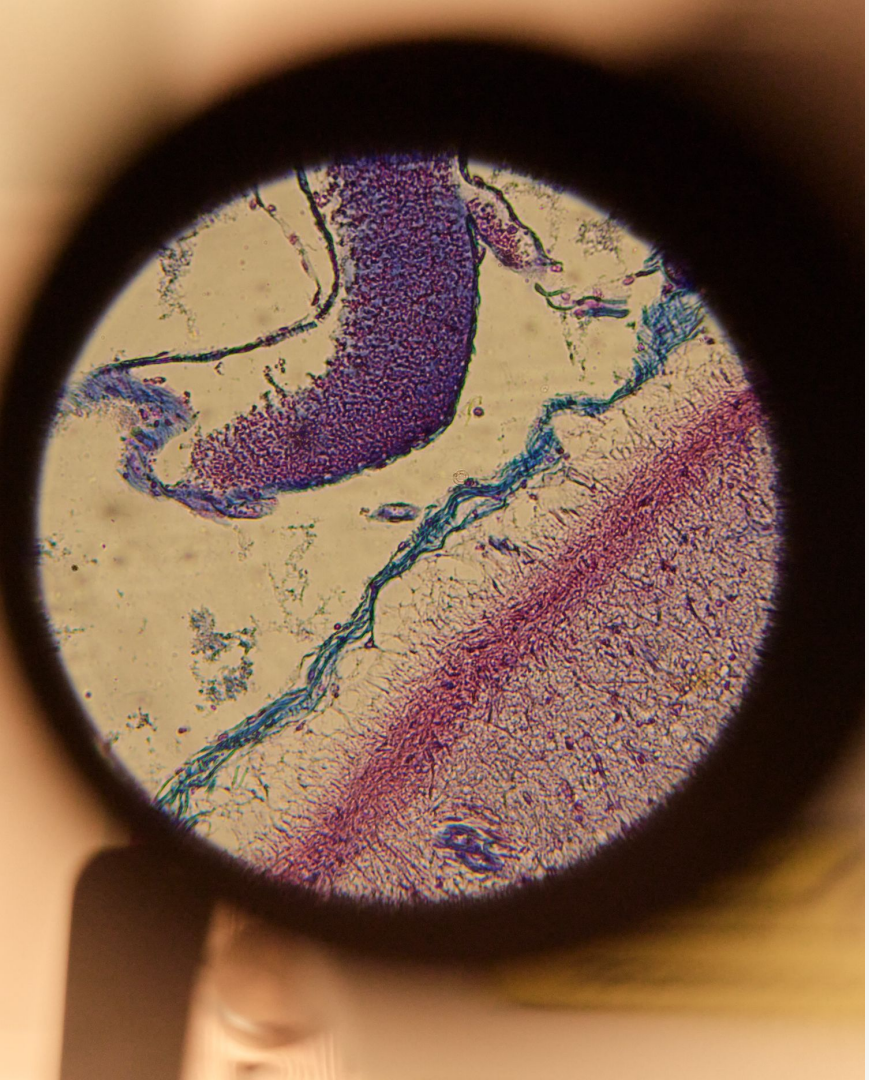


Applying the Target Trial Emulation Framework When Time 0 is Ambiguous

a case study and work in progress

Isaac Gravestock
Real World Data Science

27 October 2023 | public use

RDR Number: 1127921

COI and acknowledgments

I am an employee of F. Hoffmann-La Roche.

Views expressed in this presentation are my own and do not necessarily reflect the views of F. Hoffmann-La Roche.

This is joint work with Matt Secrest, Christophe Tchakoute, Thibaut Sanglier and other colleagues.

Thanks to the statistical and clinical colleagues for their advice on this project.

This is work in progress so the actual indication and drug of interest were removed

Table of contents

Introduction

Target Trial Framework

The **Target Trial Framework** provides a systematic design process for observational studies based on a hypothetical clinical trial.

By using a clinical trial template to design an observational study, we address many important criteria and avoid some common errors.

- Inclusion/Exclusion criteria
- Treatment arms
- Randomisation
- Outcome
- Follow-up Period

To apply the framework, we first describe a trial that would address our clinical question.

