
  - **Generalizability**: extending causal effect from RCT to RCT's original target population.
  - **Transportability**: extending causal effect from RCT to distinct population.

Thank you for your attention.

kaspar.rufibach@merckgroup.com

Slides can be downloaded on

www.kasparrufibach.ch

References I

- Hernan, M. A. and Robins, J. M. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol* **183** 758–764.
- Polito, L., Liang, Q., Pal, N., Mpofu, P., Sawas, A., Humblet, O., Rufibach, K. and Heinzmann, D. (2024). Applying the estimand and target trial frameworks to external control analyses using observational data: a case study in the solid tumor setting. *Frontiers in Pharmacology* **15**. `https://www.frontiersin.org/articles/10.3389/fphar.2024.1223858`

R version and packages used to generate these slides:

R version: R version 4.4.2 (2024-10-31 ucrt)

Base packages: stats / graphics / grDevices / utils / datasets / methods / base

Other packages: prodlim

This document was generated on 2025-03-17 at 20:07:02.