This hypothetical trial may be impractical, infeasible, or unethical.

Do not consider the observational study at this stage!

Target Trial Emulation

With a complete Target Trial protocol, we attempt to translate the elements into the observational data setting.

Some can be translated directly, some will need to be adapted.

As we describe these adaptations, we will be faced with various challenges and possible biases:

- Selection bias
- Immortal time bias
- Indication bias
- Unmeasured confounding
- Missing data

Comparing Initiators and Non-Initiators

A Common Research Question With At Least Two Major Challenges

How does *Treatment strategy* (**treatment X**) compare with alternate treatment strategy (**No Treatment X**)?

- In a clinical trial, we can use a placebo.
- In observational data, one might think we could simply use patients who don't receive **Treatment X**.

Why was a patient not treated with the indicated treatment?

- Typical setting for indication bias in observational data*
- In some setting indication bias can be reasonably mitigated

When do we start following patients?

- In a clinical trial, this is the date they received treatment or placebo
- In observational data, we have a problem:

What is the start date of **No Treatment?**

How can we define *equivalent* start times for both arms without creating *immortal time bias*?

*Bosco JL, Silliman RA, Thwin SS, Geiger AM, Buist DS, Prout MN, Yood MU, Haque R, Wei F, Lash TL. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. J Clin Epidemiol. 2010 Jan;63(1):64-74. doi: 10.1016/j.jclinepi.2009.03.001. Epub 2009 May 19. PMID: 19457638; PMCID: PMC2789188.

Ambiguous Time Zero

We can frame this problem in two ways:

1. Treatment assignment is clear (at a given time), but start of follow up is ambiguous
2. Start of follow-up is clear (eg diagnosis date) but treatment assignment is ambiguous at this time

These two conceptualisations lead to different modelling strategies:

1. **Sequence of target trials:**

   Run multiple trials starting periodically over time. Enroll all patients who meet the eligibility criteria at the beginning of each trial and assign them to whichever treatment group they are compatible with.

2. **Clone-censor-weighting:**

   Create copies of patients and assign one to each of the treatment strategies. Censor one of the clones when the observed treatment is incompatible with the assigned treatment

In both methods, we need to to account for the non-independence caused by this wizardry.

Depending on the causal estimand of interest, we will apply time-varying weights as well.

Our Target Trial

Setting

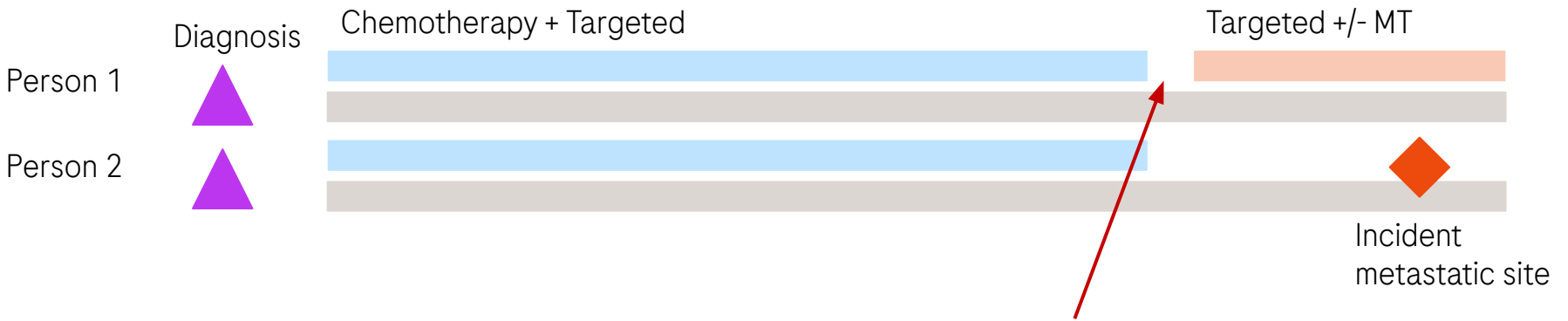
We are interested in a certain kind of **Metastatic Cancer.**

This cancer can be treated with chemotherapy and targeted therapy.

Following a successful first course of treatment, maintenance therapy (**MT**) may be initiated.

We have a hypothesis that MT mechanism of action has an impact on the incidence of metastases at a certain body site.

What is the effect of initiating MT (vs not) on the risk of incident metastasis in site X at 3 years?



Because there is an incompatibility between the chemo and MT, there is a gap between these therapies.

Setting

We are interested in a certain kind of **Metastatic Cancer.**

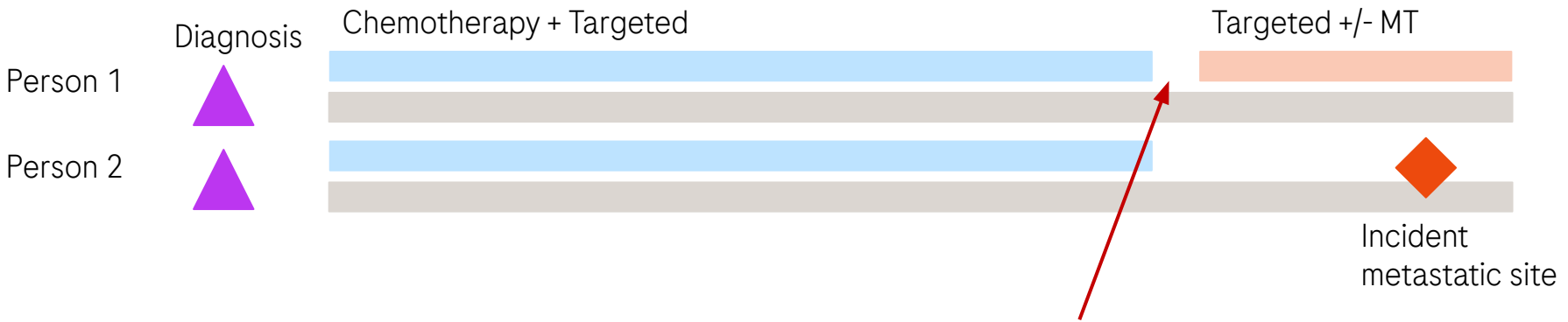
This cancer can be treated with chemotherapy and targeted therapy.

Following a successful first course of treatment, maintenance therapy (**MT**) may be initiated.

We have a hypothesis that MT mechanism of action has an impact on the incidence of metastases at a certain body site.

~~What is the effect of initiating MT (vs not) on the risk of incident metastasis in site X at 3 years?~~

What is the effect of initiating MT (vs not) on the risk of incident metastasis in site X or death at 3 years?



Because there is an incompatibility between the chemo and MT, there is a gap between these therapies.

Target Trial Design

Inclusion/Exclusion Criteria

- Specific cancer diagnosis with confirmed biomarkers
- 18 or older at diagnosis
- Received chemotherapy + targeted therapy
- Has not received MT in the metastatic setting
- No metastases at the body site of interest

Treatment Strategies

- MT should be initiated within 16 weeks of end of chemotherapy
- No MT should be initiated within 16 weeks of end of chemotherapy

Treatment Assignment

Patients will be randomly assigned to a treatment strategy at the end of their chemotherapy

Target Trial Design

Follow Up

Patients will be followed from time of treatment assignment until death, date of last contact, or administrative end of study (March 2023)

Outcome

Metastases at site of interest or death from any cause up to 3 years (metastatic-site-free-survival)

Causal Contrast

Intention-to-treat effect

Per-protocol effect

Target Trial Design

Analysis

Intention-to-treat analysis:

- Comparison of Kaplan-Meier survival curves (difference in 1-year survival, median survival)
- Pooled logistic regression model with time and MT as covariates to estimate: **hazard ratio** (HR), **risk ratio** (RR), **risk difference** (RD) effects across treatment strategies on the outcomes of interest.

Per-protocol analysis:

Same as intention-to-treat analysis, except

- Patients are censored if they do not comply with their assigned treatment strategy
- Time-varying inverse probability of censoring weights are used to adjust for the potential selection bias due to this artificial censoring

Target Trial Emulation

Thoughts on the Target Trial Protocol

Generally, we need to add to every line in the protocol “Patient has record of …”

We want to follow patients from time of **treatment assignment** until metastasis or death

What does this mean in the observational data? We only have records of

- Chemotherapy treatment dates
- MT dates
- Metastasis and death dates

In the target trial, we planned to assign patients to treatment strategies at the end of chemotherapy, but given the gap between chemo and MT we can’t know until sometime in the future which they will receive

⚠️ **ambiguity at time zero!** /!\ **immortal time bias!**

Translating the Protocol

What is required and reasonably achievable ?

Inclusion/Exclusion Criteria

- Specific cancer diagnosis codes and biomarker test results recorded
- 18 or older at diagnosis
- ≥ 5 chemotherapy cycles + targeted therapy recorded
- No record of receiving MT in the metastatic setting
- No record of metastases at the body site of interest

Treatment Strategies

- MT initiated within 16 weeks of last chemotherapy record
- No MT initiated within 16 weeks of last chemotherapy record

Treatment Assignment

Patients are classified to the treatment strategy compatible with their observed data at last chemotherapy

Translating the Protocol

Follow Up

Patients will be followed from time of treatment assignment until death, date of last contact, or administrative end of study (March 2023)

Outcome

Metastases at site of interest or death from any cause up to 3 years (metastatic-site-free-survival)

Causal Contrast

~~Intention-to-treat effect~~

Per-protocol effect(s)

We don't estimate an intention-to-treat effect as we need to apply the **clone-censor-weighting**. This means a clone of each patient is in each group, therefore an intention-to-treat effect must be 0.

Translating the Protocol

Analysis

Per-protocol analysis:

Clone

- duplicate all patients at baseline and assign to all different treatment strategies.

Censor

- censor patients as they deviate from assigned treatment strategies

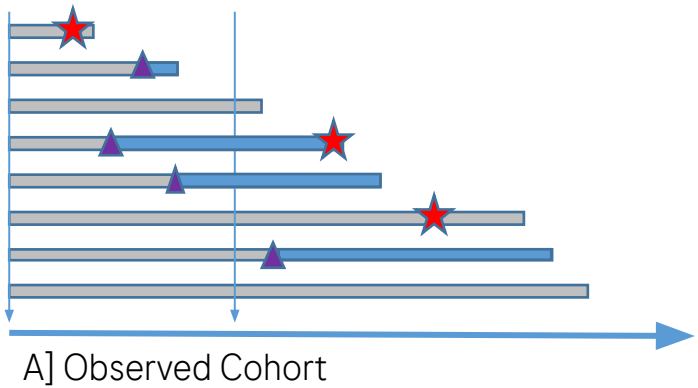
Weight

- estimate probability of not being censored over time using baseline covariates and calculate inverse probability weights

Estimate

- Using weighted data fit pooled logistic regression model with time and MT as covariates to estimate: **hazard ratio** (HR), **risk ratio** (RR), **risk difference** (RD) effects across treatment strategies on the outcomes of interest.

Cohort



A] Observed Cohort

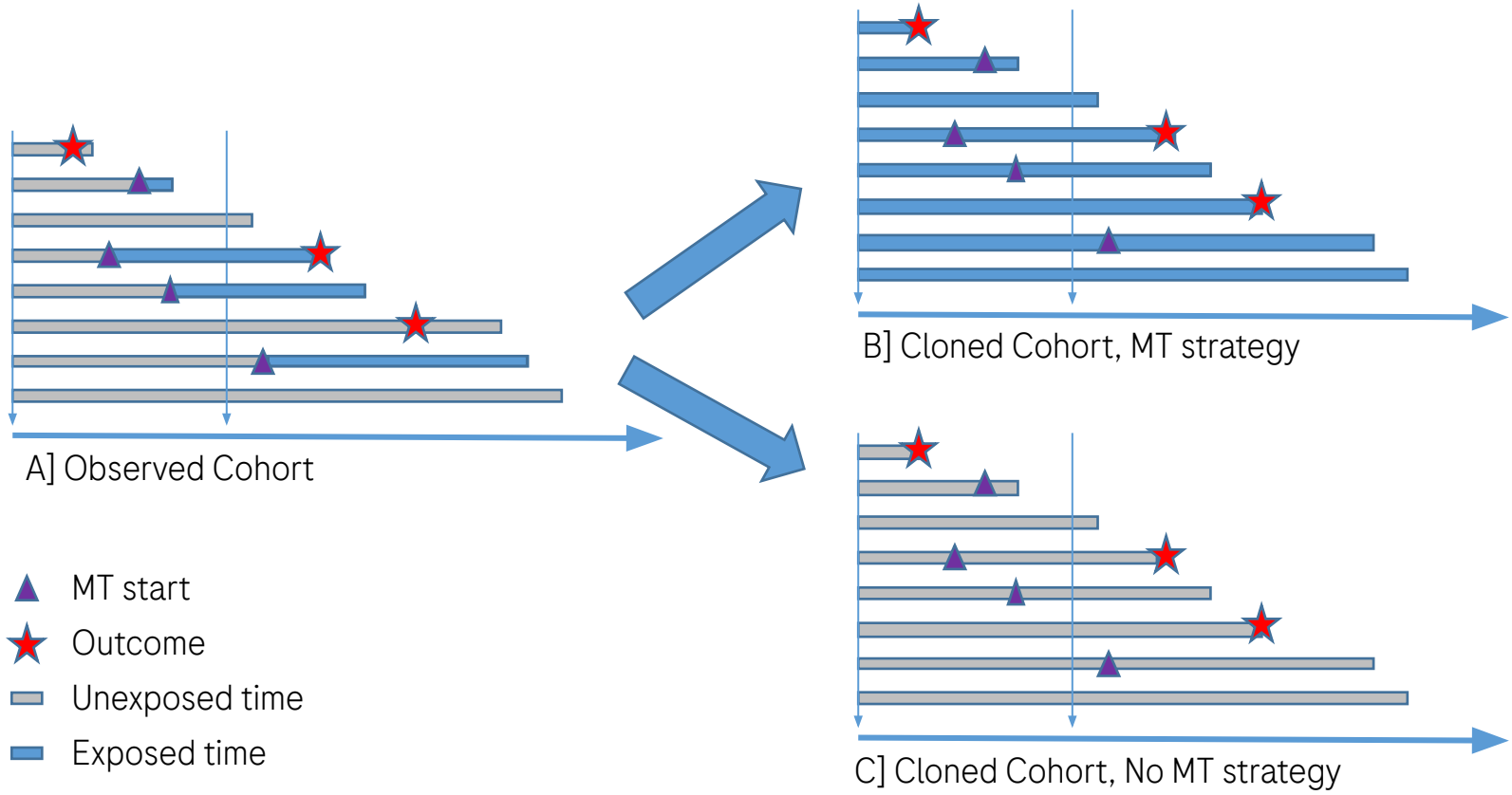
MT start

Outcome

Unexposed time

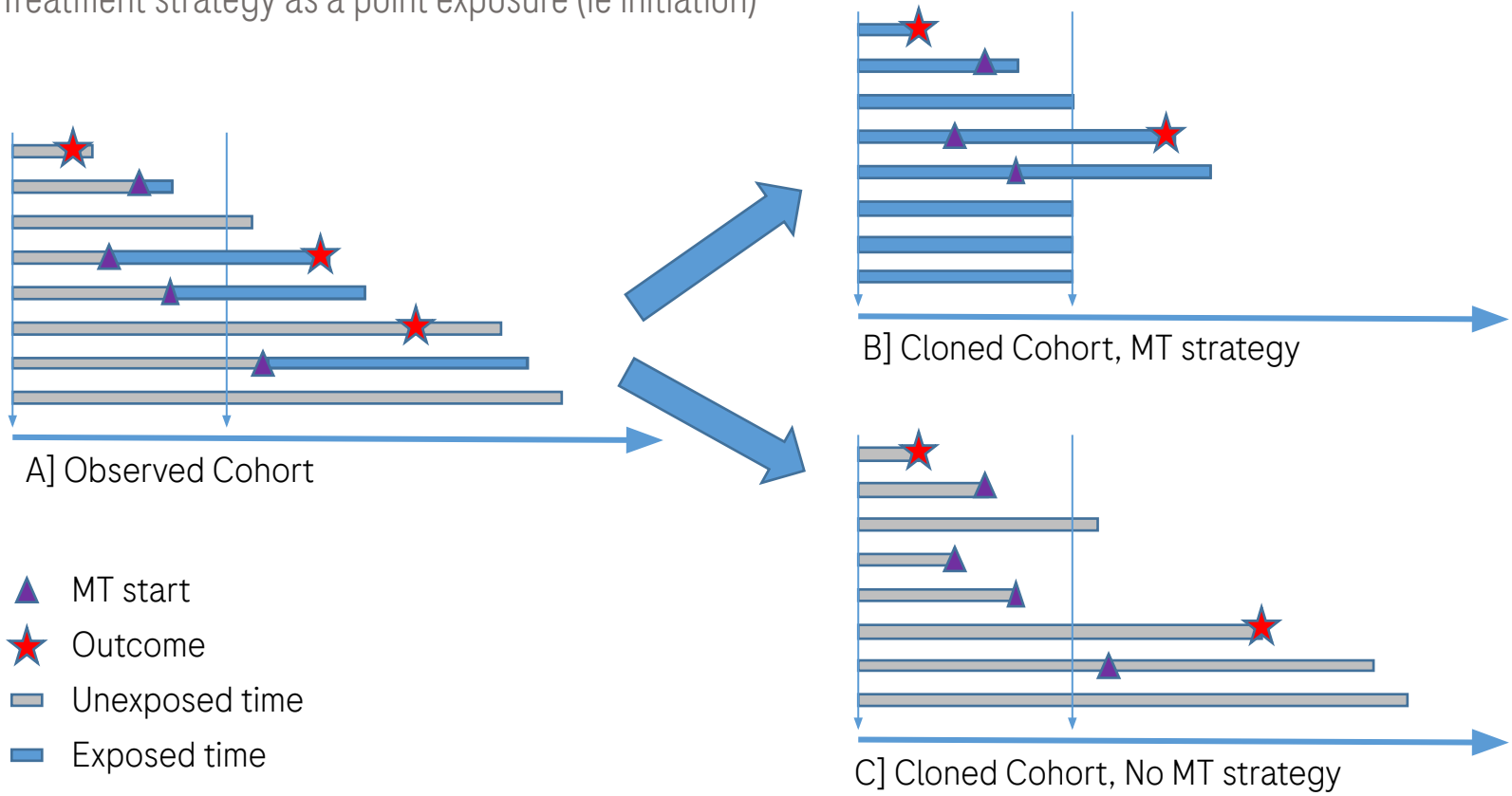
Exposed time

Cloning


A] Observed Cohort
B] Cloned Cohort, MT strategy
C] Cloned Cohort, No MT strategy
MT start
Outcome
Unexposed time
Exposed time

Censoring

Treatment strategy as a point exposure (ie initiation)



Weighting

By cloning the patients, we have removed any confounding at baseline

The trade-off is that the artificial censoring may lead to selection bias

To account for the artificial censoring that we construct **inverse probability of censoring weights** (IPCW)

We calculate the probability of not being censored over time adjusted for baseline covariates.

$$W_{t,i} = (1 - P(C \mid t, X_i))^{-1}$$

Where the probability of being censored for the No MT arm is:

$P^{No\ MT}(C \mid t, X) \sim t + age + cancer\_stage\_at\_diag + tumor\_biomarkers + n\_metastatic\_sites + time\_to\_metastatic + prev\_therapy$

and for the MT arm:

$P^{MT}(C \mid t, X) \sim I_{(t<16w)} + age + cancer\_stage\_at\_diag + tumor\_biomarkers + n\_metastatic\_sites + time\_to\_metastatic + prev\_therapy$

Analysis Models

We can now fit our pooled logistic regression model for the outcome metastasis-free survival

We combined the cloned patients into one data set and use generalised linear model function with the weights we have calculated.

Due to the cloning and weighting, the variances from the standard logistic regression model will not be correct.

We use a non-parametric bootstrap to get valid confidence intervals.

We sample patients before cloning and apply the cloning-censoring-weighting and analysis steps and then combine these results.

Further Challenges & Solutions

Indication Bias

Leverage trial data with the relevant target population and for which MT use was not allowed

In the observational data, patients are not randomised to the treatment groups.

Therefore there may be good reasons why a patient did or did not receive this therapy.

If these reasons are not captured in the data and are related to the outcome, we have a bias.

This is a risk, because this follow on therapy has become fairly standard in recent times, so patients who do not receive it may have poor prognosis.

To address this, we will use data from a clinical trial before this therapy was believed to be effective and its use is excluded by the trial protocol.

/!\ not a perfect solution as enrolment could lead to selection bias (unlikely in our case for clinical reasons)

Competing Risks

Due to the severe nature of cancer, we expect that over a few years of follow-up there will be a non-ignorable number of patient deaths, preventing the observation of incident metastasis at the site of interest.

Can we ignore these and treat death as censoring?

We need to consider death as a competing risk. To calculate a causal treatment effect that is easy to interpret is not straight forward*

We will use a composite endpoint (incident site of interest or death) at x months

*Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA. A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med. 2020 Apr 15;39(8):1199-1236. doi: 10.1002/sim.8471. Epub 2020 Jan 27. PMID: 31985089; PMCID: PMC7811594.

Negative Control Outcomes (with a twist)

Is the treatment effect specific to the site of interest (according to our biological hypothesis)?

We propose estimating the treatment effect on alternate metastatic sites with a known different biology

We claim:

- the biological hypothesis is supported if the effect of MT on the sites of interest >> effect of MT on alternate sites
- the biological hypothesis is not supported if the effect of MT on the sites of interest ≈ effect of MT on alternate sites then the biological

→ Of note due to the prior use of chemotherapy and the sustained use of targeted therapy the effect of MT on alternate sites is expected to be modest

/!\ Usually negative controls are used to detect unmeasured confounding, here they will be used to generate more evidence on the “site specific effect” of MT

Conclusions

- Challenging research questions!
- We believe t0 issue will be substantially mitigated (how much did we improve vs. traditional landmark analysis?)
- TTE and estimand frameworks allow us to design more robust studies with more clearly disclosed assumptions
- Lots of tools available to mitigate risk but their deployment can be challenging
  - **TrialEmulation** R package for sequence of target trials on CRAN, CCW to be implemented!
- Competing risk and estimands pose specific challenges
- Causal contrast vs. ICH9 addendum estimands, divergences and alignment in between RWD and RCT populations… more collaboration and education needed

Ressources

Competing risks:

Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA. A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med. 2020 Apr 15;39(8):1199-1236. doi: 10.1002/sim.8471. Epub 2020 Jan 27. PMID: 31985089; PMCID: PMC7811594.

Clone censor weighting:

Petito LC, García-Albéniz X, Logan RW, et al. Estimates of Overall Survival in Patients With Cancer Receiving Different Treatment Regimens: Emulating Hypothetical Target Trials in the Surveillance, Epidemiology, and End Results (SEER)–Medicare Linked Database. *JAMA Netw Open.* 2020;3(3):e200452. doi:10.1001/jamanetworkopen.2020.0452

Per-protocol effects:

Murray, E. J., Caniglia, E. C., & Petito, L. C. (2021). Causal survival analysis: A guide to estimating intention-to-treat and per-protocol effects from randomized clinical trials with non-adherence. Research Methods in Medicine & Health Sciences, 2(1), 39-49.
https://github.com/eleanormurray/CausalSurvivalWorkshop_2019

Sequence of target trials:

Danaei G, Rodríguez LAG, Cantero OF, Logan R, Hernán MA. Observational data for comparative effectiveness research: An emulation of randomised trials of statins and primary prevention of coronary heart disease. *Statistical Methods in Medical Research*. 2013;22(1):70-96. doi:10.1177/0962280211403603

García-Albéniz, X., Hsu, J. & Hernán, M.A. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol* 32, 495−500 (2017). https://doi.org/10.1007/s10654-017-0287-2

Doing now what patients need next